Title
A Multi-Dimensional View of the Galactic Contribution to Reionization based on Rest-UV Spectroscopic Probes at z ~ 3 - 5

Author
Pahl, Anthony

Publication Date
2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

A Multi-Dimensional View of the Galactic Contribution to Reionization based on Rest-UV Spectroscopic Probes at $z \sim 3 - 5$

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Astronomy and Astrophysics

by

Anthony Joseph Pahl

2023

ABSTRACT OF THE DISSERTATION

A Multi-Dimensional View of the Galactic Contribution to Reionization based on Rest-UV Spectroscopic Probes at $z \sim 3-5$

by

Anthony Joseph Pahl
Doctor of Philosophy in Astronomy and Astrophysics
University of California, Los Angeles, 2023
Professor Alice E. Shapley, Chair

While star-forming galaxies are likely the dominant force behind cosmic reionization, the relative contributions of different galaxy populations remains uncertain. Understanding how ionizing photons are produced and subsequently escape from these galaxies in the early Universe is crucial for reducing such uncertainties in existing models. In this dissertation, I conduct four observational studies of star-forming galaxies at $z \sim 3-5$ to build a more complete picture of ionizing-photon production and escape. These works utilize observations of the rest-UV portion of the electromagnetic spectrum to examine the interplay between massive stars, neutral gas, and dust in star-forming galaxies. These studies are conducted at the highest redshifts feasible for this type of analysis by avoiding complete attenuation of relevant observational signals from the neutral intergalactic medium. I present the redshift evolution of rest-UV spectral features out to $z \sim 5$ using a data set from the DEIMOS 10K Survey, alongside comparisons to similar analyses performed at $z \sim 2-4$. I find a relationship between Lyα emission line strength and low-ionization absorption line strength and conclude that the neutral-gas covering fraction of a galaxy simultaneously modulates the two. The

redshift evolution of this relation indicates that the ionizing photon production efficiency $\xi_{\rm ion}$ of this $z \sim 5$ sample is elevated as compared to $z \sim 2-4$ samples. I examine the effects of foreground contamination in direct measurements of the Lyman continuum (LyC) at $z \sim 3$ using data from the Keck Lyman Continuum (KLCS) Survey and novel *Hubble Space Telescope* imaging. After the removal of contamination, I measure an average escape fraction of ionizing radiation of $f_{\rm esc} = 0.06 \pm 0.01$ at $z \sim 3$, and present 13 confirmed, individual LyC leakers. Using this cleaned KLCS sample, I measure trends between $f_{\rm esc}$ and different galaxy properties in order to elucidate which galaxy populations contribute most strongly to the ionizing emissivity during reionization, while simultaneously determining which galaxy properties are most appropriate to use as a proxy of $f_{\rm esc}$ at $z > 6$. I attempt to measure the correlation between $f_{\rm esc}$ and star-formation rate surface density, an important diagnostic of $f_{\rm esc}$ found in the local Universe, and make recommendations for the total sample size required to determine whether this correlation exists out to $z \sim 3$. I present a negative correlation between $f_{\rm esc}$ and stellar mass, indicating that lower-mass galaxies ($\lesssim 10^9 M_\odot$) likely drive reionization. I also find a negative correlation between $f_{\rm esc}$ and E(B-V), underlining the importance of the configuration of neutral gas and dust in determining the escape fraction of an individual galaxy. Finally, I present a lack of significant trend between $f_{\rm esc}$ and stellar age or specific star-formation rate, indicating that recent star-formation does not strongly affect the conditions of the neutral-phase interstellar and circumgalactic media at the masses of the KLCS sample. Continued explorations of the physics behind $f_{\rm esc}$ and $\xi_{\rm ion}$ at $z \sim 3-5$ are crucial for building a comprehensive understanding of the reionization process.

The dissertation of Anthony Joseph Pahl is approved.

Tommaso Treu

Steven R. Furlanetto

Matthew A. Malkan

Alice E. Shapley, Committee Chair

University of California, Los Angeles

2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First and foremost, I am immensely grateful to my advisor, Alice Shapley, whose unwavering support, guidance, and expertise have been invaluable throughout my research journey. Alice's insightful feedback and tireless efforts in reviewing and refining my work have significantly shaped the outcome of this dissertation. I am indebted to her for her mentorship and the countless hours she invested in me throughout my graduate career.

I would also like thank my parents, my brothers (Alec and Nick), and my sister in law (Lauren) for their support and encouragement. I am deeply grateful for their love and sacrifices, as well as leaving idyllic Minnesota to visit me in Los Angeles. I am truly fortunate to have such a kind and entertaining family. Additionally, I want to express my profound appreciation to my partner, Echo, for her support and understanding throughout the course of my PhD.

I want to thank the UCLA Astronomy graduate students, who have such collective strength, which in turn gave me individual strength. I will fondly recall the recurring events that brought us closer, namely dewar, groffee, and grunch. I want to especially thank my cohort of Ronald, Veronica, Kelly, and Sanaea, who were the most supportive and thoughtful group of colleagues (now friends) I could have asked for.

Finally, I want to thank the support staff within the UCLA Physics & Astronomy department who made my graduate career possible. I particular want to thank Stephanie Krilov and Laura Quattropani for all of their thoughtful and attentive assistance.

I acknowledge that some of the 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 the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M.

Keck Foundation. I wish to extend special thanks to those of Hawaiian ancestry on whose sacred mountain we are privileged to be guests. Without their generous hospitality, most of the observations presented herein would not have been possible.

I acknowledge the results presented in this thesis are based upon published works with additional coauthors, and are reprinted here with permission. Chapter 2 of this dissertation is a version of Pahl et al. (2020) and is reproduced by permission of MNRAS. Chapter 3 of this dissertation is a version of Pahl et al. (2021) and is reproduced by permission of MNRAS. Chapter 4 of this dissertation is a version of Pahl et al. (2022) and is reproduced by permission of MNRAS. Chapter 5 of this dissertation is a version of Pahl et al. (2023) and is reproduced by permission of MNRAS.

VITA

2014–2017 **Undergraduate Research Assistant**, Department of Physics & Astronomy, University of Minnesota, Twin Cities, Minneapolis, MN

2016 **Intern**, NASA Goddard Space Flight Center, Greenbelt, MD

2017 **B.S., Physics and Astrophysics**, University of Minnesota, Twin Cities, Minneapolis, MN

2017–2019 **Teaching Assistant**, Department of Physics and Astronomy, UCLA, Los Angeles, CA. Astronomy 180 – Astrophysics Laboratory (Fall 2019), Astronomy 6 – Cosmology: Our Changing Concepts of the Universe (Winter 2018), Astronomy 4 – Black Holes and Cosmic Catastrophes (Spring and Fall 2018), Astronomy 3 - Nature of the Universe (Spring 2019), Astronomy 5 - Life in Universe (Fall 2019)

2017–present **Graduate Student Researcher**, Department of Physics and Astronomy, UCLA, Los Angeles, CA

2019 **M.S., Astronomy and Astrophysics**, UCLA, Los Angeles, CA

PUBLICATIONS

The connection between the escape of ionizing radiation and galaxy properties at $z \sim 3$ in the Keck Lyman Continuum Spectroscopic Survey

Pahl, A. J., Shapley, A., Steidel, C. C., Reddy, N. A., Chen, Y., Rudie, G. C., Strom, A. L., 2023, MNRAS, 521, 3

Searching for the connection between ionizing-photon escape and the surface density of star formation at $z \sim 3$

Pahl, A. J., Shapley, A., Steidel, C. C., Reddy, N. A., Chen, Y., 2022, MNRAS, 516, 2

An uncontaminated measurement of the escaping Lyman continuum at $z \sim 3$

Pahl, A. J., Shapley, A., Steidel, C. C., Chen, Y., Reddy, N. A., 2021, MNRAS, 505, 2447

The redshift evolution of rest-UV spectroscopic properties to $z \sim 5$

Pahl, A. J., Shapley, A., Faisst, A. L., Capak, P. L., Du, X., Reddy, N. A., Laursen, P., Topping, M. W., 2020, MNRAS, 493, 3194

The KBSS-KCWI survey: the connection between extended Lyα haloes and galaxy azimuthal angle at $z \sim 2-3$

Chen, Y., ..., Pahl, A. J., et al., 2021, MNRAS, 508, 19

The Mass-Metallicity Relation at $z \sim 1-2$ and Its Dependence on the Star Formation Rate

Henry, A., ..., Pahl, A., et al., 2021, ApJ, 919, 143

Lyman Continuum Escape Fraction of Star-forming Dwarf Galaxies at $z \sim 1$

Rutkowski, M. J., ..., Pahl, A. J., et al., 2016, ApJ, 819, 81

CHAPTER 1

Introduction

After cosmic recombination, the Universe consisted primarily of neutral-phase Hydrogen gas. The first stars and galaxies condensed out of this diffuse material, and subsequently, the Universe began to undergo a key phase transition known as reionization. During this process, the vast majority of neutral Hydrogen in the intergalactic medium (IGM) transitioned to an ionized state. From observations of the Lyα absorption features in quasar spectra, it is thought that reionization ends at $z \sim 5.5-6$, approximately a billion years into the Universe's history (e.g., Fan et al., 2006; Becker et al., 2021). Significant uncertainty remains in the comprehensive timeline and driving forces behind this process.

Reionization requires a steady input of ionizing radiation into the intergalactic medium to progress. Both active galactic nuclei (AGN) and star-forming galaxies are potential sources of these high-energy photons in the early Universe, but the total ionizing emissivity from galaxies at $z > 6$ is likely more dominant (e.g., Parsa et al., 2018). Ionizing photons are produced from massive, O/B type stars located in HII regions in the earliest galaxies, and proceed to stream into the IGM after being partially absorbed by neutral Hydrogen in the interstellar and circumgalactic media (respectively, ISM and CGM). These production and escape mechanisms naturally lead to a parameterization of the global ionizing emissivity as a function of three variables (Robertson et al., 2015). The first is the cosmic star-formation rate (SFR) density, or $\rho_{\rm SFR}$, which correlates with the total number of massive stars in the Universe per unit volume. Second, the ionizing photon production efficiency, or $\xi_{\rm ion}$, describes

the efficiency with which these massive stars produce photons with enough energy to ionize the neutral-phase IGM. Third, the escape fraction, or $f_{\rm esc}$, describes the fraction of ionizing photons that are not absorbed by the neutral Hydrogen in the ISM and CGM of a galaxy and subsequently free stream into the IGM. All three of these variables must be understood at $z > 6$ in order to constrain the evolution of reionization.

The rest-UV portion of a galactic spectrum contains a wealth of information useful for constraining our understanding of reionization. The Lyman continuum (LyC) portion of the UV spectrum, from $\sim 880 - 910$Å, represents the direct ionizing signal output from the massive stars in HII regions. The Lyα recombination line, at 1216Å, has an equivalent width that is modulated both by the ionizing efficiency of star formation and the covering fraction of neutral-phase gas. Low- and high-ionization metal absorption lines in the rest-UV spectrum contain information about the physical location and kinematics of different phases of gas within a galaxy. Thus, spectroscopic measurements of at these wavelengths can elucidate both direct measurements and indirect tracers of $f_{\rm esc}$ and $\xi_{\rm ion}$, which represent the majority of the unknown parameter space when attempting to characterize the evolution of the ionizing emissivity over cosmic time.

While the ultimate goal is understanding these quantities in the epoch of reionization itself, there are significant challenges in obtaining direct constraints at $z > 6$. Measuring $f_{\rm esc}$ is particularly difficult, considering that the IGM is largely opaque to LyC radiation down to $z \sim 3.5$ (Vanzella et al., 2015). This opacity renders direct constraints of $f_{\rm esc}$ at $z \gtrsim 3.5$ infeasible. Additionally, the Lyα line is similarly attenuated by the neutral-phase IGM, with the transmission fraction for typical galaxies at $z \sim 6.5$ dropping to 20% (Laursen et al., 2011). Thus, a powerful avenue of study exists at $z \sim 3 - 5$ where lower-redshift analogs of reionization-era galaxies can be studied in detail to understand both the physical mechanisms of ionizing-photon production and escape, and, most importantly, which galactic observables are correlated with $f_{\rm esc}$. Armed with this knowledge, the process of reionization can be

accurately modeled using assumptions of $f_{\rm esc}$ for a given population of galaxies, as well as both direct and indirect measurements of $\xi_{\rm ion}$.

Significant observational effort has been directed towards exploring the mechanisms of $f_{\rm esc}$ using lower-redshift analogs to reionization-era galaxies. The escape fraction, $f_{\rm esc}$, has been successfully measured in the local Universe using the *Hubble Space Telescope*/COS Spectrograph, with selection functions that include compact sizes, high [O III]$\lambda\lambda$4959, 5007/[O II]$\lambda\lambda$3726, 3729 ratios, and high SFR surface densities (Borthakur et al., 2014; Izotov et al., 2016, 2018, 2021; Flury et al., 2022a). Nonetheless, it is important to constrain $f_{\rm esc}$ closer to reionization at $z > 2$, where the properties of galaxies are more similar to their reionization-era counterparts. A suite of photometric (Grazian et al., 2017; Fletcher et al., 2019; Nakajima et al., 2019; Saxena et al., 2021; Begley et al., 2022) and spectroscopic (Marchi et al., 2017, 2018; Steidel et al., 2018; Begley et al., 2022) surveys have attempted to measure $f_{\rm esc}$ at $z \sim 2-4$ with varying levels of success. Hallmarks of a successful programs include sensitive enough near-IR data to confidently detect an average LyC signal, a selection function that is independent of LyC strength, careful treatment of foreground contamination (a significant concern for detecting LyC leakers at $z \sim 3$, Vanzella et al., 2012; Mostardi et al., 2015), and spectral stacking methodology to remove the variability of IGM attenuation on the LyC signal.

Armed with assumptions regarding $f_{\rm esc}$ that are built upon our current observational knowledge, researchers have developed a variety of reionization models that arrive at starkly different conclusions when predicting the evolution of the cosmic neutral fraction. Finkelstein et al. (2019) assumes that $f_{\rm esc}$ depends negatively on halo mass, and concludes that reionization starts early ($z \sim 12$) and proceeds gradually. Naidu et al. (2020) assumes that $f_{\rm esc}$ is a direct function of SFR surface density, and concludes that reionization ends late, such that neutral fractions remain at $\gtrsim 90\%$ until $z \sim 8$. We currently lack a statistical, representative, and uncontaminated sample of LyC leakers at $z \sim 3$ studied as a function of a large number of galaxy properties that can delineate which assumptions regarding $f_{\rm esc}$ in the

epoch of reionization are appropriate.

In this dissertation, I investigate spectroscopic probes of $z \sim 3-5$ galaxies in order to elucidate the physical mechanisms governing ionizing photon production and escape, as well as explore key related galaxy properties that help inform assumptions of $f_{\rm esc}$ during reionization. I perform a comprehensive study of rest-UV spectral features out to $z \sim 5$, examine the effects of foreground contamination on $f_{\rm esc}$ measurements, and measure the relationships between $f_{\rm esc}$ and a suite of galaxy properties. This work is primarily based on the Keck Lyman Continuum Spectroscopic (KLCS) survey, a large observing program that obtained deep, rest-frame ultraviolet spectra of 124 galaxies at $2.75 < z < 3.4$, described in Steidel et al. (2018). Critically, this data set provides deep spectra that cover the LyC spectral region with a selection criteria that is independent of LyC emission strength, and thus enables an unbiased determination of $f_{\rm esc}$ at $z \sim 3$. This sample can additionally be used to explore $f_{\rm esc}$ as a function of galaxy properties based on the construction of composite spectra in bins of galaxy property.

The dissertation is organized as follows. In Chapter 2, I analyze 176 galaxies at $z \sim 5$ from the DEIMOS 10K Specroscopic Survey (Hasinger et al., 2018) and study the redshift evolution of Lyα and rest-UV metal absorption lines from $z \sim 2-5$ to draw conclusions about the intrinsic evolution of the ionizing photon production efficiency. In Chapter 3, I present an uncontaminated measurement of $f_{\rm esc}$ at $z \sim 3$ after performing careful removal of foreground contamination in the KLCS survey. Chapter 4 explores the connection between SFR surface density and $f_{\rm esc}$ in the KLCS by combining its deep, rest-UV spectra; high-resolution imaging from *Hubble Space Telescope*; and multi-wavelength, ground-based photometry. In Chapter 5, I continue the analysis of $f_{\rm esc}$ and galaxy property in the KLCS as a function of stellar mass, SFR, E(B-V), specific SFR, and stellar age. Finally, Chapter 6 summarizes the results of these studies.

CHAPTER 2

The Redshift Evolution of Rest-UV Spectroscopic Properties to z∼5

2.1 Introduction

The rest-UV spectrum of a star-forming galaxy provides a uniquely rich probe of the properties of massive stars and the interstellar and circumgalactic medium (ISM and CGM; Shapley et al., 2003; Steidel et al., 2010, 2016; Leitherer et al., 2011; Senchyna et al., 2019). Understanding the interplay between these components is important for placing constraints on galaxy evolution across cosmic time. Furthermore, the gas in and around distant galaxies that is probed by rest-UV spectra also modulates the escape of ionization radiation. This gas provides important clues to the process of cosmic reionization, a key phase transition in which the intergalactic medium (IGM) transformed from neutral to ionized by roughly a billion years after the Big Bang at $z \sim 6$ (Fan et al., 2006; Robertson et al., 2015; Planck Collaboration et al., 2016).

Measuring the rest-UV spectra of distant galaxies requires optical instrumentation and long integrations. Based on observations with instruments such as the Low Resolution Imager and Spectrometer (LRIS) on the Keck I telescope (Oke et al., 1995; Steidel et al., 2004), there have been several analyses of the relationships between rest-UV spectral features and other galaxy properties out to $z \sim 4$ (e.g., Du et al., 2018; Jones et al., 2012; Marchi et al., 2019). Du et al. (2018) (hereafter, D18) performed a comprehensive evolutionary analysis

of rest-UV spectral trends of star forming galaxies from $z \sim 2$–4. This analysis centred around the hydrogen Lyα line and low- and high-ionization interstellar (LIS and HIS) metal absorption lines, which together probe the neutral gas covering fraction, gas kinematics, and ISM and CGM properties of galaxies. Both D18 and Jones et al. (2012) find a strong and non-evolving correlation between Lyα equivalent width (EW) and low-ionization interstellar absorption EW, revealing the connection between escaping Lyα photons (and, by extension, ionizing radiation; Reddy et al., 2016b; Steidel et al., 2018; Chisholm et al., 2018) and the covering fraction of neutral gas in the CGM. Building on these earlier works, we now seek to probe the rest-UV properties of galaxies at redshifts even closer to the epoch of reionization.

Approaching reionization at $z > 6$, we must attempt to understand how the rest-UV spectrum reveals the changing properties of both massive stars and the ISM/CGM at higher redshift. The evolution of the relationship between Lyα emission strength and LIS absorption strength sheds light on the intrinsic production rate of Lyα photons in high-redshift galaxies, which in turn probes the intrinisic ionizing-photon production rate. Affecting observations near reionization, the IGM becomes more optically thick to Lyα photons and attenuates the spectra of galaxies. We must correct for this attenuation to accurately determine the changing intrinsic production rate of Lyα photons. As stated in D18, the Lyα-LIS relationship also indicates the changing covering fraction of neutral gas in the CGM at higher redshift. Both intrinsic ionizing-photon production rate and neutral-gas covering fraction are key components of understanding the contribution of star-forming galaxies to the ionizing background during reionization (Robertson et al., 2015; Nakajima et al., 2018; Steidel et al., 2018; Trainor et al., 2019; Chisholm et al., 2018, 2019).

The DEIMOS 10K Spectroscopic Survey (Hasinger et al., 2018) provides an ideal dataset for performing this type of analysis, including thousands of galaxies ranging from $z \sim 0$ to $z \sim 6$ with deep, rest-UV spectroscopy. This survey was conducted in the COSMOS field (Scoville et al., 2007), which is covered by extensive broadband photometry enabling the

measurement of integrated galaxy properties. In this work, we extend the analysis of D18 to higher redshift by carefully examining a $z \sim 5$ sample of star-forming galaxies drawn from the DEIMOS 10K Spectroscopic Survey. We perform a joint analysis of the new $z \sim 5$ sample with the $z \sim 2$–4 samples of D18, using consistent methodology across all redshifts to remove any potential systematics.

In Section 2.2, we introduce our sample of $z \sim 5$ galaxies and the $z \sim 2$–4 galaxies of D18, and discuss sample completeness, redshift measurements, and spectral energy distribution (SED) fitting. In Section 2.3, we describe the methodology for measuring the equivalent widths of Lyα and interstellar absorption features, and the binning of the sample and subsequent creation of composite spectra. In Section 2.4, we present the resulting trends of rest-UV spectral features at $z \sim 5$. In Section 2.5 we analyze the sample at a finer redshift sampling and discuss the implications of our results for quantifying the changing intrinsic Lyα photon production and the contribution of star-forming galaxies to reionization. In Section 2.6, we summarize our key results and conclude.

Throughout this paper, we adopt a standard ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70\ \mathrm{km\,s^{-1}Mpc^{-1}}$.

2.2 Sample

2.2.1 Selection Criteria

We analyzed a sample of galaxies drawn from the DEIMOS 10K Spectroscopic Survey of the COSMOS field (Hasinger et al., 2018; Scoville et al., 2007). This survey provides deep optical spectroscopy taken with the DEep Imaging Multi-Object Spectrograph (DEIMOS) on the Keck II telescope on Mauna Kea. The COSMOS field is covered by a wealth of ancillary data, including multi-wavelength imaging from space-based and ground-based missions. Out of the ∼10,000 galaxies targeted in the DEIMOS 10K survey, we selected those identified

as Lyman break galaxies (LBGs), dropping out in the B_J and g^+ optical filters for $z \sim 4$, and V_J or r^+ for $z \sim 5$. Also included were objects with photometric redshifts between $z \sim 3.75$–5.75. These criteria yielded a parent sample of 416 star-forming galaxies at high-redshift, 337 selected via dropout and 79 via photometric redshifts.

We further required objects to have high-confidence spectroscopically-confirmed redshifts. The DEIMOS 10K catalog provided a quality flag Qg indicating the confidence level of the spectroscopic-redshift measurement based on the quality of the spectrum and the number of features in the redshift estimate (either Lyα or LIS lines). We examined all objects with $Qg \geq 1$ that satisfied the above photometric criteria, eliminating only those with unsuccessful redshift measurements. From this set of objects, we measured a redshift from the Lyα line, LIS lines, or both. We successfully recovered redshifts for 196 objects at $4 < z < 5.5$. For the purpose of our completeness calculations, from our parent photometric sample of 416 galaxies, we removed galaxies selected by photometric redshift that had measured spectroscopic redshifts $z < 4.0$, reducing the parent sample to 375.

2.2.2 Systemic Redshifts

Ideally, rest-optical nebular emission lines would be used to estimate the systemic redshifts of high-redshift galaxies. As these lines are shifted into the thermal infrared at $z \sim 5$, we measure the redshifts of the spectra via the redshifts of the Lyα emission line ($z_{\mathrm{Ly}\alpha}$) and the three strongest LIS absorption features (z_{LIS}): Si IIλ1260, O Iλ1302+Si IIλ1304, and C IIλ1334. Using the spectroscopic redshifts provided by DEIMOS 10K, we estimated the central wavelength of each non-resonant line and fit a Gaussian profile to the flux-calibrated spectrum. Each line was visually inspected to confirm the validity of the line detection and corresponding fit. The redshifts based on the central wavelengths of the well-fit Gaussian profiles of the three LIS lines were averaged with weights to calculate z_{LIS}. Given the asymmetry of the Lyα profile, we did not model it with a Gaussian, but rather calculated $z_{\mathrm{Ly}\alpha}$ based on the wavelength at which the spectrum reached a maximum within

the Lyα profile. We then determined the systemic redshift using the method described in Rudie et al. (2012). $z_{\mathrm{Ly}\alpha}$ was assumed to be offset from systemic by 300 km s^{-1} and z_{LIS} by -160 km s^{-1}, with these offsets empirically determined by Steidel et al. (2010). While these rules are calculated at $z \sim 2$, we assume they hold to $z \sim 5$ in the absence of other information. This calculation enabled an estimate of the systemic redshift based on a high-quality measurement of either $z_{\mathrm{Ly}\alpha}$ or z_{LIS}. If both $z_{\mathrm{Ly}\alpha}$ and z_{LIS} were measured, the systemic redshift was estimated as the average of the two. After combining 14 duplicate spectra, removing four spectra with poorly-fit lines and two spectra with background subtraction issues, we assembled a final sample of 176 objects between $4 < z < 5.5$: 68 with $z_{\mathrm{Ly}\alpha}$ only, 34 with z_{LIS} only, and 74 with both types of redshift measurement. The redshift distribution of the sample is displayed in Figure 2.1. The median redshift of our sample is $z_{\mathrm{med}} = 4.521$, with a 16th–84th percentile range of 4.172–4.930. The redshift uncertainty σ_z for each object was calculated using uncertainties on the assumed velocity offsets of Lyα and LIS, empirically determined in Steidel et al. (2010). For galaxies with both $z_{\mathrm{Ly}\alpha}$ and z_{LIS} measured, we assumed $\sigma_{v_{\mathrm{Ly}\alpha}} = 175$ km s^{-1} and $\sigma_{v_{\mathrm{LIS}}} = 115$ km s^{-1}. For galaxies with only z_{LIS} measured, we assumed $\sigma_{v_{\mathrm{LIS}}} = 165$ km s^{-1}. For galaxies with only $z_{\mathrm{Ly}\alpha}$ measured, we assumed $\sigma_{v_{\mathrm{Ly}\alpha}} = 175$ km s^{-1}. The median σ_z of the sample redshifts was 0.002. We refer to this sample as the "$z \sim 5$ sample" for simplicity and keep this sample distinct from samples featured in D18 to signify a new dataset.

For the objects with both Lyα and LIS absorption redshifts measured, we examined the distribution of velocity differences between Lyα and LIS lines in Figure 2.2. This measurement was calculated as

$$\Delta v_{\mathrm{Ly}\alpha\text{–LIS}} = c \times \frac{(z_{\mathrm{Ly}\alpha} - z_{\mathrm{LIS}})}{(1 + (z_{\mathrm{Ly}\alpha} + z_{\mathrm{LIS}})/2)}. \tag{2.1}$$

The velocity differences are distributed about a $\langle \Delta v_{\mathrm{Ly}\alpha\text{–LIS}} \rangle = 496$ km s^{-1} with a standard deviation of 222 km s^{-1}. This value is largely consistent with $\langle \Delta v_{Ly\alpha\text{–}LIS} \rangle \sim 600$ km s^{-1} of Steidel et al. (2010) at $z \sim 2$–3, and $\langle \Delta v_{Ly\alpha\text{–}LIS} \rangle = 429 \pm 229$ km s^{-1} of Faisst et al. (2016) at $z \sim 5$, the latter using a similar sample drawn from the DEIMOS 10K survey. The

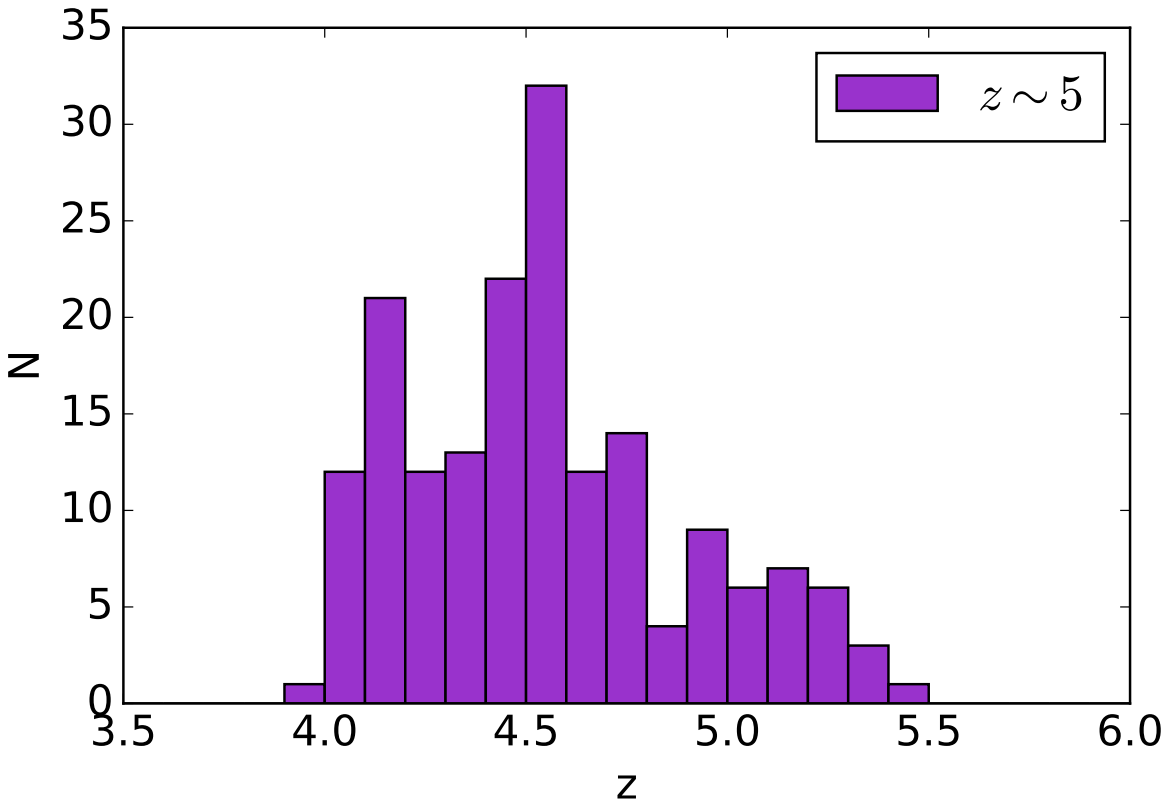

Figure 2.1: Redshift distribution of the DEIMOS 10K sample at $z \sim 5$. The sample comprises 176 objects with a median redshift of 4.521.

distribution of velocity differences of Figure 2.2 has a similar shape to that of $z \sim 3$ LBGs of Shapley et al. (2003), a sample which has been shown to have similar outflow kinematics to those of the Steidel et al. (2010) $z \sim 2$ sample. This similarity lends credence to our assumption that the outflow velocities calculated at $z \sim 2$ are applicable out to $z \sim 5$.

2.2.3 Completeness

In order to explore the completeness of the $z \sim 5$ COSMOS DEIMOS sample, we examined the number of galaxies in the parent sample that report a high-confidence redshift as a function of brightness. The results of this analysis are illustrated in the top panel of Figure

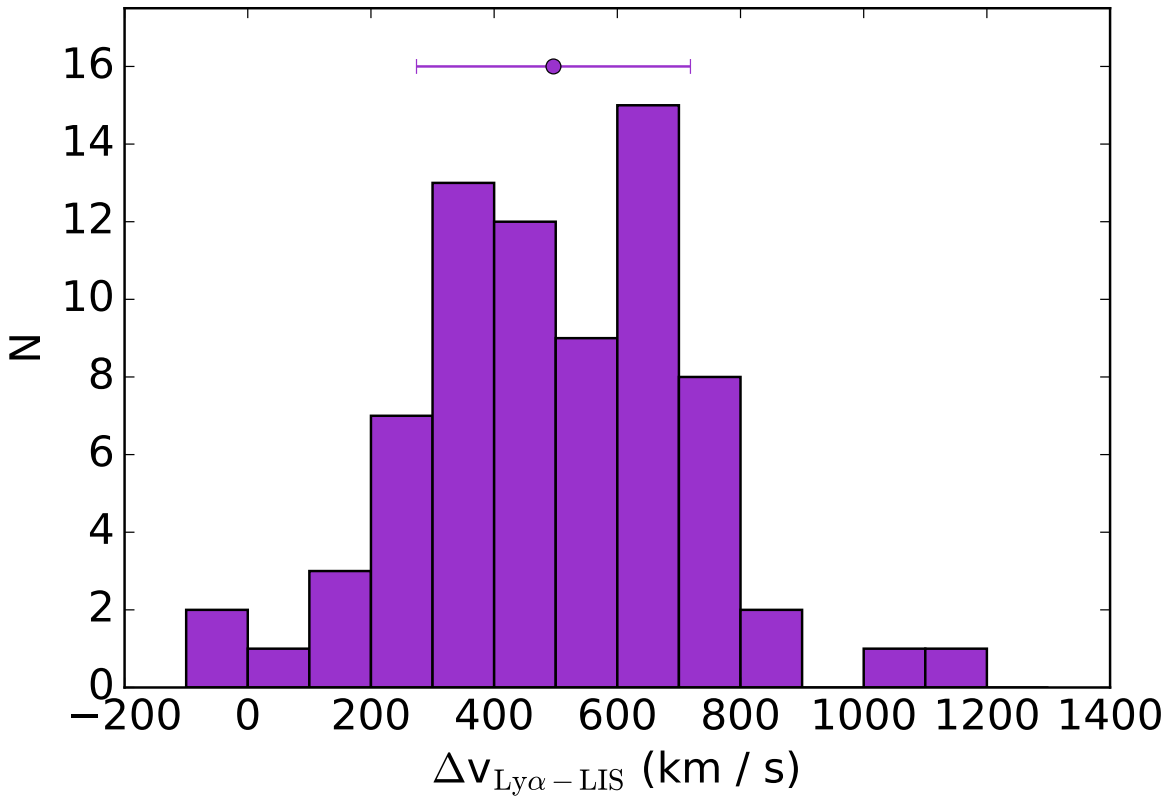

Figure 2.2: Distribution of velocity offsets between Lyα emission and LIS absorption. The velocities are distributed about $\langle \Delta v_{\mathrm{Ly}\alpha-\mathrm{LIS}} \rangle = 496 \pm 222$ km s^{-1}.

2.3, where the total number of galaxies in the parent sample and the percentage of successful redshift measurements are plotted as a function of i^+-band magnitude. At $i^+ > 25$ mag, the sample completeness drops below ~ 60 per cent, thus a more careful examination of the sample completeness must be performed.

As described above, spectroscopic redshifts at high redshift are measured using a combination of HI Lyα emission and LIS absorption lines of Si IIλ1260, O Iλ1302+Si IIλ1304, and C IIλ1334. Measuring redshifts requires the robust detections of these lines. Since we explore the *variation* of the strengths of LIS and HIS absorption as a function of Lyα emission strength, strong detections should not bias our results. It is more important to examine whether the galaxies in our $z \sim 5$ spectroscopic sample are representative of the parent photometric sample of star-forming galaxies in terms of other galaxy properties, such as color and brightness. To investigate this issue, we plotted our sample and the parent photometric sample in color-magnitude space in the bottom panels of Figure 2.3. For the galaxies selected as dropouts, all filters chosen in the color-magnitude diagram are redward of the corresponding dropout filter. For both dropouts and photo-z selected samples, all but two of the deviations in median color are smaller than 0.06 mag, and all but one of the deviations in median magnitude are smaller than 0.05 mag. The lack of significant deviation in the median between our sample and the parent photometric sample demonstrates that the sample is not significantly biased in color and brightness. The most significant deviation (0.22 mag) is between the median i^+-band magnitudes of the V_J dropouts in our sample vs. the parent sample, thus we examined our sample as a function of brightness in detail in section 2.3.1. We conclude that our spectroscopic sample is not significantly biased towards certain types of star-forming galaxies with respect to the parent photometric sample in the same redshift range.

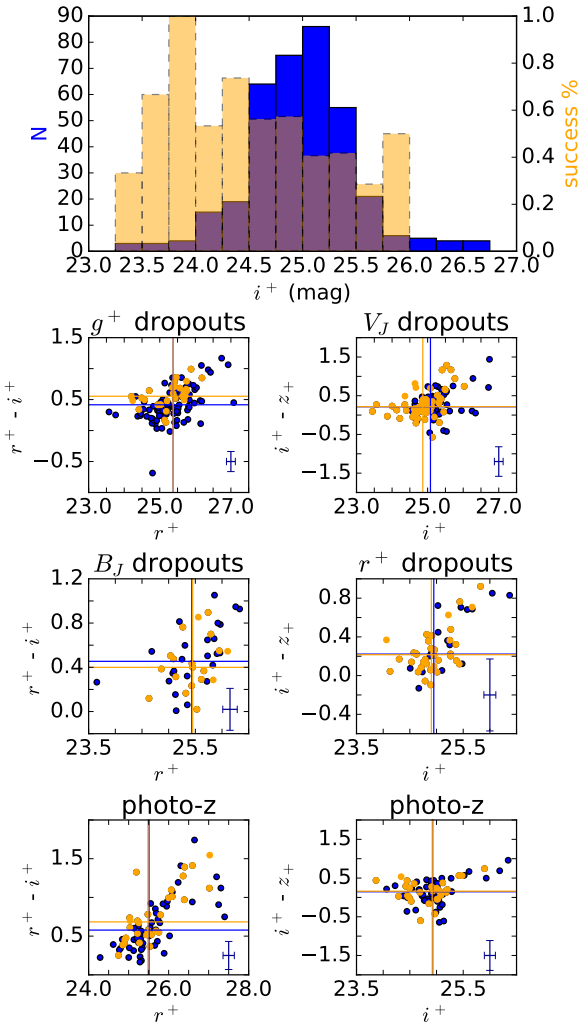

Figure 2.3: Completeness of the $z \sim 5$ sample. **Top panel:** Histogram (in orange) showing the percentage of successfully measuring a high-confidence spectroscopic redshift in the $z \sim 5$ parent sample as a function of i^+magnitude. The total i^+ magnitude distribution of the parent sample is overplotted in blue. The completeness drops below ~ 60 per cent at an i^+ magnitude of $\sim$25.0. **Middle four panels:** Color-magnitude diagrams of the dropout-selected galaxies. The parent sample is plotted in blue, with the corresponding median colors and magnitudes plotted as vertical and horizontal lines. The $z \sim 5$ sample with well-measured redshifts is plotted in orange, along with the corresponding medians. Typical error bars for the respective samples are given in the lower-right corner of each panel. No significant deviation between the parent sample and the spectroscopic sample is found in color-magnitude space. **Bottom two panels:** Color-magnitude diagrams of the galaxies selected by photometric redshifts between $3.75 < z < 5.75$. These panels follow the same color scheme as the dropout-selected galaxy panels.

2.2.4 $z \sim 2$–4 sample

For a lower-redshift comparison sample, we used the $z \sim 2$–4 LBGs described in D18. The sample consists of $z \sim 2$–3 galaxies preselected using U_nGR color cuts and followed up with LRIS on Keck I as part of the redshift surveys described in Steidel et al. (2003), Steidel et al. (2004), and Reddy et al. (2008). The $z \sim 4$ sample was drawn from Jones et al. (2012) and consists of B-band dropouts in the GOODS-S field, 42 of which were followed up with DEIMOS on Keck II. In addition, 28 had spectra from the FOcal Reducer and low dispersion Spectograph 2 (FORS2; Vanzella et al., 2005, 2006, 2008, 2009) archival catalog. The average spectral resolutions were $R_{\rm average} \sim 970$ and 1280 for the $z \sim 2$–3 and $z \sim 4$ samples respectively, compared to $R_{\rm average} \sim 2700$ for the $z \sim 5$ sample. Systemic redshifts were measured in a similar method to that described above[1], and the sample was divided into three bins of increasing redshift, with 671 $z \sim 2$ objects defined at $1.4 \leq z < 2.7$, 352 $z \sim 3$ objects at $2.7 \leq z < 3.4$, and 80 $z \sim 4$ objects at $3.4 \leq z \leq 4.5$. The full redshift distribution of the $z \sim 2$–5 galaxies is displayed in Figure 2.4. Some overlap does exist between galaxies in the $z \sim 4$ and $z \sim 5$ samples, but they are largely distinct with median redshifts of 3.856 and 4.521, respectively. In Section 2.5, we search for evolution within the $z \sim 5$ sample by dividing this sample into two distinct redshift bins with $z_{\rm med} = 4.330$ and 4.742. We note that the latter of these two more finely-sampled bins does not overlap in redshift space with the $z \sim 4$ sample from Jones et al. (2012).

2.2.5 SED fits

As we aim to compare spectral trends across a range of redshifts, we wanted to make a controlled comparison between similar types of galaxies at each redshift. To this end, we first made measurements of the galaxy stellar-population properties in each redshift bin

[1]There was a small difference in the method for estimating systemic redshifts in the case where both $z_{\rm Ly\alpha}$ and $z_{\rm LIS}$ were measured. D18 simply used $z_{\rm Ly\alpha}$ with an assumed velocity offset instead of averaging the two redshift measurements. In practice, the two estimates are very similar and this small difference in approach does not affect our results.

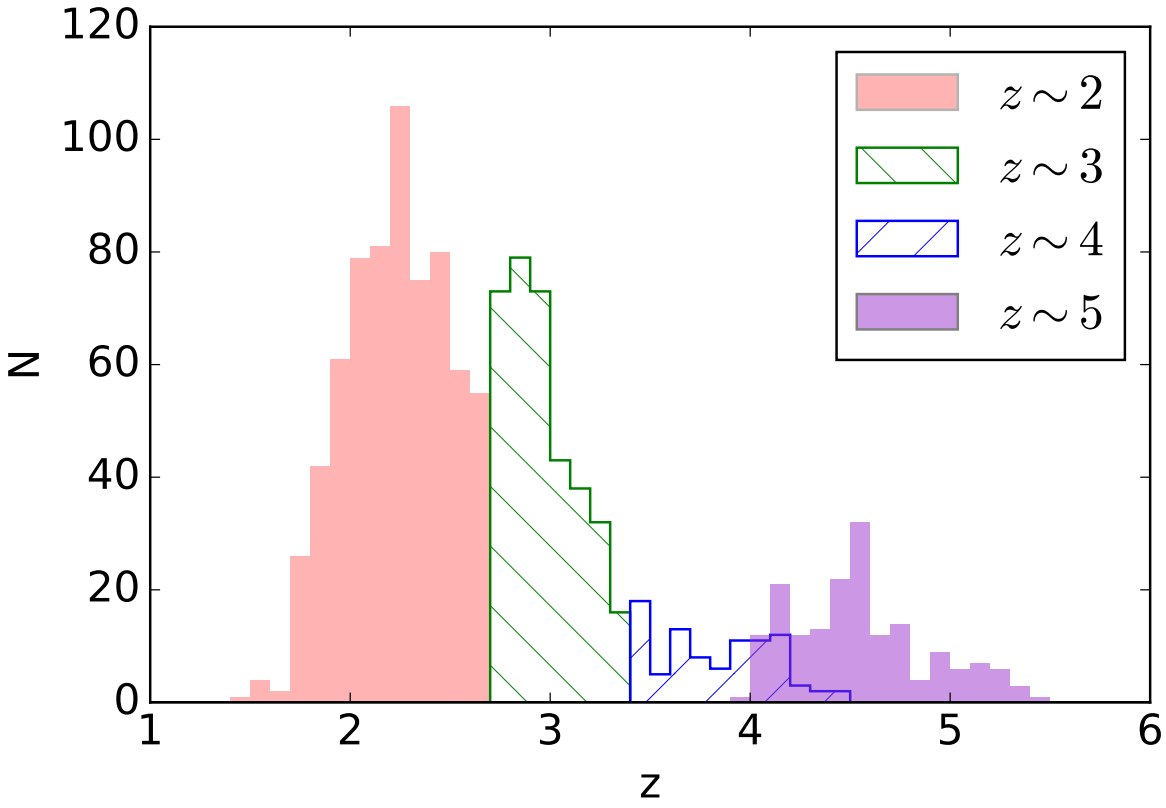

Figure 2.4: Redshift distribution of the four samples at $z \sim 2$–5. The $z \sim 2$, 3, and 4 samples in red, green, and blue are drawn from D18. The $z \sim 5$ sample from this work is shown in purple.

via stellar-population synthesis model fits to the galaxy SEDs. We used photometry from the publicly-available catalogs of COSMOS2015 (Laigle et al., 2016) for 171 of our objects, well matched in RA and Dec, and, for four objects, we used the photometric catalog of Capak et al. (2007). For one object, we used the COSMOS Intermediate and Broad Band Photometry Catalog 2008[2], but, as this object did not have sufficient coverage redward of the Balmer break, we removed it from the sample, leaving us with 175 galaxies at $z \sim 5$.

In our SED fitting, we utilized the Bruzual and Charlot (2003) stellar-population templates and assumed a Chabrier (2003) initial mass function (IMF). As these models do not include emission lines, we omitted the IRAC channel 1 band from the SEDs of galaxies at $4.00 < z < 4.99$ and the IRAC channel 2 band from the SEDs of galaxies at $5.06 < z < 5.31$. This correction removes potential SED contamination by the Hα and [O III]λ5007 rest-optical lines. As in D18, we performed the SED fits using a combination of a fixed metallicity of 1.4 times the solar metallicity ($Z = 0.02$) and a Calzetti extinction curve (Calzetti et al., 2000), or else a fixed metallicity of 0.28 $Z_{\odot}$ and an SMC extinction curve (Gordon et al., 2003). This method of SED fitting was initially described in Reddy et al. (2018b). The star-formation history was set at either constant or exponentially increasing. The fits were performed with ages > 10 Myr allowed, and then with only those > 50 Myr. With these different assumptions, we fit for stellar mass (M_*), ages younger than the age of the universe, star-formation rate (SFR), and $E(B-V)$ from 0.00 to 0.60. Given that the quality of the fits of each model was similar, we adopted the best-fitting parameters of 0.28 $Z_{\odot}$, SMC attenuation, constant star-formation history, and age > 50 Myr models. This choice was motivated by the fact that a sub-solar metallicity with an SMC extinction curve has been demonstrated to fit the IRX-β relationship better than other models at $z > 4$ (Bouwens et al., 2016; Oesch et al., 2012). Ages were constrained to > 50 Myr so as to not be shorter than the typical dynamical timescales of distant star-forming galaxies (assuming a relaxed disk and ignoring the timescsales of local gas motions, Reddy et al., 2012), and a constant

[2]Available at https://irsa.ipac.caltech.edu/data/COSMOS/ datasets.html

	$z_{\rm med}$	M_{UV}	$\log(M_*/M_\odot)$	$E(B-V)$	Sample Size
$z\sim2$	2.267	-20.51	10.00	0.09	539
$z\sim3$	2.925	-21.00	9.87	0.08	309
$z\sim4$	3.856	-21.06	9.72	0.04	91
$z\sim5$	4.521	-21.33	9.80	0.05	175

Table 2.1: Median properties of the controlled $z\sim2$–5 samples.

SFR was assumed for consistency with past work (Du et al., 2018; Reddy et al., 2012; Steidel et al., 2014; Strom et al., 2017). The SED fitting of the $z\sim4$ sample in D18 was performed in a similar way. For the $z\sim2$–3 galaxies, the 1.4 $Z_\odot$+Calzetti model was used as they produced overall lower χ^2.

In order to perform a controlled study across redshift of galaxies with comparable properties, we restricted the $z\sim2$, 3, 4, and 5 samples to span the same range in stellar mass and rest-UV luminosity. We examined the distribution of M_* determined by the SED fits and the UV absolute magnitude (M_{UV}) in Figure 2.5. The M_*-M_{UV} cuts performed in D18 of ($-22.62 < M_{UV} < -19.91$) and ($8.04 < \log(M_*/M_\odot) < 11.31$) are displayed as vertical and horizontal lines in the figure. These cuts remove low-luminosity objects in the $z\sim2$ sample. The M_* and M_{UV} of the $z\sim5$ sample span much of the same range as the $z\sim4$ sample. Since the $z\sim2$–3 samples were already truncated to match the $z\sim4$ sample, we perform no stricter M_*-M_{UV} cut on the $z\sim2$ and $z\sim3$ samples. The median M_{UV} and M_* of each sample within the D18 cuts are recorded in Table 2.1, along with other median sample properties from the SED fitting. The similarity of the M_{UV} and M_* medians further motivates that no additional cut is needed. After applying the D18 limits in M_* and M_{UV}, we find the number of objects in the $z\sim2$, 3, 4 and 5 samples are 539, 309, 91 and 175, respectively.

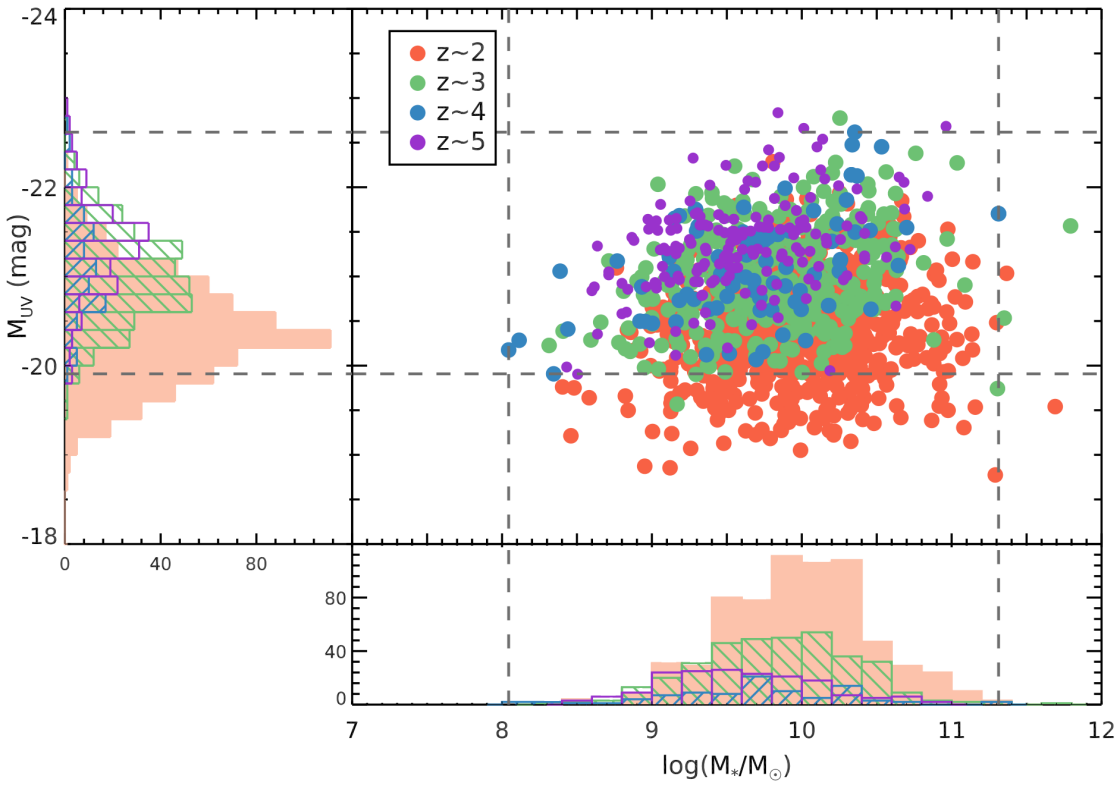

Figure 2.5: Absolute UV magnitude (M_{UV}) vs. stellar mass (M_*) for the four redshift samples. The distributions of M_{UV} and M_* are displayed in the histograms to the left and below the scatterplot. The cuts performed on the $z \sim 2$–4 data are shown as vertical and horizontal dashed lines.

2.3 Methods

Our goal is to examine the properties of the rest-UV spectral features of our different samples, including Lyα; LIS lines of Si IIλ1260, O Iλ1302+Si IIλ1304, C IIλ1334, and Si IIλ1527; and HIS lines of Si IV$\lambda\lambda$1393, 1402 and C IVλ1548, 1550. While Lyα can be well detected at $z \sim 5$ due to its large typical EW, it is difficult to detect the full suite of LIS and HIS lines in individual spectra, given their typical EWs of $\sim 1 - 2$ Å. Thus, we rely on stacks of spectra in bins of various galaxy properties to reveal the characteristic strengths of the LIS and HIS lines. The Lyα line is known to have an important role in a galaxy's spectrum, strongly correlating with rest-UV spectral morphology (Shapley et al., 2003), motivating its choice as a sorting parameter and the creation of composites in bins of Lyα EW. Lyα EW, as opposed to Lyα luminosity, is normalized to the UV continuum and probes the efficiency of Lyα photon production and escape. The methods in this section broadly follow those described in D18. In addition, we consider here for the first time the increasing degree of IGM attenuation on average as a function of increasing redshift, quantifying this effect using the models of Laursen et al. (2011).

2.3.1 Lyα Equivalent Width in Individual Spectra

In order to construct composite spectra spanning a range of Lyα EWs, we required an individual measurement of Lyα for each object. We used the fitting method of Kornei et al. (2010) to measure Lyα EW. First, the rest-frame spectra were examined individually to determine whether the spectrum at 1216Å follows an "emission," "absorption," "combination," or "noise" profile. The large majority of our sample were fit using "emission" or "noise," representing either a positively-peaked, visible feature or no discernible one, with only one object following an "absorption" profile consisting of a trough at 1216Å and four objects following a "combination" profile of a blueside trough and redside peak.

The continuum level was then determined using the following procedure: the redside contin-

uum was calculated as the average of the flux values of the spectrum between 1225Å–1255Å. If the spectrum had coverage blueward of 1120Å, the blueside continuum was estimated as the average of the flux between 1120Å–1180Å. If the spectrum only had coverage down to 1160Å, the blueside continuum was averaged between 1160Å–1180Å. If the spectrum did not have coverage to 1160Å, the blueside continuum value was estimated by multiplying the redside value by the mean ratio of blue-to-red of 0.260, as determined from the remaining 140 objects of the sample with blueside coverage. The blueside continuum is suppressed relative to the redside due to absorption by the IGM. For 14 of the objects, the redside continuum was not well detected, leading to a ∼zero continuum value well within the 1σ noise of the spectrum. In order to measure an equivalent width for these galaxies, we estimated the continuum flux photometrically. This estimate was obtained using the best-fitting SED of the object as described in Section 2.2.5, averaging in the same wavelength window of 1225Å–1255Å. These 14 spectra were also normalized over the wavelength interval 1300Å–1500Å so the continuum estimated from the spectra matched those from the photometry. For the objects with a well-detected continuum, this photometric estimation was also performed and compared to the spectroscopically-determined values to ensure consistency between the measurements.

For spectra classified as "emission," "combination," or "absorption," the Lyα flux was integrated between the wavelengths at which the full Lyα profile intersected the blueside and redside continuum values. For "noise," the spectrum was integrated over a fixed window of 1210–1220 Å. Finally, the equivalent width was measured as the ratio of the integrated Lyα flux and the redside-continuum flux density. The Lyα EW distribution of the four redshift samples is shown in Figure 2.6. The median Lyα EW of each sample is overplotted on the corresponding histogram as a vertical line, indicating the evolution of the strength of Lyα within our four samples of increasing redshift. The sample median Lyα EWs are -6.07Å, -0.08Å, 9.51Å, and 11.75Å for redshifts 2, 3, 4, and 5, respectively. This evolution must be interpreted in the context of the incompleteness of our $z \sim 5$ sample. We find that 38.6 per cent of objects at $z \sim 5$ had a systemic redshift measurement based only $z_{\mathrm{Ly}\alpha}$, compared to

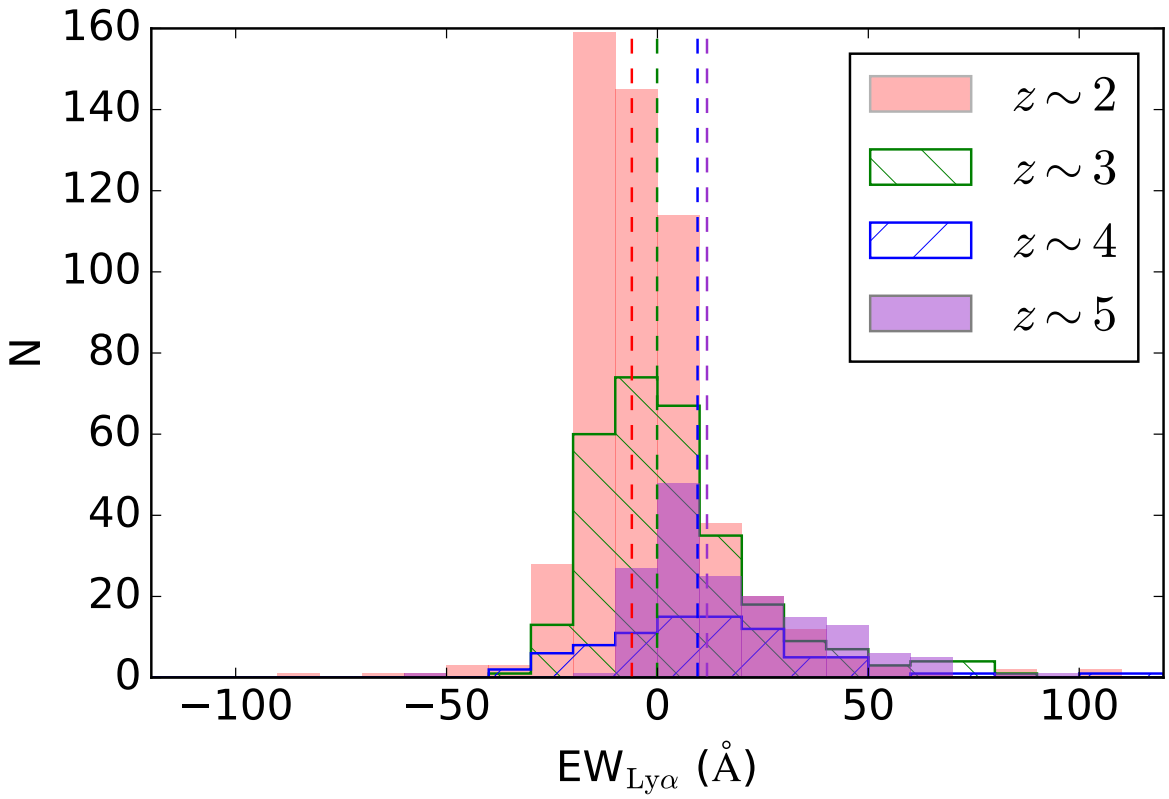

Figure 2.6: Rest-frame EW$_{\mathrm{Ly}\alpha}$ distribution as a function of redshift. The $z \sim 2$, 3, and 4 samples are from D18. The median EW$_{\mathrm{Ly}\alpha}$ for each sample is displayed as a vertical, dashed line. These medians increase towards stronger emission with increasing redshift.

56.3 per cent at $z \sim 4$. In past work, a weak evolution of Lyα emitting fraction has been found at these redshifts (Stark et al., 2010). We do not find a significant tail in the EW$_{\mathrm{Ly}\alpha}$ distribution at values less than 0 Å (i.e., absorption) at $z \sim 5$. However, we argue that, as the $z \sim 5$ sample is not significantly biased in redshift measurement method or color-color space (Figure 2.3), it is also not significantly biased with respect to the underlying Lyα EW distribution – at least compared with $z \sim 4$. See D18 for a discussion of sample completeness from $z \sim 2$ to $z \sim 4$.

2.3.2 Composite Spectra

We first created composite spectra of the spectra in each redshift sample, and then split the samples into bins of various properties and generated composites of the spectra of the objects in each bin.

We generated composite spectra by performing a bootstrapping analysis of the objects in each sample. Based on the number of objects in the sample, we drew the same number of spectra from the bin with replacement. For each draw, we took each individual galaxy spectrum and normalized it to its photometric flux density. We then shifted it into the rest frame in L_λ units based on its measured systemic redshift. The rest-frame, L_λ spectrum of each object was then perturbed according to a Gaussian distribution with a standard deviation corresponding to the value of the error spectrum at that wavelength. These spectra were combined with the IRAF tool *scombine* using the combination mode "average," with rejection of the three highest and three lowest flux values at each wavelength increment. This bootstrapping was performed 100 times, generating 100 composites, each representing a different sampling of spectra in the bin and reflecting the statistical noise in each individual spectrum. The resulting 100 spectra were then combined into a final science composite and error spectrum by taking the average L_λ and standard deviation at each wavelength increment.

We can examine the overall spectral characteristics of each sample with this composite method. The four composites of increasing redshift are presented in Figure 2.7. The most striking difference in the $z \sim 5$ spectrum is the strength of its Lyα feature, reflecting the difference in median Lyα EW of the sample. The properties of the absorption features are difficult to distinguish qualitatively and require further analysis.

We split our redshift sample into bins of various properties. In addition to Lyα, these properties include $E(B-V)$ and integrated galaxy properties such as M_*, SFR, age, and M_{UV}. Depending on the property being examined, the set of objects comprising each bin

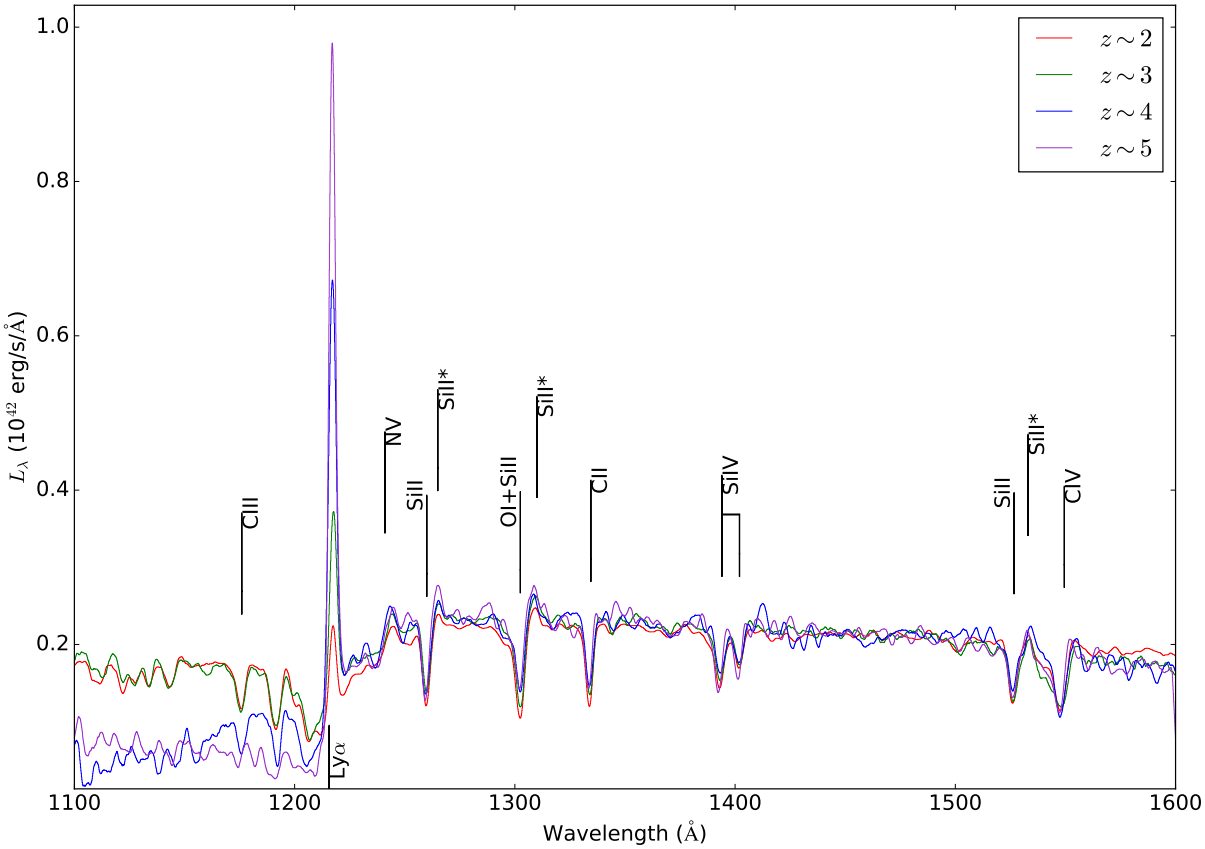

Figure 2.7: Composite spectra of the four redshift samples. These spectra were smoothed with a boxcar window of 3Å. The $z \sim 2$–4 spectra were normalized to the $z \sim 5$ spectra using their respective median flux value in the range 1450-1500Å. The rest-frame positions of various absorption and emission features are displayed.

change based on the spectral coverage of the objects. For examining LIS strength as a function of Lyα EW, we required every object in the sample to have coverage out to the most redward LIS line, Si IIλ1527. This requirement slightly reduced our sample to 160 objects. We split this sample into four bins of increasing Lyα EW with 40 objects per bin. Using an equal number of objects per bin provides a comparable S/N for each composite spectrum. To examine HIS strength as a function of Lyα EW, we required coverage of the most redward HIS line, C IVλ1548, 1550. We divided the resulting sample of 154 objects into four bins of increasing Lyα. Finally, we examined Lyα and LIS strength as a function of integrated galaxy properties using four bins of the 160 objects with coverage of Si IIλ1527.

The four composite spectra binned by Lyα EW are shown in Figure 2.8. The absorption features are clearly resolved in the composite spectra, and the variation of the Lyα EW in each bin can be seen in the profile at 1216Å. From these composites, we can remeasure the Lyα EW and measure the EW of interstellar absorption features. To determine the Lyα EW of each bin, we applied the fitting technique described in Section 2.3.1 to each of the bootstrapped spectra, representing the variation in each composite bin of Figure 2.8. The final Lyα EW and 1σ error reported for each bin was the mean and standard deviation of these 100 measurements. This measured Lyα EW from stacked spectra and the median Lyα EW of the objects composing each bin were comparable within 1σ.

To understand how our analysis is affected by the changing intrinsic brightness of the galaxies, we examined the M_{UV} of each galaxy as a function of its measured Lyα EW. All the galaxies in our $z \sim 5$ sample are plotted in M_{UV}-$EW_{Ly\alpha}$ space in Figure 2.9. As Lyα emission strength increases, the galaxies tend to be dimmer in the UV continuum. The objects are split into four bins of increasing Lyα EW, and the median M_{UV} and Lyα EW are displayed. The median M_{UV} of the lowest $EW_{Ly\alpha}$ bin is -21.44, while the highest bin has a median M_{UV} of -21.11. While we observe a trend of fainter M_{UV} with increasing $EW_{Ly\alpha}$, just as that found within the $z \sim 3-4$ samples of D18, this dependence is weak. Accordingly, it is reasonable

to interpret differences in $EW_{Ly\alpha}$ as being dominated by variations in emergent Lyα flux.

2.3.3 IGM Correction to Lyα Equivalent Widths

In order to obtain a more accurate measurement of the intrinsic Lyα production and escape through the ISM, we attempted to remove the effects of IGM attenuation on the strength of the observed Lyα line. The IGM transmission in the vicinity of the Lya feature decreases noticeably from $z \sim 2$ to $z \sim 5$. To quantify this evolution, Laursen et al. (2011) generated sightlines through high-resolution cosmological simulations of galaxy formation, where the simulated sightlines originate in samples of galaxies at different redshifts ranging from $z = 2.5$ to 6.5. These authors produced the mean IGM transmission profile in the vicinity of Lyα at each redshift, showing how its shape and depth evolve towards stronger absorption at higher redshift. We interpolated the model of Laursen et al. (2011) to produce IGM transmission curves appropriate for the median redshift of each of our four redshift samples. [3]

We then degraded the curves in a manner that represented the process of observing our galaxy spectra at moderate resolution and combining them with uncertain systemic redshifts into composite spectra. Specifically, we smoothed the model IGM transmission curve by a Gaussian kernel to match the average spectral resolution of each composite: $R_{\rm average} \sim$ 970, 1280, and 2700 for the $z \sim$ 2–3, $z \sim 4$, and $z \sim 5$ samples respectively. The more important blurring effect arises from the fact that the estimates of systemic redshift from Lyα and LIS features have associated systematic errors (i.e., they are based on mean relations between Lyα or LIS redshifts and systemic redshifts, which have intrinsic scatter). In order to simulate the effect of stacking N spectra with uncertain systemic redshifts, we generated N individual IGM transmission curves and perturbed each curve by a small $\Delta\lambda$. The value $\Delta\lambda$ was drawn from a normal distribution with zero mean and a standard deviation of

[3]We note that each IGM transmission curve has an associated uncertainty based on the range of sightline opacities found in the simulations of (Laursen et al., 2011). Given that this uncertainty for the average of ∼40 sightlines is subdominate compared to our other uncertainties, we did not take it into account.

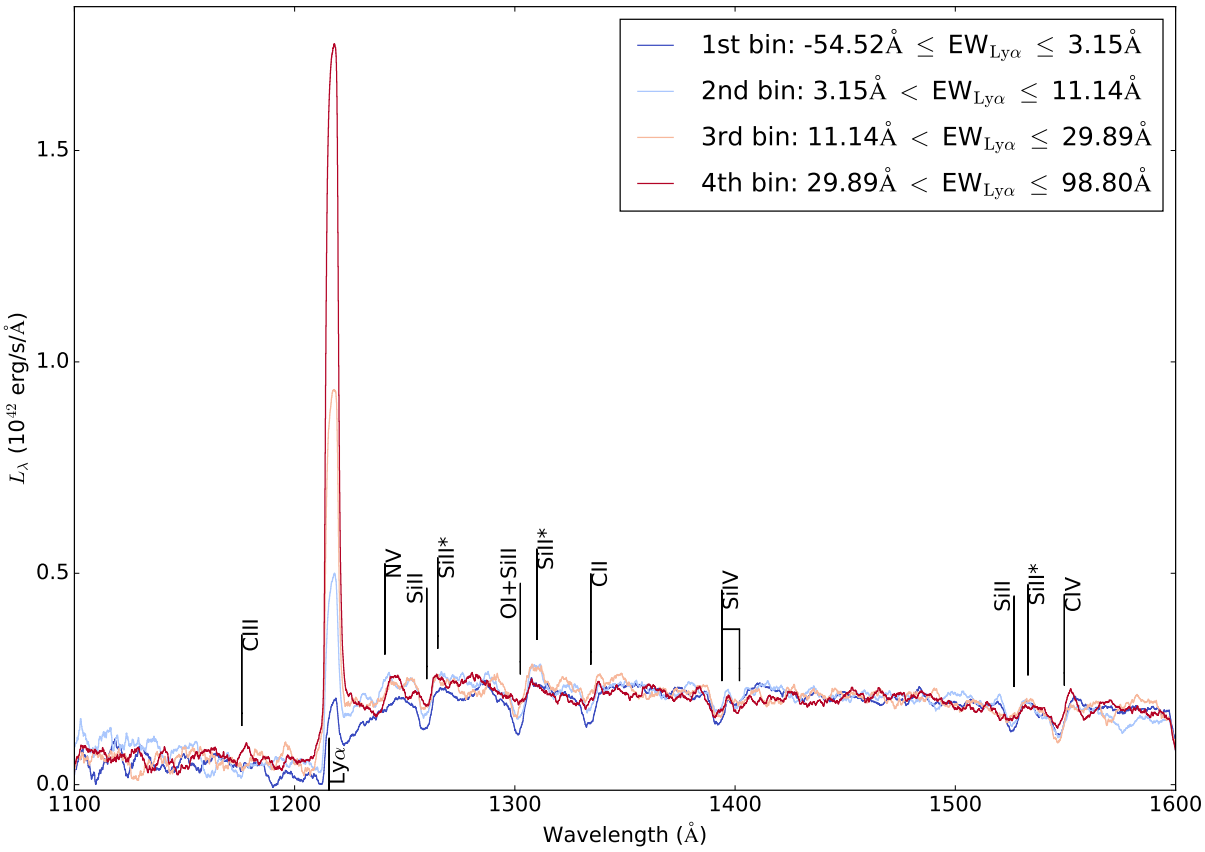

Figure 2.8: Composite spectra of the $z \sim 5$ sample in bins of increasing Lyα EW. The dark blue spectrum corresponds to the objects with the lowest EW$_{Ly\alpha}$, while the dark red contains the objects with the highest EW$_{Ly\alpha}$. Each composite spectrum is constructed from the spectra of 34 objects characterized by the EW$_{Ly\alpha}$ ranges displayed in the legend. The variation of rest-UV spectral features as a function of increasing EW$_{Ly\alpha}$ can be qualitatively observed near the wavelengths of various LIS and HIS lines labelled by solid black lines.

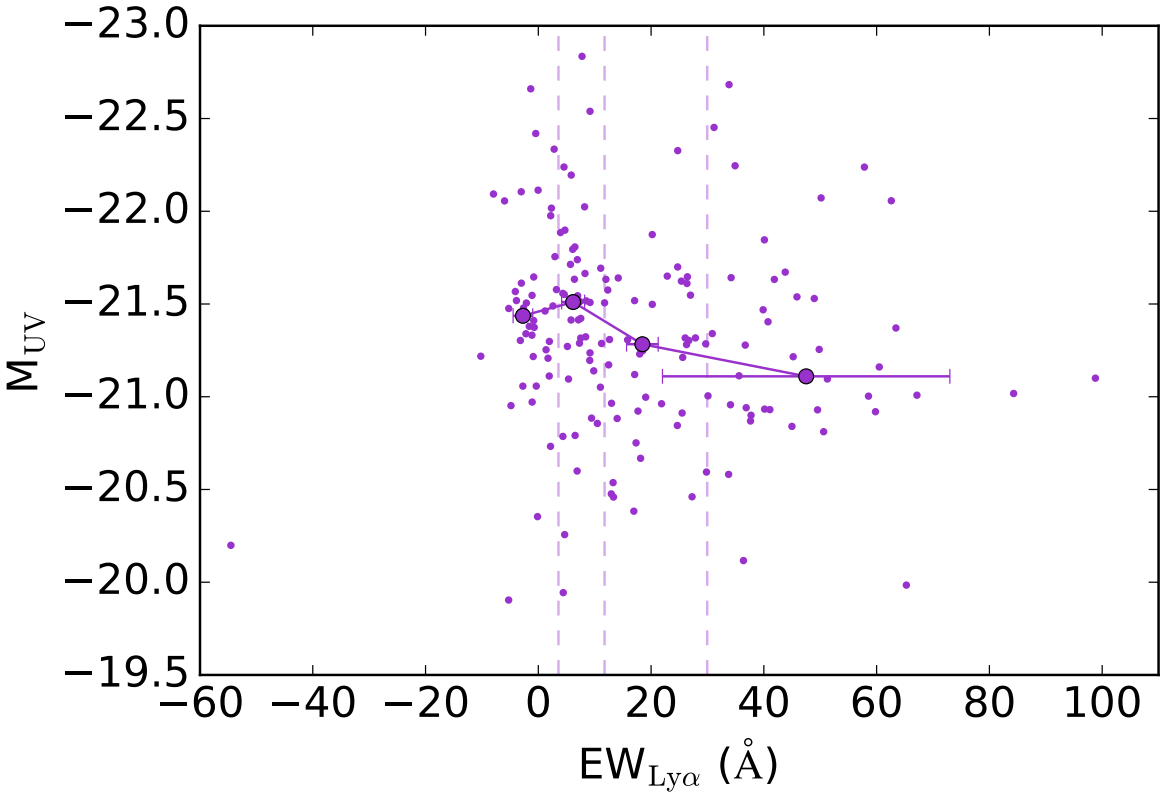

Figure 2.9: Absolute UV magnitude (M_{UV}) vs. $EW_{Ly\alpha}$. Individual detections are displayed as smaller circles. The sample is divided into bins of increasing Lyα EW, with these divisions shown as vertical, dashed lines. The median properties of the objects of each bin are plotted as larger, filled circles.

$\sigma_{\Delta\lambda} = (\sigma_z/(1+z)) \times 1215.67$Å. The redshift uncertainty σ_z was calculated through the method described in Section 2.2.2. The resulting N perturbed (in wavelength) transmission curves were then averaged together. The transmission curves calculated in this way for the $z \sim 2-5$ samples are shown in the bottom panel of Figure 2.10. The fraction of photons transmitted by the IGM at 1215.67Å decreases substantially towards $z \sim 5$.

In order to convert the transmitted Lyα profiles to those intrinsic to the galaxy, we divided the observed composite spectra by the associated degraded IGM transmission curve. Both the transmitted and intrinsic Lyα profiles of the $z \sim 2-5$ samples are shown in the top panel of Figure 2.10. The Lyα profiles in composite spectra at higher redshift receive larger positive corrections for IGM absorption. This analysis was applied to the composites of the $z \sim 2-5$ samples binned by Lyα EW and galaxy properties, and the Lyα EW of each composite was remeasured as described in section 2.3.2.

2.3.4 Absorption Line Equivalent Width

For each composite spectrum, we measured both LIS and HIS equivalent widths. First, we normalized the continua of the rest-frame, L_λ spectra using the IRAF tool *continuum*, using a fourth-order spline fit with continuum regions that avoid the spectral features defined by Rix et al. (2004). From the normalized spectra, we then fit each absorption feature with a Gaussian profile using the scipy routine *leastsq*, with initial fit parameters corresponding to the rest-frame central wavelength of the line and a linewidth of $\sigma = 2$Å, typical of composite LIS lines at $z \sim 5$. To calculate the flux of each line, we integrated the composite spectrum about the central wavelength of the Gaussian fit to the line using an integration window that spanned from $-3\sigma_v$ to $+3\sigma_v$. Here, σ_v is defined as the standard deviation in velocity space of the average of the four key LIS profiles under consideration. In order to determine σ_v, we converted the four LIS profiles from rest-frame wavelength space to velocity space and averaged them at each velocity increment. We then fit a Gaussian function to the combined absorption profile in velocity space using the scipy routine *leastsq*, leading to an

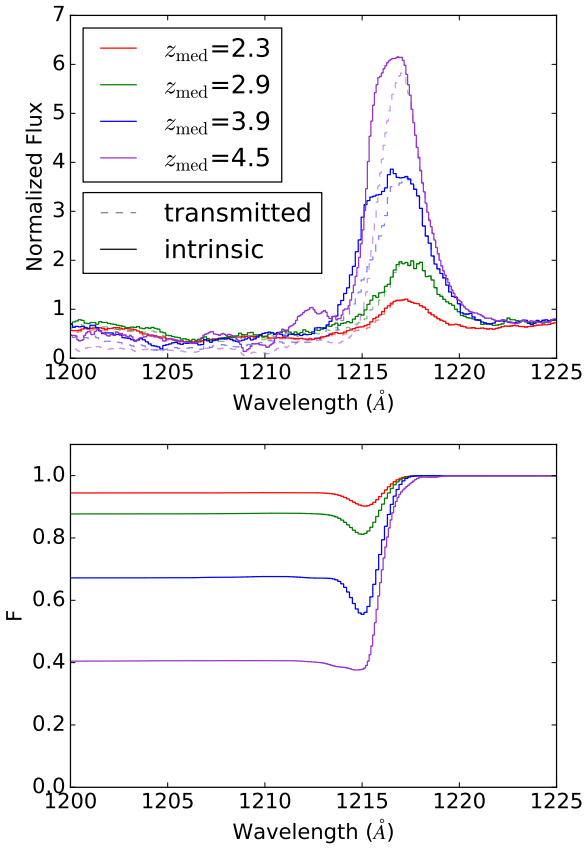

Figure 2.10: IGM correction method applied to the $z \sim 2$–5 composites spectra. **Top panel:** Lyα profiles for $z \sim 2$–5 sample composites, with and without IGM attenuation correction applied. The observed composites transmitted through the IGM are displayed as dashed lines, and the intrinsic Lyα profiles are displayed as solid lines. The intrinsic profiles were determined by dividing the transmitted profiles by the transmission curves shown in the bottom panel. **Bottom panel:** Degraded IGM transmission curves calculated at $z = 2.3, 2.9, 3.9$, and 4.5. Transmission curves were calculated at a given redshift by interpolating the models of Laursen et al. (2011) and were degraded based on the average spectral resolution of each redshift sample and the systemic redshifts of the objects contained in each sample.

estimate of σ_v. This technique yielded an integration window for each LIS feature that was informed by the average width of the LIS lines in the composite spectrum. At the same time, averaging the four LIS features in velocity space mitigated the effects of low-quality individual profile fits. We finally calculated the EW of the absorption feature by dividing this flux by the normalized continuum value in the integration window. The error on the EW measurement was determined via propagation of errors from the error spectrum in the integration window. The continuum-normalized spectra in bins of increasing Lyα EW are shown in Figure 2.11. The fits to the Gaussian profiles of the LIS lines of Si IIλ1260, O Iλ1302+Si IIλ1304, C IIλ1334, and Si IIλ1527 are displayed in detail here. The integration windows of $\pm$5Å are also displayed, with the entirety of the line profiles fitting in these windows.

For the HIS doublet of Si IV$\lambda\lambda$1393, 1402, a double Gaussian profile was fit to the line in the same method described for the single profiles, above. Additionally, the central-wavelength ratio of the doublet members was fixed at the theoretical value of (1402.770/1393.755). The integration window was set to range from $-$5Å of the centre of the bluer profile up to +5Å of the centre of the redder profile. These Si IV$\lambda\lambda$1393, 1402 fits in four bins of increasing $\mathrm{EW}_{\mathrm{Ly}\alpha}$ are shown in the left panels of Figure 2.12.

The C IVλ1548, 1550 profile is more complex: on top of the blended absorption feature, there is a stellar component produced by high-velocity winds in O and B stars (Du et al., 2016, and references therein). This stellar component must be removed before the absorption from the interstellar gas can be measured. While a stellar P-Cygni component can appear on the Si IV$\lambda\lambda$1393, 1402 absorption feature, the component is less significant within the mass ranges and corresponding metallicity ranges of our samples (Steidel et al., 2014; Sanders et al., 2018), thus does not require modeling and removal. We used the rest-frame UV model starburst spectra of Leitherer et al. (2010), smoothed to the resolution of our data, to represent the stellar component. We fit the templates, evaluated at 0.05 $Z_\odot$, 0.2 $Z_\odot$, 0.4 $Z_\odot$, $Z_\odot$, and 2 $Z_\odot$,

and combined the two bracketing (or one adjacent for the lowest- and highest-metallicities) best-fitting models. This fit is performed on the blue wing of C IVλ1548, 1550 between 1535 and 1544Å, and we refer readers to Figure 3 of Du et al. (2016) for an illustration of our fitting method. By dividing the continuum-normalized spectra by the best-fitting stellar template, we effectively removed the stellar component of C IVλ1548, 1550. We then measured the EW of the interstellar component of C IVλ1548, 1550 using the method previously described for the single LIS lines, as the doublet is blended together into a single profile at our spectral resolution. The fits to the C IVλ1548, 1550 interstellar profile after removing the stellar component are shown in the right panels of Figure 2.12.

2.4 Results

In this section, we examine the the properties of our $z \sim 5$ sample as a function of binned properties of Lyα EW, $E(B-V)$, stellar mass, M_{UV}, age, and SFR.

2.4.1 Line Strength

We are motivated by a physical picture in which the strength of LIS absorption probes the covering fraction of neutral gas in a galaxy (Reddy et al., 2016b). Thus, with the reasonable assuption of a roughly spherically symmetric gas distribution (e.g., Law et al., 2012), there will be a connection between the Lyα photons, which are scattered by any neutral gas covering the galaxy (Steidel et al., 2010), and the resulting strength of the LIS absorption features. In contrast, there is evidence that the HIS lines exist in an ionized medium, physically distinct from the neutral gas (Shapley et al., 2003; Du et al., 2016, 2018). HIS features have been found to have similar kinematics as LIS lines (Pettini et al., 2002; Heckman et al., 2015; Chisholm et al., 2016; Berg et al., 2018), however Du et al. (2016) finds that while the kinematics of HIS and LIS lines are similar, they correlate differently with galaxy properties, indicating they may exist in different ISM phases. A lack of relationship between Lyα and HIS will support this picture.

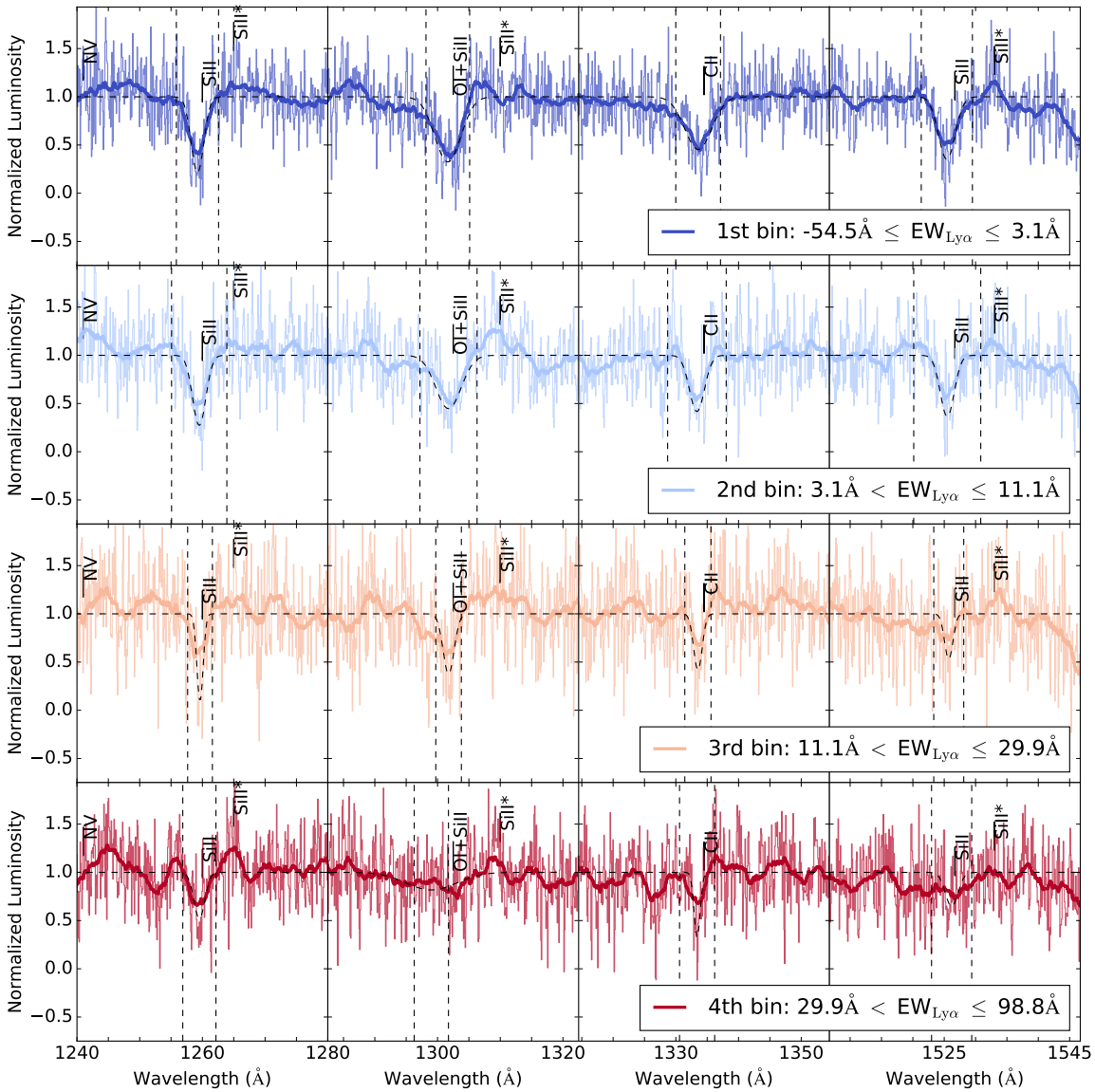

Figure 2.11: Fits to the four LIS features of the composite spectra of the $z \sim 5$ sample, split into four bins with $\mathrm{EW}_{\mathrm{Ly}\alpha}$ decreasing from the top to the bottom row. These $\mathrm{EW}_{\mathrm{Ly}\alpha}$ bins are the same as those shown in Figure 2.8. The more transparent spectra in each row are the continuum-normalized, rest-frame, L_λ spectra from which measurements were performed. The more opaque spectra are convolved with a boxcar window of width 3Å for easier viewing. The vertical, dashed lines illustrate the ±5Å integration windows from which EW measurements were made.

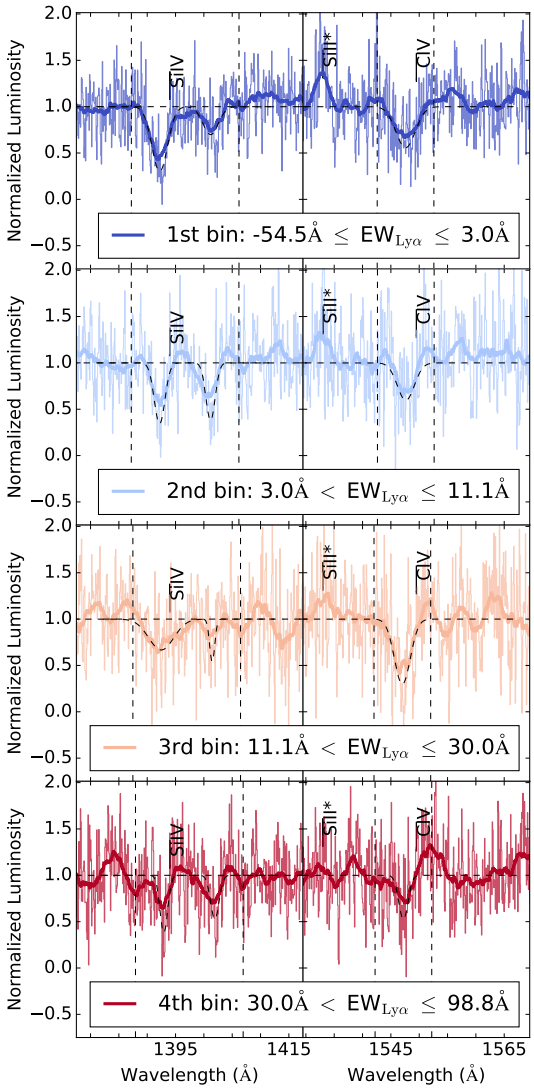

Figure 2.12: Fits to the Si IV$\lambda\lambda$1393, 1402 and C IVλ1548, 1550 features of the composite spectra of the $z \sim 5$ sample, split into four bins of increasing EW$_{Ly\alpha}$. These spectra have the stellar component of C IVλ1548, 1550 removed. The top panel corresponds to the objects with the lowest EW$_{Ly\alpha}$, while the bottom panel contains the objects with the highest EW$_{Ly\alpha}$. As in Figure 2.11, the transparent spectra are the continuum-normalized, rest-frame, L_λ spectra from which measurements were performed. The opaque spectra are convolved with a boxcar window of width 3Å for easier viewing. The vertical, dashed lines illustrate the ±5Å integration windows from which EW measurements were made.

To quantify the strength of the LIS lines, we measure EW_{LIS} as the weighted average of the Si IIλ1260, O Iλ1302+Si IIλ1304, C IIλ1334, and Si IIλ1527 EWs and their associated errors. EW_{LIS} is presented in the top left panel of Figure 2.13 as a function of increasing $EW_{Ly\alpha}$. Here, the relationship between Lyα and EW_{LIS} is presented at $z \sim 5$ using the sample from this work, while the $z \sim 2$–4 data points are taken from D18. The Lyα EWs uncorrected for IGM absorption are connected by dotted lines, while the intrinsic Lyα EWs calculated by the method described in Section 2.3.3 are connected by solid lines. It can be seen that at $z \sim 5$ the Lyα EW range skews significantly higher than that at lower z, with the lowest EW bin at -1.8 ± 1.4Å. The trends at lower redshift are largely unchanged at $z \sim 5$. This result implies that the attenuation of Lyα photons (i.e., due to scattering and dust absorption) and the strength of the LIS feature is intrinsically linked, and is fundamental across redshift. In the highest bin of $EW_{Ly\alpha}$, $EW_{Ly\alpha}$ is greater at fixed EW_{LIS} compared to lower redshift, though within 1σ. This deviation suggests intrinsically stronger Lyα at fixed ISM covering fraction (but see e.g. McKinney et al., 2019). We further explore this result in Section 2.5.

As seen in Figure 2.11, the shape of the LIS profiles is not strongly dependent on Lyα strength. The ratio of $EW_{Si\ II\lambda1260}$ / $EW_{Si\ II\lambda1527}$ can reflect the changing properties of the gas producing Si II absorption: when the ratio approaches ~ 5, the ratio of their oscillator strengths, Si II is in the optically-thin regime. The ratio of $EW_{Si\ II\lambda1260}$ / $EW_{Si\ II\lambda1527}$ in all $EW_{Ly\alpha}$ bins is roughly unity, indicating that these Si II features are in the optically-thick regime. This result is consistent with Si II line ratios in the $z \sim 2-4$ samples presented in D18.

We also examine the evolution of the average spectral properties of the composites after the IGM correction was applied, shown as stars in the top left panel of Figure 2.13. We see the overall composites following a similar linear relationship at $z \sim 2$–4, with the average Lyα emission strength increasing and LIS absorption strength decreasing with increasing redshift. We see an evolution off this linear relationship at $z \sim 5$, with the overall composite having

stronger Lyα emission and similar LIS absorption properties as the $z \sim 4$ sample. We will further examine the average and binned spectral properties in a $z \sim 5$ subsample with higher median redshift in Section 2.5.

The $\mathrm{EW_{HIS}}$ is defined as the weighted average of $\mathrm{EW_{CIV\lambda1548,1550}}$ and $\mathrm{EW_{SiIV\lambda\lambda1393,1402}}$. We display the relationship between $\mathrm{EW_{HIS}}$ in bins of Lyα line strength in the lower left panel of Figure 2.13. In D18, no significant relationship between HIS and Lyα strength was found at $z \sim 2$–4. In the $z \sim 5$ sample, we find similar strengths of $\mathrm{EW_{HIS}}$, showing no redshift evolution in the overall strengths of the features. We also find no trend of stronger $\mathrm{EW_{HIS}}$ as a function of stronger $\mathrm{EW_{Ly\alpha}}$ emission, similar to what is observed at lower redshift.

In the highest quartile of $\mathrm{EW_{Ly\alpha}}$ emission strength, we find evidence for nebular C IV emission at $z \sim 5$. This detection is shown in the bottom panel of Figure 2.12. We discuss this spectral feature and its implications for the inferred metallicity of the $z \sim 5$ sample in Section 2.5.3.

2.4.2 $E(B-V)$

Of all the galaxy properties measured by SED fitting, dust attenuation is most directly tied to the makeup of the ISM and CGM. As dust grains preferentially scatter bluer light, the amount of reddening reflects the the amount of dust along the line-of-sight. The reddening of the SED of high-redshift sources by dust is typically measured in terms of the quantity $E(B-V)$. The relationship between Lyα and $E(B-V)$ has been well demonstrated: stronger Lyα emission in galaxies correlates with a bluer SED (Shapley et al., 2003; Kornei et al., 2010; Vanzella et al., 2009; Pentericci et al., 2007). The scattering of Lyα photons through neutral gas in the ISM implies greater path lengths to interact with the dust in the galaxy. Shapley et al. (2003) explored the possibility of dust existing in the same phase as the neutral gas itself, implying a direct correlation between the neutral-gas covering fraction and dust reddening for non-resonant radiation, although uncertainty exists between whether the dust lies in outflowing gas or HII regions.

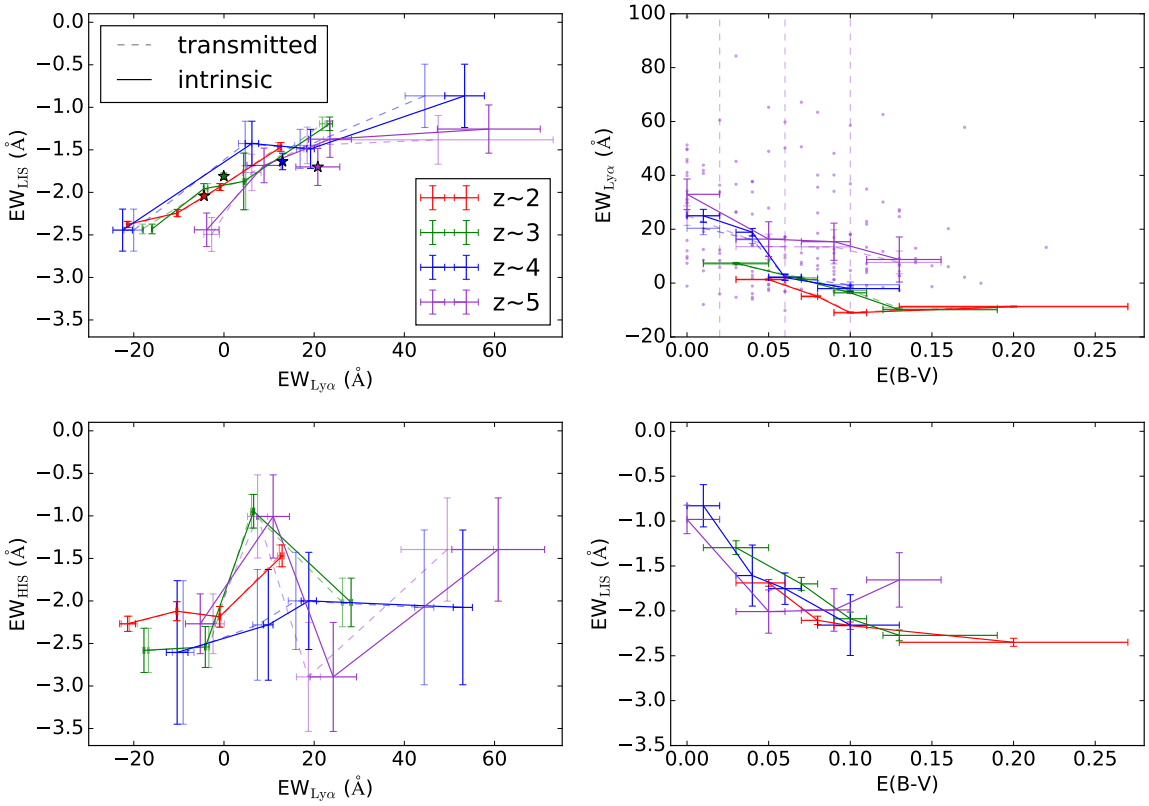

Figure 2.13: Relations between $\mathrm{EW_{Ly\alpha}}$, $\mathrm{EW_{LIS}}$, $\mathrm{EW_{HIS}}$, and $E(B-V)$. $\mathrm{EW_{Ly\alpha}}$, $\mathrm{EW_{LIS}}$, and $\mathrm{EW_{HIS}}$ were measured from composite spectra. The measurements from the observed composites are displayed in transparent points connected by dotted lines, while the corrected measurements after application of degraded IGM transmission curves are the opaque points connected by solid lines. The purple points correspond to measurements of the $z \sim 5$ sample. The red, green, and blue points are of the $z \sim 2$, 3, and 4 samples of D18. **Top left panel:** $\mathrm{EW_{LIS}}$ vs. $\mathrm{EW_{Ly\alpha}}$. The redshift samples are divided into four bins of increasing $\mathrm{EW_{Ly\alpha}}$, with $\mathrm{EW_{Ly\alpha}}$ and $\mathrm{EW_{LIS}}$ measured from composite spectra composed of the objects in each bin. The average measurements of the $z \sim 2$–5 samples are displayed as stars. **Bottom left panel:** $\mathrm{EW_{HIS}}$ vs. $\mathrm{EW_{Ly\alpha}}$. The binned samples are the same as in the top left panel, with the exception that the objects were required to have coverage of C IVλ1548, 1550. $\mathrm{EW_{HIS}}$ and $\mathrm{EW_{Ly\alpha}}$ are also measured from composite spectra, made from combining the individual spectra of the objects in each $\mathrm{EW_{Ly\alpha}}$ bin. **Top right panel:** $\mathrm{EW_{Ly\alpha}}$ vs. $E(B-V)$. The points are the redshift samples divided into four bins of increasing $E(B-V)$, with $\mathrm{EW_{Ly\alpha}}$ measured from composite spectra and $E(B-V)$ from the median of the objects in each bin. The transparent points are individual detections of the $z \sim 5$ sample. **Bottom right panel:** $\mathrm{EW_{LIS}}$ vs. $E(B-V)$. The binned samples are the same as the top right panel, with $\mathrm{EW_{LIS}}$ also measured from composite spectra.

We examine the strength of Lyα in bins of increasing $E(B-V)$ in top right panel of Figure 2.13. Individual detections of $E(B-V)$ and $\mathrm{EW}_{\mathrm{Ly}\alpha}$ are overplotted as transparent points. The trends shown by the solid purple line, corresponding to measurements post-IGM correction at $z \sim 5$, are consistent with the well-established inverse-correlation between $\mathrm{EW}_{\mathrm{Ly}\alpha}$ and $E(B-V)$ that pertains all the way down to $z \sim 0$ (Hayes et al., 2014). The $z \sim 2$–4 sample of D18 follows a similar decreasing trend, although the consistency breaks down at larger $E(B-V)$. Here, $\mathrm{EW}_{\mathrm{Ly}\alpha}$ is larger at fixed $E(B-V)$ in the $z \sim 5$ sample compared to lower redshift. We also calculate the correlation between the individual transmitted $\mathrm{EW}_{\mathrm{Ly}\alpha}$ and $E(B-V)$ measurements in our sample. We perform a Spearman correlation test, with a null hypothesis that $\mathrm{EW}_{\mathrm{Ly}\alpha}$ and $E(B-V)$ are uncorrelated in the $z \sim 5$ sample. The Spearman test produces a correlation coefficient of -0.27 with a p-value of 0.00045, strongly rejecting the null hypothesis. This test further supports that the inverse correlation between $\mathrm{EW}_{\mathrm{Ly}\alpha}$ and $E(B-V)$ continues out to $z \sim 5$.

The strength of $\mathrm{EW}_{\mathrm{LIS}}$ in bins of increasing $E(B-V)$ is presented in the lower right panel of Figure 2.13. The $z \sim 5$ trend also follows the one at $z \sim 2$–4, with larger $E(B-V)$ leading to a stronger LIS detection. This result supports the physical picture of dust grains existing in the same phase of gas as that attenuating Lyα and strengthening the LIS absorption (Reddy et al., 2016b). This trend also appears to be invariant out to $z \sim 5$. The question of which relation is most fundamental – LIS-$E(B-V)$ or LIS-Lyα – is difficult to surmise from these plots, but has important implications for the relationship between dust, neutral-phase gas, and escaping Lyα radiation from a galaxy. In Section 2.5, we re-examine the evolution of the relationships among Lyα, LIS, and $E(B-V)$ using finer redshift divisions at $z > 4$, and gain further insights into this important question.

2.4.3 Other Galaxy Properties

In addition to $E(B-V)$, we examine the strength of Lyα binned in terms of galaxy properties of M_*, SFR, age, and M_{UV}. These relationships out to $z \sim 5$ are presented in Figure 2.14.

Both the intrinsic and transmitted $EW_{Ly\alpha}$ are shown as solid and dashed lines, respectively, and the individual transmitted measurements for each object are shown as purple points.

Most of these relationships extend those that are present at $z \sim 2$–4. At $z \sim 5$, we find that stronger $EW_{Ly\alpha}$ emission occurs in fainter galaxies, albeit weakly, similar to the trend observed at $z \sim 4$. This trend is skewed towards stronger $EW_{Ly\alpha}$ emission at fixed brightness at $z \sim 5$. The relationship between M_{UV} and $EW_{Ly\alpha}$ flattens in the lower redshift samples. We find a similiar behavior in SFR space, where we find greater $EW_{Ly\alpha}$ emission in galaxies with lower SFR, and greater $EW_{Ly\alpha}$ at fixed SFR at $z \sim 5$ compared to $z \sim 2-4$. In stellar-mass space, we find stronger $EW_{Ly\alpha}$ emission in less-massive galaxies, but the trend does not appear to evolve beyond that at $z \sim 4$. Massive galaxies tend to be more dusty (McLure et al., 2018; Garn & Best, 2010; Whitaker et al., 2017), which would suppress their $EW_{Ly\alpha}$. We find no trend between $EW_{Ly\alpha}$ and age at $z \sim 5$, consistent with results at $z \sim 2$–4. This lack of trend could be explained by $EW_{Ly\alpha}$ reaching equilibrium quickly after the onset of a starburst (50-100 Myr, see Verhamme et al., 2008).

We also examine the correlation between $\Delta v_{Ly\alpha\text{–LIS}}$ and $EW_{Ly\alpha}$ and recover the anti-correlation of Shapley et al. (2003), indicating that a Lyα velocity peak closer to systemic indicates an easier pathway of escape for Lyα photons. This anti-correlation is weak, with a Spearman correlation test producing a correlation coefficient of -0.17 and a p-value of 0.14.

In summary, at $z \sim 5$, we find similar trends to those at $z \sim 4$: galaxies with stronger Lyα emission tend to have lower SFRs, less star formation, bluer UV continua, and less dust extinction. We also find stronger $EW_{Ly\alpha}$ emission at fixed galaxy property in the $z \sim 5$ sample.

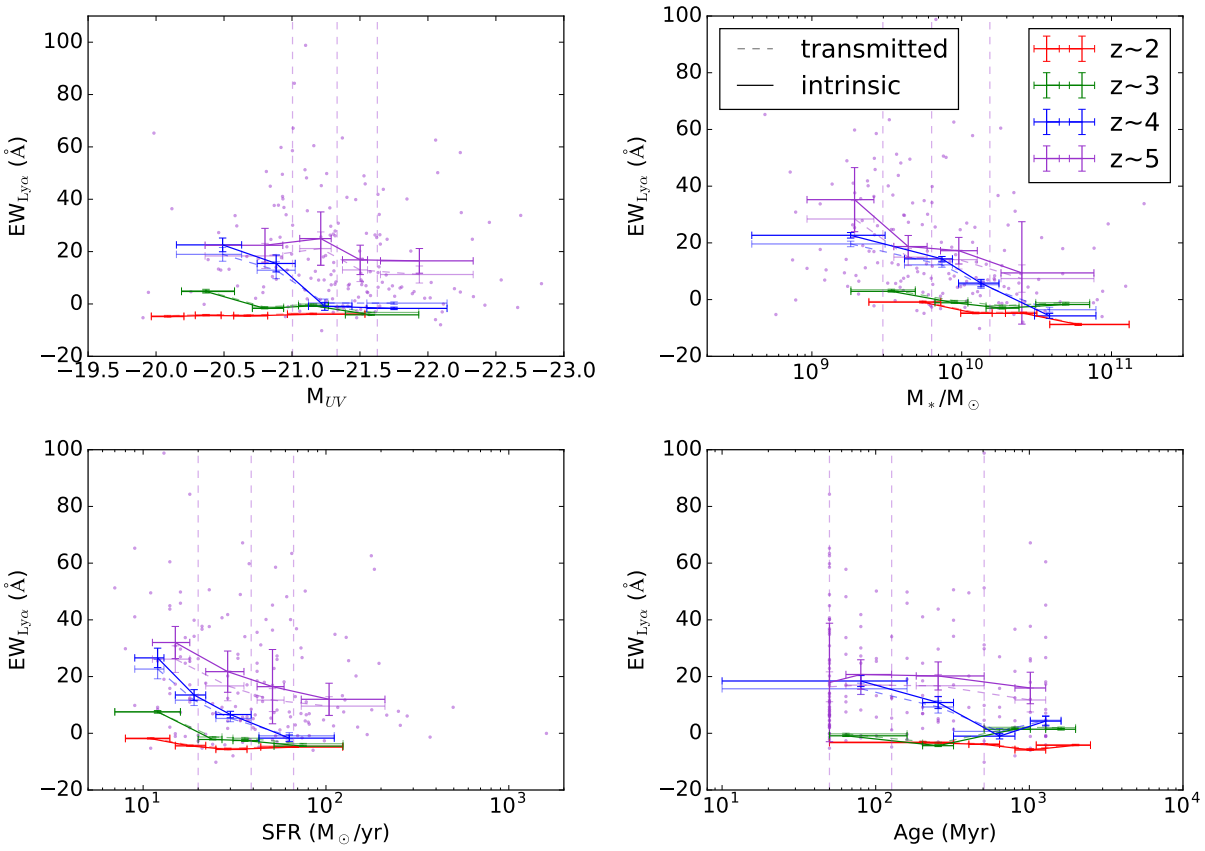

Figure 2.14: $EW_{Ly\alpha}$ as a function of integrated galaxy properties. Within each panel, each redshift sample is divided into four bins of an increasing galaxy property corresponding to that displayed on the x axis: clockwise from top left, M_{UV}, M_*, age, and SFR. The dashed lines connect measurements of $EW_{Ly\alpha}$ from the composite spectra and the median galaxy property of the objects in each bin. The solid lines represent the same measurements after applying the IGM transmission curves of Laursen et al. (2011). The purple points are the individual measurements of $EW_{Ly\alpha}$ and galaxy property determined in this work for the $z \sim 5$ sample. The vertical, purple, dashed lines represent the borders of the bins of the respective galaxy property.

2.5 Discussion

Cosmic reionization is the phase change of neutral to ionized hydrogen in the IGM, which appears to be roughly finished by $z \sim 6$ (Fan et al., 2006; McGreer et al., 2015). The leading theory is that reionization was caused by radiation from massive stars at $\lambda \leq 912$Å escaping from galaxies and ionizing the surrounding IGM (Finkelstein et al., 2012, 2019; Robertson et al., 2015). Due to the high opacity of the IGM at $z > 6$ to photons just below the Lyman limit, it is difficult to directly observe the fraction of leaking ionizing to non-ionizing radiation in high-redshift galaxies (Vanzella et al., 2012). However, the Lyα feature is well detected out to redshifts near the end of the epoch of reionization (i.e., $z \sim 6$; Stark et al., 2010; Schenker et al., 2014; De Barros et al., 2017) and while significantly more challenging, even into the epoch of reionization (e.g., Jung et al., 2018). By modeling the rest-UV spectra of LBGs to infer the intrinsic ionizing luminosity, Steidel et al. (2018) has shown that the fraction of escaping ionizing radiation in a galaxy at $z \sim 3$ is an increasing function of $\mathrm{EW}_{\mathrm{Ly}\alpha}$. This result supports a physical picture of both Lyα photons and ionizing radiation escaping through "holes" in the neutral ISM, similar to that discussed in Reddy et al. (2016b). This picture is empirically supported the Lyα profile and Lyman continuum emission of Ion2 in Vanzella et al. (2020). Thus, understanding the intrinsic Lyα strength of high-redshift galaxies can provide important clues to the evolution of leaking ionizing radiation, and consequently the epoch of reionization (Jung et al., 2018; Mason et al., 2018; Laursen et al., 2019).

2.5.1 Splitting $z \sim 5$ sample into low- and high-redshift

Given the size of the $z \sim 5$ sample (including over twice as many galaxies as in the $z \sim 4$ sample from Jones et al. (2012)) and our desire to obtain finer redshift sampling with a high-redshift bin that approaches the epoch of reionization, we isolate high-redshift subsamples within our $z \sim 5$ dataset. We divide the $z \sim 5$ sample at its median redshift. Our "low-redshift" $z \sim 5$ bin contains 87 galaxies with $z_{\mathrm{med}} = 4.34$, and our "high-redshift" $z \sim 5$ bin

contains 88 galaxies with $z_{\rm med} = 4.73$.

The strengths of interstellar metal lines and Lyα in the lower- and higher-redshift $z \sim 5$ subsamples are presented in the left panels of Figure 2.15. The purple $z \sim 5$ line of Figure 2.13 at $z_{\rm med} = 4.52$ is replaced by a magenta curve for the $z_{\rm med} = 4.34$ sample and a gold curve for the $z_{\rm med} = 4.73$ sample. At $z_{\rm med} \leq 4.34$, the positive correlation between $\rm EW_{LIS}$ and $\rm EW_{Ly\alpha}$ is preserved. This unchanging correlation reflects the fundamental connection between the amount of neutral gas and the attenuation of the Lyα line. In the $z_{\rm med} = 4.73$ sample, the trend begins to evolve at $\rm EW_{Ly\alpha} \geq 20$Å. We observe stronger Lyα emission at fixed LIS absorption strength.

The average properties of the composite spectra out to $z \sim 5$ are displayed as stars in the upper left panel of Figure 2.15. At $z_{\rm med} \leq 4.34$, the average Lyα emission strength increases and the average LIS absorption strength decreases with increasing redshift. This trend indicates that the neutral-gas covering fraction is decreasing with increasing redshift. This relationship appears to be different at $z_{\rm med} = 4.73$, where the average properties of the $z_{\rm med} = 4.73$ sample are offset from the trend of decreasing $\rm EW_{Ly\alpha}$ with increasing $\rm EW_{LIS}$ established at $z_{\rm med} \leq 4.34$. This result implies additional factors beyond neutral-gas covering fraction are influencing the average $\rm EW_{Ly\alpha}$ and $\rm EW_{LIS}$ of the $z_{\rm med} = 4.73$ subsample.

We further probe the state of the ISM near the epoch of reionization based on the mutual relationships among $\rm EW_{Ly\alpha}$, $\rm EW_{LIS}$, and $E(B-V)$. The upper right panel of Figure 2.15 shows stronger $\rm EW_{Ly\alpha}$ emission in bluer galaxies, similar to that of Figure 2.13, with the two $z \sim 5$ subsamples more clearly illustrating the evolution at $z_{\rm med} \geq 4.34$. Galaxies at $z_{\rm med} \geq 4.34$ have stronger Lyα emission at fixed $E(B-V)$. This evolution is not present in the relationship between $\rm EW_{LIS}$ and $E(B-V)$ shown in the lower right panel of Figure 2.15: stronger LIS absorption correlates with more reddening, invariant with redshift.

The redshift evolution of $\rm EW_{Ly\alpha}$ and other galaxy properties of Figure 2.14 also sheds light

on the physical differences of galaxies as we look back towards the epoch of reionization. The trends found in the single $z \sim 5$ sample analysis are more strongly demonstrated in Figure 2.16, based on the lower- and higher-redshift $z \sim 5$ subsamples. The pattern of lower $EW_{Ly\alpha}$ in brighter, more massive, highly star-forming galaxies is preserved across all redshift samples. This trend can also be seen down to $z \sim 0$ (Hayes et al., 2014). Additionally, the evolution towards greater $EW_{Ly\alpha}$ at fixed M_{UV}, SFR, and age with increasing redshift is shown more clearly with the $z_{med} = 4.34$ and $z_{med} = 4.73$ subsamples. While we don't see a correlation of increased $EW_{Ly\alpha}$ with increased specific star-formation rate (sSFR), we note that $EW_{Ly\alpha}$ is elevated in the $z_{med} = 4.73$ subsample compared to the $z_{med} = 4.34$ subsample at fixed sSFR, indicating that there are other galaxy properties affecting the strength of Lyα. On the other hand, the relationship between $EW_{Ly\alpha}$ and M_* does not evolve in the three highest redshift subsamples. The trends of $EW_{Ly\alpha}$ and M_{UV}, SFR, sSFR, and age reflect that $EW_{Ly\alpha}$ is increasing at fixed galaxy property, especially at $z_{med} \geq 4.34$. This is similar to the rest-UV spectral trends of Figure 2.15, where it is shown that $EW_{Ly\alpha}$ is increasing at fixed ISM/CGM property.

We also detect nebular C IV emission in the highest $EW_{Ly\alpha}$ quartile of the $z_{med} = 4.73$ subsample. This profile, shown in Figure 2.17, is similar to that found within the total $z \sim 5$ sample, shown in Figure 2.12. We discuss the implications of these detections for the inferred metallicity of the $z \sim 5$ sample in Section 2.5.3.

2.5.2 Intrinsic Lyα Production Evolution

The three factors that affect the observed strength of Lyα in a galaxy spectrum are (1) the intrinsic production rate of Lyα photons in HII regions, which is set by the efficiency of ionizing photon production for a given SFR, (2) the radiative transfer of Lyα photons through the ISM and CGM of the galaxy, and (3) the transfer of Lyα photons through the IGM. We have corrected for IGM opacity through application of the theoretical transmission curves of Laursen et al. (2011), which, calculated as an average of many sight lines through the IGM,

approximate the effects of the IGM on our stacked spectra composed of many galaxies. The anti-correlation between $\mathrm{EW_{Ly\alpha}}$ and $\mathrm{EW_{LIS}}$, which remains roughly constant from $z_{\mathrm{med}} = 2.3$ to $z_{\mathrm{med}} = 4.3$, can be seen as consequence of ISM and CGM radiative transfer. An increased covering fraction of neutral ISM and CGM gas both strengthens LIS absorption and weakens Lyα emission. Based on the lack of evolution in the relationship between $\mathrm{EW_{Ly\alpha}}$ and $\mathrm{EW_{LIS}}$, we argue that this correlation is driven by variations in covering fraction of interstellar and circumgalactic gas, and that the intrinsic typical Lyα production efficiency of samples at these redshifts is similar. This is supported by the distinction in Trainor et al. (2019) between quantities related to Lyα *production* and *escape* that jointly modulate observed Lyα strength. The authors argue that $\mathrm{EW_{LIS}}$ is strictly an escape-related quantity, supporting our conclusion that changing ISM/CGM properties should entirely define the $\mathrm{EW_{Ly\alpha}}$ vs. $\mathrm{EW_{LIS}}$ relationship assuming intrinsic Lyα production is constant.

The invariance of the $\mathrm{EW_{Ly\alpha}}$ vs. $\mathrm{EW_{LIS}}$ relation breaks down in the highest-redshift bin at $z_{\mathrm{med}} = 4.73$. In the highest bin of $\mathrm{EW_{Ly\alpha}}$, the $z_{\mathrm{med}} = 4.73$ sample differs significantly from the relationship defined at lower redshift. Additionally, at $z_{\mathrm{med}} = 4.73$, the average $\mathrm{EW_{Ly\alpha}}$ and $\mathrm{EW_{LIS}}$ deviate strongly from the linear relationship between $\mathrm{EW_{Ly\alpha}}$ and $\mathrm{EW_{LIS}}$ defined by the lower redshift samples. These deviations, which are at $\geq\sim 1\sigma$, suggest the presence of additional factors modulating the strength of Lyα emission, along with radiative transfer through the ISM, CGM, and IGM. Specifically, this difference suggests a significant increase in the intrinsic efficiency of Lyα photon production, and, correspondingly, ionizing radiation.

A scenario of increased intrinsic Lyα production efficiency among the strongest Lyα emitters at $z_{\mathrm{med}} = 4.73$ is additionally favored by the joint consideration of the $\mathrm{EW_{Ly\alpha}}$ vs. $\mathrm{EW_{LIS}}$ and $\mathrm{EW_{Ly\alpha}}$ vs. $E(B-V)$ relations. The dust and neutral-gas content in the ISM of a galaxy is intimately connected, with both thought to exist in the same physical location in the galaxy. The relationship between $\mathrm{EW_{LIS}}$ and $E(B-V)$ is thus expected to be invariant across redshift, as it is entirely defined by the neutral-gas covering fraction (and consequently

the amount of neutral-gas and dust) of the galaxy. The relationship in the bottom right panel of Figure 2.15 shows exactly this pattern. On the other hand, the redshift evolution towards stronger Lyα emission at fixed $E(B-V)$ shows that, in addition to the dust and neutral-gas content of the galaxy modulating the strength of Lyα, the intrinsic production rate also increases across redshift (assuming all else being equal). While our results suggest that the dust and neutral-phase gas exist co-spatially, the dust may be present in many locations throughout the ISM/CGM. Fully disentangling the interplay between $\mathrm{EW_{Ly\alpha}}$, $\mathrm{EW_{LIS}}$, and $E(B-V)$ will require further work (i.e., individual detections of $\mathrm{EW_{LIS}}$).

The observed difference in intrinsic Lyα photon production rate has important implications for understanding the contribution of galaxies to the reionization of the Universe. The comoving ionizing emissivity of the Universe can be modeled as the cosmic SFR (ρ_{SFR}) multiplied by the number of LyC photons produced per unit SFR (ξ_{ion}) multiplied by the escape fraction of LyC photons from the galaxy (f_{esc}). Physically, the EW of Lyα describes the amount of escaping Lyα emission relative to the non-ionizing UV continuum (analogous to ρ_{SFR} when averaged). The production of these Lyα photons in HII regions comes from recombination of protons and electrons, driven by LyC photons ionizing the hydrogen gas. Given that f_{esc} is directly connected to the properties of the ISM and CGM demonstrated by the relationships of $\mathrm{EW_{Ly\alpha}}$, $\mathrm{EW_{LIS}}$, and $E(B-V)$ at low redshift, an evolving $\mathrm{EW_{Ly\alpha}}$ at fixed $\mathrm{EW_{LIS}}$ and $E(B-V)$ indirectly measures the evolving ξ_{ion} with redshift. The evolution of ξ_{ion} is an important input into models of reionization (Bouwens et al., 2015; Stanway et al., 2016), and should not be taken as constant (e.g., Robertson et al., 2015).

2.5.3 Metallicity of the $z \sim 5$ Sample

The relationship between UV slope and UV luminosity evolves towards bluer colors at fixed luminosity towards higher redshift (Bouwens et al., 2014). The observed evolution towards decreasing median $E(B-V)$ from $z \sim 2$ to $z \sim 5$ in this work is consistent with these results. This evolution could reflect a fundamental change in dust and metal content of these

galaxies, or could reflect a bluer intrinsic stellar population at higher redshift. Spectroscopic observations are required to distinguish among different scenarios of dust, metallicity, or stellar population evolution.

The detection of nebular C IV emission in the highest bin of $EW_{Ly\alpha}$ at $z \sim 5$ provides independent evidence for significantly sub-solar metallicity. Nebular C IV emission is produced by extreme radiation fields in star-forming galaxies (Berg et al., 2019) and detections at lower redshift are typically found in metal-poor galaxies (Erb et al., 2010; Stark et al., 2014; Senchyna et al., 2017, 2019; Berg et al., 2019; Vanzella et al., 2016a, 2017). Stark et al. (2015) detected nebular C IV in one galaxy at $z = 7.045$, out of a sample of four $z \sim 6-8$ galaxies, whereas the typical C IV detection rate in UV-selected galaxies at $z \sim 2$–3 is only 1 per cent (Steidel et al., 2002; Hainline et al., 2011). A single, strong C IV detection in a sample of four $z \sim 6-8$ galaxies could imply a larger population of lower-metallicity, nebular C IV-emitting galaxies during the epoch of reionization. In our overall $z_{med} = 4.52$ sample, we see strong nebular C IV emission in the highest $EW_{Ly\alpha}$ bin (near rest-frame 1549Å in the bottom panel of Figure 2.12). In the $z_{med} = 4.73$ subsample, this feature is also prominent, as shown in Figure 2.17. A significant detection in a bin containing 25 per cent of the objects in the $z \sim 5$ sample indicates a changing population of galaxies at $z \sim 5$ with greater nebular C IV, harder ionizing fields, and lower metallicity. When compared to the highest $EW_{Ly\alpha}$ quartile of Steidel et al. (2002) at $z \sim 2$, we find an elevated nebular C IV to Lyα flux ratio of 4.0 ± 2.4 per cent and 2.3 ± 1.7 per cent in the $z_{med} = 4.52$ and $z_{med} = 4.73$ samples respectively, compared to ≤ 1 per cent at $z \sim 2$. The EW of this nebular C IV line was measured using the method in Section 2.3.4 as 1.98 ± 1.13Å and 1.61 ± 1.08Å in the $z_{med} = 4.52$ and $z_{med} = 4.73$ samples, respectively.

Systematically lower metallicity in galaxies near the epoch of reionization has important implications for ξ_{ion}. The connection between metallicity and ionizing photon production is more easily studied at lower redshift, where rest-optical metal lines can be observed with

ground-based optical telescopes. Shivaei et al. (2018) measured photon production efficiencies in star-forming galaxies at $z \sim 2$ and found galaxies with elevated $\xi_{\rm ion}$ have bluer UV spectral slopes, higher ionization parameters, and lower metallicities. Extreme galaxies at lower redshift with subsolar metallicity have also been demonstrated to have higher average $\mathrm{EW_{Ly\alpha}}$ (Erb et al., 2010). These observations at lower redshift paint a physical picture of galaxies with lower metallicity producing an intrinsically harder ionizing spectra per unit star-formation rate, emitting Lyα photons that are also less likely to be absorbed during multiple scatterings in less-dusty ISM/CGM. This picture is once again supported by the framing of Trainor et al. (2019), according to which the evolution of the average $\mathrm{EW_{Ly\alpha}}$ vs. $\mathrm{EW_{LIS}}$ properties across redshift reveals the changing *escape*-related properties of the ISM/CGM, while the evolving metallicity supports changing *production*-related parameters of the stellar population at high redshift. If these lower-metallicity, lower neutral-gas covering-fraction galaxies are more common at $z \sim 5$, as suggested by the evolving $E(B-V)$ and nebular C IV emission in our sample, they will contribute more strongly to reionization because of increased $\xi_{\rm ion}$ and $f_{\rm esc}$. Direct metallicity measurements at $z \geq 5$ are vital to disentangle and quantify these effects.

2.6 Summary

The rest-UV spectra of star-forming galaxies contain important information on star formation, dust and metal content of the ISM and CGM, and gas kinematics. Interpreting this information provides clues into the changing galaxy properties closer to the epoch of reionization. These properties include those related to the escape of ionizing radiation, such as neutral-gas covering fraction, and properties related to production of ionizing radiation, such as intrinsic Lyα production rate. To study these properties, we analyzed the relative strengths of Lyα, LIS, and HIS lines in a sample of $z \sim 5$ star-forming galaxies for the first time and compared them to similar analyses at $z \sim 2$–4 performed in D18. We also performed corrections for IGM attenuation on the Lyα profiles of sample composites using

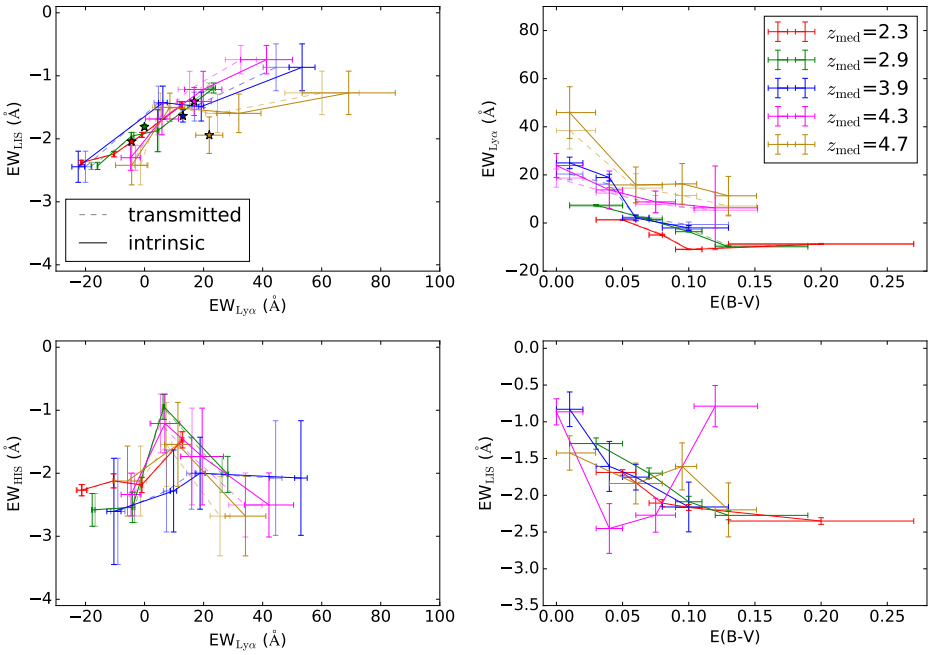

Figure 2.15: Relations between $\mathrm{EW_{Ly\alpha}}$, $\mathrm{EW_{LIS}}$, $\mathrm{EW_{HIS}}$, and $E(B-V)$ with the two redshift subsamples at $z \sim 5$. The magenta and gold points are the $z \sim 5$ sample divided into a lower and higher redshift sample of $z_{\rm med} = 4.34$ and 4.73. $\mathrm{EW_{Ly\alpha}}$, $\mathrm{EW_{LIS}}$, and $\mathrm{EW_{HIS}}$ were measured from composite spectra. The measurements from the observed composites are displayed in transparent points connected by dotted lines, while the corrected measurements after application of degraded IGM transmission curves are the opaque points connected by solid lines. The red, green, and blue points are measurements of the $z \sim 2$, 3, and 4 samples of D18. **Top left panel:** $\mathrm{EW_{LIS}}$ vs. $\mathrm{EW_{Ly\alpha}}$ with two redshift samples of $z \sim 5$. Each redshift sample was divided into four bins of increasing $\mathrm{EW_{Ly\alpha}}$. Both $\mathrm{EW_{Ly\alpha}}$ and $\mathrm{EW_{LIS}}$ are measured from composite spectra made from combining the individual spectra of the objects in each $\mathrm{EW_{Ly\alpha}}$ bin. Also included are average measurements of the composites of the $z \sim 2$–4 samples and the $z_{\rm med} = 4.34$ and 4.73 subsamples, displayed as stars. **Bottom left panel:** $\mathrm{EW_{HIS}}$ vs. $\mathrm{EW_{Ly\alpha}}$. The binned samples are the same as the top panel, with the exception that the objects were required to have coverage of C IVλ1548, 1550. $\mathrm{EW_{HIS}}$ and $\mathrm{EW_{Ly\alpha}}$ are also measured from composite spectra, made from combining the individual spectra of the objects in each $\mathrm{EW_{Ly\alpha}}$ bin. The $\mathrm{EW_{HIS}}$-$\mathrm{EW_{Ly\alpha}}$ measurement in the highest $\mathrm{EW_{Ly\alpha}}$ bin of the $z_{\rm med} = 4.73$ subsample was removed due to contamination from CIV emission. **Top right panel:** $\mathrm{EW_{Ly\alpha}}$ vs. $E(B-V)$. Each sample was divided into four bins of increasing $E(B-V)$, with $\mathrm{EW_{Ly\alpha}}$ measured from composite spectra and $E(B-V)$ from the median of the objects in each bin. The transparent points are individual detections. **Bottom right panel:** $\mathrm{EW_{LIS}}$ vs. $E(B-V)$. The binned samples are the same as the top panel, with $\mathrm{EW_{LIS}}$ also measured from composite spectra.

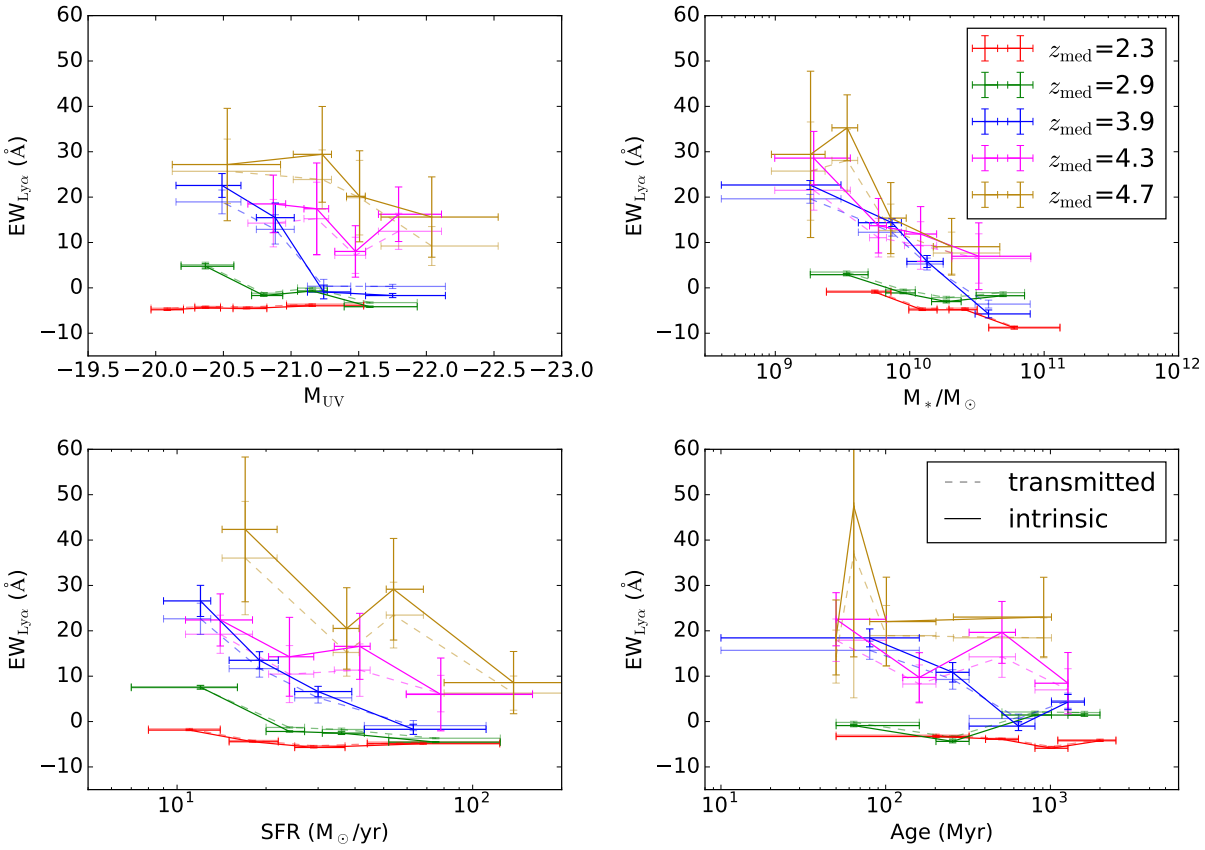

Figure 2.16: $EW_{Ly\alpha}$ as a function of integrated galaxy properties. The magenta and gold points correspond to the purple points of Figure 2.14 split into two redshift samples of $z_{med} = 4.34$ and 4.73. Within each panel, each redshift sample is divided into four bins of an increasing galaxy property corresponding to that displayed on the x axis: clockwise from top left, M_{UV}, M_*, age, and SFR. The dashed lines connect measurements of $EW_{Ly\alpha}$ from the composite spectra and the median galaxy property of the objects in each bin. The solid lines represent the same measurements after applying the IGM transmission curves of Laursen et al. (2011).

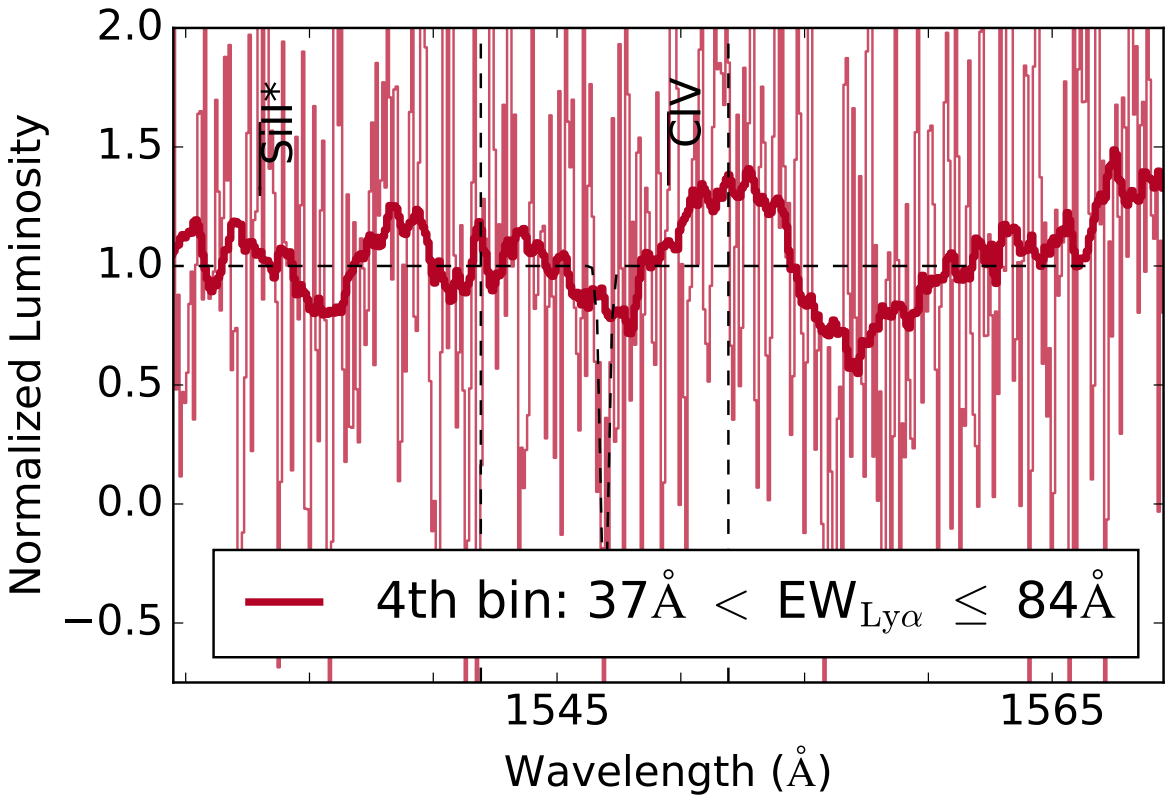

Figure 2.17: Nebular C IV emission bluewards of the interstellar C IV absorption feature of the $z_{\rm med} = 4.73$ subsample. Displayed is composite of the objects in the highest $\mathrm{EW}_{\mathrm{Ly}\alpha}$ quartile. As in Figure 2.12, the thin curve represents the continuum-normalized, rest-frame composite spectrum. The opaque spectrum is convolved with a boxcar window of width 3Å for easier viewing. The vertical, dashed lines illustrate the ±5Å integration windows from which EW measurements would have been made, but due to contamination from the nebular C IV emission, the interstellar C IV measurement was removed from the analysis.

transmission curves of Laursen et al. (2011). The key results are as follows:

1. We find that the redshift invariance of the relationship of EW_{LIS} vs. $EW_{Ly\alpha}$ deviates at $z_{med} = 4.52$, and even more significantly at $z_{med} = 4.73$, implying an evolving intrinsic Lyα photon production rate at $z \sim 5$ (assuming all else being equal). The EW_{LIS} vs. $EW_{Ly\alpha}$ relation illustrates the effects of the neutral-gas covering fraction on the rest-UV spectrum, and greater $EW_{Ly\alpha}$ at fixed EW_{LIS} at $z \sim 5$ implies a greater intrinsic Lyα production rate.

2. We find that the relationship of $EW_{Ly\alpha}$ vs. $E(B-V)$ evolves at higher redshift, with increased $EW_{Ly\alpha}$ at fixed $E(B-V)$ at $z \sim 5$, but no redshift evolution in the relationship of EW_{LIS} vs. $E(B-V)$, suggesting greater intrinsic Lyα production at fixed ISM/CGM properties using the non-evolution of EW_{LIS} and $EW_{Ly\alpha}$ as a control. The invariance of EW_{LIS} vs. $E(B-V)$ reflects the fundamental connection between neutral-phase gas content and dust in the ISM and CGM of star-forming galaxies at high redshift. However, we find greater $EW_{Ly\alpha}$ at fixed $E(B-V)$ at $z \sim 5$, implying for a given amount of dust and neutral gas, there are more Lyα photons being produced.

3. The robust detection of nebular C IV in the highest quartile of $EW_{Ly\alpha}$ objects within the $z_{med} = 4.52$ and the higher-redshift $z_{med} = 4.73$ subsample supports a scenario of significantly subsolar metallicity in ~ 25 per cent of our sample of star-forming galaxies near the epoch of reionization. Among non-AGNs, nebular C IV is typically found in a lower-metallicity population in galaxies from $z \sim 0-2$. Additionally, we find evolving UV slopes seen through the changing median $E(B-V)$ from z$\sim 2-5$, also consistent with a more metal-poor population of galaxies at higher-redshift.

4. The relationship of EW_{HIS} vs. $EW_{Ly\alpha}$ remains invariant with redshift out to $z \sim 5$, implying that ionized-phase and neutral-phase gas in the ISM and CGM remains in physically-distinct regions out to higher redshift.

5. Higher $EW_{Ly\alpha}$ is correlated with fainter M_{UV}, lower M_*, lower SFR, and younger age at $z \sim 5$, similar to the trends found at $z \sim 4$. We also see an evolution of greater $EW_{Ly\alpha}$ at fixed SFR, M_{UV}, and age at $z \sim 5$. This evolution of greater $EW_{Ly\alpha}$ at fixed galaxy property is consistent with the conclusion that the intrinsic Lyα production rate is increasing at $z \sim 5$.

These results have important implications for the understanding of the evolving galaxy population just after the completion of reionization. Higher intrinsic Lyα photon production rate implies a greater ξ_{ion} at $z \sim 5$, a vital input parameter to models of reionization. The analyses performed here are especially revealing *in* the epoch of reionization, in which the star-forming galaxies are thought to be the most significant contributers to the comoving ionizing emissivity of the universe. The upcoming launch of the *James Webb Space Telescope*, with its groundbreaking near-IR spectroscopic capabilities, will allow for the measurement of rest-UV spectral features out to $z \sim 10$, and we will discover whether the trends at $z \sim 5$ continue into the epoch of reionziation.

CHAPTER 3

An uncontaminated measurement of the escaping Lyman continuum at $z \sim 3$

3.1 Introduction

Reionization, the last major phase transition of the Universe, describes the ionization of the Hydrogen gas in the inter-galactic medium (IGM) in the first billion years of the Universe's history. Since the reionization process ends by $z \sim 6$ (Fan et al., 2006; Robertson et al., 2015; Planck Collaboration et al., 2016) and the number density of luminous QSOs drops off sharply beyond $z \sim 2$ (Shen et al., 2020), the prevailing notion is that ionizing radiation leaking from early star-forming galaxies drives the process of reionization (e.g., Bouwens et al., 2015; Parsa et al., 2018).

A number of observational constraints and challenges prevent a straightforward analysis of the galaxies that contribute strongly to the ionizing background during reionization. While we can chart the abundance of star-formation rates (SFRs) as a function of redshift (Madau & Dickinson, 2014), the difficulty lies in converting the cosmic SFR density ($\rho_{\rm SFR}$) into an ionizing emissivity, a key factor in dictating the evolution of the IGM neutral fraction (Bouwens et al., 2015; Robertson et al., 2015). One of the essential parameters for converting between $\rho_{\rm SFR}$ and the ionizing emissivity is $f_{\rm esc,abs}$, the absolute escape fraction of ionizing radiation produced in H II regions that evades absorption by neutral-phase gas and dust in the ISM. Unfortunately, it is not possible to constrain $f_{\rm esc,abs}$ directly based on measurements

during the epoch of reionization considering the drastic drop-off of transmission of ionizing rest-UV continuum through the IGM past $z > 3.5$ (Vanzella et al., 2012; Steidel et al., 2018). Instead, there has been a focus on sources analogous to the $z > 6$ reionizing population but observed at redshifts where the IGM is still transparent to Lyman Continuum (LyC) radiation at $\lambda \leq 912$ Å.

Many of the direct $f_{\rm esc,abs}$ measurements in the local Universe come from compact galaxies with large [O III]/[O II] ratios and high [O III]λ5007 equivalent widths, referred to as "Green Peas" (HST/COS, Borthakur et al., 2014; Izotov et al., 2016, 2018, 2021). At higher redshifts ($z \sim 3$), the properties of galaxies and their circumgalactic environments begin to align more with those in the reionization era. Meanwhile, the IGM opacity is still at a level that allows for the direct detection of LyC emission from star-forming galaxies. These aspects make this redshift range compelling for LyC surveys, but the sightline-to-sightline variability in IGM opacity makes the interpretations of individual LyC detections difficult. Thanks to advances in deep optical spectroscopy and narrowband imaging, there have been a handful of individual LyC detections at $z \sim 3$ (e.g., Mostardi et al., 2015; Shapley et al., 2016; De Barros et al., 2016; Bian et al., 2017; Vanzella et al., 2016b, 2017). In order to understand the LyC of the full population of galaxies at $z \sim 3$, however, larger and representative ensembles of galaxies with sensitive LyC measurements are required, which can smooth out the deviations in the transmission of the IGM through sample averaging.

To this end, the Keck Lyman Continuum Spectroscopic (KLCS) survey (Steidel et al., 2018, hereafter S18) utilized the LRIS double spectrograph on the Keck I telescope (Oke et al., 1995; Steidel et al., 2004) to observe 124 $2.75 \leq z \leq 3.4$ galaxies in the rest-UV with the goal of constructing spectral composites that represent global properties of $z \sim 3$ galaxies. S18 reported a sample-averaged absolute escape fraction of $f_{\rm esc,abs} = 0.09 \pm 0.01$ for their sample after performing spectroscopic contamination rejection as well as fitting Binary Model and Stellar Synthesis (BPASS; Eldridge et al., 2017) templates and physically-motivated ISM

models to rest-frame UV composite spectra. This sample-averaged $f_{\rm esc,abs}$ has a robust correction for IGM and CGM absorption due to the large sample number that went into the calculation. These results are complemented by other large-scale LyC surveys at $z \sim 3-4$, such the LACES survey of 61 $z \sim 3.1$ Lyα-emitting galaxies using WFC3/UVIS F336W imaging on *HST* (Fletcher et al., 2019) and the survey of 201 star-forming $z \sim 4$ galaxies drawn from the VUDS survey using VLT/VIMOS spectroscopy (Marchi et al., 2018).

One remaining limitation in interpreting the results of high-redshift LyC surveys is the risk of non-ionizing UV light from foreground objects masquerading as an apparent ionizing signal. From the ground, a $z \sim 3$ galaxy may appear as a coherent object with strong LyC flux density, but higher-resolution imaging can reveal a lower-redshift source nearby along the line of sight that contributes to the observed-frame ∼3600 Å spectrum. Vanzella et al. (2012) demonstrated the potential pitfalls of line-of-sight contamination in seeing-limited ground-based LyC observations. Building on these ideas, Mostardi et al. (2015, hereafter M15) examined 16 apparent $z \sim 2.85$ LyC leakers identified through ground-based narrowband imaging. Multi-band *Hubble Space Telescope* (*HST*) imaging was used to examine the photometric-redshifts of individual sub-components within the seeing-limited extent of each target. M15 found that 15 out of 16 of the apparent LyC detections were contaminated with foreground signal. The demonstrated efficacy of this methodology and the propensity for line-of-sight contamination highlights the importance of high-resolution, space-based follow up of LyC detection candidates at $z \sim 3$.

In the current work, we attempt to increase the confidence of the results of one of the largest $z \sim 3$ LyC surveys, the KLCS survey, by examining the strongest LyC sources with *HST* imaging. We specifically targeted the 15 KLCS galaxies with significant (3σ) individual detections of LyC flux in deep Keck/LRIS spectroscopy, as they significantly affect the global statistics of the sample. We simultaneously examined 24 additional galaxies in the sample that fall on the *HST* mosaics but lack individual LyC detections. In addition to searching

for foreground contamination in the KLCS sample, this work also serves as an update to S18 by revisiting the key results of the paper using a cleaned sample. Accordingly, our analysis yields a statistically robust, uncontaminated description of the LyC properties of $z \sim 3$ star-forming galaxies.

In Section 3.2, we describe the KLCS parent sample, the ground-based observation methodology, and the details of the *HST* observations. In Section 3.3, we present the reduction procedures, photometric analysis including the segmentation of objects with complex morphologies, and subsequent photometric measurements. In Section 3.4, we detail the search for foreground contamination and present the results of the contamination analysis. Finally, in Section 3.5, we report updated measurements of the LyC properties of the KLCS sample after contamination removal. We summarize our results in Section 3.6.

Throughout this paper, we adopt a standard ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70\ \mathrm{km\,s^{-1}Mpc^{-1}}$. We also employ the AB magnitude system.

3.2 Sample and Methodology

The KLCS survey (S18) utilized the Keck/LRIS spectrograph (Oke et al., 1995; Steidel et al., 2004) to obtain deep rest-UV spectra of an initial target sample of 137 galaxies selected as Lyman Break Galaxies (LBGs) at $2.75 \leq z \leq 3.4$ (for a description of $U_{\rm n}GR$ selection methodology, see Steidel et al., 2003; Adelberger et al., 2004; Reddy et al., 2012). The LRIS spectra of KLCS galaxies cover the Lyman continuum region (LyC, 880-910 Å), the Lyα feature, and far-UV interstellar metal absorption features redward of Lyα. The LRIS observations were taken between 2006 and 2008 across nine different survey fields with integration times per mask of $\sim$ 10 hr. Of the 137 galaxies observed, eight objects were removed due to instrumental defects or clear evidence for blending with nearby sources. The remaining 2D and 1D spectra were examined for evidence of spectral blending, where an additional redshift along the line of sight could indicate contamination in the apparent

LyC flux of a higher-redshift source. Five galaxies were removed due to apparent spectral blending, resulting in a final sample of 124 galaxies for analysis. S18 examined the ratio of ionizing to non-ionizing UV flux density within the KLCS sample, defined as the average flux density within $880 \leq \lambda_0/\text{Å} \leq 910$ (f_{900}) divided by the average flux density within $1475 \leq \lambda_0/\text{Å} \leq 1525$ (f_{1500}), or $\langle f_{900}/f_{1500} \rangle_{\rm obs}$. Fifteen galaxies were significantly detected ($f_{900} > 3\sigma_{900}$, where σ_{900} is the f_{900} measurement uncertainty) and were defined as the LyC detection sample. The remaining 109 galaxies with $f_{900} < 3\sigma_{900}$ were defined as the LyC non-detection sample.

Despite significant efforts to remove foreground contamination from the KLCS sample through examination of the 1D and 2D spectra, high-resolution imaging has previously been required to vet individual detections of leaking LyC (Vanzella et al. 2012; M15). In $z \sim 3$ LyC detections, apparent LyC leakage at ~ 3600 Å may actually originate from a low-redshift component that lies in projection within the angular extent of a clumpy galaxy morphology (e.g., Vanzella et al., 2016b; Siana et al., 2015). Thus, we require high-resolution, space-based imaging for spatially-resolved photometric-redshift analysis of KLCS LyC detections. These observations enable the consideration of individual sub-components as potential contaminants. We require imaging in multiple filters to judge the redshift of individual, extracted components. The *HST* $V_{606}J_{125}H_{160}$ filter set is ideal for this type of analysis (M15). At $z = 3.07$, the median redshift of the KLCS detection sample, the J_{125} and H_{160} band filters are situated on either side of the Balmer break. These filters are shown in Figure 3.1 along with the SED typical of $z \sim 3$ LBGs from the BPASS set of stellar-population models (Stanway & Eldridge, 2018) redshifted to $z = 3.07$. Also displayed is the same BPASS SED at $z = 1.50$, representing a low-redshift interloper. At $z \sim 3$, the $J_{125} - H_{160}$ color is expected to reflect the Balmer break, while at lower redshift, $J_{125} - H_{160}$ color is expected to be flatter as both bands probe redward of the break. Additionally, at $z \sim 3$, $V_{606} - J_{125}$ is sensitive to the rest-frame UV spectral slope, which provides information on the stellar age and reddening of the galaxy. At lower redshift, the Balmer break passes between the V_{606} and J_{125} bands.

We observed 14 KLCS LyC detection candidates with ACS/F606W (V_{606}), WFC3/F125W (J_{125}), and WFC3/F160W (H_{160}) on *HST* as a part of Cycle 25 Program ID 15287 (PI: Shapley). The observations took place between 2017 and 2019 in five separate LBG survey fields (Q0933, Q1422, DSF2237b, Q1549, and Westphal) across seven ACS pointings and 11 WFC3 pointings. The observations within each field were designed to cover the 14 LyC detection candidates with apparent LyC detections but no multi-band HST data, whereas Q1549-C25 had already been analysed by Shapley et al. (2016). We found the most efficient combination of pointings in each field that covered all apparent LyC detections with V_{606}, J_{125}, and H_{160}. Coincidentally, we covered with all three filters an additional 24 KLCS sources lacking individual LyC detections. An additional 11 KLCS LyC non-detections were covered by at least one filter. Each pointing was observed for three orbits per filter, or ~ 7000 s of exposure time. The total duration of the program was 87 orbits, required to ensure $>10\sigma$ detections in V_{606} for potential contaminant components, based on the faintest f_{900} (i.e., observed 3600 Å flux density) values of the detection sample and assuming (at worst) a flat SED between ~ 3600 Å and V_{606}. We executed each orbit using 4-point DITHER-BOX dither patterns for ACS/WFC pointings and 4-point DITHER-LINE patterns for WFC3/IR pointings. For one LyC detection, Q1549-D3, we included previously-acquired *HST* $V_{606}J_{125}H_{160}$ imaging from Cycle 20 Program ID 12959 (PI Shapley, M15). We included 5 orbits of V_{606} in a single pointing and 3 orbits of J_{125} and H_{160} each in a single pointing to cover this detection. We re-reduced these exposures with the same techniques applied to the more recent Cycle 25 data to provide consistency across the datasets.

3.3 Reduction

In this section, we detail the steps necessary to convert raw *HST* imaging to aligned images and photometric measurements of our targets. Many of the procedures in this section follow those detailed in M15.

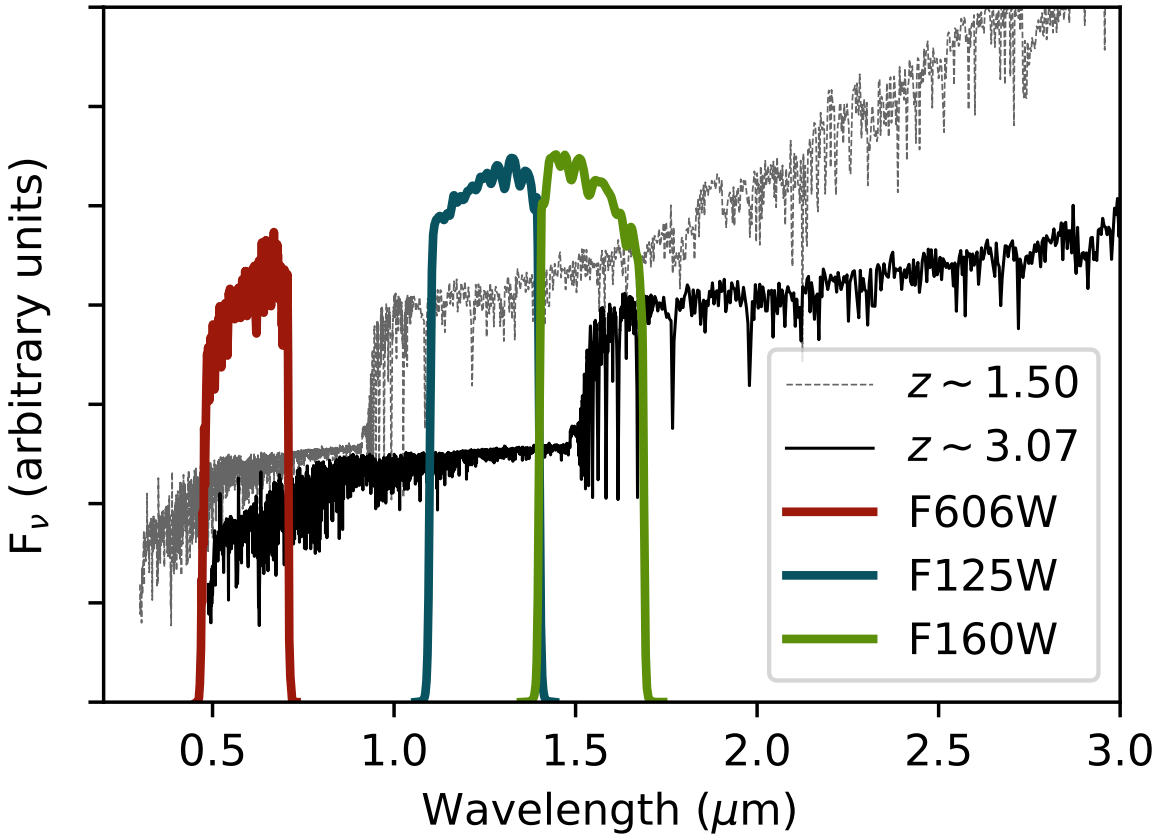

Figure 3.1: *HST* $V_{606}J_{125}H_{160}$ filter set in wavelength space in relation to a characteristic SED from BPASS (Stanway & Eldridge, 2018) shown at different redshifts. This SED was generated with 100Myr constant star-formation history and solar metallicity, and was reddened using $E(B-V) = 0.2$ and assuming a Calzetti et al. (2000) extinction curve. The black curve indicates the SED shifted to $z = 3.07$, the median redshift of the LyC detection sample. At this redshift, the $J_{125} - H_{160}$ color probes the 4000 Å Balmer break. The grey curve shows the SED shifted to $z = 1.5$, a redshift typical of low-redshift interlopers. At lower redshift, the $V_{606} - J_{125}$ probes the Balmer break.

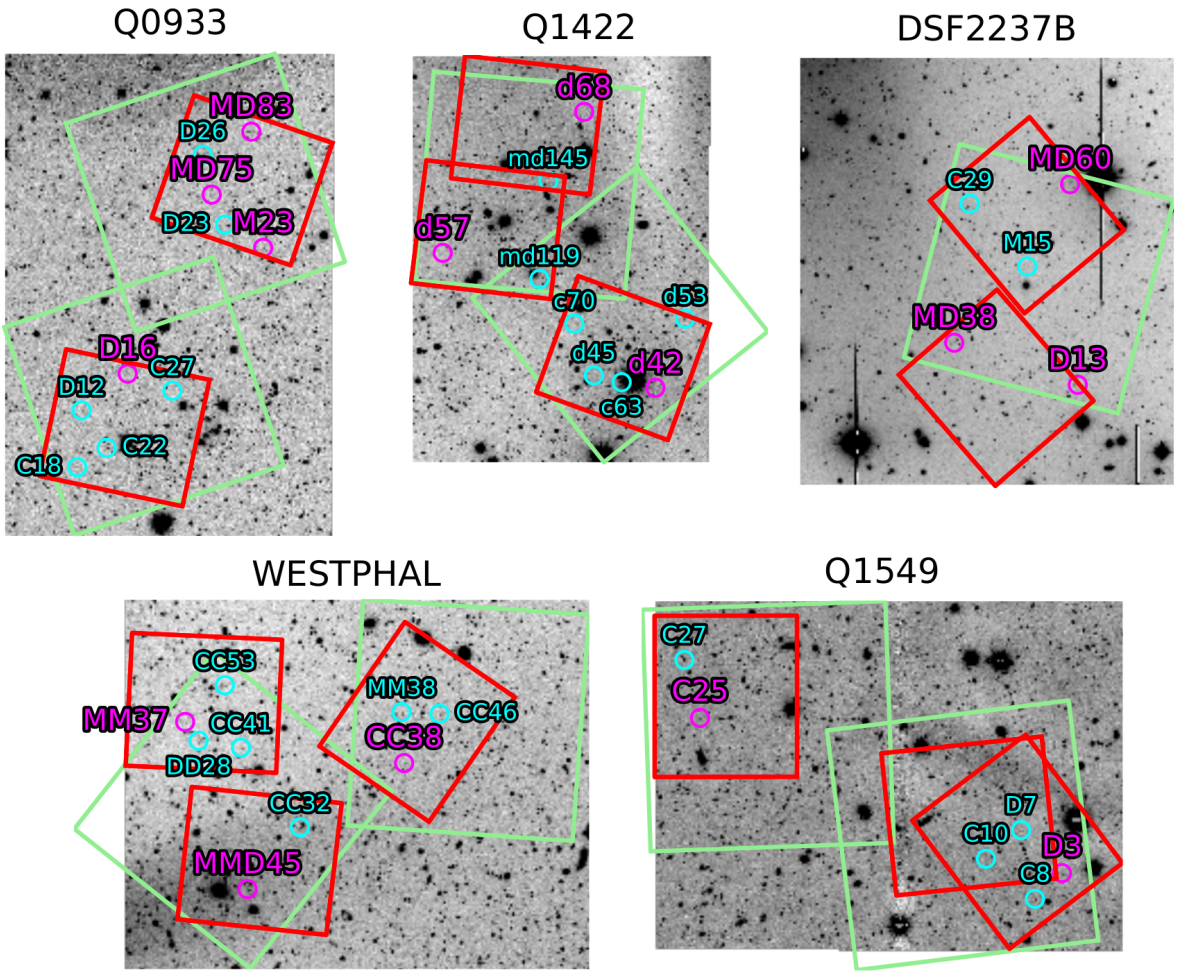

Figure 3.2: The five fields targeted by our HST observing program. In green are the $202'' \times 202''$ footprints of the ACS pointings, with imaging taken in the V_{606} band. In red are the $136'' \times 123''$ footprints of the WFC3-IR pointings, with imaging taken in J_{125} and H_{160} bands. The magenta circles represent the positions of the 14 LyC detection candidates, while the cyan circles represent the non-detections that lie in the same fields. One LyC detection candidate (Q1549-C25) was not the target of this observing program; the *HST* $V_{606}J_{125}H_{160}$ images are described in Mostardi et al. (2015) (also see Shapley et al., 2016).

3.3.1 Mosaic generation and PSF matching

First, the calibrated, flux-fielded, and charge-transfer-efficiency-corrected (for ACS) exposures were combined and aligned to produce a single mosaic for each filter and field observed. This reduction step was completed using the *DrizzlePac* software package (Fruchter et al., 2010). Contained in the package is the task *AstroDrizzle*, which performs automatic cosmic-ray rejection and sky subtraction during image combination, as well as the alignment task *TweakReg*. We used *TweakReg* to calculate residual shifts in RA and Dec. by inputting coordinate lists of extended sources that were first measured using *SExtractor* (Bertin & Arnouts, 1996). After calculating the shifts by comparing the coordinate lists, *TweakReg* performs corrections by re-registering the WCS of each exposure. The individual exposures within each pointing and filter did not require significant re-alignment to one another, save for the F160W imaging of the northern WFC3/IR pointing in the DSF2237b field, in which HST experienced a reaction-wheel failure partway through the exposures and required a $\sim$0.05″ adjustment.

After averaged mosaics were generated for each pointing and filter combination, the mosaics within each of the five target fields were aligned to one another and registered on the same WCS. The alignment RMS in both RA and Dec. between ACS images of different pointings was $\sim$0.005″ and the alignment RMS between WFC3/IR and ACS pointings was $\sim$0.01″. The final V_{606}, J_{125}, and H_{160} mosaics produced by *AstroDrizzle* were generated with matching pixel grids across filters at a pixel scale of 0.03″/px and a drop-to-pixel ratio (`final_pixfrac`) of 0.7 to maximize pixel resolution of the morphological features in the V_{606} imaging. These final mosaics were registered to the WCS of the ground-based UGR images (decribed in Reddy et al., 2012) available for these fields using the task *ccmap* in IRAF and manually-generated starlists of $\sim$50 unsaturated stars. The typical RMS in alignment between the *HST* mosaics and the ground-based G imaging in both RA and Dec. was $\sim$0.06″. The LyC detection candidates in both ground-based imaging and aligned *HST* imaging are shown in

the Figure 3.3.

For one object, Westphal-CC38, a small area of decreased sensitivity, or "blob" (Sunnquist, 2018), was in the vicinity of the object in the WFC3/IR images. To accurately measure the photometry of this object, we used the sky flats of Mack et al. (2021) that included corrections to the blobs present on the IR detector.

Next, the V_{606} and J_{125} mosaics were smoothed to the same angular resolution as the corresponding H_{160} mosaic to allow for accurate measurement of colors. Hereafter, we refer to the smoothed V_{606} and J_{125} mosaics as $sm(V_{606})$ and $sm(J_{125})$. The point-spread functions (PSFs) for each field and filter were generated using a set of 10-15 bright, unsaturated point sources identified manually within each mosaic. We combined these stellar profiles into PSFs using the *Astropy* routine *Photutils* (Bradley et al., 2020). The average Gaussian full width at half maximum (GFWHM) of the V_{606}, J_{125}, and H_{160} PSFs across all fields were $0.100''$, $0.177''$, and $0.183''$, respectively. We used the IRAF routine *psfmatch* to generate and apply kernels matching the V_{606} and J_{125} mosaics to the resolution of the H_{160} PSF. Curve of growth analysis of the $sm(V_{606})$ and $sm(J_{125})$ PSFs with the original H_{160} PSF demonstrated that the fraction of enclosed flux for all three agrees to within 1% at $0.5''$. We display the PSF-matched $sm(V_{606})$ and $sm(J_{125})$ and original-resolution H_{160} postage stamps of the LyC detection candidates as well as $V_{606}J_{125}H_{160}$ (for blue, green, and red) false-color images in Figure 3.3. Postage stamps of the non-detections can be found in the Appendix.

3.3.2 Photometry

In order to make comparable photometric measurements of each detection in multiple filters, we first must define regions corresponding to the distinct morphological features within each object. To this end, we used *SExtractor* on the PSF-matched V_{606} image to produce segmentation maps for each object, separating the light profile into different components. We set the detection threshold at 4 times the background RMS, but reduced this parameter

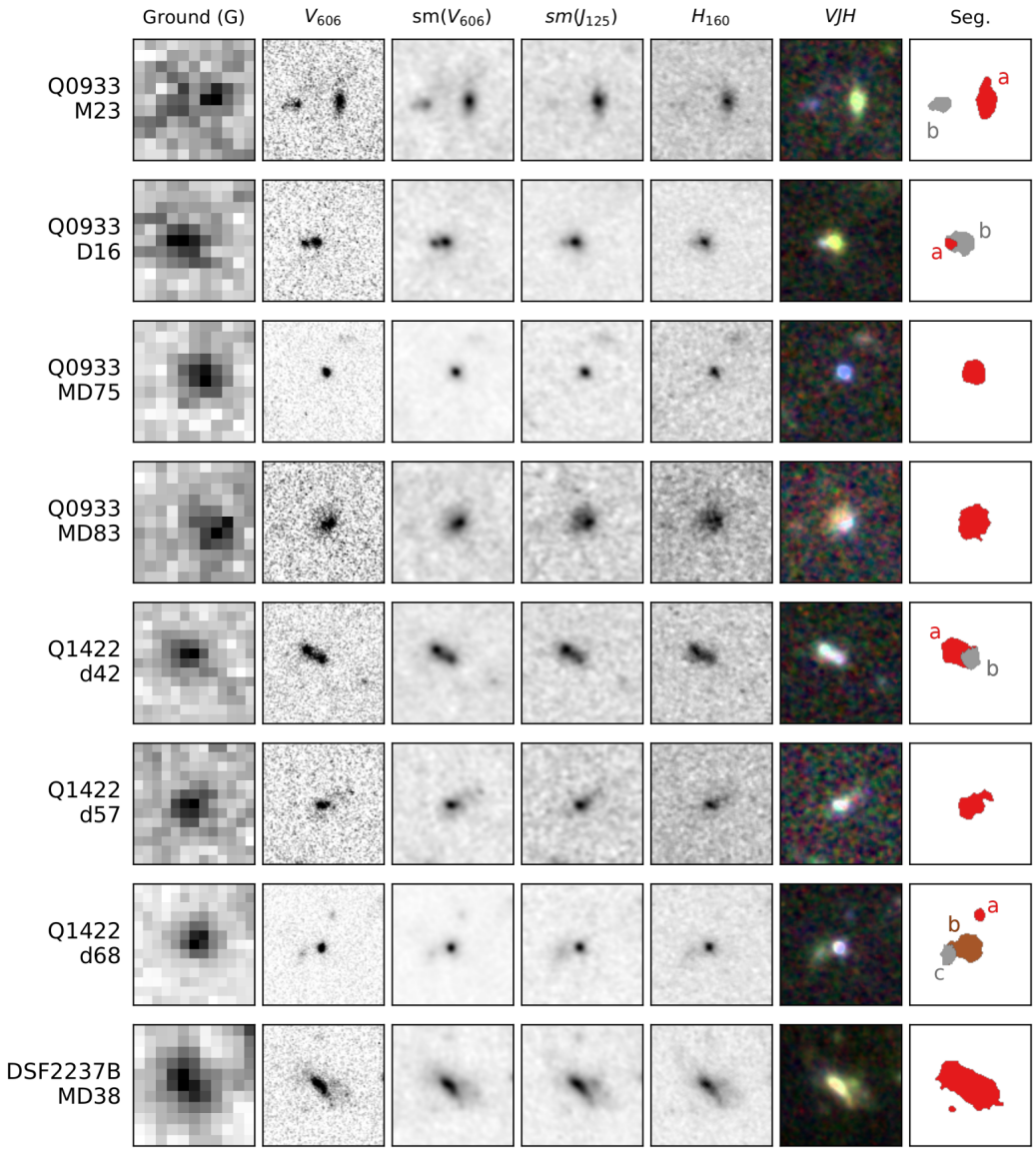

Figure 3.3: $3'' \times 3''$ postage stamps of the 15 LyC detection candidates targeted by HST. **First column:** Ground-based G (Reddy et al., 2012). **Second column**: V_{606} at the original resolution. **Third and fourth columns:** V_{125} and J_{125} smoothed to the lower resolution of H_{160}. **Fifth Column:** Original-resolution H_{160}. **Sixth Column:** False-color postage stamps. The $sm(V_{606})$, $sm(J_{125})$, and H_{160} images are represented by blue, green, and red, respectively. **Seventh Column:** segmentation map generated by SExtractor. Separate components extracted by the program are represented by different-colored regions.

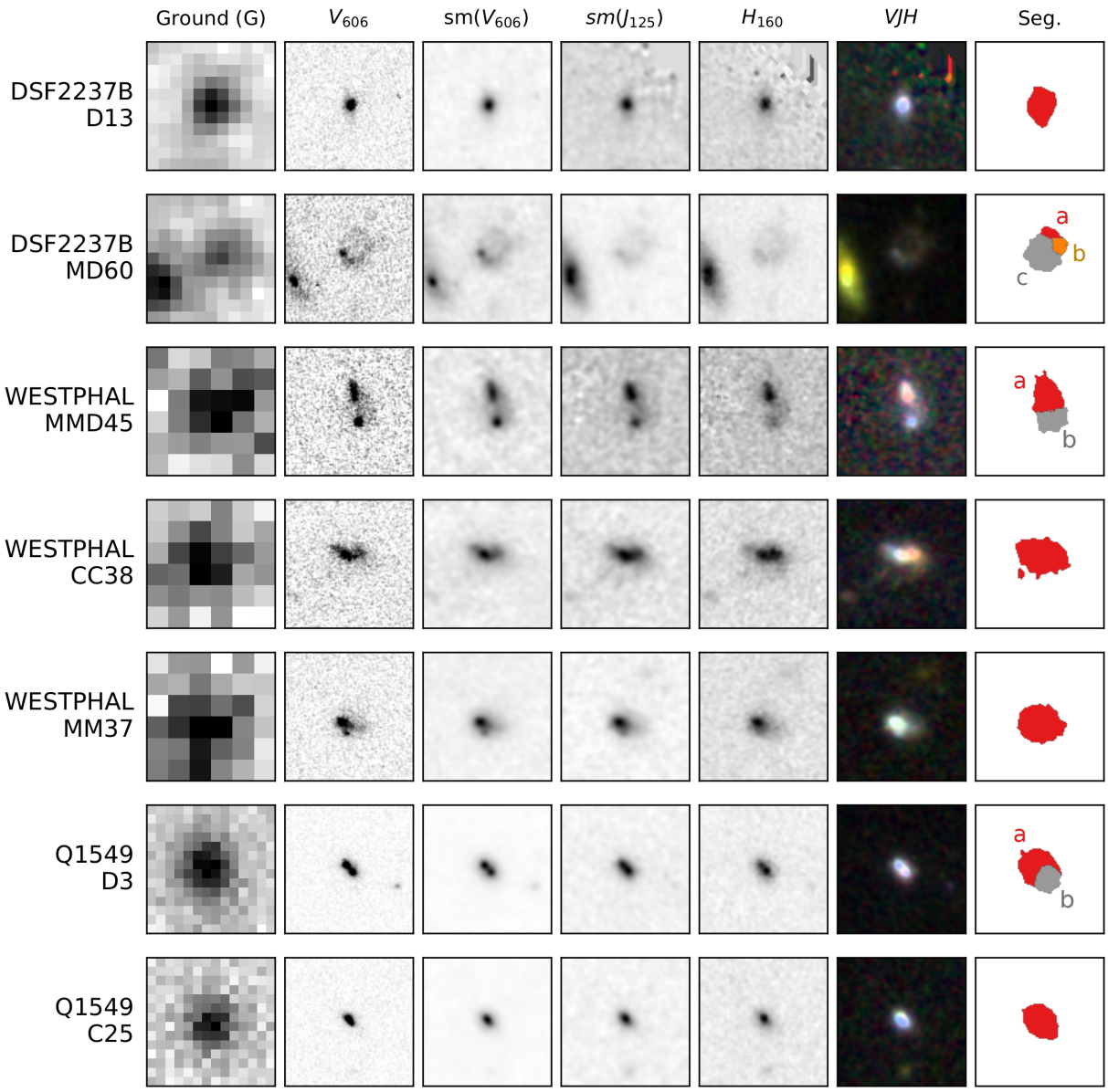

Figure 3.3: Continued.

down to 3 times the background RMS to achieve extraction of sub-components in individual cases. For background subtraction, we used the "local" method which uses a rectangular annulus around the source to estimate the sky level. We also set the minimum contrast parameter for deblending (`DEBLEND_MINCONT`) to 0 to maximize the number of identified segments.

We placed further requirements on what could be considered a separate component for our analysis. In order to remove spurious detections, we required a minimum area of 30 pixels, corresponding to fully sampling the PSF FWHM at the resolution of H_{160}. We also required each component to be brighter than or equal to 28.5 mag in V_{606} to eliminate all noisy and low ($< 10\sigma$) S/N detections. The segmentation maps generated for the LyC detection candidates are shown in the the last column of Figure 3.3. For the detection sample, eight were characterized by single component morphology, five were separated into two subcomponents, while the final two were characterized by three subcomponents. In the non-detection sample, 13 have a single component, 8 have double components, and 3 have three or more components.

For each subcomponent, we measured $V_{606}J_{125}H_{160}$ photometry from the PSF-matched mosaics. The AB photometric zeropoints of each field and filter combination were calculated using the inverse sensitivity and pivot wavelengths listed in the header of the exposures and were dust corrected using the Galactic reddening maps from Schlegel et al. (1998). The V_{606} photometry was measured isophotally by summing the V_{606} flux of the smoothed image within each region in the $sm(V_{606})$ segmentation map. We used the V_{606} segmentation map to measure the $J_{125}H_{160}$ photometry with *SExtractor* in dual-image mode.

We determined robust estimates of photometric error empirically. Analytical estimates of error, such as summing the error image calculated from an *AstroDrizzle* exposure-time weight map, do not take into account the covariance between adjacent pixels in averaged mosaics. To reflect this covariance, M15 adopted the empirical approach of Förster Schreiber et al.

(2006). The authors created blank apertures across the mosaics and measured the average signal RMS as a function of radius as a direct estimate of photometric errors. This method accurately captures correlated noise in our mosaics, but doesn't take into account the spatial deviations in the error image due to differences in exposure times and sky levels. We build upon the M15 method using the covariance correction of Law et al. (2016). We generated 1000 blank circular apertures in each mosaic in the range of sizes of our components with radii ranging from 0.2″ up to 1.0″. Within each aperture, we measured the ratio between the standard deviation of the signal within the blank aperture and the sum of the smoothed error image. This ratio is calculated as a function of angular size and indicates the correction of the theoretical error value from *AstroDrizzle* to an empirical one, taking into account correlated noise in the over-sampled and smoothed mosaics. For a given subcomponent, we take the ratio computed at the subcomponent's angular size and multiply it by the summed-in-quadrature error image within the $sm(V_{606})$ segmentation map as the 1σ error on the photometric measurement. These photometric measurements for each component in the detection sample can be found in Table 3.1 as well as the systemic redshift ($z_{\rm sys}$) from S18. Descriptions of the individually-undetected sample can be found in the Appendix. The average 3σ limiting magnitudes across our pointings for the $V_{606}J_{125}H_{160}$ mosaics calculated at 1.5x the average GFWHM of the PSF were 30.18, 29.12, and 28.72 mag, respectively.

3.4 Contamination Rejection

As demonstrated by M15, analysis of LyC detection candidates based on high-resolution multi-band *HST* imaging is an effective method for identifying low-redshift interlopers. In M15, photometric redshifts were estimated form the $U_{336}V_{606}J_{125}H_{160}$ magnitudes of each spatially-resolved component in the vicinity of a $z \sim 3$ LBG. The identification of foreground interlopers based on these component photometric redshifts led to the removal of 15 apparent LyC leakers at $z \sim 3$ from a sample of 16. These LBGs were all photometric LyC detections that had not yet been cleaned using deep spectroscopy, as described in Section 3.2

Table 3.1: Photometric measurements of the LyC detection subcomponents.

ID[a]	R.A.	Dec.	$z_{\rm sys}$[b,c]	V_{606}	J_{125}	H_{160}
DSF2237B-D13	22:39:27.42	+11:49:47.23	2.922	$24.52^{+0.03}_{-0.03}$	$24.95^{+0.04}_{-0.04}$	$24.72^{+0.04}_{-0.04}$
DSF2237B-MD38	22:39:35.49	+11:50:27.50	3.328	$24.00^{+0.03}_{-0.03}$	$23.78^{+0.02}_{-0.02}$	$23.57^{+0.02}_{-0.02}$
DSF2237B-MD60a	22:39:27.91	+11:53:00.23	3.141	$27.25^{+0.12}_{-0.11}$	$27.04^{+0.08}_{-0.07}$	$26.68^{+0.07}_{-0.07}$
DSF2237B-MD60b	22:39:27.90	+11:52:59.86	3.141	$26.96^{+0.11}_{-0.10}$	$26.78^{+0.07}_{-0.07}$	$26.46^{+0.07}_{-0.06}$
DSF2237B-MD60c	22:39:27.93	+11:52:59.65	3.141	$24.97^{+0.05}_{-0.05}$	$25.00^{+0.04}_{-0.04}$	$24.73^{+0.04}_{-0.03}$
Q0933-D16a	09:33:30.59	+28:44:53.54	3.047	$27.17^{+0.06}_{-0.06}$	$27.03^{+0.06}_{-0.06}$	$26.83^{+0.07}_{-0.07}$
Q0933-D16b	09:33:30.58	+28:44:53.58	...	$25.90^{+0.05}_{-0.05}$	$25.04^{+0.02}_{-0.02}$	$24.88^{+0.03}_{-0.03}$
Q0933-M23a	09:33:20.62	+28:46:54.76	...	$25.66^{+0.06}_{-0.05}$	$24.99^{+0.03}_{-0.03}$	$24.93^{+0.05}_{-0.05}$
Q0933-M23b	09:33:20.71	+28:46:54.70	3.289	$26.77^{+0.09}_{-0.08}$	$27.57^{+0.21}_{-0.18}$	$27.11^{+0.22}_{-0.18}$
Q0933-MD75	09:33:24.41	+28:47:46.24	2.913	$25.21^{+0.03}_{-0.03}$	$25.76^{+0.05}_{-0.05}$	$25.70^{+0.08}_{-0.08}$
Q0933-MD83	09:33:21.51	+28:48:46.70	2.880	$25.42^{+0.06}_{-0.05}$	$25.20^{+0.05}_{-0.05}$	$24.88^{+0.06}_{-0.06}$
Q1422-d42a	14:24:27.75	+22:53:51.01	3.137	$25.47^{+0.05}_{-0.04}$	$25.56^{+0.05}_{-0.05}$	$25.45^{+0.06}_{-0.05}$
Q1422-d42b	14:24:27.73	+22:53:50.84	3.137	$26.04^{+0.05}_{-0.05}$	$26.29^{+0.07}_{-0.06}$	$26.07^{+0.07}_{-0.06}$
Q1422-d57	14:24:43.25	+22:56:06.67	2.946	$25.68^{+0.06}_{-0.06}$	$25.82^{+0.08}_{-0.07}$	$25.65^{+0.08}_{-0.07}$
Q1422-d68a	14:24:32.92	+22:58:29.84	3.287	$27.76^{+0.12}_{-0.11}$	$28.52^{+0.30}_{-0.23}$	$28.50^{+0.39}_{-0.29}$
Q1422-d68b	14:24:32.94	+22:58:29.03	3.287	$24.86^{+0.03}_{-0.03}$	$25.15^{+0.04}_{-0.04}$	$24.91^{+0.04}_{-0.04}$
Q1422-d68c	14:24:32.98	+22:58:28.89	3.287	$26.81^{+0.09}_{-0.08}$	$26.49^{+0.07}_{-0.07}$	$26.65^{+0.10}_{-0.09}$
Q1549-C25	15:52:06.07	+19:11:28.44	3.153	$24.55^{+0.02}_{-0.02}$	$24.86^{+0.03}_{-0.03}$	$24.68^{+0.04}_{-0.03}$
Q1549-D3a	15:51:43.72	+19:09:12.50	2.937	$24.65^{+0.02}_{-0.02}$	$24.91^{+0.03}_{-0.03}$	$24.67^{+0.04}_{-0.03}$
Q1549-D3b	15:51:43.71	+19:09:12.36	2.937	$24.96^{+0.02}_{-0.02}$	$25.24^{+0.03}_{-0.03}$	$24.87^{+0.03}_{-0.03}$
Westphal-CC38	14:18:03.81	+52:29:07.23	3.073	$24.82^{+0.03}_{-0.03}$	$24.43^{+0.04}_{-0.04}$	$24.04^{+0.04}_{-0.04}$
Westphal-MM37	14:18:26.24	+52:29:45.03	3.421	$24.60^{+0.02}_{-0.02}$	$24.31^{+0.02}_{-0.02}$	$24.17^{+0.02}_{-0.02}$
Westphal-MMD45a	14:18:19.81	+52:27:09.45	2.936	$25.64^{+0.04}_{-0.04}$	$25.54^{+0.05}_{-0.05}$	$25.04^{+0.04}_{-0.04}$
Westphal-MMD45b	14:18:19.80	+52:27:08.77	2.936	$26.16^{+0.05}_{-0.05}$	$26.18^{+0.07}_{-0.06}$	$25.96^{+0.08}_{-0.07}$

[a] The field the object is located in, the object name, and a letter corresponding to the subcomponent in Figure 3.3. A subcomponent label is omitted in the case of single-component morphology. [b] Systemic redshift from S18. [c] Systemic redshifts are omitted for components determined to be at low redshift based on their $V_{606}J_{125}H_{160}$ colors.

for the KLCS sample. For the current analysis, our photometric measurements are limited to $V_{606}J_{125}H_{160}$, thus preventing us from making well-constrained photometric-redshift fits for each component. However, we are motivated to use $V_{606}J_{125}H_{160}$ photometry to predict contamination by the clear separation of low-redshift contaminants and confirmed high-redshift components of the M15 sample in $V_{606}J_{125}H_{160}$ (i.e., $V_{606} - J_{125}$ vs. $J_{125} - H_{160}$) color-color space (Figure 3.4, upper left), and the positioning of the $V_{606}J_{125}H_{160}$ filters relative to the Balmer break at $z \sim 3$ (Section 3.2).

3.4.1 $V_{606}J_{125}H_{160}$ color-color diagram

In order to explore the validity of using the $V_{606}J_{125}H_{160}$ color-color diagram to predict foreground contamination, we examined the distribution of low- and high-redshift galaxies in $V_{606}J_{125}H_{160}$ color-color space using the photometric catalogs of the 3D-HST Survey (Brammer et al., 2012; Skelton et al., 2014). We first analysed galaxies with spectroscopically-confirmed redshifts in the redshift range of our KLCS *HST* sample, $2.75 < z_{spec} < 3.4$, as a proxy for uncontaminated $z \sim 3$ objects. We considered all galaxies with spectroscopic redshifts lower than our KLCS *HST* sample, $z_{spec} < 2.75$, as a representation of potential foreground contaminants. We display the low-redshift and $z \sim 3$ populations alongside our detection and non-detection subcomponents in $V_{606}J_{125}H_{160}$ color-color space in the upper-right panel of Figure 3.4. Similar to the contaminant and $z \sim 3$ populations of M15, the low-redshift and $z \sim 3$ samples of 3D-HST fall in distinct regions of $V_{606}J_{125}H_{160}$ color-color space. By using the 3D-HST sample to examine where different redshift populations of galaxies lie, we can predict whether an individual component in our analysis is a low-redshift interloper or the $z \sim 3$ source measured by KLCS spectroscopy.

We adapted an empirical approach to rejecting low-redshift contamination in our sample of LyC detections using different 3D-HST redshift samples. Here, we used a sample of galaxies with either photometric or spectroscopic redshifts in 3D-HST, prioritizing spectroscopic redshifts if a galaxy had both. Including galaxies with photometric redshift measurements

better samples the $V_{606}J_{125}H_{160}$ dynamic range of our targets and is more complete than a spectroscopic sample weighted towards either UV selected galaxies (Steidel et al., 2003; Reddy et al., 2012) or those with strong rest-optical emission lines measured through the WFC3 grism (Momcheva et al., 2016; Bezanson et al., 2016). We again built "contaminant" and "$z \sim 3$" samples with the redshift ranges of $2.75 < z < 3.4$ and $z < 2.75$, respectively. These 3D-HST samples are shown with the KLCS *HST* subcomponents overlaid in the lower-left panel of Figure 3.4. We translated these sample distributions into a quantitative, predictive model using the Gaussian kernel-density estimation (KDE) method (Parzen, 1962; Rosenblatt, 1956). This analysis assigns each $V_{606}J_{125}H_{160}$ datapoint a 2D Gaussian profile in color-color space. These Gaussians are summed together to estimate the underlying 2D probability density function (PDF) of the distribution of data points. This PDF can be used to estimate the likelihood that a random galaxy drawn from the sample would have a specific $V_{606}J_{125}H_{160}$. We constructed PDFs using the KDE method for our contaminant and $z \sim 3$ samples using the *SciPy* routine *gaussian_kde* with bandwidth determined by Scott's Rule (Virtanen et al., 2020; Scott, 2015). These PDFs are displayed in the lower-right panel of Figure 3.4, each normalized to one.

We additionally examined the effects of the photometric errors of our measurements and the 3D-HST samples in the constructed PDFs. We generated 100 realizations of the 3D-HST contaminant and $z \sim 3$ samples randomly perturbed by their $V_{606}J_{125}H_{160}$ errors. We evaluated the resulting PDFs at the $V_{606}J_{125}H_{160}$ colors of our subcomponents, resulting in 100 probabilities from the contaminant and $z \sim 3$ PDFs for each component. We then took the median of these probabilities and compared them on an component-to-component basis. If the median likelihood drawn from the contaminant PDF was higher than that from the $z \sim 3$ PDF, the component was assigned as a contaminant. If the $z \sim 3$ median likelihood was higher, the component was confirmed to be at $z \sim 3$.

We found that two apparent LyC detections (Q0933-M23 and Q0933-D16) contained com-

ponents that had a higher likelihood of being drawn from the contaminant sample than the $z \sim 3$ sample. We also found one LyC non-detection (Q1422-md145) that was similarly consistent with being contaminated. As a sanity check, we examined the predictions of our targets that were least likely to be contaminated, the LyC non-detections with single-component morphology. All of these objects had components consistent with being at $z \sim 3$. We also examined the predictions from estimating PDFs only from 3D-HST galaxies with spectroscopically-confirmed redshifts and found them to be consistent with predictions from the larger photometric-redshift 3D-HST samples.

Using the $V_{606}J_{125}H_{160}$ PDF estimated from Gaussian KDE to predict sample membership is built on a number of assumptions. By normalizing each PDF to one and comparing the probability values at a given $V_{606}J_{125}H_{160}$ directly, we assumed that each component has equal likelihood of being drawn from either distribution. We consider this assumption to be conservative, considering at least one component must be contributing spectroscopic signal consistent with $z \sim 3$, and the rigorous spectroscopic-blending rejection performed by S18.

We also assumed that the 3D-HST contaminant and $z \sim 3$ distributions we've constructed are representative of possible foreground contaminants and uncontaminated $z \sim 3$ KLCS galaxies. To examine this assumption, we constructed $V_{606}J_{125}H_{160}$ PDFs that were individually tailored to the properties of each component. For each subcomponent, a $z \sim 3$ and contaminant 3D-HST sample were constructed with galaxies of similar effective radii and V_{606} magnitude. 3D-HST effective radii measurements were drawn from Van Der Wel et al. (2012) with areas required to be within 2σ of the size of the *SExtractor* V_{606} isophotal area of the subcomponent, while V_{606} magnitudes were required to be within 1σ of the subcomponent V_{606} magnitude. These cuts, defined using the Van Der Wel et al. (2012) errors on effective radius and our photometric-error estimate of Section 3.3.2, were set to ensure that both the $z \sim 3$ and contaminant PDFs were well sampled. We bootstrapped the sample to generate median probability predictions in the method described above. The effects of

cutting the 3D-HST samples based on effective radius and V_{606} magnitude are shown for the two-component object Q0933-D16 in Figure 3.5. Here, we restricted the 3D-HST samples to have a median V_{606} magnitude of 27.17 mag and 25.90 mag and median effective radii of $0.17''$ and $0.29''$ for subcomponents a and b, motivated by the properties of the components. These cuts changed the shape of the contaminant and $z \sim 3$ PDFs, but ultimately did not affect the classification of component a at $z \sim 3$ and component b as a contaminant. We found that the only object to be classified differently by this method was the non-detection Q1422-d53, now predicted to contain low-redshift components. We display the 3D-HST $z \sim 3$ and low-redshift PDFs after Q1422-d53 brightness and size cuts in Figure 3.6. Using these modified PDFs, components d, e, and f were predicted to be low-redshift. We considered this method to be more physically-motivated and remove Q1422-d53 from the non-detection sample.

Finally, we consider the potential effects of contamination from $2.75 < z \lesssim 3.0$ interlopers of galaxies with spectroscopic redshifts at the high-redshift end of our sample ($3.0 < z < 3.4$). Out of the 13 individual LyC detection candidates remaining once Q0933-D16 and Q0933-M23 are removed, this type of contamination is relevant for the three multicomponent galaxies, DSF2237b-MD60, Q1422-d42, and Q1422-d68 (i.e., the other 10 galaxies are either single component or at $z < 3.0$). We rebuild the 3D-HST "contaminant" sample to include higher-redshift galaxies, changing the redshift range to $z < (z_{\rm sys} - 0.1)$, where $z_{\rm sys}$ is the systemic redshift of the LBG in question. Subsequently, the 3D-HST "$z \sim 3$" sample was modified to have a redshift range of $(z_{\rm sys} - 0.1) < z < 3.4$. Using this method, we found no difference in the classification of subcomponents for DSF2237b-MD60, Q1422-d42, and Q1422-d68. Specifically, all subcomponent colors are consistent with being drawn from the $(z_{\rm sys} - 0.1) < z < 3.4$ population.

In summary, we removed two apparent LyC detections (Q0933-D16 and Q0933-M23) and two LyC non-detections (Q1422-md145 and Q1422-d53) from the KLCS sample based on evidence

of contamination from the $V_{606}J_{125}H_{160}$ colors of their morphological subcomponents.

3.4.2 2D spectra

In addition to examining the positions of the KLCS *HST* components in $V_{606}J_{125}H_{160}$ color-color space, we analysed the high-resolution, unsmoothed V_{606} images in tandem with the 2D LRIS spectra to search for new evidence of contamination. We examined the 2D spectra of the four objects predicted to be contaminated in Section 3.4.1. For Q0933-D16, Q1422-md145, and Q1422-d53, the subcomponents were either aligned along the slit or narrowly separated, making spectral deblending impossible. In Q0933-M23, we found evidence for foreground contamination in the LRIS spectrum shown in Figure 3.7. Here, the Lyα feature used for spectroscopic redshift measurement is clearly offset from the continuum towards the component predicted to lie at $z \sim 3$ by our color-color method. There also exists an additional spectral feature offset toward the predicted low-redshift component, which was classified as Lyα at $z = 3.380$ by S18. Based on the high-resolution photometric information offered by *HST* and the lack of other offset spectral features, it is more likely this emission line is [O II]λ3727 at $z = 0.43$.

3.5 Re-measurement of global properties

In this section, we revisit the results and conclusions of S18 with a $z \sim 3$ rest-UV spectroscopic sample free from significant foreground contamination in the LyC spectral range. With the removal of two apparent LyC detections, Q0933-D16 and Q0933-M23, whose emission is actually contaminated by lower-redshift interlopers, global properties of the sample will be unbiased by line-of-sight contamination and will present a more accurate picture of the global ionizing properties of galaxies in the $z \sim 3$ Universe. We also remove the LyC non-detections Q1422-md145 and Q1422-d53 due to evidence of low-redshift subcomponents within their *HST* morphology. Our rate of significant LyC detections for individual objects ($f_{900} > 3\sigma_{900}$) has been modified from 15/124 in S18 to 13/124 in this work, with a fore-

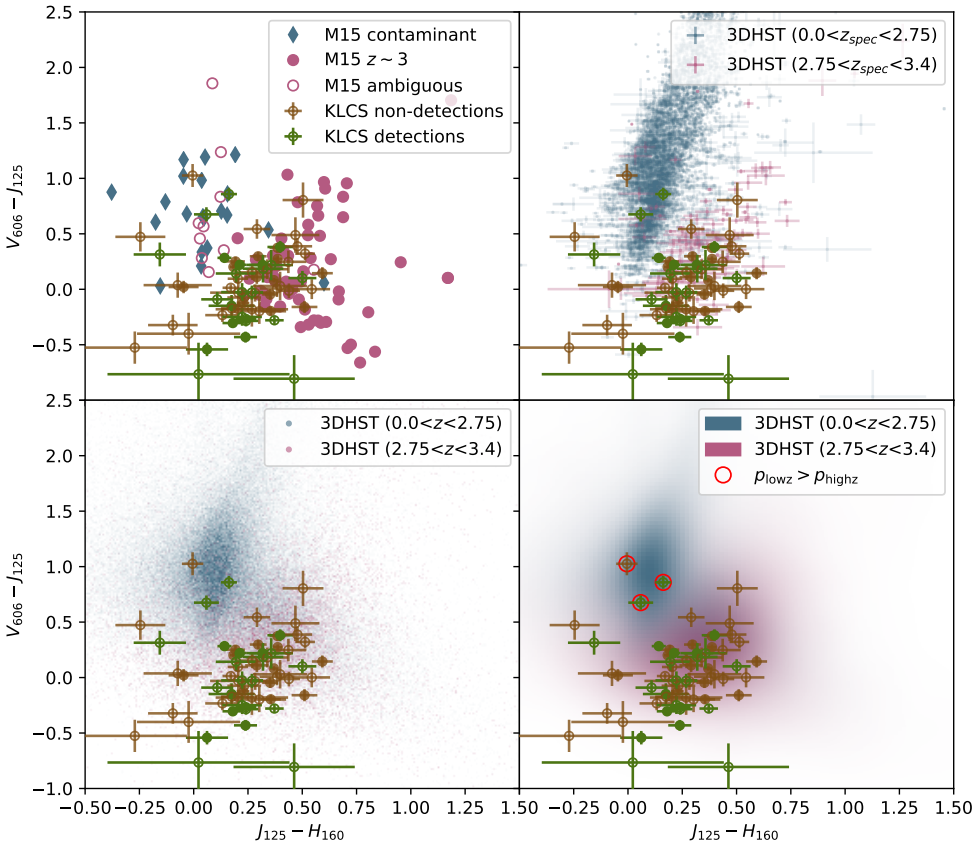

Figure 3.4: **Upper Left:** $V_{606}J_{125}H_{160}$ color-color diagram of LyC detection candidates from Mostardi et al. (2015). The contaminants, selected by photometric-redshift fitting, are displayed as blue diamonds. Subcomponents predicted at $z \sim 3$ are displayed as solid pink circles, while components with ambiguous SEDs are displayed as unfilled pink circles. Also displayed are all subcomponents extracted from the photometry of the KLCS *HST* sample, separated into LyC detections in green and non-detections in brown. Open circles represent subcomponents from multi-component objects, while filled circles represent single-component objects. **Upper Right:** The subcomponents of the KLCS *HST* sample overlaid with spectroscopic redshift samples of 3D-HST. The contaminant sample ($z_{spec} < 2.75$) is displayed in blue while the $z \sim 3$ ($2.75 < z_{spec} < 3.4$) sample is displayed in pink. **Lower Left:** The subcomponents of the KLCS *HST* sample overlaid with photometric and spectroscopic redshift samples of 3D-HST. The redshift ranges of the two samples are identical to the spectroscopic redshift samples. Error bars of the 3D-HST samples are removed for clarity. **Lower Right:** The subcomponents of the KLCS *HST* sample overlaid with the PDFs estimated from the 3D-HST samples of the lower-left panel. Each PDF was constructed using the Gaussian KDE method and were normalized to one. Components of the LyC detections Q0933-M23 and Q1422-D16 and a component of the LyC non-detection Q1422-md145 can be seen in the peak region of the contaminant (blue) PDF and are highlighted in red.

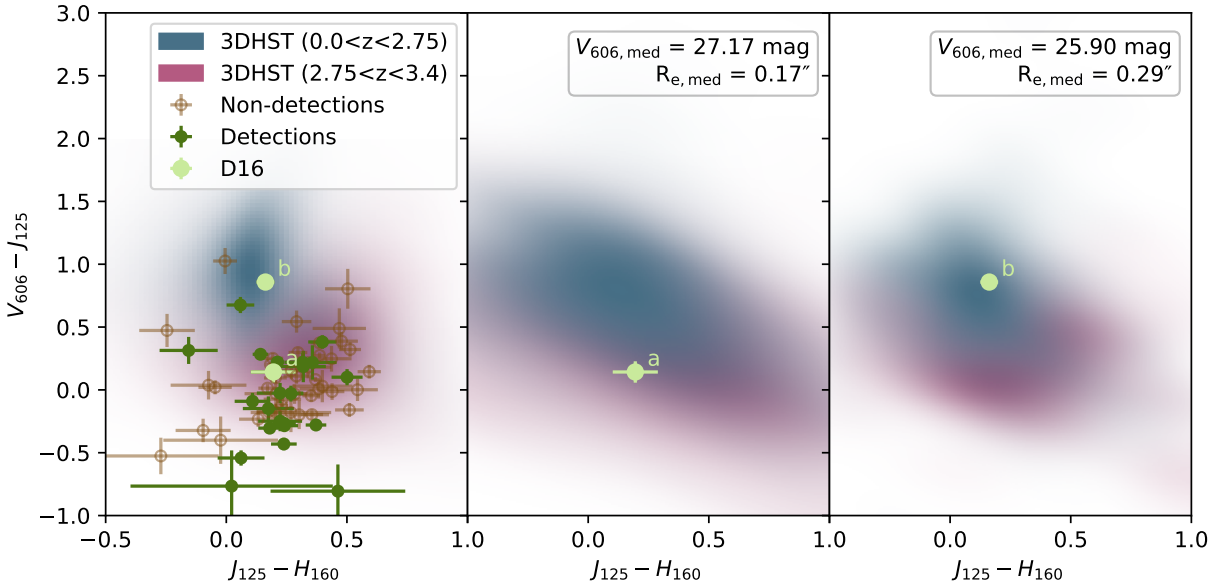

Figure 3.5: PDFs generated by the Gaussian KDE method on samples of 3D-HST galaxies at similar UV magnitudes and effective radii to Q0933-D16a and Q0933-D16b. **Left:** PDFs generated from the full 3D-HST redshift samples in $V_{606}J_{125}H_{160}$ color-color space. The Q0933-D16a and Q0933-D16b subcomponents are highlighted. **Middle and Right:** PDFs generated from 3D-HST redshift samples with similar V_{606} apparent magnitude and effective radii as the subcomponents of Q0933-D16. The median V_{606} apparent magnitude and R_e are listed in the legend and are comparable to those of the components (see Table 3.1). The same redshift ranges as the left panel were used for the low-redshift (blue) and $z \sim 3$ (pink) distributions.

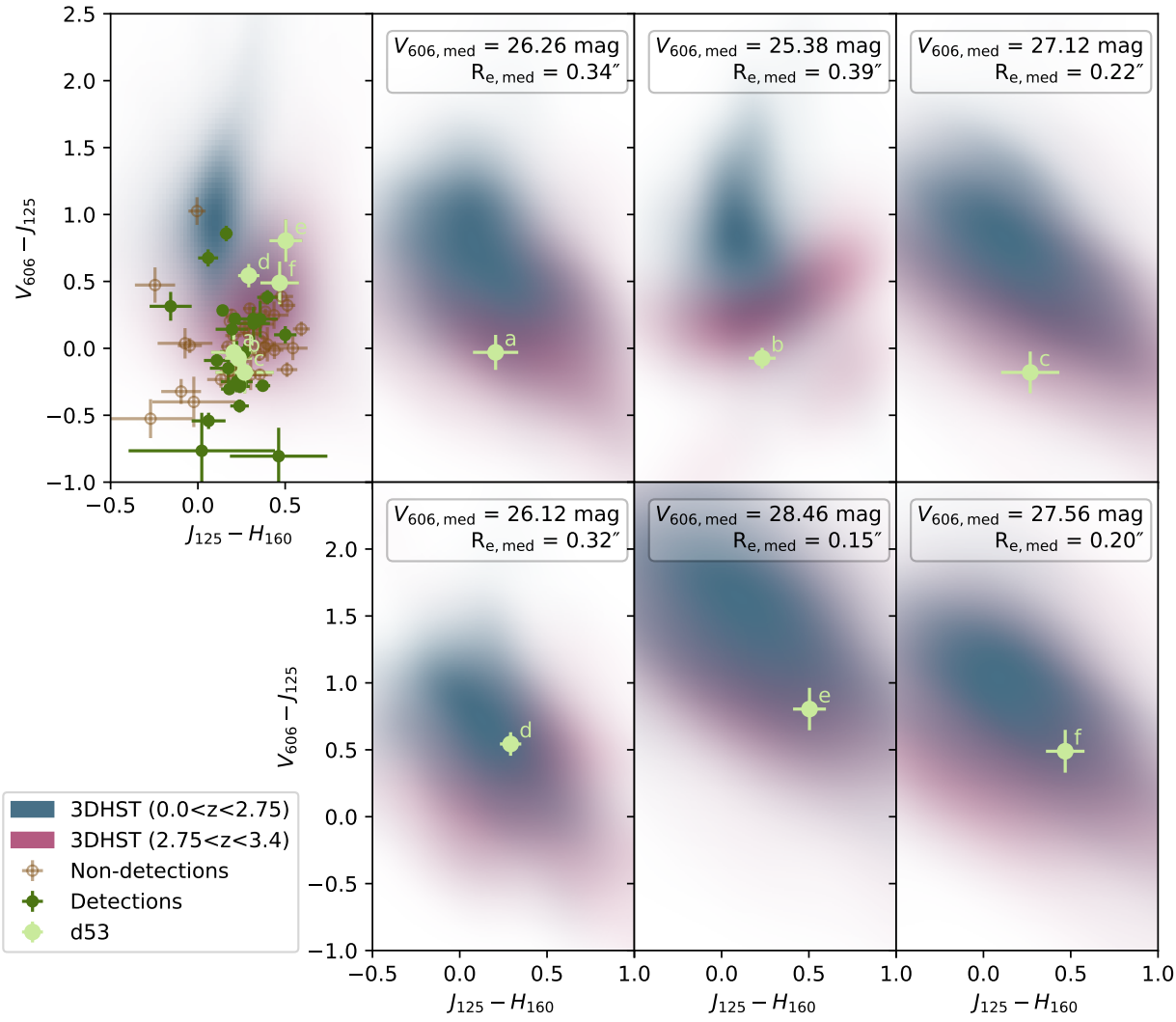

Figure 3.6: PDFs generated by the Gaussian KDE method on samples of 3D-HST galaxies at similar UV magnitudes and effective radii to the components of Q1422-d53. **Left:** PDFs generated from the full 3D-HST redshift samples in $V_{606}J_{125}H_{160}$ color-color space. The Q1422-d53 subcomponents are highlighted. Using these PDFs, all Q1422-d53 subcomponents are predicted at $z \sim 3$. **Right:** PDFs generated from 3D-HST redshift samples with similar V_{606} apparent magnitude and effective radii as the subcomponents of Q1422-d53. The median V_{606} apparent magnitude and R_e are listed in the legend and are comparable to those of the components (see Table 3.4). Components *d*, *e*, and *f* are predicted to be low redshift by these modified PDFs.

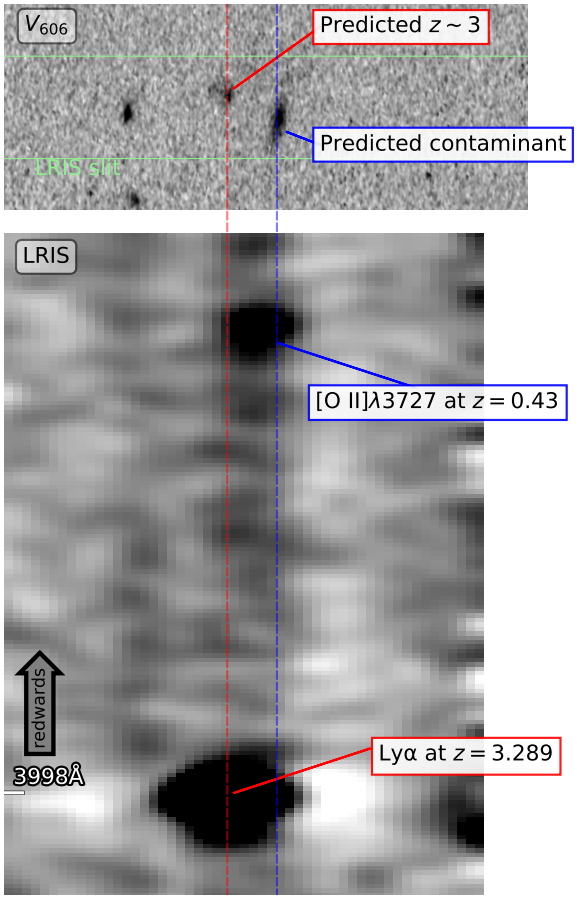

Figure 3.7: *HST* V_{606} imaging and LRIS spectrum of Q0933-M23. **Top:** Unsmoothed V_{606} postage stamp of Q0933-M23. The two components in the center of the image are associated with the extracted 1D spectrum of Q0933-M23. Color-color analysis indicates that the left component, labeled in red, is at $z \sim 3$, while the right component is found to be a low-redshift interloper. Displayed in light green is the location of the LRIS slit that covered this source. **Bottom:** 2D LRIS spectrum of Q0933-M23 oriented such that wavelength increases upwards. The horizontal position has been matched with the top panel. In the center, the blended continuum for Q0933-M23 is shown, complete with the horizontal positions of the $z \sim 3$ and low-redshift component in dashed lines. The brighter $z = 3.289$ Lyα feature is offset towards the $z \sim 3$ component, while a spurious feature is offset towards the low-redshift component. The continuum to the left is from an unrelated object.

ground contamination rate of 2/15 for galaxies with individual apparent LyC detections. We detail how the removal of these contaminated objects affects sample-averaged measurements of $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and $f_{\rm esc,abs}$ as well as relationships between $\langle f_{900}/f_{1500}\rangle_{\rm out}$, $f_{\rm esc,abs}$ and galaxy properties. We also explore the change in ionizing emissivity in light of the cleaned sample.

Save for a small modification to the calculation of the ionizing emissivity, the methods in this section are identical to those of S18.

3.5.1 $\langle f_{900}/f_{1500}\rangle_{out}$

First, we analyse the ratio of ionizing to non-ionizing UV flux density in the uncontaminated KLCS sample. Following S18, in order to average out variations of IGM transmission in different sightlines across the sky, we build a composite spectrum from the rest-frame, f_{1500}-normalized spectra of the sample, where f_{1500} is the average flux density in the range 1475-1525 Å. We combine the spectra using a spline-interpolated common wavelength grid, computing the mean flux density at each wavelength increment after 3σ rejection of outlier flux densities.

The ratio $\langle f_{900}/f_{1500}\rangle_{\rm obs}$ is defined as the average observed flux density of the composite spectrum in the range 880-910 Å, f_{900}, divided by f_{1500}, both in units of ergs s^{-1} cm^{-2} Hz^{-1}. Based on our cleaned sample of 120 galaxies, we estimate $\langle f_{900}/f_{1500}\rangle_{\rm obs} = 0.015 \pm 0.002$, which is lower than the value of $\langle f_{900}/f_{1500}\rangle_{\rm obs} = 0.019 \pm 0.002$ in S18. However, since f_{900} is heavily influenced by IGM and CGM transmission at $z \sim 3$, we perform transmission simulations required to estimate the ratio of ionizing to non-ionizing flux density directly outside the galaxy's "LyC photosphere" at $r_{\rm gal} = 50$ (proper) kpc. This corrected ratio is defined as $\langle f_{900}/f_{1500}\rangle_{\rm out}$. We generate 1000 estimates of t_{900}, the IGM+CGM transmission at 880-910 Å, where each simulation comprises 120 sight lines that have an identical redshift distribution to that of our sample. We then divide $\langle f_{900}/f_{1500}\rangle_{\rm obs}$ by the average IGM+CGM transmission in the LyC region, $\langle t_{900}\rangle$, to obtain $\langle f_{900}/f_{1500}\rangle_{\rm out}$, whose uncertainty includes

the uncertainty in $\langle t_{900} \rangle$. We calculate $\langle f_{900}/f_{1500} \rangle_{\rm out} = 0.040 \pm 0.006$ for the uncontaminated KLCS sample, compared to the ratio of $\langle f_{900}/f_{1500} \rangle_{\rm out} = 0.057 \pm 0.006$ presented in S18. This $\sim 30\%$ discrepancy represents the level of contamination introduced by the strong ~ 3500 Å flux density of low-redshift interlopers in Q0933-D16 and Q0933-M23, masquerading as rest-frame LyC emission at $z \sim 3$.

HST images are available only for the 24 LyC non-detections that lie on the pointings designed to target the apparent detection sample, thus we attempt to model the effects of potential contamination within the 85 KLCS non-detections without *HST* imaging. We build two composite spectra. The first contains the original 24 LyC non-detections imaged by *HST*. The second removes the two galaxies with low-redshift subcomponents, Q1422-md145 and Q1422-d53, for a un-contaminated *HST* non-detection sample of 22 galaxies. We then re-measure $\langle f_{900}/f_{1500} \rangle_{\rm out}$ from sample composite spectra in order to analyse any bias in f_{900} that may arise from contamination in the full non-detection sample. We measure $\langle f_{900}/f_{1500} \rangle_{\rm out}$ of the original non-detection *HST* sample and the uncontaminated non-detection *HST* sample to be 0.003 ± 0.006 and 0.002 ± 0.006, respectively. Due to the lack of significant bias in $\langle f_{900}/f_{1500} \rangle_{\rm out}$ from the inclusion of non-detections with low-redshift subcomponents, we conclude that the lack of *HST* imaging for the larger KLCS non-detection sample will not significantly bias our global measurements of LyC flux density. Nevertheless, a larger sample of non-detections with *HST* imaging will better quantify the level of contamination in the full sample.

We build additional composites to examine how LyC leakage correlates with a variety of galaxy properties. In addition to our full-sample composite of 120 galaxies ("All"), we build a composite with the 13 galaxies with significant f_{900} ("All, detected") and a composite with the 107 galaxies with f_{900} non-detections ("All, not detected"). We split the KLCS sample into four equivalent bins of increasing Lyα equivalent width (W_λ(Lyα) (Q1-Q4)) and decreasing UV luminosity ($L_{\rm UV}$ (Q1-Q4)). Each of these composites includes $120/4 = 30$ galaxies,

ensuring that the uncertainty due to variation in $\langle t_{900} \rangle$ is only $\sim$ 10% (see discussion in S18). We also construct composites by bisecting the sample at $W_\lambda(\mathrm{Ly}\alpha) = 0$ and $L_{\mathrm{UV}} = L^*_{\mathrm{UV}}$ ($M_{\mathrm{UV}} = -21.0$, Reddy & Steidel, 2009) to create "$W_\lambda(\mathrm{Ly}\alpha) > 0$," "$W_\lambda(\mathrm{Ly}\alpha) < 0$," "$L_{\mathrm{UV}} < L^*_{\mathrm{UV}}$," and "$L_{\mathrm{UV}} > L^*_{\mathrm{UV}}$." These composites, as well as all other composites defined by S18, are reconstructed based on our cleaned sample and corrected for IGM and CGM transmission in the same manner as "All." We present the $\langle f_{900}/f_{1500} \rangle_{\mathrm{out}}$ measurements for these composites in Table 3.2 along with the corresponding values from S18 for comparison.

3.5.2 Trends with W_λ(Lyα) and L_{UV}

A deep, uncontaminated spectroscopic sample covering the LyC at $z \sim 3$ presents a unique opportunity to study how LyC leakage correlates with galaxy properties. This type of analysis is vital for understanding the physics of LyC escape and inferring a population of LyC leakers in the epoch of reionization.

We re-examine the galaxy characteristics most strongly associated with $\langle f_{900}/f_{1500} \rangle_{\mathrm{out}}$ from S18: $W_\lambda(\mathrm{Ly}\alpha)$ and L_{UV}. In S18, a strong correlation between $\langle f_{900}/f_{1500} \rangle_{\mathrm{out}}$ and $W_\lambda(\mathrm{Ly}\alpha)$ was measured in the KLCS sample using composite spectra in bins of $W_\lambda(\mathrm{Ly}\alpha)$. In a similar manner, a strong anti-correlation between $\langle f_{900}/f_{1500} \rangle_{\mathrm{out}}$ and L_{UV} was measured. We re-evaluate these trends using uncontaminated composites binned with respect to $W_\lambda(\mathrm{Ly}\alpha)$ and L_{UV} and present them in Figure 3.8 alongside a measurement from the "All" composite. We recover the trends of increasing $\langle f_{900}/f_{1500} \rangle_{\mathrm{out}}$ with lower L_{UV} and larger $W_\lambda(\mathrm{Ly}\alpha)$ of S18, although the slopes are slightly shallower with the removal of contaminants with strong apparent f_{900}. We quantify this change in $\langle f_{900}/f_{1500} \rangle_{\mathrm{out}}$ vs. $W_\lambda(\mathrm{Ly}\alpha)$ by performing a linear fit to the values of $\langle f_{900}/f_{1500} \rangle_{\mathrm{out}}$ in the four independent quartiles of $W_\lambda(\mathrm{Ly}\alpha)$, with the assumption that $\langle f_{900}/f_{1500} \rangle_{\mathrm{out}}$ tends to zero at $W_\lambda(\mathrm{Ly}\alpha) = 0$. With the cleaned sample, we find

$$\langle f_{900}/f_{1500} \rangle_{\mathrm{out}} = 0.28(W_{\mathrm{Ly}\alpha}/110\text{Å}), \tag{3.1}$$

which has a shallower slope than in the relationship $\langle f_{900}/f_{1500} \rangle_{\mathrm{out}} \sim 0.36(W_{\mathrm{Ly}\alpha}/110$ Å)

Table 3.2: $< f_{900}/f_{1500} >_{\rm out}$ measurements of KLCS composite spectra after contamination removal.

Sample[a]	$< f_{900}/f_{1500} >_{\rm out}$ (S18)	$< f_{900}/f_{1500} >_{\rm out}$ (this work)
All	0.057 ± 0.006	0.040 ± 0.006
All, detected[b]	...	...
All, not detected[b]	...	...
$W_\lambda(\mathrm{Ly}\alpha) > 0$	0.086 ± 0.010	0.063 ± 0.009
$W_\lambda(\mathrm{Ly}\alpha) < 0$	0.019 ± 0.008	0.016 ± 0.011
LAEs	0.175 ± 0.026	0.107 ± 0.023
Non-LAEs	0.032 ± 0.008	0.030 ± 0.008
$L_{UV} > L^*_{UV}$	0.006 ± 0.008	0.005 ± 0.008
$L_{UV} < L^*_{UV}$	0.113 ± 0.014	0.085 ± 0.012
L_{UV} (Q1)	0.005 ± 0.008	0.011 ± 0.008
L_{UV} (Q2)	0.000 ± 0.011	0.000 ± 0.006
L_{UV} (Q3)	0.114 ± 0.018	0.075 ± 0.016
L_{UV} (Q4)	0.138 ± 0.024	0.111 ± 0.022
z (Q1)	0.053 ± 0.018	0.050 ± 0.018
z (Q4)	0.056 ± 0.011	0.050 ± 0.014
$W_\lambda(\mathrm{Ly}\alpha)$ (Q1)	0.013 ± 0.011	0.005 ± 0.010
$W_\lambda(\mathrm{Ly}\alpha)$ (Q2)	0.033 ± 0.011	0.048 ± 0.014
$W_\lambda(\mathrm{Ly}\alpha)$ (Q3)	0.047 ± 0.015	0.033 ± 0.014
$W_\lambda(\mathrm{Ly}\alpha)$ (Q4)	0.166 ± 0.025	0.103 ± 0.020
$(G-R)_0$ (Q1)	0.055 ± 0.013	0.050 ± 0.012
$(G-R)_0$ (Q2)	0.059 ± 0.017	0.032 ± 0.016
$(G-R)_0$ (Q3)	0.080 ± 0.016	0.089 ± 0.017
$(G-R)_0$ (Q4)	0.029 ± 0.016	0.000 ± 0.013

[a] Full composite descriptions can be found in S18. [b] Due to the uncertainty in the IGM+CGM correction, these entries are omitted.

presented in S18.

These correlations imply that LyC escape occurs in galaxies with $L_{\rm UV} < L^*_{\rm UV}$ and strong observed Lyα emission relative to their UV continuum. While both relationships are compelling, we note that W_λ(Lyα) transmission has been found to be a strong function of the H I covering fraction, both in the local Universe and at $z \sim 3$ (e.g., Gazagnes et al., 2018; Reddy et al., 2016b). To deepen our physical picture of a LyC leaking galaxy, we must consider descriptions of the structure of the neutral-phase ISM.

3.5.3 $f_{\rm esc,abs}$

While $\langle f_{900}/f_{1500}\rangle_{\rm out}$ is a particularly effective observational parameter for informing the global ionizing budget using only the observed far-UV (FUV) luminosity function (LF) at high-redshift (e.g., S18; Ouchi et al., 2009), it is also important to understand how ionizing radiation escapes galaxies in relation to their intrinsic SFRs. To this end, we must simultaneously model the intrinsic UV stellar spectrum as well as the H I and dust in the ISM that these photons encounter. All of these factors are required to describe the absolute escape fraction ($f_{\rm esc,abs}$), defined here as the ratio of escaping LyC photons to those produced by stars for the assumed stellar-population synthesis (SPS) model.

We estimate $f_{\rm esc,abs}$ in the uncontaminated KLCS composites using SPS model fits to the IGM+CGM-corrected, composite spectra. We use identical SPS models to those in S18 (also see Steidel et al., 2016), with stellar metallicity $Z_* = 0.001$, IMF slope $\alpha = -2.35$, and upper stellar mass limit of 300 $M_\odot$ from BPASS (BPASS v2.1, Eldridge et al., 2017). We also assume a continuous star-formation history and an age of $t = 10^8$ yr. We allow the continuum reddening to vary across $0 \leq E(B-V) < 1.0$ for a variety of attenuation relations (Reddy et al., 2015, 2016a; Gordon et al., 2003). We simultaneously perform geometric ISM modeling for each composite, quantifying the covering fraction ($f_{\rm c}$) of the optically-thick H I gas attenuating the FUV continuum. We adopt the "holes" model of S18 to describe the

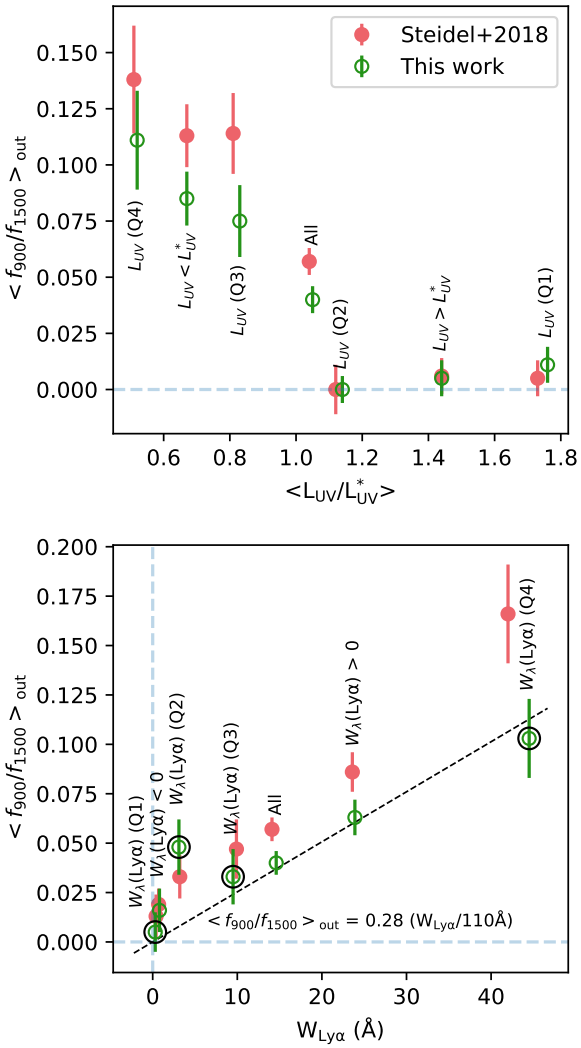

Figure 3.8: Trends of $\langle f_{900}/f_{1500}\rangle_{\rm out}$ as a function of $L_{\rm UV}$ and $W_\lambda(\mathrm{Ly}\alpha)$ from KLCS composites. Steidel et al. (2018) composite measurements are shown with solid red circles, while updated measurements from this work are shown with open green circles. **Top:** Composites in bins of $L_{\rm UV}$, including four in independent quartiles of $L_{\rm UV}$ and two bisecting the sample on $L^*_{\rm UV}$. Also displayed is the full-sample composite, "All." A decreasing, almost bimodal relationship is seen preserved between $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and $L_{\rm UV}$. **Bottom:** Composites in bins of $W_\lambda(\mathrm{Ly}\alpha)$, including four independent quartiles of $W_\lambda(\mathrm{Ly}\alpha)$ and two bisecting the sample on $W_\lambda(\mathrm{Ly}\alpha)= 0$, and the "All" composite. Also displayed is the increasing, linear fit to the four independent quartiles, which are highlighted with black circles.

physical arrangement of H I and dust around the H II regions of our galaxies, where neutral H I exists in a picket-fence configuration and dust is only located where neutral gas is (also see Reddy et al., 2016b). Accordingly, f_c varies between zero and one and the "holes" are free of both H I and dust. This model is consistent with correlations found between Lyα emission strength and low-ionization interstellar absorption strength in star-forming galaxies at intermediate redshifts (Shapley et al., 2003; Steidel et al., 2010; Du et al., 2018; Pahl et al., 2020). S18 defines the attenuated observed spectrum as

$$S_{\nu,\mathrm{obs}} = S_{\nu,\mathrm{int}} \left[(1 - f_c) + f_c\, e^{-\tau}(\lambda) 10^{-A_\lambda/2.5}\right], \tag{3.2}$$

where $S_{\nu,\mathrm{int}}$ is the intrinsic stellar spectrum, $A_\lambda = k_\lambda(E(B-V)_{\mathrm{cov}})$ where $E(B-V)_{\mathrm{cov}}$ is the continuum reddening in the foreground gas (and k_λ is the wavelength-dependent dust attenuation law), and $e^{-\tau}(\lambda)$ is the transmission function due to line and continuum absorption (a function of H I column density N_{HI}).

We fit for $E(B-V)_{\mathrm{cov}}$, $\log(N_{\mathrm{HI}})$ (cm^{-2}), and f_c for each composite and tabulate the results in Table 3.3. In the "holes" model, the H I and dust are optically thick to ionizing light save for the holes in the ISM where extreme UV light is transmitted entirely. Thus, the relation between $f_{\mathrm{esc,abs}}$ and f_c in this ISM model is simply

$$f_{\mathrm{esc,abs}} = 1 - f_c. \tag{3.3}$$

We list the estimated $f_{\mathrm{esc,abs}}$ for each composite in Table 3.3 along with the values from S18 for comparison.

A key result from this analysis is the sample-averaged $f_{\mathrm{esc,abs}}$ of the KLCS survey. Based on the "All" composite, we estimate $f_{\mathrm{esc,abs}} = 0.06 \pm 0.01$, lower than the sample-averaged $f_{\mathrm{esc,abs}} = 0.09 \pm 0.01$ of S18. The value of $f_{\mathrm{esc,abs}}$ is an essential input to models of the contribution of star-forming galaxies to reionization. Due to increased attenuation from the neutral-phase IGM at $z > 3$, observing this quantity at higher redshifts becomes increasingly difficult (Vanzella et al. 2012; S18). Well-constrained direct measurements of model-

independent $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and model-dependent $f_{\rm esc,abs}$ at $z\sim3$ inform our picture of LyC escape at high redshift and have direct consequences for the contribution of star-forming galaxies to the ionizing background. Foreground contamination in only two galaxies in a sample of 124 can elevate $f_{\rm esc,abs}$ by $\sim30\%$, underlining the need for high-resolution, multi-band imaging of LyC detection candidates at $z\sim3$ to vet sample-averaged measurements (Vanzella et al. 2012; M15).

We emphasize that $f_{\rm esc,abs}$ is calculated here relative to BPASS SPS models that are characterized by a high ionizing photon production efficiency ($\xi_{\rm ion}=25.5$). If we instead, for example, assume an SPS model with $\xi_{\rm ion}=25.2$ (e.g., Robertson et al., 2013, 2015) we would infer $f_{\rm esc,abs}$ a factor of ∼two larger for our full sample composite. In contrast, $\langle f_{900}/f_{1500}\rangle_{\rm out}$ is a (mostly) model-independent measure of escaping ionizing radiation that only depends on a mean IGM+CGM correction, and does not rely on assumptions about the underlying stellar population, including stellar age.

We again connect our composite measurements of LyC escape with galaxy characteristics, now in terms of the relationship between $f_{\rm esc,abs}$ and $W_\lambda({\rm Ly}\alpha)$ within the uncontaminated KLCS sample. We display the $f_{\rm esc,abs}$ and $W_\lambda({\rm Ly}\alpha)$ measurements for each composite in Figure 3.9, including 2σ upper limits for measurements with less than 2σ significance. We highlight the four independent quartiles of $W_\lambda({\rm Ly}\alpha)$ in black and fit a linear trend to these points, fixing $f_{\rm esc,abs}=0$ at $W_\lambda({\rm Ly}\alpha)=0$, and treating the upper limits as $f_{\rm esc,abs}=0$ with the appropriate Gaussian error. We recover the relationship

$$f_{\rm esc,abs}=0.58(W_{{\rm Ly}\alpha}/110{\rm \AA}), \tag{3.4}$$

slightly shallower than $f_{\rm esc,abs}=0.75(W_{{\rm Ly}\alpha}/110$ Å) as presented in S18. This relationship between $f_{\rm esc,abs}$ and $W_\lambda({\rm Ly}\alpha)$ supports the assertion that the escape of LyC photons is directly connected with the spatially-resolved distribution of H I in the ISM that governs $W_\lambda({\rm Ly}\alpha)$ measurements at $z\sim3$ (e.g., Scarlata et al., 2009; Rivera-Thorsen et al., 2015). Similar relationships between $f_{\rm esc,abs}$ and $W_\lambda({\rm Ly}\alpha)$ have been found in other large-scale LyC

surveys at $z \sim 3-4$ (Marchi et al., 2017, 2018; Fletcher et al., 2019). f_c and Lyα escape fraction have also been determined to predict LyC escape fractions at low redshift (Chisholm et al., 2018; Gazagnes et al., 2018). Additionally, *HST* analysis of a spatially-resolved far-UV color map has shown that bluer colors are coincident with LyC escape, demonstrating that favorable H I and dust configurations lead to LyC escape (Ji et al., 2020). Finally, the Lyα kinematics of a strongly-leaking $z \sim 3$ galaxy have demonstrated the consistent geometry of LyC and Lyα escape (Vanzella et al., 2020).

3.5.4 Ionizing emissivity and implications

The removal of apparent LyC leakers with evidence of foreground contamination has lowered our estimates of $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and $f_{\rm esc,abs}$ compared to previous values based on the composites of S18. Additionally, the relationships between $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and $L_{\rm UV}$,$\langle f_{900}/f_{1500}\rangle_{\rm out}$ and W_λ(Lyα), and $f_{\rm esc,abs}$ and W_λ(Lyα) have been recovered, but with slightly shallower correlation slopes.

We explore the effects of these updated results on the predictions of the galactic contributions to the ionizing background at $z \sim 3$ by re-tracing the global emissivity calculations of (S18). The ionizing emissivity of galaxies can be estimated by

$$\epsilon_{\rm LyC} \simeq \int_{L_{\rm UV,min}}^{L_{\rm UV,max}} \langle f_{900}/f_{1500}\rangle_{\rm out}(L_{\rm UV}) \times L_{\rm UV}\Phi(L_{\rm UV})dL_{\rm UV}, \tag{3.5}$$

where $\Phi(L_{\rm UV})$ is the galaxy LF evaluated in the rest-frame FUV (1500-1700 Å). We can perform a simple estimate by expressing $\langle f_{900}/f_{1500}\rangle_{\rm out}$ as a function of $L_{\rm UV}$, taking $\langle f_{900}/f_{1500}\rangle_{\rm out} = 0$ for galaxies with $M_{\rm UV} \leq -21.0$ and $\langle f_{900}/f_{1500}\rangle_{\rm out} = 0.085$ for galaxies with $-21 < M_{\rm UV} \leq -19.5$ (see "$L_{UV} < L^*_{UV}$" and "$L_{UV} > L^*_{UV}$" composites in Table 3.2). Using the UV LF from Reddy & Steidel (2009), we calculate a corrected $\epsilon_{\rm LyC} \simeq 9.78 \times 10^{24}$ erg s^{-1} Hz^{-1} Mpc^{-3}, compared to the S18 value of $\epsilon_{\rm LyC} \simeq 13.0 \times 10^{24}$ erg s^{-1} Hz^{-1} Mpc^{-3}.

As in S18, we examine an alternate estimate of the ionizing emissivity by assuming that $\langle f_{900}/f_{1500}\rangle_{\rm out}$ varies with W_λ(Lyα). This estimation is supported by the strong relationship

Table 3.3: Spectral-fitting parameters from ISM modeling using the "holes" configuration of Steidel et al. (2018).

Sample[a]	Att	$E(B-V)_{\rm cov}$	$\log(N_{\rm HI})$ (cm^{-2})	f_c	$f_{\rm esc,abs}$ (S18)	$f_{\rm esc,abs}$ (this work)
All	SMC	0.161	20.56	0.94	0.09 ± 0.01	0.06 ± 0.01
All, detected[b,c]	SMC	0.080	18.59	0.85	0.31 ± 0.03	0.21 ± 0.03
All, not detected	R16	0.160	20.62	0.95	0.05 ± 0.01	0.05 ± 0.01
$W_\lambda(\mathrm{Ly}\alpha) > 0$	SMC	0.056	20.12	0.88	0.14 ± 0.02	0.12 ± 0.02
$W_\lambda(\mathrm{Ly}\alpha) < 0$	R16	0.193	20.98	0.97	<0.03	<0.03
LAEs	SMC	0.052	19.95	0.76	0.29 ± 0.03	0.25 ± 0.03
Non-LAEs	R16	0.170	20.69	0.96	0.04 ± 0.02	0.04 ± 0.01
$L_{UV} > L^*_{UV}$	R16	0.166	20.72	0.96	<0.04	<0.04
$L_{UV} < L^*_{UV}$	SMC	0.064	20.38	0.88	0.13 ± 0.03	0.12 ± 0.03
L_{UV} (Q1)	R16	0.064	20.61	0.93	<0.04	<0.05
L_{UV} (Q2)	R16	0.082	20.79	0.94	<0.04	0.06 ± 0.02
L_{UV} (Q3)	SMC	0.146	20.23	0.92	0.13 ± 0.03	0.08 ± 0.02
L_{UV} (Q4)	SMC	0.153	20.49	0.89	0.16 ± 0.03	0.12 ± 0.02
z (Q1)	SMC	0.076	20.59	0.91	0.08 ± 0.01	0.09 ± 0.01
z (Q4)	R16	0.098	20.36	0.89	0.12 ± 0.02	0.11 ± 0.02
$W_\lambda(\mathrm{Ly}\alpha)$ (Q1)	R16	0.185	21.06	0.97	<0.03	<0.03
$W_\lambda(\mathrm{Ly}\alpha)$ (Q2)	R16	0.202	20.79	0.96	<0.04	0.04 ± 0.02
$W_\lambda(\mathrm{Ly}\alpha)$ (Q3)	SMC	0.056	19.97	0.93	0.07 ± 0.02	0.07 ± 0.02
$W_\lambda(\mathrm{Ly}\alpha)$ (Q4)	SMC	0.049	20.06	0.77	0.27 ± 0.02	0.23 ± 0.02
$(G-R)_0$ (Q1)	SMC	0.021	20.36	0.84	0.15 ± 0.02	0.16 ± 0.02
$(G-R)_0$ (Q2)	R16	0.143	20.56	0.93	0.06 ± 0.02	0.07 ± 0.02
$(G-R)_0$ (Q3)	R16	0.196	20.47	0.95	<0.08	0.05 ± 0.01
$(G-R)_0$ (Q4)	R16	0.274	20.85	0.97	<0.06	<0.03

[a] Full composite descriptions can be found in Steidel et al. (2018). [b] Due to the unique IGM+CGM corrections described in Section 3.5.1, $f_{\rm esc,abs}$ is calculated with an additional term for LyC photons traveling through an ISM with order-unity LyC optical depth. [c] Attenuation assumed to be in the lowest 12% of the expected values.

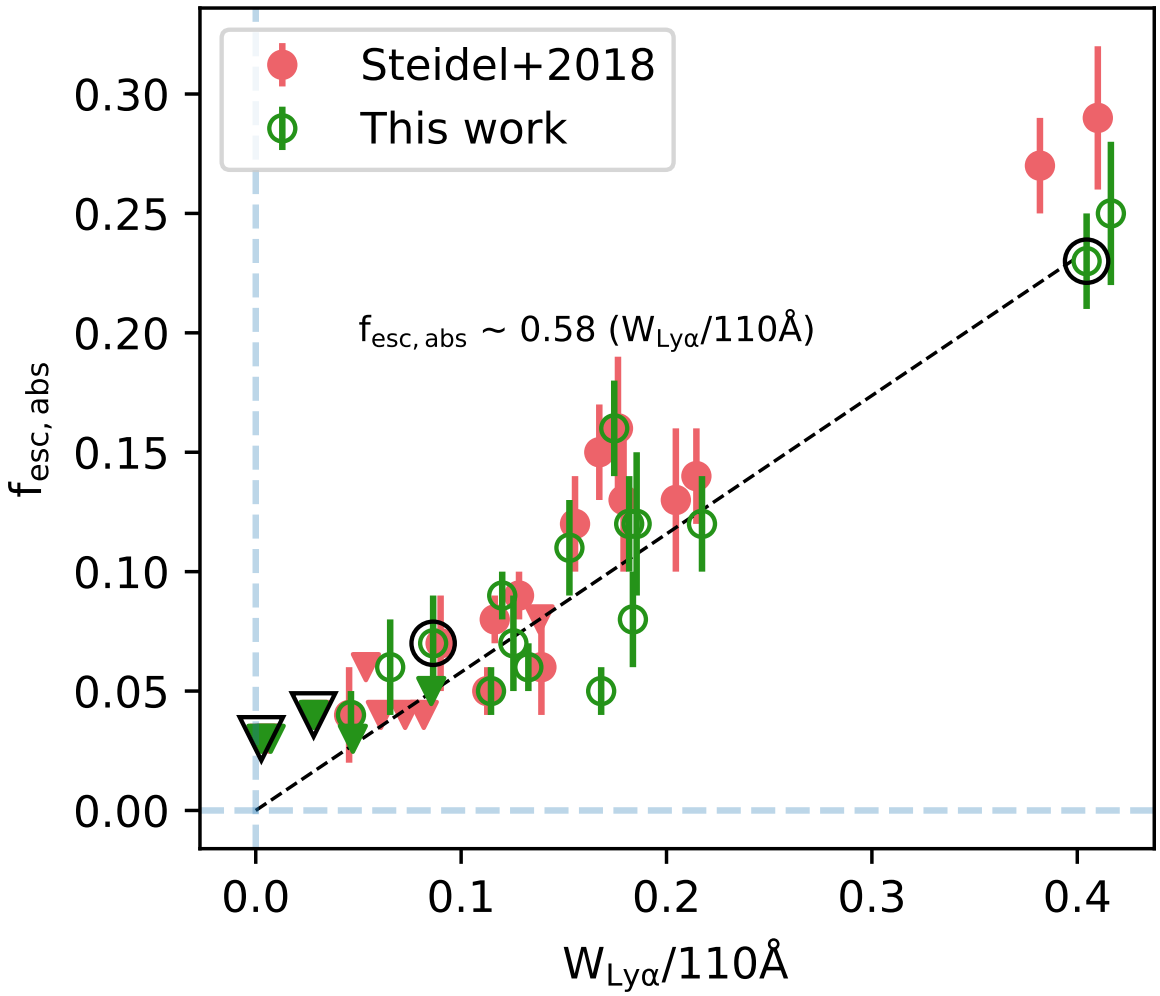

Figure 3.9: Updated $f_{\rm esc,abs}$ measurements for the KLCS composites as a function of $W_\lambda(\mathrm{Ly}\alpha)$ alongside values from Steidel et al. (2018). Each point represents a composite from Table 3.3. Triangles represent 2σ upper limits on $f_{\rm esc,abs}$. We fit a linear trend to the four independent quartiles of $W_\lambda(\mathrm{Ly}\alpha)$, highlighted in black, fixing $f_{\rm esc,abs} = 0$ at $W_\lambda(\mathrm{Ly}\alpha) = 0$.

between the two properties in Figure 3.8 and can be extrapolated to higher redshifts if $W_\lambda(\mathrm{Ly}\alpha)$ is corrected for IGM opacity. For this calculation, we move $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ outside the integral in Equation 3.5, and estimate an average $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ weighted by $W_\lambda(\mathrm{Ly}\alpha)$. We multiply this average $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$, $\mathrm{E}(\langle f_{900}/f_{1500}\rangle_{\mathrm{out}})$, by the integral over the non-ionizing UV LF, ϵ_{UV}. To find the $W_\lambda(\mathrm{Ly}\alpha)$-weighted average value of $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$, we combine our relationships between $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ and $W_\lambda(\mathrm{Ly}\alpha)$ in Equation 3.1 and the distribution function $n(W_\lambda(\mathrm{Ly}\alpha))$.

While it is possible to combine Equation 3.1 and $n(W_\lambda(\mathrm{Ly}\alpha))$ directly, we introduce physical descriptions of Lyα escape to extrapolate the behavior of $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ outside of the range of $W_\lambda(\mathrm{Ly}\alpha)$ in our sample. S18 connected the modulation of $W_\lambda(\mathrm{Ly}\alpha)$ by f_{c} through two separate processes: the scattering of Lyα photons by foreground gas and the decrease in the Lyα source function when H II regions are no longer optically thick to ionizing radiation. These processes connect through the equation

$$(1 - f_{\mathrm{c}}) = 0.5 - [|0.5W' - 0.25|]^{1/2}, \tag{3.6}$$

where $W' \equiv W_\lambda(\mathrm{Ly}\alpha)/W_\lambda(\mathrm{Ly}\alpha,\ \mathrm{Case\ B})$ and $W_\lambda(\mathrm{Ly}\alpha,\ \mathrm{Case\ B})$=110 Å in our assumed SPS model. We can join this physically-motivated parameterization with the empirical relations in Equations 3.1 and 3.4 to express $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ in terms of W':

$$\langle f_{900}/f_{1500}\rangle_{\mathrm{out}} \sim 0.24 - [|0.117W' - 0.058|]^{1/2}. \tag{3.7}$$

To calculate the expectation value of $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ weighted by W', we now include the relative incidence of $W_\lambda(\mathrm{Ly}\alpha)$,

$$n(W_\lambda) \propto \exp(-W_\lambda/23.5\text{Å}), \tag{3.8}$$

consistent with S18 and spectroscopic samples analysed in Shapley et al. (2003) and Kornei et al. (2010). To find the expectation value of $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$, $\mathrm{E}(\langle f_{900}/f_{1500}\rangle_{\mathrm{out}})$, we simplify the method of S18 by integrating Equation 3.8 with a change of variables from Equation

3.7. After including an assumption that 40% of galaxies have $W_\lambda(\mathrm{Ly}\alpha) \leq 0$ and therefore $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}} \simeq 0$ for such sources, we find that the integral becomes

$$E(\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}) = \frac{0.6\int_0^1 [0.24 - [|0.117W' - 0.058|]^{1/2}] \times \exp(-W' \times (110/23.5))dW'}{\int_0^1 exp(-W' \times (110/23.5))dW'} \tag{3.9}$$

and evaluates to $E(\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}) = 0.032$. Based on the equivalent values relating $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$, $f_{\mathrm{esc,abs}}$, and $W_\lambda(\mathrm{Ly}\alpha)$ from S18, this method estimates $E(\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}) = 0.042$. We then multiply $E(\langle f_{900}/f_{1500}\rangle_{\mathrm{out}})$ by ϵ_{UV} to find $\epsilon_{\mathrm{LyC}} \simeq 5.5\times10^{24}$ erg s^{-1} Hz^{-1} Mpc^{-3}. For comparison, when we apply this methodology to the contaminated S18 results, we find a higher emissivity of $\epsilon_{\mathrm{LyC}} \simeq 7.2\times10^{24}$ erg s^{-1} Hz^{-1} Mpc^{-3}. Our value remains consistent with other recent estimates, such as those in Jones et al. (2021): $\epsilon_{\mathrm{LyC}} \simeq 10.0^{+10.0}_{-5.0} \times 10^{24}$ erg s^{-1} Hz^{-1} Mpc^{-3} at $z \sim 2.5$ and $\epsilon_{\mathrm{LyC}} \simeq 15.8^{+34.3}_{-10.8} \times 10^{24}$ erg s^{-1} Hz^{-1} Mpc^{-3} at $z \sim 2.9$.

The reduction of ionizing emissivity from star-forming galaxies found using the uncontaminated composites of this work is significant. However, galaxies still provide a comparable contribution to the ionizing background at $z \sim 3$ to that of active galactic nuclei (AGNs), which are estimated to contribute between 1.6 and 10×10^{24} erg s^{-1} Hz^{-1} Mpc^{-3} at these redshifts (Hopkins et al., 2007; Cowie et al., 2009; Kulkarni et al., 2019; Shen et al., 2020). Furthermore, the unique size, purity, and data products of the sample in this work can inform the details of reionization models and the physics of LyC escape at $z > 3$. Authors have recently described competing models of reionization as "democratic," where reionization is driven by numerous faint sources with high $f_{\mathrm{esc,abs}}$ (e.g., Finkelstein et al., 2019), and "oligarchical," where <5% of galaxies contribute >80% of the reionization budget (Naidu et al., 2020). Both of these interpretations depend on base assumptions of the evolution and value of $f_{\mathrm{esc,abs}}$ across cosmic time and the degree of LyC leakage associated with a range of galaxy properties.

Finkelstein et al. (2019) used a halo-mass dependent $f_{\mathrm{esc,abs}}$ in their fiducial model. These authors also considered a fixed $f_{\mathrm{esc,abs}} = 0.09 \pm 0.01$, as found in S18, and accordingly re-

port tension with observational constraints on the progress of reionization. However, in order to consider the implications of different $f_{\rm esc,abs}$ values on the evolution of the ionizing emissivity, careful attention must be paid to the manner in which key quantities are empirically estimated, modeled, or assumed. Given that the most robust quantity we estimate is $\langle f_{900}/f_{1500}\rangle_{\rm out}$, the only fair way to incorporate our $f_{\rm esc,abs}$ value into the Finkelstein et al. (2019) model is also to assume the same intrinsic LyC luminosity for similar galaxies. Additionally, the fiducial model of Finkelstein et al. (2019) predicts that AGNs dominate the ionizing budget even extrapolated to $z \sim 3$, a determination in tension with this work and S18. Keeping in mind the same caveats about incorporating an $f_{\rm esc,abs}$ determined from observations into reioniziation models, we turn to Naidu et al. (2020), who built two empirical models, one that fits for a constant $f_{\rm esc,abs} = 0.21^{+0.06}_{-0.04}$ during reionization and one that parameterizes $f_{\rm esc,abs}$ as a function of SFR surface density ($\Sigma_{\rm SFR}$). For the second model specifically, these authors include $f_{\rm esc,abs} = 0.09 \pm 0.01$ from S18 as a constraint in the fit, and predict $\sim 10\%$ of sources have $f_{\rm esc,abs} > 20\%$ at $z \sim 4$. The Naidu et al. model would have to be tweaked further to incorporate a lower average $f_{\rm esc,abs}$ at $z \sim 3$, but the rough percentage of sources with large escape fractions remains consistent with the KLCS detection fraction. This model is also built upon the assumption that the leakage of LyC is correlated with increasing $\Sigma_{\rm SFR}$, a determination that can now be tested with the KLCS sample using dust-corrected SFR values and *HST* sizes. In future work, we will focus on the properties of the LyC leaking galaxies in the uncontaminated KLCS survey, including stellar mass, $E(B-V)$, $\Sigma_{\rm SFR}$, and age.

3.6 Summary

Due to the opacity of the IGM to ionizing photons for $z > 3.5$, direct detections of the LyC in galaxies at lower redshifts are vital for drawing inferences about the ionizing spectra of reionization-era galaxies. With the goal of increasing the number of confirmed individual LyC detections and removing contamination from sample-averaged measurements at $z \sim 3$, we

examined LyC detection candidates of the KLCS survey for foreground contamination. By testing each morphological sub-component for low-redshift signatures with the $V_{606}J_{125}H_{160}$ color-color diagram, we removed contaminated objects from the KLCS sample and revisited the measurements of S18 after building un-contaminated composite spectra. The key results are as follows:

1. We find evidence of foreground contamination in the subcomponents of two LyC detection candidates, Q0933-M23 and Q0933-D16, as well as two non-detections, Q1422-d53 and Q1422-md145. The fraction of contaminated sources within the apparent LyC detections was 2/15, and the subsequent removal of these sources brought the LyC detection fraction for individual objects of the KLCS survey from 15/124 to 13/124.

2. We re-measured the sample-averaged ratio of ionizing to non-ionizing rest-UV flux density after performing corrections for IGM and CGM attenuation, finding $\langle f_{900}/f_{1500}\rangle_{\rm out} = 0.040\pm0.006$. This value has been reduced by $\sim$30% compared to the original $\langle f_{900}/f_{1500}\rangle_{\rm out} = 0.057 \pm 0.006$ of S18, indicating the contribution of low-redshift flux density to the KLCS composite spectra. This significant reduction underscores the necessity for high-resolution, multi-band imaging to remove bias in $z \sim 3$ sample-averaged quantities related to the LyC spectrum.

3. The strong dependence of $\langle f_{900}/f_{1500}\rangle_{\rm out}$ on W_λ(Lyα) and $L_{\rm UV}$ within the KLCS sample remains in the contamination-corrected sample. $\langle f_{900}/f_{1500}\rangle_{\rm out}$ increased from 0.005 ± 0.008 to 0.085 ± 0.014 in composites built from the UV-brightest and dimmest halves of the KLCS sample. We also measured a positive, linear correlation between $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and W_λ(Lyα) from composite spectra, tying together the escape of LyC and Lyα photons in the ISM of our galaxies.

4. We estimated a sample-averaged absolute escape fraction of $f_{\rm esc,abs} = 0.06\pm0.01$ at $z \sim 3$ after performing stellar-population fits and ISM modeling of the rest-UV spectrum.

Similar to $\langle f_{900}/f_{1500}\rangle_{\rm out}$, this uncontaminated value is lower than the S18 $f_{\rm esc,abs}$ of 0.09 ± 0.01 for the same assumptions concerning the SPS model. We also recovered the positive, linear relationship between $f_{\rm esc,abs}$ and W_λ(Lyα). This relationship can be interpreted as the H I covering fraction modulating the strengths of both Lyα emission and LyC leakage, supported by relationships between W_λ(Lyα) and $f_{\rm c}$ that have been found locally (e.g., Gazagnes et al., 2018).

5. Based on our modified results, we re-estimate the ionizing emissivity of $L_{UV} > 0.3L^*_{UV}$ galaxies at $z \sim 3$, determining $\epsilon_{\rm LyC} \simeq 5.5 \times 10^{24}$ erg s^{-1} Hz^{-1} Mpc^{-3}. We perform this calculation using the observed functional dependence of $\langle f_{900}/f_{1500}\rangle_{\rm out}$ on W_λ(Lyα), the W_λ(Lyα) distribution function, and the far-UV luminosity function. Based on this uncontaminated estimate, we conclude that the ionizing background at $z \sim 3$ contains comparable contributions of both galaxies and AGNs.

These results have important implications for models of reionization. Mapping the evolution of the IGM neutral fraction to the change in galaxy populations with redshift requires assumptions of $f_{\rm esc,abs}$ that must be consistent with direct, sample-averaged results such as those from this work. Moreover, the degree to which $f_{\rm esc,abs}$ and $\langle f_{900}/f_{1500}\rangle_{\rm out}$ can be affected by just a few low-redshift interlopers demonstrates the utility of high-resolution follow-up of any individual or statistical detection of $z \sim 3$ LyC. As the sample of confirmed LyC detections widens through deep, spectroscopic and narrow-band imaging surveys, our understanding of the ionizing characteristics of reionization-era galaxies will follow.

3.7 Appendix 2A: Non-detections

Here we describe the properties of the 24 LyC non-detections targeted by *HST*. In Figure 3.10, we show the $V_{606}J_{125}H_{160}$ postage stamps of each non-detection alongside ground-based images, *HST* false-color images, and segmentation maps in a manner similar to Figure 3.3.

We report the photometric measurements of each subcomponent displayed in Figure 3.10 in Table 3.4.

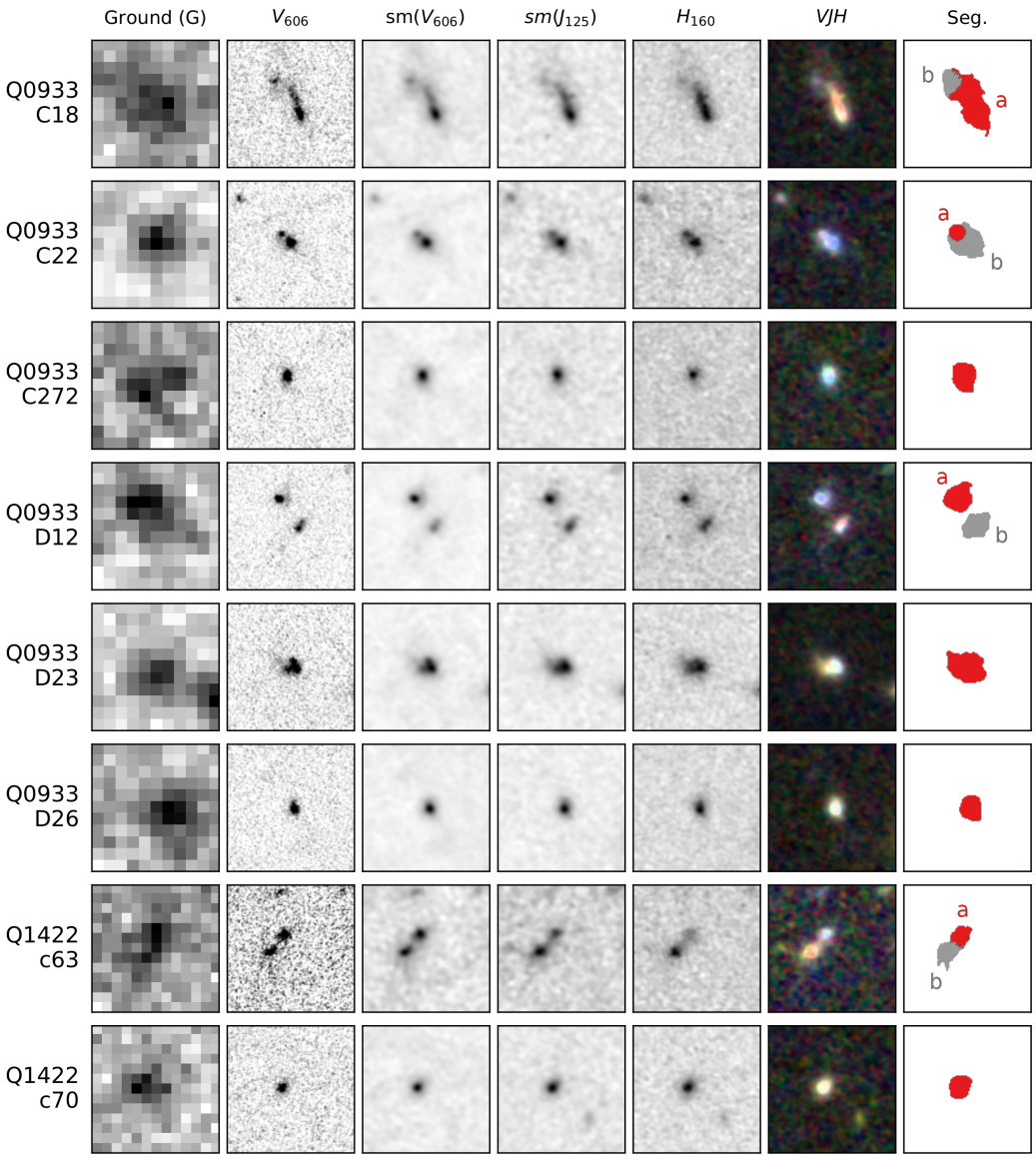

Figure 3.10: $3'' \times 3''$ postage stamps of the 24 LyC non-detections targeted by HST. **First column:** Ground-based G (Steidel et al., 2003). **Second column**: V_{606} at the original resolution. **Third and fourth columns:** V_{125} and J_{125} smoothed to the lower-resolution of H_{160}. **Fifth Column:** Original-resolution H_{160}. **Sixth Column:** False-color postage stamps. The $sm(V_{606})$, $sm(J_{125})$, and H_{160} images are represented by blue, green, and red, respectively. **Seventh Column:** segmentation map generated by SExtractor. Separate components extracted by the program are represented by different-colored regions.

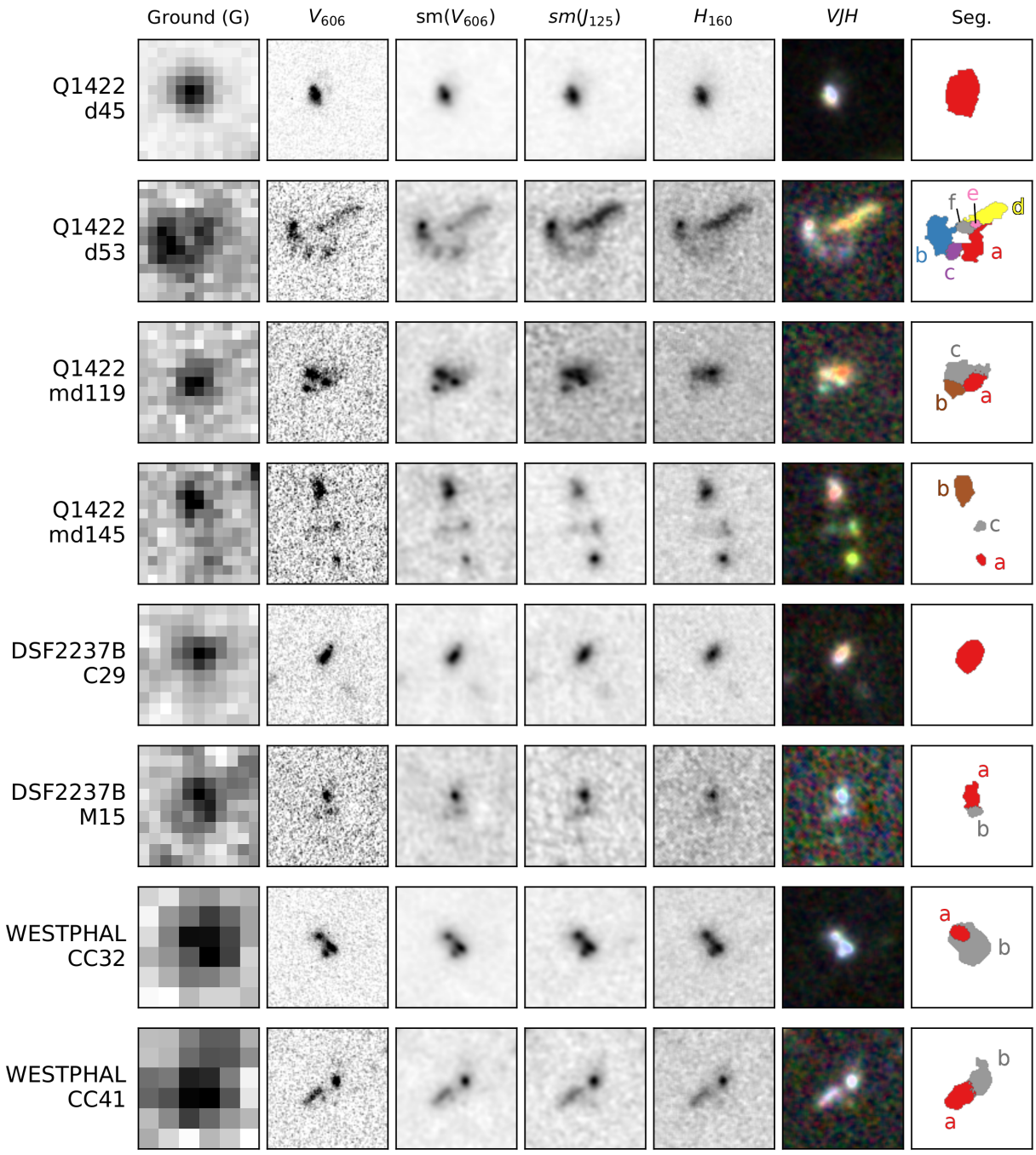

Figure 3.10: Continued.

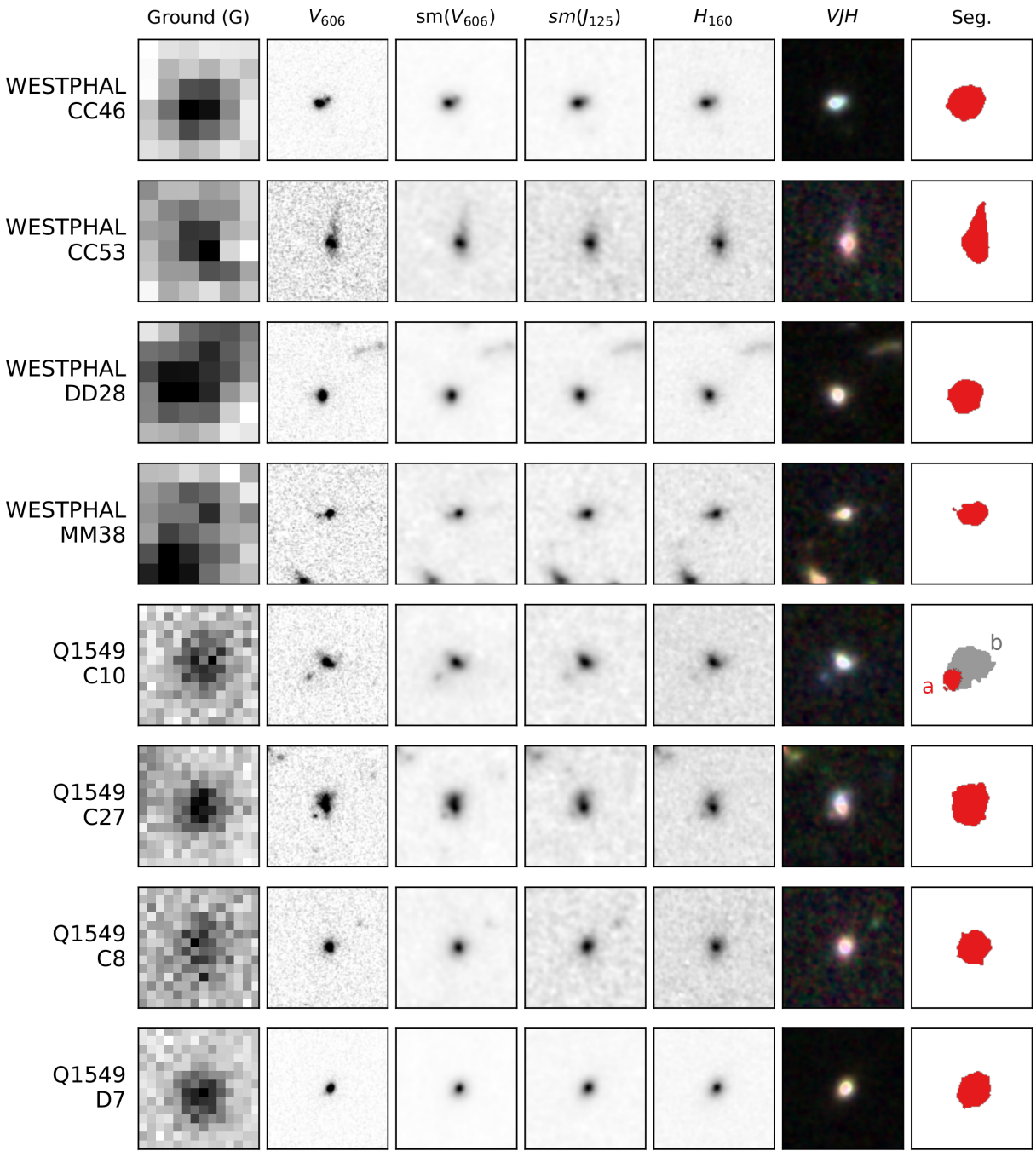

Figure 3.10: Continued.

Table 3.4: Photometric measurements of the LyC non-detection subcomponents.

ID[a]	R.A.	Dec.	$z_{\rm sys}$[b,c]	V_{606}	J_{125}	H_{160}
DSF2237B-C29	22:39:34.50	+11:52:41.15	3.100	$24.83^{+0.04}_{-0.04}$	$24.88^{+0.03}_{-0.03}$	$24.53^{+0.03}_{-0.02}$
DSF2237B-M15a	22:39:30.71	+11:51:40.93	3.403	$26.02^{+0.07}_{-0.06}$	$26.34^{+0.07}_{-0.06}$	$26.44^{+0.10}_{-0.09}$
DSF2237B-M15b	22:39:30.71	+11:51:40.57	3.403	$27.45^{+0.12}_{-0.11}$	$27.85^{+0.16}_{-0.14}$	$27.87^{+0.20}_{-0.17}$
Q0933-C18a	09:33:34.24	+28:43:23.39	2.926	$24.89^{+0.05}_{-0.05}$	$24.57^{+0.04}_{-0.03}$	$24.06^{+0.03}_{-0.03}$
Q0933-C18b	09:33:34.28	+28:43:23.98	2.926	$26.69^{+0.09}_{-0.09}$	$26.66^{+0.10}_{-0.09}$	$26.26^{+0.09}_{-0.08}$
Q0933-C22a	09:33:32.12	+28:43:42.26	3.164	$26.53^{+0.34}_{-0.26}$	$26.56^{+0.48}_{-0.33}$	$26.28^{+0.51}_{-0.35}$
Q0933-C22b	09:33:32.10	+28:43:42.08	3.164	$25.21^{+0.03}_{-0.03}$	$25.44^{+0.04}_{-0.04}$	$25.31^{+0.05}_{-0.05}$
Q0933-C272	09:33:27.30	+28:44:37.37	3.546	$25.42^{+0.04}_{-0.04}$	$25.40^{+0.04}_{-0.04}$	$25.45^{+0.06}_{-0.06}$
Q0933-D12a	09:33:33.95	+28:44:18.83	2.924	$25.38^{+0.04}_{-0.04}$	$25.56^{+0.05}_{-0.05}$	$25.36^{+0.06}_{-0.06}$
Q0933-D12b	09:33:33.92	+28:44:18.19	2.924	$25.95^{+0.06}_{-0.06}$	$25.95^{+0.07}_{-0.06}$	$25.40^{+0.06}_{-0.05}$
Q0933-D23	09:33:23.44	+28:47:17.11	3.224	$24.71^{+0.03}_{-0.03}$	$24.41^{+0.02}_{-0.02}$	$24.12^{+0.03}_{-0.03}$
Q0933-D26	09:33:25.08	+28:48:24.41	3.266	$25.34^{+0.03}_{-0.03}$	$25.09^{+0.03}_{-0.03}$	$24.90^{+0.04}_{-0.03}$
Q1422-c63a	14:24:30.19	+22:53:56.52	3.059	$26.51^{+0.07}_{-0.07}$	$26.70^{+0.09}_{-0.09}$	$26.40^{+0.09}_{-0.08}$
Q1422-c63b	14:24:30.22	+22:53:56.14	3.059	$26.42^{+0.08}_{-0.08}$	$26.18^{+0.07}_{-0.06}$	$25.74^{+0.05}_{-0.05}$
Q1422-c70	14:24:33.65	+22:54:55.27	3.129	$25.72^{+0.05}_{-0.05}$	$25.61^{+0.04}_{-0.04}$	$25.32^{+0.04}_{-0.04}$
Q1422-d45	14:24:32.23	+22:54:03.16	3.072	$23.77^{+0.02}_{-0.02}$	$23.90^{+0.02}_{-0.02}$	$23.67^{+0.02}_{-0.02}$
Q1422-d53a	14:24:25.53	+22:55:00.28	3.086	$26.25^{+0.10}_{-0.09}$	$26.28^{+0.09}_{-0.09}$	$26.08^{+0.10}_{-0.09}$
Q1422-d53b	14:24:25.59	+22:55:00.71	3.086	$25.38^{+0.06}_{-0.05}$	$25.45^{+0.06}_{-0.05}$	$25.22^{+0.06}_{-0.05}$
Q1422-d53c	14:24:25.57	+22:55:00.24	3.086	$27.13^{+0.11}_{-0.10}$	$27.31^{+0.12}_{-0.11}$	$27.04^{+0.13}_{-0.11}$
Q1422-d53d	14:24:25.50	+22:55:01.22	...	$26.12^{+0.08}_{-0.07}$	$25.57^{+0.05}_{-0.04}$	$25.28^{+0.04}_{-0.04}$
Q1422-d53e	14:24:25.53	+22:55:00.92	...	$28.48^{+0.15}_{-0.13}$	$27.67^{+0.07}_{-0.07}$	$27.17^{+0.06}_{-0.06}$
Q1422-d53f	14:24:25.54	+22:55:00.83	...	$27.56^{+0.14}_{-0.13}$	$27.07^{+0.09}_{-0.08}$	$26.61^{+0.07}_{-0.07}$

Continued on next page

Table 3.4: Continued.

ID[a]	R.A.	Dec.	$z_{\rm sys}$[b,c]	V_{606}	J_{125}	H_{160}
Q1422-md119a	14:24:36.19	+22:55:40.33	2.751	$26.44^{+0.06}_{-0.05}$	$26.26^{+0.07}_{-0.07}$	$25.92^{+0.07}_{-0.06}$
Q1422-md119b	14:24:36.22	+22:55:40.17	2.751	$26.74^{+0.06}_{-0.06}$	$26.70^{+0.10}_{-0.09}$	$26.78^{+0.13}_{-0.12}$
Q1422-md119c	14:24:36.21	+22:55:40.50	2.751	$25.73^{+0.05}_{-0.05}$	$25.34^{+0.05}_{-0.05}$	$24.86^{+0.04}_{-0.04}$
Q1422-md145a	14:24:35.52	+22:57:18.53	...	$27.78^{+0.10}_{-0.09}$	$26.75^{+0.03}_{-0.03}$	$26.76^{+0.04}_{-0.04}$
Q1422-md145b	14:24:35.56	+22:57:20.26	2.800	$25.76^{+0.05}_{-0.05}$	$25.78^{+0.04}_{-0.04}$	$25.34^{+0.03}_{-0.03}$
Q1422-md145c	14:24:35.53	+22:57:19.36	2.800	$27.92^{+0.12}_{-0.11}$	$27.45^{+0.06}_{-0.06}$	$27.69^{+0.10}_{-0.09}$
Q1549-C10a	15:51:48.45	+19:09:24.67	3.189	$26.84^{+0.07}_{-0.07}$	$27.37^{+0.13}_{-0.12}$	$27.64^{+0.23}_{-0.19}$
Q1549-C10b	15:51:48.42	+19:09:25.01	3.189	$24.73^{+0.03}_{-0.03}$	$24.90^{+0.03}_{-0.03}$	$24.72^{+0.04}_{-0.04}$
Q1549-C27	15:52:07.05	+19:12:19.31	2.926	$24.42^{+0.02}_{-0.02}$	$24.62^{+0.03}_{-0.03}$	$24.26^{+0.03}_{-0.03}$
Q1549-C8	15:51:45.39	+19:08:49.85	2.937	$24.99^{+0.03}_{-0.03}$	$25.15^{+0.05}_{-0.05}$	$24.64^{+0.04}_{-0.04}$
Q1549-D7	15:51:46.25	+19:09:50.10	2.936	$24.18^{+0.02}_{-0.01}$	$24.10^{+0.01}_{-0.01}$	$23.73^{+0.01}_{-0.01}$
Westphal-CC32a	14:18:14.48	+52:28:07.26	3.192	$26.01^{+0.03}_{-0.03}$	$25.92^{+0.03}_{-0.03}$	$25.71^{+0.04}_{-0.03}$
Westphal-CC32b	14:18:14.46	+52:28:06.99	3.192	$24.68^{+0.02}_{-0.02}$	$24.67^{+0.02}_{-0.02}$	$24.49^{+0.03}_{-0.03}$
Westphal-CC41a	14:18:20.55	+52:29:21.14	3.027	$26.01^{+0.05}_{-0.05}$	$26.00^{+0.06}_{-0.06}$	$25.62^{+0.05}_{-0.05}$
Westphal-CC41b	14:18:20.49	+52:29:21.48	3.027	$26.00^{+0.05}_{-0.05}$	$25.84^{+0.05}_{-0.05}$	$25.57^{+0.05}_{-0.05}$
Westphal-CC46	14:18:00.20	+52:29:53.02	3.261	$23.96^{+0.01}_{-0.01}$	$23.76^{+0.01}_{-0.01}$	$23.58^{+0.01}_{-0.01}$
Westphal-CC53	14:18:22.15	+52:30:19.70	2.807	$25.22^{+0.04}_{-0.04}$	$25.08^{+0.04}_{-0.04}$	$24.48^{+0.03}_{-0.03}$
Westphal-DD28	14:18:24.87	+52:29:27.32	3.021	$24.49^{+0.02}_{-0.02}$	$24.22^{+0.02}_{-0.02}$	$23.83^{+0.01}_{-0.01}$
Westphal-MM38	14:18:04.07	+52:29:54.94	2.925	$25.53^{+0.03}_{-0.03}$	$25.16^{+0.03}_{-0.03}$	$24.77^{+0.03}_{-0.03}$

[a] The field the object is located in, the object name, and a letter corresponding to the subcomponent in Figure 3.10. A subcomponent label is omitted in the case of single-component morphology. [b] Systemic redshift from S18. [c] Systemic redshifts are omitted for components predicted as low redshift based on their $V_{606}J_{125}H_{160}$ colors.

CHAPTER 4

Searching for the connection between ionizing-photon escape and the surface density of star formation at $z \sim 3$

4.1 Introduction

Reionization is a key phase transition of the Universe in which the neutral hydrogen in the intergalactic medium (IGM) becomes ionized. While the reionization process concludes by $z \sim 6$ (Fan et al., 2006), the overall evolution of the hydrogen neutral fraction depends on our understanding of the population of the first star-forming galaxies that are thought to emit the bulk of the ionizing photons into the IGM (e.g, Parsa et al., 2018). The ionizing emissivity as a function of cosmic time can be simply parameterized as the product of the comoving UV luminosity density ($\rho_{\rm UV}$), ionizing photon efficiency ($\xi_{\rm ion}$), and escape fraction of ionizing photons that go on to ionize the surrounding H I gas ($f_{\rm esc}$) (Robertson et al., 2015). These parameters and their redshift dependence must be characterized beyond $z \sim 6$ to construct a complete picture of reionization.

The most difficult of the three parameters to constrain is $f_{\rm esc}$, for which estimates rely on direct measurement of the Lyman continuum (LyC; $\lambda < 912$Å) spectral region. Due to the large cross section of H I to LyC photons, even a small neutral fraction ($< 1\%$) can attenuate this signal, thus direct observations become prohibitive at $z \gtrsim 4$ (Vanzella et al., 2012; Steidel et al., 2018). We must use lower-redshift galaxies to obtain direct measurements of escaping LyC emission at the highest redshift where the measurements are possible, and,

simultaneously, connect their properties with those of ionizing sources at $z > 6$. To this end, large-scale LyC surveys have attempted to measure the average $f_{\rm esc}$ of galaxies at $z \sim 2-4$ and to understand the mechanisms of LyC escape. These surveys either attempt to detect LyC photometrically (Grazian et al., 2017; Fletcher et al., 2019; Nakajima et al., 2019; Saxena et al., 2021) or spectroscopically (Marchi et al., 2017, 2018). Of particular note is the Keck Lyman Continuum Spectroscopic Survey (KLCS), which includes 124 objects selected as Lyman Break Galaxies (LBGs) with $\langle f_{\rm esc} \rangle = 0.09 \pm 0.01$ measured through deep, rest-UV spectroscopy (Steidel et al., 2018). A subset of 39 KLCS galaxies was followed up with *Hubble Space Telescope* (*HST*) multi-band imaging, including all galaxies with apparent individual LyC detections. These observations were designed to eliminate the effects of foreground contamination on the average measurements, a significant concern for detecting LyC leakers at $z \sim 3$ (Vanzella et al., 2012; Mostardi et al., 2015). In Pahl et al. (2021), we reported that four KLCS galaxies (two of which had been identified as significant leakers) imaged by *HST* had multi-band photometry consistent with low-redshift interlopers, and updated the average $f_{\rm esc}$ of the KLCS sample to 0.06 ± 0.01 after removing these four galaxies.

Within these studies, one of the parameters that positively correlates with LyC leakage is the equivalent width of Lyα emission (W_λ(Lyα)), in the sense that galaxies with higher W_λ(Lyα) have higher $f_{\rm esc}$ on average (Marchi et al., 2017, 2018; Steidel et al., 2018; Pahl et al., 2021). This correlation suggests that Lyα and LyC leakage in star-forming galaxies are connected because of their similar escape pathways through the interstellar and circumgalactic media (ISM and CGM). Since Lyα emission has been found to be inversely correlated with the neutral gas covering fraction both in the local Universe and at $z \sim 3$ (Reddy et al., 2016b; Gazagnes et al., 2020), it follows that a "picket fence" arrangement of H I would simultaneously modulate LyC emission along with Lyα. The galaxy properties that favor this configuration are an active area of study, and further characterization of the dependence of $f_{\rm esc}$ on galaxy properties is necessary to disentangle this physical picture.

Recently, Naidu et al. (2020) proposed a model of reioniziation in which a galaxy's $f_{\rm esc}$ is solely a function of its star-formation rate (SFR) surface density ($\Sigma_{\rm SFR}$). This model implies that the bulk of ionizing photon leakage at $z > 6$ can be attributed to relatively luminous and massive galaxies ($M_{\rm UV} < -18$ and $\log(M_*/M_\odot) > 8$), in stark contrast to models that favor the dominant role of dwarf galaxies at $M_{\rm UV} > -15$ (e.g., Finkelstein et al., 2019). The parameterization of $f_{\rm esc}$ as a function of $\Sigma_{\rm SFR}$ is motivated by the above-average $\Sigma_{\rm SFR}$ found in many individual LyC detections at $z \sim 2-4$ (De Barros et al., 2016; Vanzella et al., 2016a; Shapley et al., 2016; Bian et al., 2017; Vanzella et al., 2017), as well as in the HST/COS sample of LyC detections at $z \sim 0.3$ (Borthakur et al., 2014; Izotov et al., 2016, 2018, 2021). Additionally, the survey of Marchi et al. (2018) found that stacks of galaxies with small UV spatial size had a significantly higher average LyC flux than those with larger size, implying that compact galaxies have conditions favorable for LyC escape. The spectroscopic survey of Reddy et al. (2022) found Lyα escape can be aided by a high $\Sigma_{\rm SFR}$, especially in low stellar mass galaxies. Finally, hydrodynamic simulations including radiative transfer, such as those of Ma et al. (2016), predict that feedback from massive stars carves channels in the ISM that allow for efficient leakage of LyC. This effect is predicted to be most pronounced in galaxies containing a high density of star formation. The $f_{\rm esc}$-$\Sigma_{\rm SFR}$ paradigm has important implications for testing the model of reionization presented by Naidu et al. (2020), and for informing the evolution of the H I fraction, but has not been confirmed observationally in a large-scale survey with selection criteria independent of $\Sigma_{\rm SFR}$.

In this work, we attempt the first LBG-selected study of $f_{\rm esc}$ as a function of $\Sigma_{\rm SFR}$ using the subset of KLCS galaxies covered by HST imaging. Thanks to the high-resolution HST imaging offered by Pahl et al. (2021), well-constrained size measurements are possible for the first time for 35 of the galaxies in KLCS, including the 13 with significant individual LyC detections. Combined with the ability to estimate SFRs from the wealth of multi-band photometry available for KLCS, this subsample is ideal for the first observational test of of the connection between ionizing photon escape and $\Sigma_{\rm SFR}$. Using the deep, rest-frame

UV spectra of KLCS, which include coverage of LyC at rest-frame $880 - 910$Å at $z \sim 3$, these spectra can be stacked as a function of $\Sigma_{\rm SFR}$ to explore characteristics of the LyC spectral region. While stacking reduces the uncertainty due to line-of-sight variations in IGM transmission, it remains to investigate whether this KLCS *HST* subsample is adequate to constrain any dependence of $f_{\rm esc}$ on $\Sigma_{\rm SFR}$.

The structure of the paper is as follows: in Section 4.2, we describe the observations and data available for the KLCS *HST* sample. In Section 4.3, we describe the SFR and size measurements, which we present in Section 4.4 alongside descriptions and measurements of the composite spectra constructed in two bins of $\Sigma_{\rm SFR}$. In Section 4.5, we examine the robustness of the KLCS *HST* sample for quantifying the correlation between $f_{\rm esc}$ and $\Sigma_{\rm SFR}$, and perform simulations to determine the minimum sample composition for analyzing the relationship between the escape of ionizing radiation and galaxy properties. We summarize our conclusions in Section 4.6.

Throughout this paper, we adopt a standard ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70$ km s^{-1}Mpc^{-1}. The $f_{\rm esc}$ values reported in this paper are absolute escape fractions, equivalent to $f_{\rm esc,abs}$ in Steidel et al. (2018), and defined as the fraction of all H-ionizing photons produced within a galaxy that escapes into the IGM. We also employ the AB magnitude system (Oke & Gunn, 1983).

4.2 Sample

The KLCS (Steidel et al., 2018) includes an analysis of the hydrogen-ionizing spectra of 137 galaxies selected as LBGs (for details of selection see Steidel et al., 2003; Adelberger et al., 2004; Reddy et al., 2012). To this end, deep optical spectra were obtained using the Low-Resolution Imaging Spectrometer (LRIS; Oke et al., 1995; Steidel et al., 2004) on the Keck I telescope, covering the LyC spectral region, as well as the Lyα feature and far-UV metal absorption lines. These data were taken with total exposure times per mask of

~ 10 hrs. Thirteen galaxies were removed from the sample due to clear evidence of blending, the presence of instrumental defects, or with evidence of multiple spectroscopic redshifts, for a final analysis sample of 124 objects. Steidel et al. (2018) characterized the escaping ionizing signal from each galaxy using the average flux density in the wavelength region 880Å $-$ 910Å, or f_{900}. Fifteen galaxies with significant LyC escape (defined as $f_{900} > 3\sigma_{900}$, where σ_{900} is the error on the flux density in the LyC wavelength region) were defined as the LyC detection sample. The remaining 109 galaxies were characterized as individual LyC non-detections. Despite the best efforts to clean the sample of interloper contamination, high-resolution imaging is necessary for identifying all low-redshift interlopers along the line-of-sight that frequently contaminate apparent LyC flux for $z \sim 3$ galaxies (Vanzella et al., 2012; Mostardi et al., 2015).

In Pahl et al. (2021), the 15 LyC detection candidates from KLCS were followed up as part of a Cycle 25 *HST* Program ID 15287 (PI: Shapley) in order to explore the effects of foreground contamination in the KLCS sample. While these LyC detection candidates were specifically targeted as the most likely sources of contamination due to their significant individual f_{900} measurements, 24 LyC individual non-detections were also observed as they fell within the field of view of the *HST* pointings. These 39 objects were observed with ACS/F606W (V_{606}), WFC3/F125W (J_{125}), and WFC3/F160W (H_{160}) between 2017 and 2019 in five KLCS survey fields (Q0933, Q1422, DSF2237b, Q1549, and Westphal). Additional *HST* data were included from Mostardi et al. (2015) and Shapley et al. (2016) for one object, Q1549-C25. The seven ACS pointings and 11 WFC3 pointings were observed for three orbits per filter. We used $V_{606}J_{125}H_{160}$ photometry to characterize the contamination likelihood of individual subcomponents within the LyC detection sample, following the methodology of Mostardi et al. (2015). To this end, two LyC detections were determined to be likely contaminated and were subsequently removed from the sample. Similar contamination analysis was performed on the objects with individual LyC non-detections for completeness. Two galaxies with individual non-detections had subcomponents with $V_{606}J_{125}H_{160}$ photometry consistent with

low-redshift interlopers, thus these objects were also removed, for a remaining KLCS *HST* sample of 35.

We use this "KLCS *HST*" sample of 35 objects that were determined to be free of contamination for the analysis presented here. While all KLCS objects have multi-band photometry that can be used to estimate the SFR through spectral-energy distribution (SED) fitting, only those targeted by *HST* have the higher resolution imaging necessary to measure accurate rest-UV sizes. In this KLCS *HST* sample, there are 13 galaxies with individual LyC detections, and 22 galaxies with individual LyC non-detections, as defined by Steidel et al. (2018). Thus, the KLCS *HST* subsample sample is biased in favor of LyC detections compared to the full sample, which contains 13 individual LyC detections out of a total of 120 galaxies. We discuss the effects of this selection bias in Section 4.4.

4.3 Methodology

In this section, we present the steps necessary to measure $\Sigma_{\rm SFR}$ for the KLCS *HST* sample. This measurement requires sizes, estimated by fitting the two-dimensional light profiles in *HST* imaging, and SFRs, calculated from SED fitting.

4.3.1 Sizes

We estimated the sizes of objects in our sample from model fits to the light distributions in the V_{606} images, which have the highest resolution within our *HST* dataset. For this modeling, we used the software GALFIT (Peng et al., 2002, 2010), which uses parametric fits of one or more Sérsic profiles and a uniform sky background to the two-dimensional postage stamp of a galaxy. At the redshifts of our sample galaxies, single Sérsic profiles have been demonstrated to reasonably constrain the effective sizes of distinct morphological components in *HST* imaging (Van Der Wel et al., 2012; Law et al., 2012; Shibuya et al., 2015; Gillman et al., 2020). We follow Van Der Wel et al. (2012), who presented robust

measurements using Sérsic profiles of basic morphological properties such as half-light radii for galaxies in the CANDELS survey down to $H_{160} \sim 24.5$.

The free parameters of the fits are central position (x, y), magnitude, Sérsic index (n), position angle, axis ratio (b/a), and half-light radius along the major axis (r) for each profile, as well as a uniform sky background level. One of the inputs to the GALFIT modeling is a bad-pixel mask for each object based on the segmentation map produced by *SExtractor* (Bertin & Arnouts, 1996). The *SExtractor* detection threshold was set at 2.5 times the background RMS, and the background was estimated using the "local" setting, which estimates the sky contribution using a rectangular annulus. All components unrelated to the central object were masked, ensuring that the flux from surrounding sources does not affect the fit. Deconvolution was performed using point-spread functions (PSFs) described in Pahl et al. (2021). These PSFs were constructed for each field by averaging $10 - 15$ unsaturated stars in the V_{606} mosaic using the *Astropy* routine *Photutils* (Bradley et al., 2020).

We initialized the positions of the Sérsic profiles for each object using the central coordinates of up to three of the brightest components deblended with *SExtractor*. We set an initial guess of n=4 and r=5 pixels. The initial guesses for the magnitude of each source and the sky background level were drawn from the *SExtractor* catalogs. After fitting, the residuals were inspected by eye to ensure that the background levels were properly subtracted and the models accurately described the radial extent of the light profile. In the top three panels of Figure 4.1, the fitting process is detailed for the galaxy Q1549-D3, with the *HST* V_{606} image, model fit, and residual displayed from left to right.

To estimate an effective size for objects with single-component fits, we calculated the circularized effective radius (r_e) by combining the best-fit r and b/a parameters from GALFIT according to $r_e = r\sqrt{b/a}$. We defined the effective area of the galaxy as πr_e^2, which corresponds to the projected area containing half of the total light of the galaxy.

For objects with multi-component fits, we estimated the effective area using the best-fit model produced by GALFIT. By using the best-fit model without the PSF convolution, we effectively recreated the deconvolved image for each object. We integrated this deconvolved model to find a total flux F, then sorted the pixels in brightness. A number $n_{\rm pix}$ of the brightest pixels were chosen, starting at the brightest pixel and continuing towards fainter pixels, such that the integrated flux contained within $n_{\rm pix}$ was $F/2$. The effective area of an object was then found by applying the *HST* V_{606} pixel scale of $0.03''$/pix, which we subsequently converted to a circularized radius using $n_{\rm pix} \times (0.03(''/{\rm pix}))^2 = \pi r_{\rm e}^2$.

We pursued an empirical approach for estimating the errors on the effective sizes of the galaxies in our sample. We randomly perturbed each pixel of the science image by a value drawn from a Gaussian distribution with standard deviation corresponding to the pixel value in the error image generated from the reduction process of the *HST* V_{606} imaging (See Section 3.2; Pahl et al., 2021). To ensure our error bars encapsulated any systematic error induced by different initial guesses, we also randomly perturbed the initial guesses of all model parameters. Starting with the best-fit parameters from the unperturbed fit, we perturbed the x and y positions by a Gaussian distribution with width 2 pixels, and multiplied r by a factor drawn from a Gaussian distribution with a mean of 1.0 and a standard deviation of 0.3. We randomized n uniformly between 0.5 and 4.0, and a/b between 0.0 and 1.0. We then fit the perturbed science image with these randomized initial guesses following the methods described above and subsequently measured $r_{\rm e}$. After 100 realizations, the standard deviation of the $r_{\rm e}$ distribution was taken as the error on the object's $r_{\rm e}$ measurement. These empirically-determined errors are comparable to those calculated internally by GALFIT for the objects with single-Sérsic fits.

We present the postage stamps and sizes of the 35 KLCS HST objects in Figure 4.1. Displayed in cyan are curves containing the effective areas: for single-component morphologies, these are defined by the Sérsic model parameters r and a/b, and, for multi-component mor-

phologies, the cyan curves enclose the pixels corresponding to the half-light area as described above. In the bottom-right corner of each panel, we present the circularized r_e for each object in both angular ($''$) and physical (kpc) size. Displayed in black is the LRIS slit for each object: all components that contribute to the r_e measurement are contained within the slit for KLCS *HST* targets.

4.3.2 Star-formation rates

We fit multi-band photometry available for our objects with stellar population synthesis models in order to derive galaxy properties such as stellar mass (M_*), $E(B-V)$, stellar age, and SFR. While the SFR measurement is the focus of this work, we will explore other derived galaxy properties of KLCS in future work.

We measured $V_{606}J_{125}H_{160}$ photometry for the KLCS *HST* objects based on *HST* images reduced and described in Pahl et al. (2021). We followed a similar measurement procedure to that of Pahl et al. (2021), however, for consistency with the ground-based measurements, we set the detection threshold used with *SExtractor* to 2.5 times the background RMS, and removed any deblending to get singular measurements for each object.

For additional input photometry that constitute the KLCS *HST* SEDs, we used a multitude of data available for the objects in the survey fields (e.g., Reddy et al., 2012). At minimum, all five fields were covered by ground-based U_nGR_s optical, K_s near-IR, and *Spitzer*/IRAC Channel 1 and 2 infrared imaging. Some fields had additional photometric information, as summarized in Table 4.1. For objects with a clear K_s excess, we performed corrections to the broadband flux based on measured fluxes of the strong emission lines of [OIII] and Hα (Steidel et al., 2014).

We measured the SFRs for each object using SED fitting based on the methods described in Reddy et al. (2018b). Specifically, we used Bruzual & Charlot (2003) stellar population synthesis models and assumed a Salpeter (1955) initial mass function (IMF). We note that,

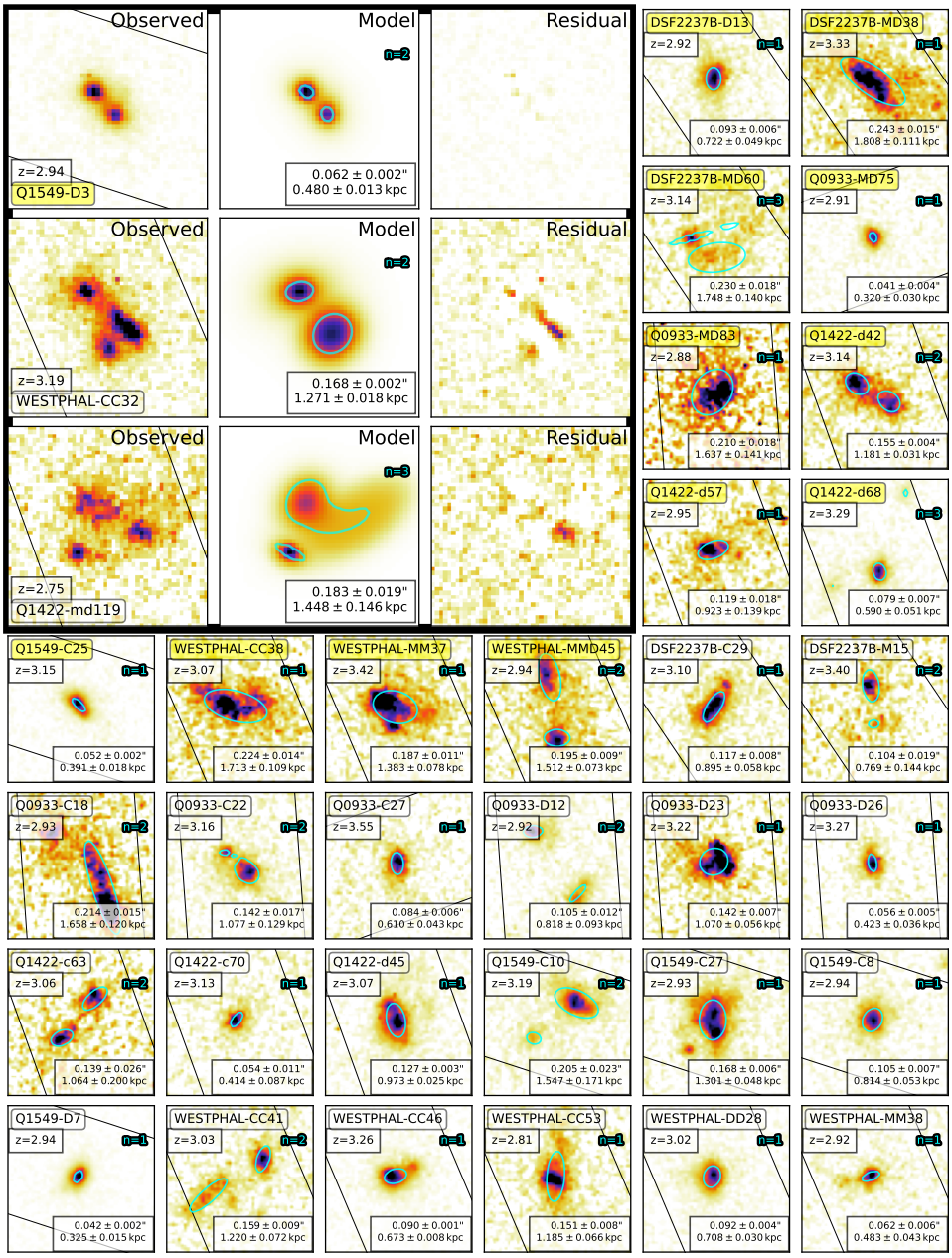

Figure 4.1: V_{606} postage stamps of the 35 galaxies in KLCS *HST*. The panels in the upper left show the HST V_{606} image, the GALFIT model, and the residual, from left to right, for three objects to demonstrate the fitting process. The cyan curve shows the effective area of each object and the black lines show the LRIS slit (1.2″ wide) for each object. Displayed in the bottom right of each panel is the circularized r_e in both angular size (″) and physical (kpc) size. Displayed in the middle right of each panel in cyan are the number of Sérsic profiles used in the model fit. Objects with individual LyC detections have object names highlighted in yellow. All panels are 1.5″ per side.

Table 4.1: Photometric bands used in SED modeling.

Fields	Photometric bands
DSF2237b	U_n^a, G^a, R_s^a, I^a, J^b, K_s^b, IRAC1, IRAC2, V_{606}, J_{125}, H_{160}
Q0933	U_n^a, G^a, R_s^a, I^a, J^b, K_s^b, IRAC1, IRAC2, V_{606}, J_{125}, H_{160}
Q1422	U_n^c, G^c, R_s^c, K_s^b, IRAC1, IRAC2, V_{606}, J_{125}, H_{160}
Q1549	U_n^d, G^d, R_s^d, J^f, H^f, $K_s^{b,f}$, IRAC1, IRAC2, IRAC3, V_{606}, J_{125}, H_{160}
Westphal	u*^{g}, g'^{g}, r'^{g}, i'^{g}, z'^{g}, J^h, H^h, K_s^h, IRAC1, IRAC2, V_{606}, J_{125}, H_{160}

[a] Observed with the COSMIC prime focus imager on the Palomar 5.08 m telescope (see Steidel et al., 2003). [b] Observed with the Multi-Object Spectrograph for Infra-Red Exploration (MOSFIRE) on the Keck I telescope. [c] Observed with the Prime Focus Imager on the William Herschel 4.2m telescope (WHT) (see Steidel et al., 2003). [d] Observed with Keck/LRIS. [f] Observed with FourStar at the Magellan Baade 6.5m telescope. [g] From the Canada-France-Hawaii Telescope (CFHT) Legacy Survey. [h] Observed with CFHT/WIRCam as part of the WIRCam Deep Survey (Bielby et al., 2012).

using the stellar population synthesis models of BPASS (BPASS v2.1, Eldridge et al., 2017) and assuming the same IMF, we find excellent agreement for the inferred SFRs. A constant star-formation history with stellar ages no younger than 50 Myr was assumed, such that stellar ages would not extend to younger ages than the typical dynamical timescales of star-forming galaxies at $z \sim 3$ (Reddy et al., 2012). We fit the SEDs using two sets of assumptions for metallicity and dust attenuation curve: 0.28 times solar metallicity and an SMC extinction curve (i.e., $0.28Z_{\odot}$+SMC, Gordon et al., 2003), and 1.4 times solar metallicity and a Calzetti et al. (2000) extinction curve (i.e., $1.4Z_{\odot}$+Calzetti). For two galaxies with $\log(M_*/M_{\odot})>10.65$ measured using the $0.28Z_{\odot}$+SMC models, we adopted the $1.4Z_{\odot}$+Calzetti model for the final fit, based on systematically better fits to the observed SEDs in Du et al. (2018) and for consistency with the mass-metallicity relation (Sanders et al., 2015; Steidel et al., 2014; Onodera et al., 2016). With the exception of the fits to these two galaxies, we adopted $0.28Z_{\odot}$+SMC models.

To ensure robust SED fits, we individually examined the input photometry, best-fit model, and resulting unreduced χ^2 for each object. Due to the higher S/N of *HST*, we removed MOSFIRE or WIRCAM J, H, and K_s data that were inconsistent with *HST* J_{125}, as long as the IRAC photometry constrained the SED redward of the Balmer break. We also removed IRAC photometry that appeared to be blended with other nearby sources. Additionally, we removed any ground-based photometry if the ground-based extraction may have included contamination from nearby objects identified with the higher-resolution *HST* imaging. After eliminating clearly problematic photometry, we achieved SED fits with unreduced $\chi^2 < 30$ for all 35 objects.

4.4 Results

In this section, we present the $\Sigma_{\rm SFR}$ measurements for the KLCS *HST* sample and the ionizing and non-ionizing spectral properties of the sample as a function of $\Sigma_{\rm SFR}$.

We present the r_e and SFR measurements for our sample in the left panel of Figure 4.2. The median r_e is 0.97 kpc (for a median $M_{UV} = -20.7$), comparable to $1.054^{+1.154}_{-0.436}$ kpc found for $z \sim 3-4$ star-forming galaxies at $M_{UV} = -21.0$ (Shibuya et al., 2015). The median SFR of the sample is 14.0 $M_\odot\,\mathrm{yr}^{-1}$. For each object, we calculate the Σ_{SFR} by assuming half of the star formation occurs within the half-light area, such that

$$\Sigma_{SFR} = \mathrm{SFR}/(2 \times \pi r_e^2), \tag{4.1}$$

where SFR is inferred from the best-fit SED model described in Section 4.3.2, and r_e is the circularized half-light radius measured for each object in Section 4.3.1. We present the resulting Σ_{SFR} distribution in the right panel of Figure 4.2. The median Σ_{SFR} is 2.81 $M_\odot\,\mathrm{yr}^{-1}\,\mathrm{kpc}^{-2}$, which is within the 16th and 84th percentile of the Σ_{SFR} relation from Shibuya et al. (2015) with comparable redshift and M_{UV}.

In order to examine the ionizing and non-ionizing spectral properties of the KLCS *HST* sample as a function of Σ_{SFR}, we split the sample into two bins: $\Sigma_{SFR,low}$, which contains the 17 galaxies with the lowest Σ_{SFR}, and $\Sigma_{SFR,high}$, which contains the 18 galaxies with the highest Σ_{SFR}. The two bins are differentiated by black and orange color for $\Sigma_{SFR,low}$ and $\Sigma_{SFR,high}$, respectively, in Figure 4.2. The median properties of the samples are summarized in Table 4.2.

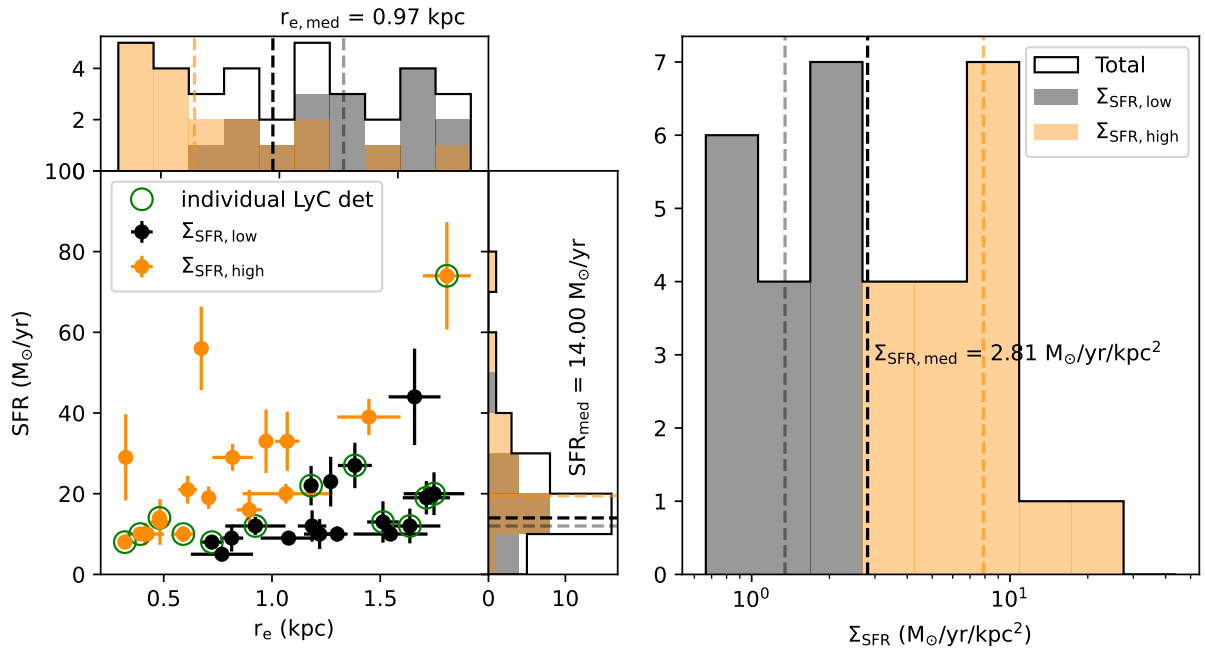

Figure 4.2: r_e, SFR, and Σ_{SFR} measurements for the KLCS *HST* sample. The 35 objects are split into two bins of Σ_{SFR}: $\Sigma_{SFR,low}$ contains the 17 objects with the lowest Σ_{SFR}, and $\Sigma_{SFR,high}$ contains the 18 objects with the highest Σ_{SFR}. Objects in the $\Sigma_{SFR,low}$ bin are displayed as black points in the scatter plot and black bars in the histograms, while objects in the $\Sigma_{SFR,high}$ bin are displayed as orange points and orange bars. The thirteen objects with individual LyC detections are highlighted with a green circle. **Left:** The SFR-r_e distribution of the sample. The dark, black line in the vertical and horizontal histograms show the total r_e and SFR distributions. The sample medians are shown by dashed lines, with the $\Sigma_{SFR,low}$ medians in grey, the $\Sigma_{SFR,high}$ medians in orange, and the total sample median in black. **Right:** The Σ_{SFR} distribution of the sample, calculated according to Equation 4.1. The color convention follows the histograms in the left panel.

Table 4.2: Median properties of the two $\Sigma_{\rm SFR}$ subsamples and resulting composite spectra.

	$r_{\rm e,med}$ (kpc)	$\rm SFR_{med}$ ($\rm M_\odot yr^{-1}$)	$\Sigma_{\rm SFR,med}$ ($\rm M_\odot yr^{-1} kpc^{-2}$)	W_λ(Lyα) (Å)	W_λ(LIS) (Å)	W_λ(HIS) (Å)	$\langle f_{900}/f_{1500}\rangle_{\rm out}$	$f_{\rm esc,abs}$
$\Sigma_{\rm SFR,low}$	1.27 ± 0.11	12.0 ± 2.4	1.35 ± 0.40	10.7 ± 4.9	1.78 ± 0.21	2.81 ± 0.32	0.104 ± 0.013	0.120 ± 0.009
$\Sigma_{\rm SFR,high}$	0.64 ± 0.12	19.5 ± 4.6	7.89 ± 1.53	19.1 ± 10.8	1.66 ± 0.20	3.20 ± 0.32	0.076 ± 0.013	0.122 ± 0.010

We then created composite spectra as described in Steidel et al. (2018) and Pahl et al. (2021) in order to explore spectral differences between the $\Sigma_{\rm SFR,low}$ and $\Sigma_{\rm SFR,high}$ samples and constrain the dependence of LyC ionizing photon escape on $\Sigma_{\rm SFR}$. The rest-frame UV, LRIS spectra of Steidel et al. (2018) covering the LyC region are normalized to the average flux density within the spectral region $1475 \leq \lambda_0/\text{Å} \leq 1525$ (f_{1500}). We use a spline interpolation to place the spectra on a common wavelength grid, then use a mean combination at each wavelength increment to create the final composite after applying 3σ outlier rejection.

Building spectral composites smooths out variations in intergalactic medium (IGM) sight lines. However, the aggregate effect of the IGM and circumgalactic medium (CGM) strongly attenuates the LyC region at $z \sim 3$. To account for this attenuation, we correct the composites using IGM+CGM transmission functions calculated in a manner consistent with that in Steidel et al. (2018). Notably, for the KLCS *HST* objects, we choose a transmission function drawn from the third quartile of IGM transparency, as opposed to the average transmission used for the composites of Steidel et al. (2018). This decision is based on the oversampling of objects with individual LyC detections in the KLCS *HST* sample (13 detections out of 35 objects) compared to the full KLCS sample (13 detections out of 120 objects, see Section 4.2). Preferentially targeting individually-detected LyC candidates has the approximate effect of choosing IGM sight-lines that are more transparent to LyC photons, on average, considering that part of the variation in observed ionizing signal at $z \sim 3$ is due to differences in IGM opacity (Steidel et al., 2018). Assuming the LyC individual detections are drawn from the top 11% of IGM transmission[1] and the 22 LyC non-detections have average sight-lines, we consider the third-quartile transmission function as a reasonable approximation of the IGM transmission for the KLCS *HST* sample bins. Below, we discuss the results based on a different assumption for IGM transmission, demonstrating that our key conclusions do not change.

[1]The 13 LyC individual detections had the highest f_{900} out of the parent sample of 120.

The $\Sigma_{\rm SFR,low}$ and $\Sigma_{\rm SFR,high}$ composite spectra, corrected for IGM+CGM attenuation, are shown in Figure 4.3. The LyC region ($880 \leq \lambda_0/\text{Å} \leq 910$) is visually demarcated by dotted-red lines. We calculated the average ionizing flux density of each composite within this region and defined it as $\langle f_{900} \rangle$. The ratio of ionizing to non-ionizing flux density, with f_{900} corrected for effects of the IGM and CGM, is $\langle f_{900}/f_{1500} \rangle_{\rm out} = 0.104 \pm 0.013$ for the $\Sigma_{\rm SFR,low}$ composite and $\langle f_{900}/f_{1500} \rangle_{\rm out} = 0.076 \pm 0.013$ for the $\Sigma_{\rm SFR,high}$ composite. Both of these $\langle f_{900}/f_{1500} \rangle_{\rm out}$ measurements are significantly elevated from the full, uncontaminated KLCS sample measurement of $\langle f_{900}/f_{1500} \rangle_{\rm out} = 0.015 \pm 0.02$ (Pahl et al., 2021). Again, this difference is expected due to the larger proportion of individually-detected LyC galaxies in KLCS *HST*. We summarize the sample composition of each $\Sigma_{\rm SFR}$ bin in Table 4.3 and the properties of each KLCS *HST* galaxy. In $\Sigma_{\rm SFR,low}$, 8/17 galaxies are LyC individual detections and in $\Sigma_{\rm SFR,high}$, 5/18 galaxies are individually detected.

The property $\langle f_{900}/f_{1500} \rangle_{\rm out}$ remains a useful empirical measurement for estimating the ionizing photon budget with knowledge of the UV luminosity function. In order to estimate the quantity more closely tied to reionization models, $f_{\rm esc}$, we model the underlying stellar populations of these galaxies as well as the structure of the neutral-phase inter-stellar medium (ISM). Following Steidel et al. (2018) and Pahl et al. (2021), we use stellar-population synthesis models from BPASS (BPASS v2.1, Eldridge et al., 2017) together with the SMC attenuation relation (Gordon et al., 2003) at a range of $E(B-V)$ to estimate the intrinsic production rate of ionizing photons, together with geometric ISM modeling according to the "holes" model (also see Reddy et al., 2016b, 2022). While the free parameters in the fit include continuum reddening in the foreground gas $E(B-V)_{\rm cov}$, H I column density $\log(N_{\rm HI}/{\rm cm}^{-2})$, and covering fraction of neutral gas and dust $f_{\rm c}$, in this work we are specifically interested in $f_{\rm esc}$, defined as $f_{\rm esc} = 1 - f_{\rm c}$ for the "holes" ISM model. Best-fit values of $\langle f_{\rm esc} \rangle = 0.120 \pm 0.009$ for $\Sigma_{\rm SFR,low}$ and $\langle f_{\rm esc} \rangle = 0.122 \pm 0.010$ for $\Sigma_{\rm SFR,high}$ are shown as green points in Figure 4.4, presented as a function of median $\Sigma_{\rm SFR}$ (see right panel of Figure 4.2). Here, the $f_{\rm esc}$ estimates assuming an SMC attenuation relation and third-quartile IGM+CGM transparency

Table 4.3: Properties of the KLCS *HST* sample, split into two bins of $\Sigma_{\rm SFR}$.

	$z_{\rm sys}$	$r_{\rm e}$ (kpc)	SFR ($M_\odot yr^{-1}$)	$\Sigma_{\rm SFR}$ ($M_\odot yr^{-1}kpc^{-2}$)	W_λ(Lyα) (Å)	$L_{\rm UV}/L^*_{\rm UV}$	$\langle f_{900}/f_{1500}\rangle_{\rm obs}$
				$\Sigma_{\rm SFR,low}$			
Q1549-C10	3.1894	1.55 ± 0.17	10.0 ± 1.5	0.66 ± 0.18	23.2	0.75	0.022 ± 0.016
Q0933-MD83	2.8800	1.64 ± 0.14	12.0 ± 4.3	0.71 ± 0.28	3.8	0.60	0.150 ± 0.041
Westphal-MMD45	2.9357	1.51 ± 0.07	13.0 ± 5.1	0.90 ± 0.37	−19.1	1.42	0.086 ± 0.023
Q1549-C27	2.9256	1.30 ± 0.05	10.0 ± 1.9	0.94 ± 0.19	14.5	0.88	-0.014 ± 0.019
Westphal-CC38	3.0729	1.71 ± 0.11	19.0 ± 4.1	1.03 ± 0.26	8.6	1.00	0.056 ± 0.013
DSF2237b-MD60	3.1413	1.75 ± 0.14	20.0 ± 5.3	1.04 ± 0.32	7.2	0.67	0.090 ± 0.018
Westphal-CC41	3.0268	1.22 ± 0.07	10.0 ± 3.7	1.07 ± 0.42	9.5	0.62	-0.031 ± 0.022
Q0933-C22	3.1639	1.08 ± 0.13	9.0 ± 0.7	1.23 ± 0.31	57.6	0.73	-0.008 ± 0.020
DSF2237b-M15	3.4034	0.77 ± 0.14	5.0 ± 0.9	1.35 ± 0.55	6.7	0.65	-0.010 ± 0.031
Westphal-CC53	2.8070	1.19 ± 0.07	12.0 ± 3.9	1.36 ± 0.46	−12.3	0.49	0.013 ± 0.029
Q1549-C8	2.9373	0.81 ± 0.05	9.0 ± 3.3	2.16 ± 0.85	−10.4	0.54	0.017 ± 0.030
Q1422-d57	2.9461	0.92 ± 0.14	12.0 ± 2.2	2.24 ± 0.79	52.8	0.30	0.294 ± 0.103
Westphal-MM37	3.4215	1.38 ± 0.08	27.0 ± 5.6	2.25 ± 0.53	7.4	1.23	0.046 ± 0.009
Westphal-CC32	3.1924	1.27 ± 0.02	23.0 ± 6.2	2.27 ± 0.61	15.6	1.40	0.009 ± 0.009
DSF2237b-D13	2.9216	0.72 ± 0.05	8.0 ± 1.6	2.44 ± 0.58	−3.8	0.58	0.077 ± 0.019
Q1422-d42	3.1369	1.18 ± 0.03	22.0 ± 4.9	2.51 ± 0.57	−10.5	0.47	0.142 ± 0.038
Q0933-C18	2.9261	1.66 ± 0.12	44.0 ± 12.0	2.55 ± 0.78	−13.0	0.76	0.015 ± 0.030
				$\Sigma_{\rm SFR,high}$			
Q1422-c63	3.0591	1.06 ± 0.20	20.0 ± 2.4	2.81 ± 1.11	−0.1	0.28	-0.005 ± 0.069
Q1422-md119	2.7506	1.45 ± 0.15	39.0 ± 4.5	2.96 ± 0.69	−5.0	0.51	0.084 ± 0.084
DSF2237b-C29	3.0999	0.90 ± 0.06	16.0 ± 5.0	3.18 ± 1.07	27.8	0.64	-0.036 ± 0.018
DSF2237b-MD38	3.3278	1.81 ± 0.11	74.0 ± 13.3	3.60 ± 0.78	−13.8	1.47	0.075 ± 0.009
Q1422-d68	3.2865	0.59 ± 0.05	10.0 ± 2.0	4.57 ± 1.19	153.4	0.88	0.200 ± 0.028
Q0933-D23	3.2241	1.07 ± 0.06	33.0 ± 7.3	4.58 ± 1.13	−14.9	1.29	0.003 ± 0.016
Q1422-d45	3.0717	0.97 ± 0.02	33.0 ± 7.9	5.55 ± 1.35	0.3	1.38	0.000

Continued on next page.

Table 4.3: Continued.

	$z_{\rm sys}$	$r_{\rm e}$ (kpc)	SFR ($M_\odot yr^{-1}$)	$\Sigma_{\rm SFR}$ ($M_\odot yr^{-1}kpc^{-2}$)	W_λ(Lyα) (Å)	$L_{\rm UV}/L^*_{\rm UV}$	$\langle f_{900}/f_{1500}\rangle_{\rm obs}$
Westphal-DD28	3.0206	0.71 ± 0.03	19.0 ± 2.8	6.04 ± 1.02	-14.6	1.62	-0.036 ± 0.017
Q0933-D12	2.9242	0.82 ± 0.09	29.0 ± 3.3	6.90 ± 1.76	3.5	1.13	0.028 ± 0.025
Westphal-MM38	2.9247	0.48 ± 0.04	13.0 ± 5.7	8.88 ± 4.19	-7.0	0.64	-0.000
Q0933-D26	3.2662	0.42 ± 0.04	10.0 ± 1.2	8.88 ± 1.87	16.0	0.87	0.052 ± 0.027
Q0933-C27	3.5463	0.61 ± 0.04	21.0 ± 3.5	8.97 ± 1.94	46.6	0.59	-0.042 ± 0.021
Q1422-c70	3.1286	0.41 ± 0.09	10.0 ± 2.8	9.28 ± 4.66	6.3	0.42	0.020 ± 0.056
Q1549-D3	2.9373	0.48 ± 0.01	14.0 ± 1.7	9.66 ± 1.31	8.7	1.16	0.056 ± 0.009
Q1549-C25	3.1526	0.39 ± 0.02	10.0 ± 1.3	10.41 ± 1.67	16.6	0.74	0.083 ± 0.019
Q0933-MD75	2.9131	0.32 ± 0.03	8.0 ± 0.9	12.40 ± 2.74	76.9	0.89	0.095 ± 0.023
Westphal-CC46	3.2608	0.67 ± 0.01	56.0 ± 10.3	19.66 ± 3.66	1.2	2.33	0.004 ± 0.006
Q1549-D7	2.9362	0.32 ± 0.01	29.0 ± 10.7	43.82 ± 16.63	11.5	1.07	0.001 ± 0.018

Individual detections are highlighted in bold.

are adopted as fiducial. The estimates of $f_{\rm esc}$ and $\langle f_{900}/f_{1500}\rangle_{\rm out}$ for our fiducial model are also provided in Table 4.2 for the $\Sigma_{\rm SFR,low}$ and $\Sigma_{\rm SFR,high}$ composites.

We do not find a significant trend of $f_{\rm esc}$ with $\Sigma_{\rm SFR}$ within the KLCS *HST* sample. We compare this trend to the $f_{\rm esc}$-$\Sigma_{\rm SFR}$ relation from "Model II" of Naidu et al. (2020). This model used Copernicus complexio Low Resolution (COLOR) dark-matter simulations (Sawala et al., 2016; Hellwing et al., 2016) together with assumptions of SFR (Tacchella et al., 2018) and a relation between galaxy size and halo size, and was constrained using a variety of observational data on reionization including the contaminated $f_{\rm esc} = 0.09 \pm 0.01$ measurement from Steidel et al. (2018). This relation from Naidu et al. (2020) is $f_{esc} = 1.6\Sigma^{0.4}_{\rm SFR}$. We display this relation as a black dashed line in Figure 4.4. Despite the elevated $f_{\rm esc}$ measurements drawn from the KLCS *HST* sample compared to an unbiased KLCS, our $\Sigma_{\rm SFR,high}$ measurement of $f_{\rm esc}$ falls significantly below the value predicted by the Naidu et al. (2020) relation for the same $\Sigma_{\rm SFR}$. This 10σ deviation may be caused by low-redshift interlopers

contributing to an elevated $f_{\rm esc}$ from Steidel et al. (2018), which Naidu et al. (2020) used as a constraint for the model fit, however the difference between the Steidel et al. (2018) $f_{\rm esc}$ and the uncontaminated measurement from Pahl et al. (2021) is only 30%.

We additionally measure $f_{\rm esc}$ using an alternative attenuation relation and average (as opposed to third-quartile) transparency of the IGM+CGM. These results, including attenuation relations from Reddy et al. (2016a), are displayed alongside the fiducial model in Figure 4.4. This second set of assumptions also does not result in a significant trend in $f_{\rm esc}$ across the two $\Sigma_{\rm SFR}$ bins.

While the measurements of the ionizing spectral region do not indicate a dependence of ionizing photon escape on $\Sigma_{\rm SFR}$, the longer rest-UV wavelength spectral features in the lower panels of Figure 4.3 appear qualitatively distinct for the two bins in $\Sigma_{\rm SFR}$. We measure the equivalent widths of prominent rest-UV features such as Lyα, low-ionization interstellar absorption lines (Si IIλ1260, O Iλ1302+Si IIλ1304, C IIλ1334, and Si IIλ1527), and high-ionization interstellar absorption lines (Si IV$\lambda\lambda$1393, 1402 and C IVλ1548, 1550) following the procedure of Pahl et al. (2020) and Kornei et al. (2010). These methods bootstrap each subsample 500 times with replacement, then perturb each individual spectrum at each pixel by a Gaussian with standard deviation set by the error spectrum. Thus, the error bars on the equivalent width measurements include both sample variance and measurement errors. We measure the Lyα equivalent width, W_λ(Lyα), using the "emission" method from Kornei et al. (2010)[2], and measure absorption line equivalent widths by integrating the absorption profile between $\pm$5Å of the line center after performing continuum normalization using the IRAF routine *continuum*. We present the equivalent-width measurements for both composites in Table 4.2. W_λ(Lyα) is elevated in the $\Sigma_{\rm SFR,high}$ composite compared to the $\Sigma_{\rm SFR,low}$ composite, while W_λ(LIS) decreases and W_λ(HIS) increase. However, none of the differences

[2]Lyα flux was integrated between the wavelength values that the Lyα profile intersected the blue-side continuum value (average flux density between 1220Å $-$ 1180Å) and the red-side continuum value (average flux density between 1225Å $-$ 1255Å).

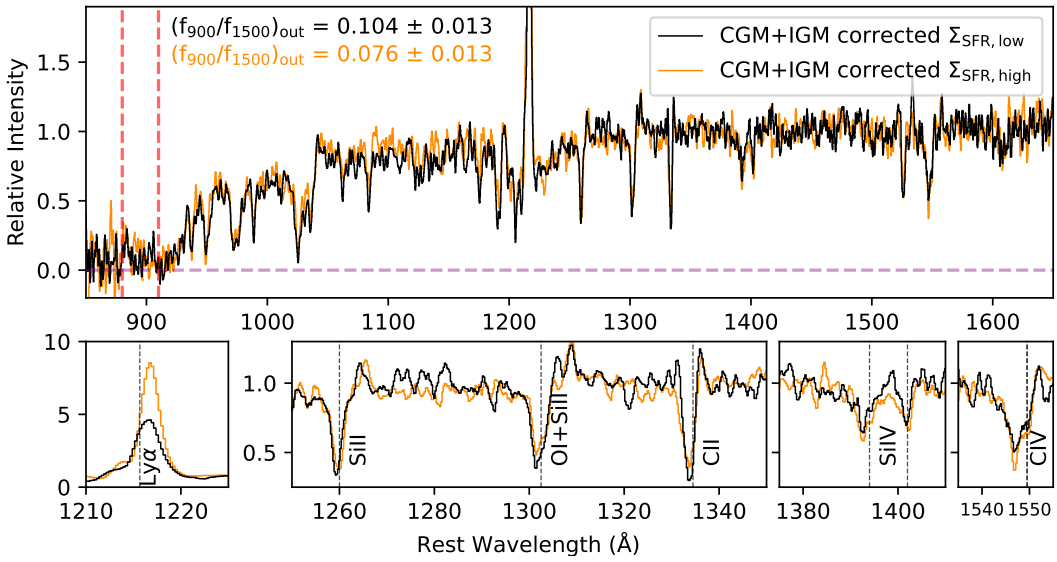

Figure 4.3: Composite spectra of the two KLCS *HST* subsamples binned in $\Sigma_{\rm SFR}$. The $\Sigma_{\rm SFR,low}$ composite, containing objects with the lowest $\Sigma_{\rm SFR}$, is displayed in black, while $\Sigma_{\rm SFR,high}$, containing objects with the highest $\Sigma_{\rm SFR}$, is displayed in orange. The LyC spectral region, which defines the f_{900} measurement, is within the dashed, red lines. **Top:** The full composites, normalized to f_{1500}. **Bottom:** The same composites as above with wavelength and intensity limits set to highlight specific spectral features. From left to right, the panels highlight Lyα; Si IIλ1260, O Iλ1302+Si IIλ1304 and C IIλ1334; Si IV$\lambda\lambda$1393, 1402; and C IVλ1548, 1550.

is significant. When the bootstrapping is removed from the error determinations to test the contribution of sample variance, the elevation in W_λ(Lyα) and W_λ(HIS) in $\Sigma_{\rm SFR,high}$ become significant. The lack of correlation between $f_{\rm esc}$ and $\Sigma_{\rm SFR}$, despite the trends in W_λ(Lyα), as well as the large sample variance in the equivalent width measurements, motivate a closer examination of how representative the KLCS *HST* sample is and its ability to recover a significant relation between $f_{\rm esc}$ and $\Sigma_{\rm SFR}$.

4.5 Discussion

In this section, we judge our ability to recover trends between galaxy properties and ionizing photon escape using the KLCS *HST* sample, and perform simulations to determine the

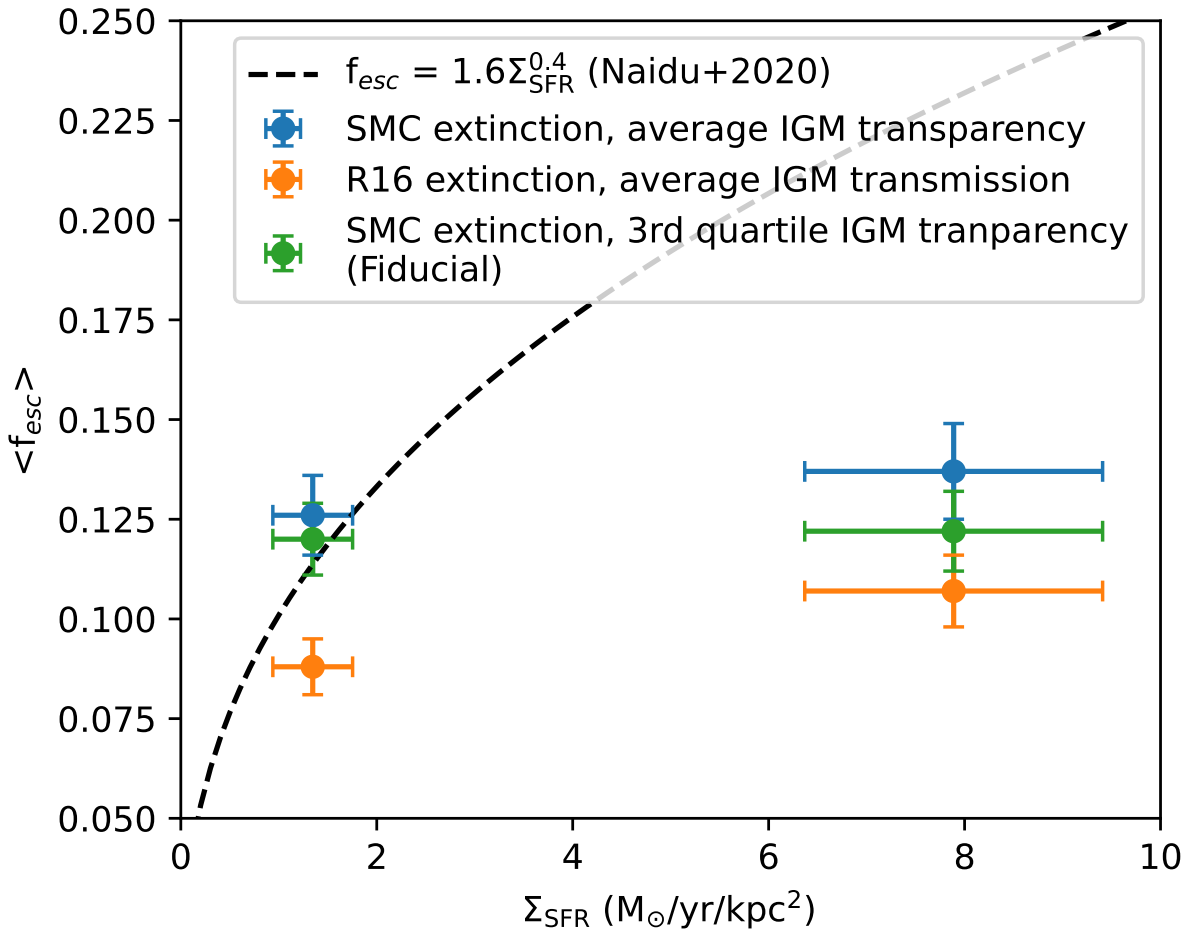

Figure 4.4: $f_{\rm esc}$ measurements for the two $\Sigma_{\rm SFR}$ composite spectra. The measurements assuming SMC extinction and third quartile IGM+CGM transparency, our fiducial model, are displayed in green. Additional measurements including a Reddy et al. (2016a) attenuation relation and average IGM+CGM transparency are shown as blue and orange trend lines. The $f_{\rm esc}$-$\Sigma_{\rm SFR}$ relation from "Model II" of Naidu et al. (2020) is shown as a dashed, black line.

minimum sample size for performing these types of analyses.

$W_\lambda(\mathrm{Ly}\alpha)$ is a well-established tracer of LyC escape in $z \sim 3$ galaxies. In the full KLCS sample, both $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ and f_{esc} are strongly correlated with $W_\lambda(\mathrm{Ly}\alpha)$, measured across four equal-sized bins of $W_\lambda(\mathrm{Ly}\alpha)$ (Steidel et al., 2018; Pahl et al., 2021). This correlation has also been found in other LyC surveys at $z \sim 3-4$ (Marchi et al., 2017, 2018; Fletcher et al., 2019), and is motivated physically by the similar modulation of escaping LyC and Lyα photons by the neutral gas covering fraction (Reddy et al., 2016b; Gazagnes et al., 2020; Steidel et al., 2018). "Holes" in the ISM and CGM, free of H I and dust, would provide clear channels for both LyC and Lyα photons to escape in the observer's direction. Nonetheless, Lyα photons may also escape off-resonance, explaining the scatter in the ratio of escaping Lyα to escaping LyC. For these reasons, we use the relationship between $W_\lambda(\mathrm{Ly}\alpha)$ and $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ as an indicator of the robustness of the KLCS *HST* sample, such that, if we can recover the correlations between $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ and $W_\lambda(\mathrm{Ly}\alpha)$, the composition and size of the sample is appropriate for measuring the correlation between f_{esc} and Σ_{SFR} presented in Section 4.4.

We split the KLCS *HST* sample into two bins of $W_\lambda(\mathrm{Ly}\alpha)$, identical in size to those of $\Sigma_{\mathrm{SFR,low}}$ and $\Sigma_{\mathrm{SFR,high}}$. $W_\lambda(\mathrm{Ly}\alpha)_{\mathrm{low}}$ contains the 17 objects with the lowest $W_\lambda(\mathrm{Ly}\alpha)$ measurements ($W_\lambda(\mathrm{Ly}\alpha)_{\mathrm{med}} = -7.0 \pm 3.8$), and $W_\lambda(\mathrm{Ly}\alpha)_{\mathrm{high}}$ contains the 18 objects with the highest $W_\lambda(\mathrm{Ly}\alpha)$ measurements ($W_\lambda(\mathrm{Ly}\alpha)_{\mathrm{med}} = 15.8 \pm 6.1$). We constructed composite spectra and performed corrections for the effects of the IGM+CGM as described in Section 4.4. The resulting $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ measurements are 0.11 ± 0.02 and 0.11 ± 0.02 for $W_\lambda(\mathrm{Ly}\alpha)_{\mathrm{low}}$ and $W_\lambda(\mathrm{Ly}\alpha)_{\mathrm{high}}$, respectively. This lack of correlation between $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ and $W_\lambda(\mathrm{Ly}\alpha)$ indicates that the KLCS *HST* sample is not capable of recovering well-established relationships between LyC escape and key galaxy properties (i.e., $W_\lambda(\mathrm{Ly}\alpha)$). We repeat this experiment with two bins of increasing L_{UV}, which is found to be strongly anti-correlated with $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ in the full KLCS (Steidel et al., 2018; Pahl et al., 2021). We measured

$\langle f_{900}/f_{1500}\rangle_{\rm out} = 0.12 \pm 0.02$ from the composite made of the 17 galaxies with the lowest $L_{\rm UV}$, and $\langle f_{900}/f_{1500}\rangle_{\rm out} = 0.12 \pm 0.02$ from the composite made of the 18 galaxies with the highest $L_{\rm UV}$, supporting the assertion that the KLCS *HST* sample is not suitable for performing correlation analysis with ionizing-photon escape and galaxy property. Finally, we bin the sample as a function of $r_{\rm e}$, as a correlation between LyC leakage and rest-UV size has been found in a sample of 201 galaxies from the Vimos Ultra Deep Survey at $3.5 \leq z \leq 4.3$ (Marchi et al., 2018). We measured $\langle f_{900}/f_{1500}\rangle_{\rm out} = 0.11 \pm 0.02$ and $\langle f_{900}/f_{1500}\rangle_{\rm out} = 0.11 \pm 0.02$ from the lower- and higher-$r_{\rm e}$ composites, respectively, finding no significant correlation between $r_{\rm e}$ and $\langle f_{900}/f_{1500}\rangle_{\rm out}$ in the KLCS *HST* sample.

The *HST* program that defines the KLCS *HST* sample was designed to cover all 15 galaxies that were candidates for individual LyC detections in KLCS (Steidel et al., 2018). These *HST* observations yielded 22 additional KLCS galaxies with individual LyC non-detections that fell within the footprint of the *HST*/ACS and WFC3 pointings (see Section 4.2). Given that these LyC non-detections are included in KLCS *HST* only on the basis of their proximity to individual apparent LyC detections, they should be representative of the population of 106 galaxies with LyC non-detections in the full sample. Briefly, we consider the sample properties of the 22 individual non-detections in KLCS *HST* compared to the 84 not covered by *HST* imaging. The median W_λ(Lyα) of the KLCS *HST* non-detections is 4.9 ± 2.8Å, while the median W_λ(Lyα) of the non-*HST* non-detections is 3.7±2.9Å. The *HST* subsample does not appear different from the parent sample in W_λ(Lyα), neither in sample median nor sample distribution. The median $L_{\rm UV}/L^*_{\rm UV}$ of the *HST* and non-*HST* non-detections are 0.7 ± 0.1 and 1.04 ± 0.04, respectively. The galaxies with individual LyC non-detections observed by *HST* appear significantly fainter than the parent sample, which could affect the dynamic range of galaxy properties that the sample is probing. At the same time, the lower $L_{\rm UV}$ of the *HST* non detections means the constraints on $\langle f_{900}/f_{1500}\rangle_{\rm out}$ (due to the lower average f_{1500}) are weaker for the *HST* sample. Additionally, the LyC individual detections have a median $L_{\rm UV}/L^*_{\rm UV} = 0.88 \pm 0.14$, fainter than that of the objects not covered by

HST imaging. This difference in the individual LyC detections is expected considering the inverse correlation with $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and $L_{\rm UV}$ found in the full sample, but further contributes to KLCS *HST* representing a biased sampling of KLCS in $L_{\rm UV}$. Upon stacking the *HST* non-detections, we find the resulting $\langle f_{900}/f_{1500}\rangle_{\rm out}$ is comparable to that of the non-*HST* non-detections, despite the inverse correlation between $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and $L_{\rm UV}$. In addition to being systematically fainter, the KLCS *HST* sample does not appear to follow the canonical relationship between $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and $L_{\rm UV}$.

Despite the failure of KLCS *HST* to recover established trends with $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and galaxy properties due to its sample construction, we can still use the full KLCS sample to understand how to design future observing programs of sufficient size to recover robust trends. To inform future observing programs that select a subset of the KLCS sample to analyze, we created simulations that explore the feasibility of recovering trends between $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and W_λ(Lyα) with a range of sample sizes. Here, we assume that a trend between $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and W_λ(Lyα) must be recovered in order to confidently constrain trends between $\langle f_{900}/f_{1500}\rangle_{\rm out}$ other galaxy properties, considering the robustness of the $\langle f_{900}/f_{1500}\rangle_{\rm out}$-$W_\lambda$(Ly$\alpha$) connection in large-scale LyC studies (Steidel et al., 2018; Gazagnes et al., 2020). If a $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and W_λ(Lyα) trend is not recovered, the sample is unlikely to be able to make confident claims about how $\langle f_{900}/f_{1500}\rangle_{\rm out}$ depends on other variables.

In the following set of simulations, we assumed that the 13 objects individually detected in LyC are always included in the analysis, similar to the design of the *HST* program described in Pahl et al. (2021). We began by randomly selecting a number of objects n with individual LyC non-detections from KLCS. We added all 13 confirmed LyC individual detections to the sample, then split the sample into two equal bins of increasing W_λ(Lyα), with a bin size of $(13+n)/2$.[3] After constructing composite spectra, we performed IGM+CGM cor-

[3]In the cases where the total sample size was odd, the higher W_λ(Lyα) bin had one more object than the lower W_λ(Lyα) bin.

rections using average IGM transparency for simplicity, then measured $\langle f_{900}/f_{1500}\rangle_{\rm out}$ from the IGM+CGM corrected composites. We repeated this experiment 1000 times for different random draws of n galaxies with individual LyC non-detections in KLCS, always including the 13 confirmed LyC individual detections. For each iteration, we determined whether the subsequent $\langle f_{900}/f_{1500}\rangle_{\rm out}$ measurements recovered a significant increase in $\langle f_{900}/f_{1500}\rangle_{\rm out}$ across the two bins of increasing W_λ(Lyα) as reported in Steidel et al. (2018) and Pahl et al. (2021). A 1σ increase was defined as

$$(\langle f_{900}/f_{1500}\rangle_{\rm out,1} - \langle f_{900}/f_{1500}\rangle_{\rm out,2}) > \sqrt{(\sigma_{fout,1})^2 + (\sigma_{fout,2})^2}, \tag{4.2}$$

where the subscript "1" denotes measurements associated with the $W_\lambda(\text{Ly}\alpha)_{\rm high}$ bin and the subscript "2" denotes measurements of the $W_\lambda(\text{Ly}\alpha)_{\rm low}$ bin, and σ_{fout} is the error on the corresponding $\langle f_{900}/f_{1500}\rangle_{\rm out}$ measurement. If this condition was met, we counted the trial as a "success." The proportion of trials that successfully recovered a 1σ increase in $\langle f_{900}/f_{1500}\rangle_{\rm out}$ for a range of n is presented as the blue, dashed curve in the upper panel of Figure 4.5. For $n = 22$, identical to the number of non-detections included in KLCS *HST* (and plotted as a vertical, purple, dashed line), a 1σ increase in $\langle f_{900}/f_{1500}\rangle_{\rm out}$ across two bins of increasing W_λ(Lyα) was found in 86.6% of random draws. In order to ensure that a 1σ increase is recovered >95% of the time, $n = 45$ was required, for a total sample size of 58. For each iteration, we also measured the ratio of ionizing to non-ionizing flux density before IGM+CGM corrections, or $\langle f_{900}/f_{1500}\rangle_{\rm obs}$, to explore the effects of the uncertainty of the IGM and CGM corrections on the simulations. The percentage of trials that successfully measured an increase in $\langle f_{900}/f_{1500}\rangle_{\rm obs}$ is presented as a function of n in the red, dashed line in Figure 4.5. As expected, because of the reduced uncertainty, removing corrections for the attenuation of the IGM+CGM induced "significant" differences in $\langle f_{900}/f_{1500}\rangle_{\rm obs}$ across the two W_λ(Lyα) bins more often. We additionally include solid curves in Figure 4.5 that represent identical simulations with a stricter success threshold of 2σ (i.e., a difference in $\langle f_{900}/f_{1500}\rangle_{\rm out}$ for the two bins that exceeds twice the error on the difference). For $n = 22$, a 2σ increase in $\langle f_{900}/f_{1500}\rangle_{\rm out}$ was recovered for 49.3% of the iterations, while $n = 91$ (total

sample size of 104) was required for a 2σ increase to be recovered for >95% of iterations.

We performed a similar set of simulations without the stipulation that objects with individual LyC detections in the KLCS are preferentially observed. In this case, the subsamples for analysis were drawn randomly from the 120 galaxies in the full KLCS sample. This set of simulations is comparable to a survey that is blind to the ionizing properties of the galaxies before analysis and stacking. Here, we chose n galaxies from KLCS such that n is the complete sample size of the subsample. The other steps of this set of simulations were identical to those described above, including the binning as a function of $W_\lambda(\mathrm{Ly}\alpha)$, generation of composite spectra, correcting for the IGM+CGM, and measuring $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$. The percentage of trials that recovered a 1σ increase in $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ across the two bins of increasing $W_\lambda(\mathrm{Ly}\alpha)$ is presented as a function of n using a dashed, blue curve in the bottom panel of Figure 4.5. Here, $n = 35$ would be comparable to an KLCS *HST*-like program of 13 LyC detections and 22 LyC non-detections, but in this case the individual LyC detections are not oversampled. For $n = 35$, a 1σ increase in $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ across the two bins of increasing $W_\lambda(\mathrm{Ly}\alpha)$ was only seen in 53.7% of the iterations. The difference between this set of simulations and the preceding one indicates that preferentially observing LyC detections increases the ability to measure trends with $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ and galaxy property at fixed sample size. To recover a 1σ difference in 95% of trials, $n = 90$ objects were required. Similar to the upper panel, we also include curves that test correlation of $\langle f_{900}/f_{1500}\rangle_{\mathrm{obs}}$ with $W_\lambda(\mathrm{Ly}\alpha)$ and a stricter 2σ calculation of trial success.

In summary, despite the tendency for KLCS *HST*-like samples to recover expected trends between $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ and $W_\lambda(\mathrm{Ly}\alpha)$, we do not find a significant correlation between $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ and $W_\lambda(\mathrm{Ly}\alpha)$ in KLCS *HST*. This result is somewhat unexpected, considering this lack of correlation is only recovered in $\sim$ 13% of trials at similar sample size. The fact that we cannot recover the connection between $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$ and $W_\lambda(\mathrm{Ly}\alpha)$ in the KLCS *HST* sample calls into question the robustness of the lack of correlation found between f_{esc} and Σ_{SFR}

in Section 4.4. Our simulations show that we require a larger sample size to ensure that a significant correlation is recovered between $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and W_λ(Lyα), which we have assumed is a requirement to discern a significant trend between $f_{\rm esc}$ and $\Sigma_{\rm SFR}$. Accordingly, a larger sample size is required to make robust statements about the connection between $f_{\rm esc}$ and $\Sigma_{\rm SFR}$. The simulations presented in this section can be used to dictate future *HST* LyC observing programs. However, these results assume a parent sample with spectra of similar quality to KLCS, and having the same intrinsic correlation. With better quality spectra or a sample that covers a larger dynamic range of galaxy property (e.g., $\Sigma_{\rm SFR}$, $L_{\rm UV}$), the necessary sample size may decrease. Nonetheless, these results provide insight into how measurement uncertainty, IGM sightline variability, and sample size can affect eventual measurements of the connection between $f_{\rm esc}$ and galaxy properties.

4.6 Summary

The timeline of reionization remains uncertain, as the main sources that drive the process are still debated. Estimating $f_{\rm esc}$ in samples of lower-redshift analogues and studying how their properties connect with $f_{\rm esc}$ is a fruitful endeavor to distinguish among various models of reionization. In order to explore the potential connection between $f_{\rm esc}$ and $\Sigma_{\rm SFR}$, which is well motivated by the physics of star-formation feedback (e.g., Sharma et al., 2016; Ma et al., 2016; Kakiichi & Gronke, 2021), we examined a sample of 35 galaxies drawn from the KLCS survey which have been shown to be free of line-of-sight contamination based on *HST* imaging. We measured $r_{\rm e}$ using Sérsic profile fits to high-resolution, V_{606} images, and SFR from SED fits to ground- and space-based photometry. We split the sample into two bins of $\Sigma_{\rm SFR}$, and constructed composite rest-UV spectra to examine whether their ionizing and non-ionizing rest-UV spectral features depend on the sample median $\Sigma_{\rm SFR}$. The main results of this work are as follows:

1. We measure no significant difference in $f_{\rm esc}$ for spectral composites constructed from

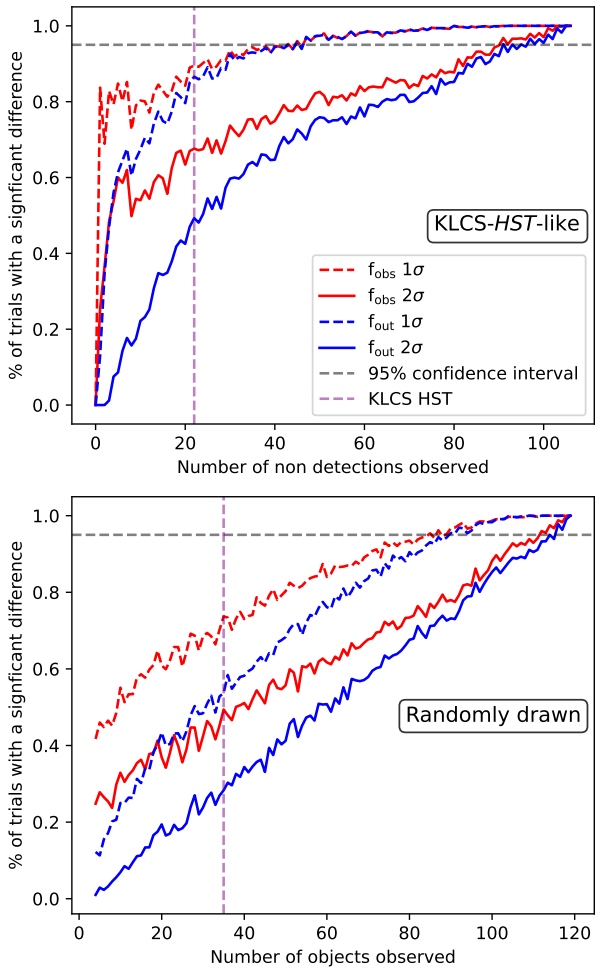

Figure 4.5: Simulations to determine if a subsample of a given size is sufficient to recover an increase of $\langle f_{900}/f_{1500}\rangle$ across two equal bins of $W_\lambda(\mathrm{Ly}\alpha)$. A successful trial was defined as a case in which the $\langle f_{900}/f_{1500}\rangle$ measurement of the upper $W_\lambda(\mathrm{Ly}\alpha)$ composite was greater than the corresponding measurement from the lower $W_\lambda(\mathrm{Ly}\alpha)$ composite at a 1σ confidence level for the dashed lines, and 2σ for the solid lines. The blue curves represent the likelihood of significant correlations recovered for $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$, which includes corrections for attenuation from the IGM+CGM attenuation, while the red curves use simply $\langle f_{900}/f_{1500}\rangle_{\mathrm{obs}}$. For each sample size, 1000 random draws were performed to determine the percent chance of success. **Top:** KLCS-*HST*-like sample construction, which includes 13 individually-detected LyC objects for every trial and a number of n non-detections, for a total sample size of $13+n$. The number of non-detections included in KLCS *HST* is 22, represented by a purple, vertical, dashed line. **Bottom:** Sample construction includes n randomly drawn galaxies from KLCS. Here, KLCS *HST* is represented by a purple, vertical, dashed line at 35, ignoring the oversampling of individually-detected objects in KLCS *HST*.

the KLCS *HST* sample split into two bins of $\Sigma_{\rm SFR}$. This null result is confirmed for a range of assumptions regarding IGM transparency and dust attenuation. Redward of LyC, we measure an elevated W_λ(Lyα) and W_λ(HIS) in the composite characterized by the larger $\Sigma_{\rm SFR}$, although the differences were not significant in the face of sample variance.

2. We measure no significant difference in $f_{\rm esc}$ across two composites binned as a function of W_λ(Lyα), despite $f_{\rm esc}$-W_λ(Lyα) relationships found in the full KLCS, and other LyC surveys (e.g., Fletcher et al., 2019; Gazagnes et al., 2020). Repeating the experiment using two bins of $L_{\rm UV}$, we also find no significant difference in $f_{\rm esc}$ between composite spectra. These results indicate that the KLCS *HST* sample is unable to robustly constrain correlations between $f_{\rm esc}$ and other galaxy properties, including $\Sigma_{\rm SFR}$.

3. We find that a sample of >58 galaxies (13 galaxies individually detected in LyC, 45 individually undetected) would be required to recover a significant positive change in $\langle f_{900}/f_{1500}\rangle_{\rm out}$ among in two subsamples binned by W_λ(Lyα) > 95% of the time. This recommendation is based on stacking simulations using the full KLCS, randomly drawing 45 individual LyC non-detections while always including the 13 LyC detections, similar to the construction of the KLCS *HST* sample. With a KLCS-*HST*-like sample of 13 LyC detections and 22 randomly-drawn non-detections, a significant correlation is recovered in only 86.6% of realizations. Here, the actual KLCS *HST* sample corresponds to one of the $\sim$ 13% of trials in which a correlation would not be recovered.

4. We repeat the stacking simulations without the assumption that the 13 individual LyC detections are preferentially observed, and find that a sample of >90 objects is required to recover a 1σ increase in $\langle f_{900}/f_{1500}\rangle_{\rm out}$ across two bins of increasing W_λ(Lyα) in >95% of trials. These simulations are performed by drawing 90 objects randomly from KLCS before binning and stacking, and did not include all detections in every mock sample. These simulations can be used to inform future LyC surveys aimed at

studying dependencies of $f_{\rm esc}$ on other galaxy properties, although the recommended numbers depend on spectral SNR and dynamic range of the sample.

We have demonstrated that the KLCS HST sample is unfortunately not appropriate for recovering possible trends between $f_{\rm esc}$ and $\Sigma_{\rm SFR}$, due to a combination of small sample size and non-representativeness in terms of the distribution of quantities such as $L_{\rm UV}$ at fixed $W_\lambda({\rm Ly}\alpha)$. We have also outlined what must be done to further explore the $f_{\rm esc}$-$\Sigma_{\rm SFR}$ connection: more objects are needed with HST-resolution imaging appropriate for accurate size estimates and contamination rejection, combined with deep, rest-UV spectra to estimate $f_{\rm esc}$ in stacked composites with sufficient objects in each bin. We will better constrain these important trends using KLCS after more individual LyC non-detections are followed up by HST.

CHAPTER 5

The connection between the escape of ionizing radiation and galaxy properties at $z \sim 3$ in the Keck Lyman Continuum Spectroscopic Survey

5.1 Introduction

Reionization is a significant phase transition in the Universe, during which hydrogen in the intergalactic medium (IGM) transitions from neutral to ionized. This transformation appears to end at $z \sim 6$ (Fan et al., 2006), but we still lack a comprehensive description of the physical processes responsible for the reionization history. At $z > 6$, the ionizing background of the Universe is thought to be dominated by ionizing photons produced by O/B stars in star-forming galaxies (Bouwens et al., 2015; Parsa et al., 2018). Different models of reionization attempt to use integrated measurements of galaxy populations and independent constraints on the IGM hydrogen neutral fraction to infer the comprehensive evolution of reionization from the early Universe to $z \sim 6$, but can arrive at remarkably different answers. Reionization may start "late," such that neutral fractions remain at $\gtrsim 90\%$ until $z \sim 8$ (e.g., Naidu et al., 2020), or reionization may be "gradual," such that the neutral fraction decreases slowly from $z \sim 12$ until $z \sim 6$ (e.g., Finkelstein et al., 2019).

To draw conclusions about the evolution of the neutral fraction from observations of galaxies, one must attempt to understand the ionizing emissivity of galaxies as a function of cosmic

time. This quantity is commonly parameterized as a function of three variables: the UV luminosity function ($\rho_{\rm UV}$), the ionizing photon production efficiency ($\xi_{\rm ion}$), and the fraction of ionizing luminosity that escapes the interstellar and circumgalactic medium (ISM and CGM) and proceed to ionize the IGM ($f_{\rm esc}$) (Robertson et al., 2015). While constraints are available for both $\rho_{\rm UV}$ and $\xi_{\rm ion}$ well into the epoch of reionization (Madau & Dickinson, 2014; Stark et al., 2015, 2017), $f_{\rm esc}$ is uniquely difficult to ascertain in the early Universe. Estimating $f_{\rm esc}$ requires direct observations of the ionizing radiation from galaxies in the Lyman continuum (LyC) spectral region. The transmission through the general IGM of LyC photons escaping a galaxy depends sensitively on emission redshift and decreases rapidly beyond $z \sim 3.5$ (Vanzella et al., 2012). This drop off is due to LyC absorption from trace amounts of H I and makes direct determinations of $f_{\rm esc}$ impossible during the epoch of reionization itself.

Models of reionization are distinguished by their assumptions about $f_{\rm esc}$. Finkelstein et al. (2019) present a "democratic" model for reionization that assumes the process is driven by faint sources with high $f_{\rm esc}$ values. In contrast, the "oligarchical" model of Naidu et al. (2020) concludes that massive, luminous ($M_{\rm UV} < -18$ and $\log(\mathrm{M}_*/M_\odot) > 8$) galaxies provide the bulk of ionizing photons during reionization. For testing assumptions of $f_{\rm esc}$ during reionization, $z \sim 3-4$ galaxies provide an essential laboratory, and can discern the fundamental properties that govern $f_{\rm esc}$ at the highest redshifts these measurements can be made. Critically, these galaxies may be closer analogs to reionization era galaxies than those in the local universe (e.g., Flury et al., 2022a,b).

A number of LyC observational surveys at $z \sim 3-4$ have attempted to measure average $f_{\rm esc}$ values and potential correlations between $f_{\rm esc}$ and the properties of galaxies. Success has been found by stacking deep observations of the LyC either photometrically (e.g., Begley et al., 2022) or spectroscopically (Marchi et al., 2017). Here we focus on the Keck Lyman Spectroscopic (KLCS) survey, which included deep Keck/Low Resolution Imaging Spectrom-

eter (LRIS; Oke et al., 1995; Steidel et al., 2004) spectra of Lyman break galaxies (LBGs) at $z \sim 3$ (Steidel et al., 2018). Steidel et al. (2018) reported an average $f_{\rm esc} = 0.09 \pm 0.01$ from 124 LBGs, estimated by stacking their rest-UV spectra with coverage of the LyC region. After careful treatment of line-of-sight contamination using *HST* imaging, the average KLCS $f_{\rm esc}$ was corrected to 0.06 ± 0.01 upon removal of four apparently-contaminated galaxies from the sample (Pahl et al., 2021). This significant reduction in sample-averaged $f_{\rm esc}$ highlights the importance of foreground decontamination in studies of LyC at high redshift. Galaxy properties correlated with inferred $f_{\rm esc}$ in the KLCS are determined to be Lyα equivalent width (W_λ(Lyα)) and UV luminosity ($L_{\rm UV}$), such that galaxies with stronger Lyα emission and lower $L_{\rm UV}$ luminosities tend to have higher $f_{\rm esc}$ (Steidel et al., 2018; Pahl et al., 2021). In the 2018 paper, we explained why W_λ(Lyα) is more fundamental in its correlation with $f_{\rm esc}$, as W_λ(Lyα) is modulated by the neutral gas covering fraction of a galaxy (Reddy et al., 2016b; Gazagnes et al., 2020), which similarly modulates the escape of ionizing radiation. Trends between $f_{\rm esc}$, W_λ(Lyα), and $L_{\rm UV}$ have also been recovered in complementary LyC surveys at $z \sim 3$ (Marchi et al., 2017, 2018; Begley et al., 2022), but $L_{\rm UV}$ and W_λ(Lyα) ultimately represent a limited parameter space from which to construct a comprehensive picture of LyC escape in star-forming galaxies.

Promising indirect indicators of $f_{\rm esc}$ may surface from the feedback of star formation and its effect on the ISM and CGM of a galaxy. Cosmological zoom-in simulations coupled with radiative transfer calculations indicate that feedback from recent, dense star-formation can induce favorable channels in the ISM and CGM that allow ionizing photons to escape (Ma et al., 2020). Understanding $f_{\rm esc}$ as a function of the surface density of star-formation ($\Sigma_{\rm SFR}$), stellar age, or specific star-formation rate would allow observational comparison to these simulations, and empirically connect the history of star formation to $f_{\rm esc}$. Additionally, dust attenuation is intricately linked to the neutral gas covering fraction in the ISM and CGM, but the relationship between $f_{\rm esc}$ and E(B-V) at $z \sim 3$ has thus far only been investigated using rest-UV observations (Reddy et al., 2016a,b; Steidel et al., 2018). Thanks to multi-

band photometry available for the KLCS, which can constrain stellar population parameters, we can explore these relationships at high redshift, many for the first time.

In Pahl et al. (2022), we began by examining 35 galaxies from the KLCS that were covered by *HST* imaging, enabling measurements of rest-UV sizes. Together with SFR estimates from fits to multi-band photometry, we measured $\Sigma_{\rm SFR}$ and attempted to constrain $f_{\rm esc}$ vs. $\Sigma_{\rm SFR}$. We ultimately determined that the limited KLCS subsample with *HST* imaging was too small and unrepresentative to determine trends with $f_{\rm esc}$ and galaxy property. In the present work, we extend the analysis of Pahl et al. (2022) by examining SED-modeled measurements of stellar mass (M_*), E(B-V), stellar age, and SFR instead, which allow nearly the entire KLCS sample to be utilized. By performing stacking of rest-UV spectra as a function of galaxy property, we investigate the dependence of $f_{\rm esc}$ on these galaxy properties. Significant correlations will test existing reionization models and strongly inform future ones.

We organize the paper as follows: in Section 5.2, we review the spectroscopic observations of the KLCS sample and its ancillary photometric measurements, and provide an overview of the SED and spectral fitting methodology. In Section 5.3, we present the SED-modeled parameters for individual galaxies and estimates of ionizing escape from binned subsamples. In Section 5.4, we explore similar observational analyses from the literature, connections to cosmological zoom-in simulations, and implications for reionization. We summarize our main conclusions in Section 5.5.

Throughout this paper, we adopt a standard ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70$ km s^{-1}Mpc^{-1}. The $f_{\rm esc}$ values reported in this paper are absolute escape fractions, equivalent to $f_{\rm esc,abs}$ in Steidel et al. (2018), and defined as the fraction of all H-ionizing photons produced within a galaxy that escapes into the IGM. We also employ the AB magnitude system (Oke & Gunn, 1983).

5.2 Sample and Methodology

In order to understand how ionizing photon escape is tied to measurable characteristics of galaxy stellar populations or the spatial distribution of the interstellar gas, we require integrated photometric measurements that sample a wide wavelength baseline as well as direct constraints on the LyC emission. Both types of measurements are available for KLCS galaxies. In this section, we outline the data included in our analysis, featuring an overview of the KLCS sample, associated spectra, and the multi-band photometry available for KLCS galaxies. We explain the methodology of SED fitting to determine galaxy properties and spectral modeling to estimate parameterizations of ionizing-photon escape such as $f_{\rm esc}$.

5.2.1 Uncontaminated KLCS

The primary goal of the KLCS was to examine the hydrogen-ionizing spectra of star-forming galaxies at $z \sim 3$ (Steidel et al., 2018). To this end, 137 galaxies were observed with LRIS on the Keck I telescope on Mauna Kea, Hawai'i. Each object was observed for 8.2 hours at minimum. These observations began in 2006 and were concluded in 2008. Of 137 targets, 13 galaxies were removed due to either instrumental defects or spectroscopic evidence of contamination by foreground galaxies. The final sample presented in Steidel et al. (2018) numbered 124 galaxies. Of these, 15 galaxies apparently had significant flux density in the LyC spectral region, defined as having $f_{900} > 3\sigma_{900}$, where f_{900} is the average flux density between $880 - 910$Å in the rest frame, and σ_{900} is the standard deviation of flux densities in the same spectral region. Objects meeting this criteria were defined as individual LyC "detections," with the remaining 109 galaxies labeled as LyC "non-detections."

Despite the efforts of Steidel et al. (2018) to produce a clean sample of LyC leakers at $z \sim 3$ by looking for evidence of spectral blending, foreground contamination remains a significant concern for surveys of LyC at high redshift (Vanzella et al., 2012; Mostardi et al., 2015). A galaxy along the line-of-sight to a $z \sim 3$ source can provide non-ionizing photons that

masquerade as rest-frame LyC assuming a single redshift of $z \sim 3$. As such, Pahl et al. (2021) presented new *HST* measurements of the 15 individual LyC detections in the KLCS sample, which were the objects most likely to be significantly contaminated by foreground light. These data were taken across five survey fields, including seven ACS/F606W (V_{606}) pointings and 11 WFC3/F125W (J_{125}) and WFC3/F160W (H_{160}) pointings. Each pointing was observed for three orbits in each filter. Pahl et al. (2021) also utilized existing *HST* data for one object (Q1549-C25) from Mostardi et al. (2015) and Shapley et al. (2016). Based on the $V_{606}J_{125}H_{160}$ colors of the subcomponents in the high-resolution, *HST* light profiles, two individual LyC detections were determined to be likely contaminated. An additional 24 LyC non-detections were included in the aforementioned *HST* pointings and were also analyzed. Two of these were found to have associated subcomponents with colors consistent with foreground sources. In total, four galaxies were removed from the KLCS, for a final sample size of 120, including 13 galaxies individually detected in LyC.

5.2.2 Photometry and SED fits

Several galaxy properties can be estimated from broadband photometric measurements. In addition to the U_nGR images used for original photometric selection of $z \sim 3$ candidates (see Steidel et al., 2003), longer wavelength photometry of the KLCS has been obtained. We summarize the photometric information available for the objects in KLCS in Table 5.1. A subset of these measurements are also summarized and analyzed in Pahl et al. (2022). Specifically, optical, near-IR, and mid-IR data were available for the majority of galaxies in KLCS. We required at least one photometric measurement entirely redward of the Balmer break in order to accurately constrain the stellar populations. The filters that fulfilled this requirement for the KLCS were the H, K_s, and Spitzer/IRAC bands. Thirteen objects do not have sufficient IR measurements (i.e., did not have any photometric measurements entirely redward of the Balmer break) and were removed from our sample. In addition, we removed two objects with significant scattered light in their ground-based light profiles from nearby

objects, and one galaxy identified with multiple redshifts in the original KLCS spectrum.

HST photometry was included for the 35 objects observed in the *HST* $V_{606}J_{125}H_{160}$ pointings presented in Pahl et al. (2021). In addition to these 35 objects, 11 objects were covered by at least one *HST* filter, without the full $V_{606}J_{125}H_{160}$ dataset required for contamination analysis. *HST* H_{160} imaging was also available for four objects in the Q0100 field and two objects in Q1009 (Law et al., 2012). We re-measured integrated photometry for all objects in KLCS with $V_{606}J_{125}H_{160}$ *HST* data available largely following the methodology of Pahl et al. (2021) and Pahl et al. (2022). In an effort to improve consistency between all photometric measurements, we adopted *HST* magnitudes from the same MAG_AUTO parameter from SEXTRACTOR (Bertin & Arnouts, 1996) that was employed by the ground-based measurements, rather than the isophotal *HST* magnitudes adopted by Pahl et al. (2021) and Pahl et al. (2022). Finally, we note that *HST* H_{140} measurements are available for three objects in Q0100 and five in Q1009, which were included in our analysis (Chen et al., 2021).

We attempted to correct the photometry from potential biases resulting from strong emission lines that lie in the bandpass of individual filters. Notably, we used existing Keck/MOSFIRE spectra with coverage of [O II]$\lambda\lambda$3726, 3729, Hβ, and[O III]$\lambda\lambda$4959, 5007 rest-optical lines to correct broadband H and K_s flux measurements, depending on the wavelength of the observed line. We identified eight objects with neither Keck/MOSFIRE spectra nor additional photometry redward of the Balmer break aside from H or K_s. We removed these objects from the sample to ensure all galaxies in our analysis had at least one trustworthy photometric measurement redward of the Balmer break, free of potential emission-line bias. In addition, we used Lyα equivalent widths presented in Steidel et al. (2018) to correct broadband G and V_{606} flux measurements if the observed wavelength of the line was contained in the respective bandpass.

The final sample with sufficient multi-band photometry for robust stellar population model-

Table 5.1: Photometric bands used in SED modeling.

Fields	Photometric bands
Q0100	U_n^a, B^a, G^a, R_s^a, J^b, H^b, K_s^c, IRAC1, IRAC2, IRAC3, H_{140}, H_{160}
Q0256	U_n^d, G^d, R_s^d, J^c, K_s^c, IRAC1
B20902	U_n^e, $G^{d,e}$, $R_s^{d,e}$, J^f, K_s^f, IRAC1, IRAC2
Q0933	U_n^d, G^d, R_s^d, I^d, J^c, K_s^c, IRAC1, IRAC2, V_{606}, J_{125}, H_{160}
Q1009	U_n^a, G^a, R_s^a, J^g, K_s^g, IRAC1, H_{140}, H_{160}
Westphal	u*^{h}, g'^{h}, r'^{h}, i'^{h}, z'^{h}, J^i, H^i, $K_s^{i,f}$, IRAC1, IRAC2, V_{606}, J_{125}, H_{160}
Q1422	U_n^e, G^e, R_s^e, K_s^c, IRAC1, IRAC2, V_{606}, J_{125}, H_{160}
Q1549	U_n^a, G^a, R_s^a, J^b, H^b, $K_s^{b,c}$, IRAC1, IRAC2, IRAC3, V_{606}, J_{125}, H_{160}
DSF2237b	U_n^d, G^d, R_s^d, I^d, $J^{c,f}$, $K_s^{c,f}$, IRAC1, IRAC2, V_{606}, J_{125}, H_{160}

[a] Observed with Keck/LRIS. [b] Observed with FourStar at the Magellan Baade 6.5m telescope. [c] Observed with the Multi-Object Spectrometer for Infra-Red Exploration (MOSFIRE) on the Keck I telescope. [d] Observed with the COSMIC prime focus imager on the Palomar 5.08 m telescope (see Steidel et al., 2003). [e] Observed with the Prime Focus Imager on the William Herschel 4.2m telescope (WHT) (see Steidel et al., 2003). [f] Observed with NIRC on the Keck I telescope (Shapley et al., 2001). [g] Observed with the Wide Field Infrared Camera (WIRC) on the Palomar 5.08m telescope. [h] From the Canada-France-Hawaii Telescope (CFHT) Legacy Survey. [i] Observed with CFHT/WIRCam as part of the WIRCam Deep Survey (Bielby et al., 2012).

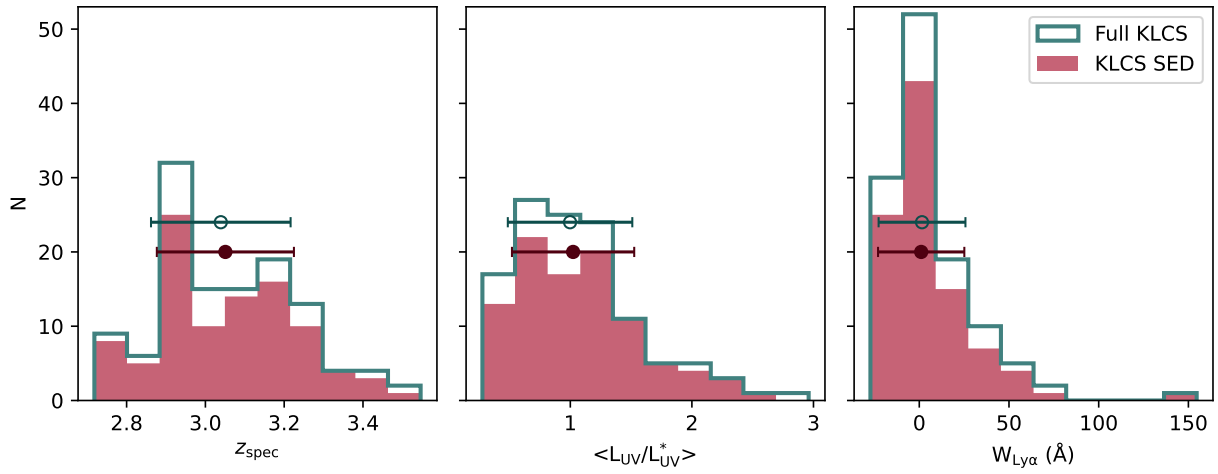

Figure 5.1: Distributions of $z_{\rm spec}$, $L_{\rm UV}$, and W_λ(Lyα) for the KLCS SED sample and the full KLCS. The full KLCS sample includes 120 galaxies from Pahl et al. (2021) while the KLCS SED sample contains the 96 galaxies with photometry appropriate for SED fitting. The median and standard deviation of each distribution are presented respectively as single data points and error bars. The filled circle is the sample median of the KLCS SED sample while the open circle is the sample median of the full KLCS.

ing consisted of 96 galaxies, of which 12 were individual LyC detections, which we define as the "KLCS SED" sample. In Figure 5.1, we display the KLCS SED sample as a function of key observables from Steidel et al. (2018), including spectroscopic redshift $z_{\rm sys}$, Lyα equivalent width W_λ(Lyα), and UV luminosity ($L_{\rm UV}/L^*_{\rm UV}$, where the characteristic luminosity $L^*_{\rm UV}$ corresponds to $M^*_{\rm UV} = -21.0$). We simultaneously present the characteristics of the parent KLCS sample of 120 galaxies. Median $z_{\rm sys}$, $L_{\rm UV}$, and W_λ(Lyα) values for KLCS SED are consistent with those for the full KLCS sample.

In order to estimate stellar-population parameters such as stellar mass (M_*), star-formation rate (SFR), stellar age, and E(B-V) for the galaxies in the KLCS SED sample, we employed SED fits to the multi-band photometry available for these objects. We broadly followed the fitting methodology of Reddy et al. (2022) (also see Pahl et al., 2022). In brief, we utilized BPASS stellar-population synthesis models (BPASS v2.2.1, Eldridge et al., 2017) assuming a Chabrier (2003) initial-mass function. We assumed a constant star-formation history (SFH)

with stellar population ages greater than 50 Myr, such that stellar ages would not be less than typical dynamical-timescales of $z \sim 3$ star-forming galaxies (Reddy et al., 2012). We adopted constant SFHs as they have been shown to reproduce independent measurements of SFR for galaxies $z \geq 1.5$ (Reddy et al., 2012). Constant SFHs may also provide a better description of galaxies at the stellar masses of our sample ($\sim 10^9 - 10^{10.5} M_\odot$), which may have less bursty SFHs than galaxies at lower masses (e.g. Domínguez et al., 2015). We adopted assumptions of metallicity of 0.14 times solar and an SMC dust attenuation curve (Gordon et al., 2003). We examined each SED fit individually for outlier photometric measurements, and dropped Spitzer/IRAC data with clear evidence of blending from nearby sources. Given our SED fitting methodology, we note that galaxies fit with larger masses tended to have higher SFRs and stellar ages. Over the mass range of our sample, we do not find a significant trend with M_* and E(B-V), which may be expected considering the weak relationship that has been recovered between M_* and E(B-V) when assuming the SMC dust extinction curve in the interpretation of galaxy colors vs. M_* SED (McLure et al., 2018).

5.2.3 Binning strategy and spectral modeling

While f_{900} can be measured for each object individually, constraining the LyC leaking in the vicinity of a galaxy requires an understanding of the attenuation of the signal from neutral hydrogen along the line of sight in the IGM and CGM. The transmission of LyC emission varies significantly from sightline to sightline at the redshifts of our sample, introducing large uncertainties on individual LyC measurements (Steidel et al., 2018). To circumvent this sightline to sightline variability, we used binned subsamples and composite spectra that reflect average effects of IGM and CGM attenuation on the LyC spectral region as in Steidel et al. (2018). In order to understand how ionizing-spectral properties vary with the properties produced by SED fits described in the previous section, we binned the KLCS SED sample as a function of M_*, SFR, E(B-V), age, and specific star-formation rate (sSFR; sSFR$\equiv$SFR/M_*). We created three bins for each property, each containing 32 galaxies, to ensure that the

mean IGM+CGM transmission is known with $\lessapprox 10\%$ uncertainty (Steidel et al., 2018) in subsequent composite spectra. We also created an "all" sample, containing all 96 galaxies from KLCS SED, and binned subsamples for $W_\lambda(\mathrm{Ly}\alpha)$ and L_{UV}.

For each subsample, we generated composite spectra representing the average spectral properties of the component galaxies. Following the methodology of Steidel et al. (2018) (also see Pahl et al., 2021, 2022), each individual spectrum is first normalized to the average flux density in the non-ionizing UV spectral region, $1475 - 1525$Å in the rest frame. Using the set of normalized spectra for each binned sample, we then computed the sigma-clipped mean of the distribution of flux densities at each rest-frame wavelength increment, with clipping applied at 3σ. We did not apply sigma clipping to the Lyα spectral region ($1200 - 1230$Å) in order to conserve the inferred composite Lyα profile. The error on the mean flux density at each wavelength was propagated from the values of individual error spectra.

For each composite spectrum, we computed $\langle f_{900}/f_{1500}\rangle_{\mathrm{obs}}$, which is the ratio between the average flux densities in the LyC region ($880-910$Å, f_{900}) and the non-ionizing UV continuum ($1475 - 1525$Å, f_{1500}). While this ratio is useful for discerning the average observed ionizing photon leakage relative to the non-ionizing ultraviolet luminosity density, as discussed above, the spectra must be corrected for lowered transmission from the IGM in the LyC region in order to understand the average effect of LyC leakage has on its environment. We corrected the spectra using average "IGM+CGM" transmission functions from Steidel et al. (2018), calculated at the mean redshift of each composite subsample, and based on the statistics of H I absorption systems along QSO sightlines presented by Rudie et al. (2012) and Rudie et al. (2013). To demonstrate the characteristics of the composite spectra used in our analysis, we display the "all" composite before and after the IGM+CGM transmission correction in the upper panel of Figure 5.2. Using corrected spectra, we repeated the measurement of the ratio of f_{900} to f_{1500}, defined as $\langle f_{900}/f_{1500}\rangle_{\mathrm{out}}$, which applies to the ratio that would be observed at 50 proper kpc from galaxy center (see Steidel et al., 2018).

While $\langle f_{900}/f_{1500}\rangle_{\rm out}$ is a useful empirical measurement of leaking LyC, $f_{\rm esc}$ remains extensively used in reionization modeling. In order to calculate the average $f_{\rm esc}$ for each subsample, we require both an understanding of the intrinsic UV spectrum of the galaxies and the average effects from any intervening gas in the ISM. Thus, $f_{\rm esc}$ is dependent on the assumed stellar population synthesis model, and we follow the well-motivated choices for such models discussed in Steidel et al. (2018). We introduced consistency between our multi-wavelength and spectroscopic modeling by again using the BPASS stellar-population synthesis models of Eldridge et al. (2017). We coupled these models with an SMC extinction curve (Gordon et al., 2003) and a range of E(B-V) from 0.0 to 0.6, and assumptions of metallicity of 0.07 times solar. This metallicity is similar to that assumed for the SED fitting and is consistent with the spectral modeling of Steidel et al. (2018) and Pahl et al. (2021). We model the ISM geometrically using the "holes" approach, which assumes LyC light escapes through a patchy neutral-phase gas (Zackrisson et al., 2013; Reddy et al., 2016b, 2022). The free parameters of the fit included the neutral gas covering fraction $f_{\rm c}$, the column density of neutral hydrogen $N_{\rm HI}$, and the dust attenuation from the foreground gas $E(B\text{-}V)_{\rm cov}$ (i.e., the uncorrected portion is assumed to be dust free). In general, $f_{\rm esc}$ is defined from $f_{\rm c}$, where $f_{\rm esc} = 1 - f_c$. To demonstrate the fitting process, we display a fit to the corrected full-sample composite in the lower panel of Figure 5.2. Here, the modeled spectrum in green is split into an unattenuated (blue) and attenuated (pink) portion, representing the light that either escaped through clear sightlines in the ISM or was partially reduced by intervening material, respectively.

In order to estimate the uncertainty in average escape parameters for a given set of galaxies, we must understand the level of variability induced from sample construction, while simultaneously including the uncertainty on galaxy property measurements. We perturbed each individual measurement randomly by a Gaussian characterized by a width equal to its error, and constructed 100 modified parameter histograms for each galaxy property. We then binned each perturbed sample distribution into three subsamples of increasing galaxy

property, with 32 galaxies in each subsample. Finally, with the goal of understanding how sample variance affects generated composite spectra, we used bootstrap resampling of the galaxies of each binned subsample. For each of the 100 sets of 32 galaxies (generated from the modified parameter distributions), we drew one new subsample with replacement. We subsequently created composite spectra and measured ionizing-photon escape for each random draw, using the process described earlier in this section. The mean and standard deviation of the $\langle f_{900}/f_{1500}\rangle_{\rm obs}$, $\langle f_{900}/f_{1500}\rangle_{\rm out}$, and $f_{\rm esc}$ distributions generated from the 100 composite spectra were used as the fiducial value and error estimate for the corresponding binned sample. The errors determined from this bootstrap resampling were larger than those associated with measurement uncertainty, average IGM+CGM transmission variability, and errors from individual measurements.

5.3 Results

Based on SED fits, we estimated M_*, SFR, stellar age, E(B-V) and sSFR for each galaxy. We present the distribution of these SED-modeled parameters for the KLCS SED sample in Figure 5.3. The median and standard deviation of each respective measurement distribution are displayed, respectively, as dark points and error bars. In the figure, we use dashed, vertical lines to indicate the edges of the three equal-sized (n=32) samples that comprise the bins for generating composite spectra.

A composite spectrum was generated for each binned sample detailed in Figure 5.3, and three estimates of ionizing-photon escape were measured, as described in Section 5.2.3. The first, $\langle f_{900}/f_{1500}\rangle_{\rm obs}$, is the ratio of ionizing to non-ionizing flux density directly observed in the composite generated from individual spectra. The second, $\langle f_{900}/f_{1500}\rangle_{\rm out}$, is the same ratio instead measured from a composite corrected for mean line-of-sight attenuation from the IGM and CGM. Finally, $f_{\rm esc}$ is a parameter estimated via stellar-population synthesis and ISM modeling of the full rest-UV composite. We display the three measurements of

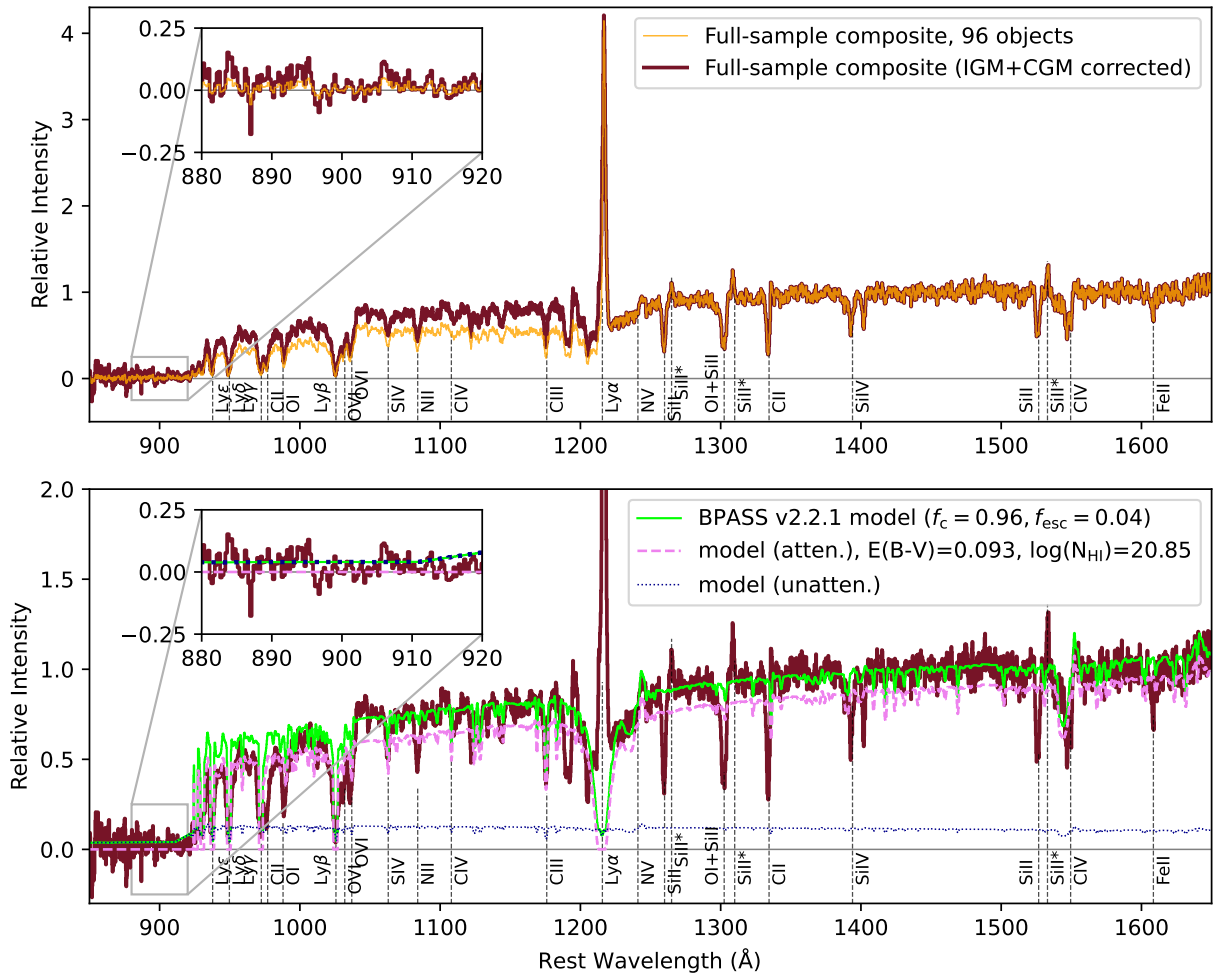

Figure 5.2: Composite spectrum for the "all" (full KLCS SED, 96 galaxy) sample. **Top:** The "all" composite alongside the same spectrum corrected from the average attenuation from the IGM and CGM at the mean redshift of the composite, $z_{\rm mean} = 3.05$. The uncorrected spectrum is shown with a thin, orange curve, while the corrected composite is shown with a thick, maroon curve. An inset is included to highlight the LyC spectral region. **Bottom:** IGM- and CGM-corrected composite spectrum alongside the best-fit spectrum from the modeling process. The corrected composite is again shown with the thick, maroon curve. The best-fit BPASS model is presented as a thin, green curve. This model is summed from two component spectra, an attenuated portion displayed as a dashed, pink line, and an unattenuated portion displayed as a dotted, blue line, as per the "holes" model of Steidel et al. (2018). An inset is included to highlight the LyC spectral region. The free parameters values of the fit that produced the model curves are $f_{\rm c} = 0.96$, E(B-V)= 0.093, and $\log(N_{\rm HI}/{\rm cm}^{-2})$= 20.85.

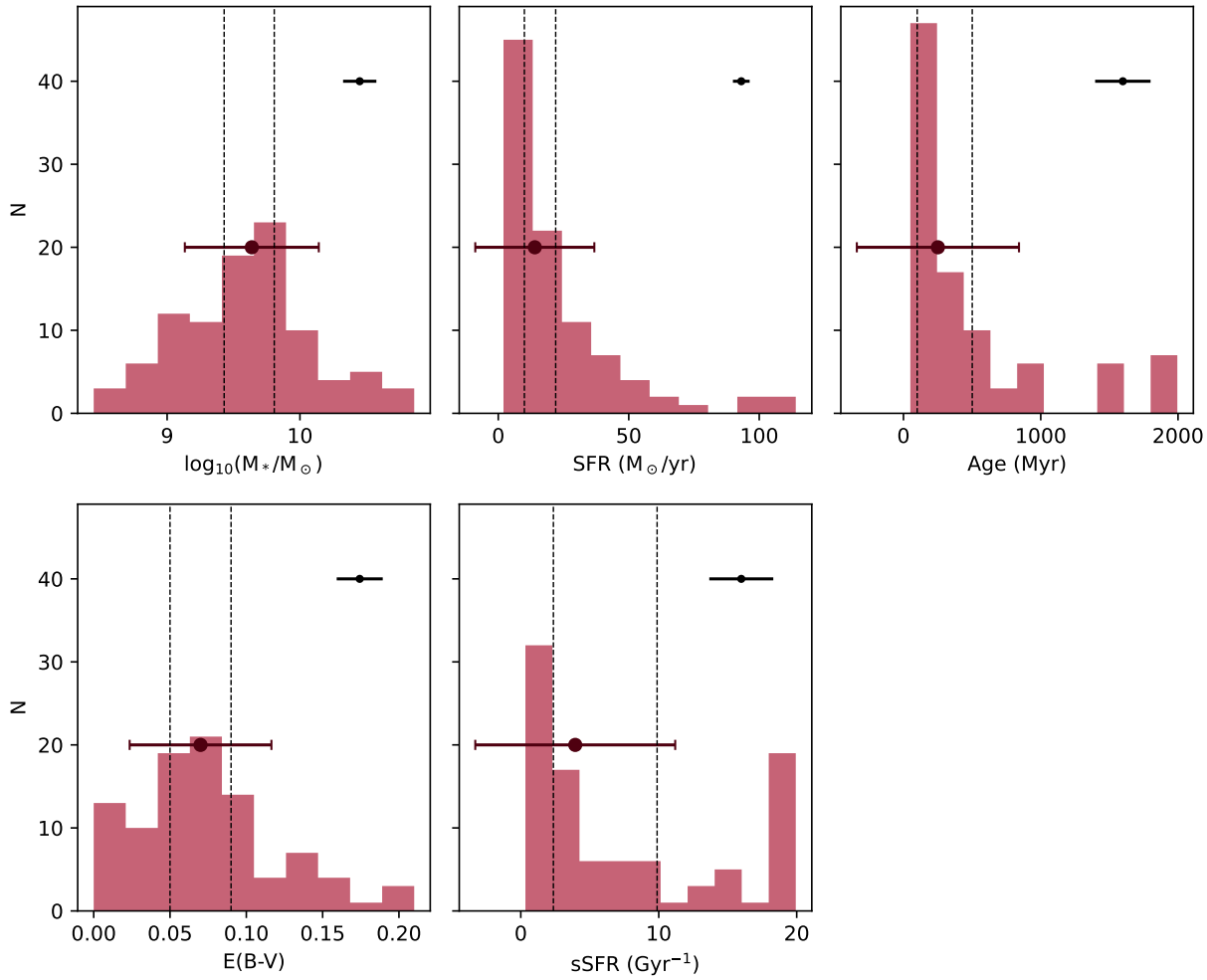

Figure 5.3: Galaxy property distributions of the KLCS SED sample. The five properties displayed here were inferred from SED fits performed for each galaxy. The median and standard deviations with respect to a given measurement are presented as large data points with capped error bars, while the typical (median) error on individual measurements are presented as smaller data points in the upper left with uncapped error bars. The edges of bins used for generation of composite spectra are shown as vertical, dashed lines. The full sample was sorted according to each galaxy property and divided into three equal-sized bins (n=32), which were then used to generate composite spectra.

ionizing escape and their respective errors for each subsample binned as a function of galaxy property in Figure 5.4. These results are tabulated in Table 5.2, which includes the median galaxy properties of each subsample.

To determine whether ionizing-photon escape is correlated with measured galaxy properties, we define a "significant" correlation as fulfilling two criteria: the escape parameter varies monotonically across the three bins, and the difference between the escape parameter in the highest and lowest bins was $> 1\sigma$. Using $f_{\rm esc}$ as an example, we define a $> 1\sigma$ difference as

$$|f_{\rm esc,highest} - f_{\rm esc,lowest}| > \sqrt{(\sigma_{f_{\rm esc,highest}})^2 + (\sigma_{f_{\rm esc,lowest}})^2}, \tag{5.1}$$

where $f_{\rm esc,highest}$ is measured from the third tertile of a given galaxy property, $f_{\rm esc,lowest}$ is from the first tertile, and $\sigma_{f_{\rm esc,highest}}$ and $\sigma_{f_{\rm esc,lowest}}$ are their corresponding errors derived as described in Section 5.2.3.

As seen in the upper left panel of Figure 5.4, both $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and $f_{\rm esc}$ are significantly, negatively correlated with M_* in the KLCS SED sample, such that lower-mass galaxies have higher escape fractions. The two measures of LyC escape are also significantly, negatively correlated with SFR, shown in the upper central panel. While we find no significant correlation with $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and E(B-V), we do find a significant, negative correlation with $f_{\rm esc}$ and E(B-V), as displayed in the middle left panel. Discrepancies between these two parameters of LyC escape arise from the fact that the fitting process to determine $f_{\rm esc}$ incorporates additional information from the composite spectrum, including the Lyman series absorption features and the UV spectral shape. Correlation between ionizing-escape parameters and E(B-V) is expected in our analysis considering that neutral gas and dust are spatially associated (Reddy et al., 2016b; Du et al., 2018; Pahl et al., 2020). We also see an anti-correlation between $f_{\rm esc}$ and E(B-V) when inferring E(B-V) from rest-UV spectral modeling. Finally, we find no significant correlation between $\langle f_{900}/f_{1500}\rangle_{\rm out}$ or $f_{\rm esc}$ and stellar age or sSFR, seen in the middle left and center panels of Figure 5.4, respectively. We additionally computed these trends using a modified SED fitting process that assumed a Calzetti et al.

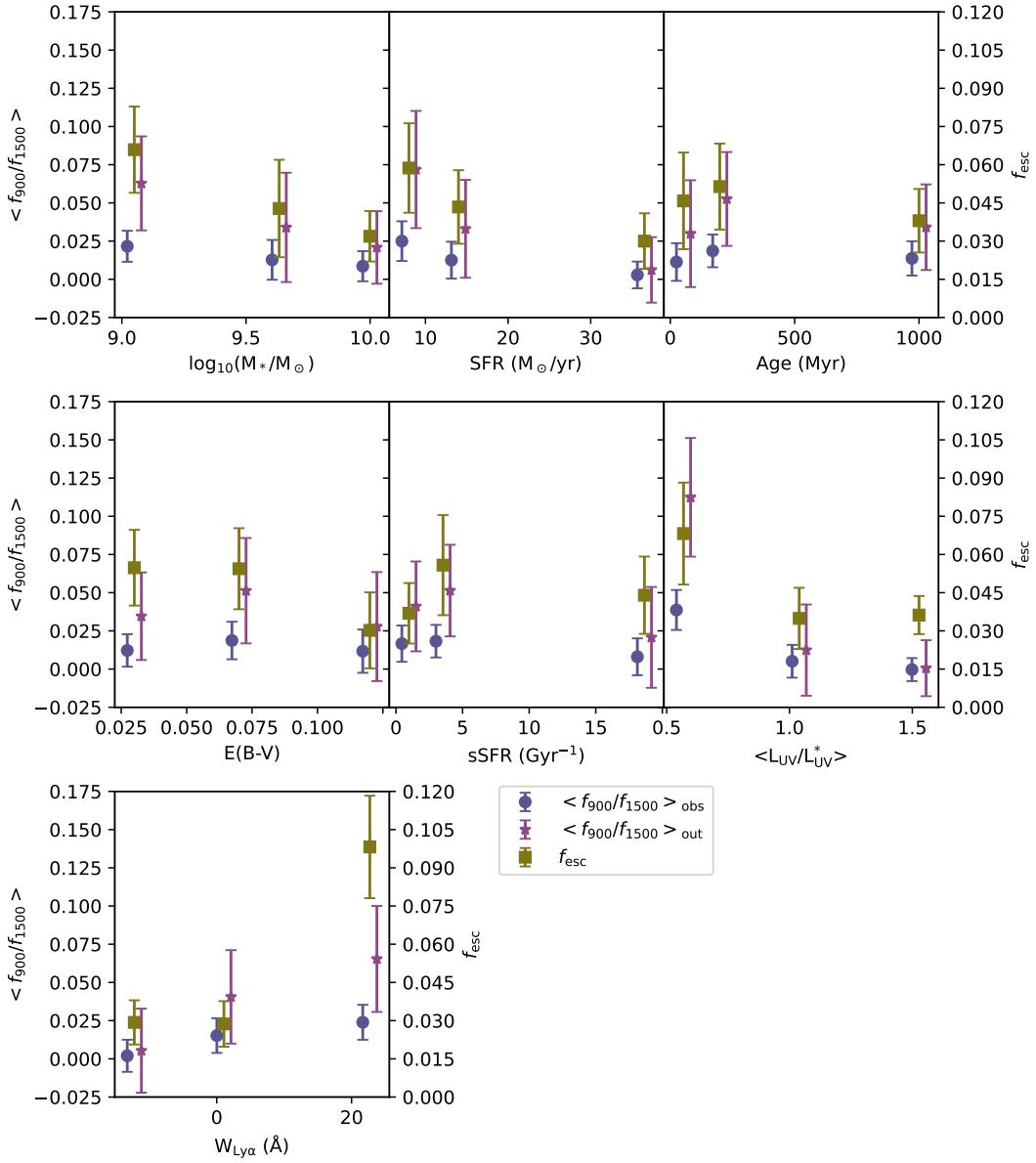

Figure 5.4: Different measures of ionizing escape for the KLCS SED sample binned according to several galaxy properties. Each escape parameter is presented as a function of the median of the respective galaxy property of the binned subsample. Measurements of $\langle f_{900}/f_{1500}\rangle_{\rm obs}$ are presented as blue circles and are based on uncorrected, composite spectra. Measurements of $\langle f_{900}/f_{1500}\rangle_{\rm out}$ are presented as purple stars and are based on composite spectra corrected for attenuation of the IGM+CGM; estimates of $f_{\rm esc}$ are presented as yellow squares and are based on SPS model fits and modeling of the ISM according to the "holes" model (Steidel et al., 2018; Reddy et al., 2016b, 2022). Blue circles and purple stars are shifted left and right, respectively, for visual clarity.

(2000) dust extinction curve, and found qualitative agreement in the trends between $f_{\rm esc}$ and galaxy properties, although we note that a stronger $f_{\rm esc}$—E(B-V) trend was recovered when using the SMC curve. We consider our SED modeling process that utilizes the SMC curve fiducial, considering greater consistency has been found between SED-fit SFR measurements and Hα-based SFRs when assuming an SMC dust extinction curve rather than Calzetti et al. (2000) at $z \sim 2$ (Reddy et al., 2022). Additionally, Reddy et al. (2018a) demonstrated that $z \sim 2$ star-forming galaxies have an IRX—β relation consistent with predictions only when assuming an SMC dust curve.

We also note that $f_{\rm esc}$ and $\langle f_{900}/f_{1500}\rangle_{\rm out}$ are significantly correlated with W_λ(Lyα) and $L_{\rm UV}$ in this analysis, seen in the middle right and bottom panels of Figure 5.4, mirroring the results of the full KLCS presented in Steidel et al. (2018) and Pahl et al. (2021). The positive trend between $f_{\rm esc}$ and W_λ(Lyα) has also been confirmed in additional $z \sim 3-4$ LyC surveys (Marchi et al., 2017, 2018; Fletcher et al., 2019; Begley et al., 2022). In Pahl et al. (2022), we argued that recovering these well-established spectral trends is important for determining whether a sample is sufficiently large and representative for examining relationships between $f_{\rm esc}$ and other galaxy properties. Considering the KLCS SED sample has both the size (n=96) and dynamic range of galaxy properties to confidently recover trends between $f_{\rm esc}$ and W_λ(Lyα)/$L_{\rm UV}$, we conclude that the KLCS SED sample is sufficient and representative, fulfilling the requirements for determining the trends between $f_{\rm esc}$ and galaxy property presented in this section. We note that if subtle correlations do exist between $f_{\rm esc}$ and age or sSFR, we may require a larger sample to discern these trends.

5.4 Discussion

The connections between $f_{\rm esc}$ and galaxy properties at $z \sim 3$ provide key insights into the physics of ionizing-photon escape, and also indicate the most appropriate assumptions for $f_{\rm esc}$ at even higher redshift, during the epoch of reionization. We find that galaxies with

higher $f_{\rm esc}$ tend to have lower E(B-V), which is consistent with a physical picture in which dust is spatially coincident with neutral-phase gas in a galaxy, such that an ISM with a higher neutral-gas covering fraction will be both dustier and have lower associated $f_{\rm esc}$ (Reddy et al., 2016b; Du et al., 2018; Pahl et al., 2020). In addition, ionizing photons are more attenuated by dust than non-ionizing photons (Reddy et al., 2016a). We find a negative trend between both $f_{\rm esc}$ and $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and M_*, highlighting the fact that more massive galaxies at $z \sim 3$ have conditions that are less conducive to LyC escape. This relationship is likely due to the fact that more massive galaxies tend to be dustier (e.g., Whitaker et al., 2017; McLure et al., 2018). Finally, the lack of trend between $f_{\rm esc}$ and either stellar age or sSFR is in tension with the physical picture advanced in simulations that bursts of recent star formation induce favorable channels in the ISM and CGM for ionizing photons to escape (Ma et al., 2020). In this section, we introduce comparisons between our results and recent LyC surveys both at $z \sim 3$ and in the local Universe. We also connect our ionizing-photon escape trends or lack thereof with radiative transfer modeling of simulated galaxies and the predictions from reionization models.

5.4.1 Comparison to related observational surveys

Direct comparisons can be made between our reported trends between $f_{\rm esc}$ and galaxy property and those found in recent LyC surveys at $z \sim 3$. Of particular note are the recent photometric LyC measurements of 148 galaxies from the VANDELS survey at $3.35 \leq z \leq 3.95$ (Begley et al., 2022). These authors constrained the average $f_{\rm esc}$ of the sample as $\langle f_{\rm esc}\rangle = 0.07 \pm 0.02$, consistent with $\langle f_{\rm esc}\rangle = 0.06 \pm 0.01$ measured from the uncontaminated KLCS (Pahl et al., 2021). The VANDELS LyC sample was binned in two as a function of a variety of galaxy properties. A positive correlation between $f_{\rm esc}$ and W_λ(Lyα) and a negative correlation between $f_{\rm esc}$ and $L_{\rm UV}$ reported in the VANDELS analysis aligns with the correlations found in the KLCS SED sample and the full KLCS (Steidel et al., 2018; Pahl et al., 2021). Best-fit $f_{\rm esc}$ values were also calculated for two bins of increasing M_* in

the VANDELS sample, which we display alongside our $f_{\rm esc}$ vs. M_* measurements for the KLCS SED sample in Figure 5.5. A weak anti-correlation was observed between $f_{\rm esc}$ and M_* in the VANDELS analysis when utilizing maximum-likelihood estimation (MLE) to determine $f_{\rm esc}$, consistent with trends found using indirect measurements of $f_{\rm esc}$ in VANDELS (Saldana-Lopez et al., 2022a). No correlation was found when using a Bayesian estimate of $f_{\rm esc}$. The trend between the VANDELS MLE $f_{\rm esc}$ values and M_* is remarkably consistent with our results. In addition, the reported $f_{\rm esc}$ values for both Bayesian and MLE methods from VANDELS are consistent with our $f_{\rm esc}$ constraints at comparable M_*, however we note that the VANDELS results use modeling that more closely resembles the "screen" model of Steidel et al. (2018), rather than the "holes" model used in this work. Using the "screen" model results in $\sim 30\%$ higher $\langle f_{\rm esc} \rangle$ than using the "holes" model in the KLCS (Steidel et al., 2018), which is still consistent with the VANDELS results. The VANDELS analysis also recovered a significant anti-correlation between $f_{\rm esc}$ and UV dust attenuation, where dust attenuation was quantified in terms of the UV slope, β. These results are qualitatively consistent with the anti-correlation between $f_{\rm esc}$ and E(B-V) we present in the central left panel of Figure 5.4.

Stellar population parameters have also been explicitly correlated with $W_\lambda(\mathrm{Ly}\alpha)$ in galaxy surveys at $z \sim 2-5$. These trends are informative for interpreting LyC escape considering that the strength of Lyα emission is similarly modulated by the neutral gas covering fraction

Table 5.2: Different measurements of ionizing-photon escape for the KLCS SED sample, binned according to several galaxy properties.

	$\log_{10}(M_{*,med}/M_\odot)^a$	SFR_{med} $(M_\odot yr^{-1})^a$	Age_{med} $(Myr)^a$	$E(B\text{-}V)_{med}{}^a$	$sSFR_{med}$ $(Gyr^{-1})^a$	$<L_{UV,med}/L^*_{UV}>^a$	$W_{Ly\alpha,med}$ $(\AA)^a$
$M_{*,low}$	$\mathbf{9.05 \pm 0.06}$	13.0 ± 1.9	63 ± 13	0.075 ± 0.009	15.5 ± 2.5	0.71 ± 0.09	2.7 ± 5.5
$M_{*,mid}$	$\mathbf{9.63 \pm 0.03}$	14.5 ± 4.8	251 ± 72	0.070 ± 0.011	3.9 ± 1.3	1.06 ± 0.11	1.1 ± 2.8
$M_{*,high}$	$\mathbf{10.00 \pm 0.09}$	15.5 ± 4.4	815 ± 268	0.065 ± 0.011	1.2 ± 0.4	1.29 ± 0.13	-2.1 ± 3.3
SFR_{low}	9.60 ± 0.11	$\mathbf{8.0 \pm 0.5}$	475 ± 169	0.050 ± 0.010	2.0 ± 0.8	0.62 ± 0.10	2.5 ± 2.5
SFR_{mid}	9.62 ± 0.16	$\mathbf{14.0 \pm 0.9}$	316 ± 107	0.070 ± 0.009	3.1 ± 1.8	0.93 ± 0.12	-2.5 ± 4.1
SFR_{high}	9.69 ± 0.07	$\mathbf{36.5 \pm 4.1}$	100 ± 37	0.110 ± 0.015	9.9 ± 3.3	1.32 ± 0.10	-1.2 ± 1.9
Age_{low}	9.27 ± 0.12	19.0 ± 6.0	$\mathbf{53 \pm 14}$	0.090 ± 0.009	17.6 ± 2.7	0.96 ± 0.16	0.7 ± 4.5
Age_{mid}	9.59 ± 0.09	15.0 ± 3.7	$\mathbf{199 \pm 54}$	0.070 ± 0.009	4.6 ± 1.4	1.02 ± 0.15	0.3 ± 3.0
Age_{high}	9.94 ± 0.10	11.0 ± 1.9	$\mathbf{1000 \pm 225}$	0.050 ± 0.011	1.0 ± 0.2	1.08 ± 0.11	1.6 ± 2.7
$E(B\text{-}V)_{low}$	9.70 ± 0.10	9.0 ± 0.8	501 ± 134	$\mathbf{0.030 \pm 0.004}$	1.9 ± 0.5	0.97 ± 0.08	3.8 ± 2.5
$E(B\text{-}V)_{mid}$	9.59 ± 0.14	14.0 ± 1.2	251 ± 77	$\mathbf{0.070 \pm 0.005}$	3.9 ± 1.3	0.86 ± 0.13	1.1 ± 3.3
$E(B\text{-}V)_{high}$	9.66 ± 0.07	33.0 ± 5.9	100 ± 24	$\mathbf{0.120 \pm 0.011}$	10.0 ± 2.6	1.18 ± 0.10	-1.8 ± 2.5
$sSFR_{low}$	9.90 ± 0.10	10.0 ± 1.9	1000 ± 241	0.050 ± 0.010	$\mathbf{1.0 \pm 0.2}$	1.08 ± 0.11	1.6 ± 2.7
$sSFR_{mid}$	9.63 ± 0.06	13.5 ± 3.0	284 ± 57	0.065 ± 0.009	$\mathbf{3.5 \pm 0.7}$	0.97 ± 0.14	1.0 ± 2.4
$sSFR_{high}$	9.08 ± 0.14	22.0 ± 6.3	50 ± 4	0.090 ± 0.009	$\mathbf{18.7 \pm 1.0}$	0.98 ± 0.16	0.2 ± 5.8
$L_{UV,low}$	9.22 ± 0.13	9.0 ± 1.1	125 ± 82	0.070 ± 0.008	7.7 ± 2.7	$\mathbf{0.57 \pm 0.03}$	2.9 ± 5.9
$L_{UV,mid}$	9.65 ± 0.06	14.0 ± 1.4	284 ± 90	0.070 ± 0.011	3.5 ± 1.2	$\mathbf{1.04 \pm 0.05}$	2.6 ± 3.5
$L_{UV,high}$	9.85 ± 0.07	28.0 ± 4.9	251 ± 134	0.075 ± 0.012	3.9 ± 2.1	$\mathbf{1.53 \pm 0.04}$	-2.2 ± 1.5
$W_{Ly\alpha,low}$	9.72 ± 0.13	14.0 ± 1.8	251 ± 111	0.080 ± 0.006	3.9 ± 2.3	1.08 ± 0.13	$\mathbf{-12.2 \pm 1.1}$
$W_{Ly\alpha,mid}$	9.72 ± 0.09	17.2 ± 3.0	316 ± 112	0.060 ± 0.010	3.1 ± 1.7	1.32 ± 0.15	$\mathbf{1.1 \pm 0.8}$
$W_{Ly\alpha,high}$	9.57 ± 0.13	12.5 ± 3.0	199 ± 94	0.070 ± 0.012	4.9 ± 4.2	0.84 ± 0.08	$\mathbf{22.7 \pm 3.7}$

Continued on next page.

Table 5.2: Continued.

	$< f_{900}/f_{1500} >_{\rm obs}$	$< f_{900}/f_{1500} >_{\rm out}$	$f_{\rm esc}$
$M_{*,\rm low}$	0.022 ± 0.010	0.063 ± 0.031	0.07 ± 0.02
$M_{*,\rm mid}$	0.013 ± 0.013	0.034 ± 0.036	0.04 ± 0.02
$M_{*,\rm high}$	0.009 ± 0.010	0.021 ± 0.024	0.03 ± 0.01
$\rm SFR_{low}$	0.025 ± 0.013	0.072 ± 0.038	0.06 ± 0.02
$\rm SFR_{mid}$	0.013 ± 0.012	0.033 ± 0.032	0.04 ± 0.01
$\rm SFR_{high}$	0.003 ± 0.009	0.006 ± 0.021	0.03 ± 0.01
$\rm Age_{low}$	0.011 ± 0.012	0.030 ± 0.035	0.05 ± 0.02
$\rm Age_{mid}$	0.019 ± 0.011	0.053 ± 0.031	0.05 ± 0.02
$\rm Age_{high}$	0.014 ± 0.011	0.034 ± 0.028	0.04 ± 0.01
$\rm E(B\text{-}V)_{low}$	0.012 ± 0.011	0.035 ± 0.029	0.05 ± 0.01
$\rm E(B\text{-}V)_{mid}$	0.019 ± 0.012	0.051 ± 0.034	0.05 ± 0.02
$\rm E(B\text{-}V)_{high}$	0.012 ± 0.014	0.028 ± 0.036	0.03 ± 0.01
$\rm sSFR_{low}$	0.017 ± 0.012	0.041 ± 0.029	0.04 ± 0.01
$\rm sSFR_{mid}$	0.018 ± 0.011	0.051 ± 0.030	0.06 ± 0.02
$\rm sSFR_{high}$	0.008 ± 0.012	0.021 ± 0.033	0.04 ± 0.02
$L_{\rm UV,low}$	0.039 ± 0.013	0.112 ± 0.039	0.07 ± 0.02
$L_{\rm UV,mid}$	0.005 ± 0.011	0.012 ± 0.030	0.03 ± 0.01
$L_{\rm UV,high}$	-0.000 ± 0.008	0.001 ± 0.018	0.04 ± 0.01
$W_{\rm Ly\alpha,low}$	0.002 ± 0.011	0.005 ± 0.028	0.03 ± 0.01
$W_{\rm Ly\alpha,mid}$	0.015 ± 0.011	0.041 ± 0.031	0.03 ± 0.01
$W_{\rm Ly\alpha,high}$	0.024 ± 0.011	0.065 ± 0.035	0.10 ± 0.02

[a] The medians and standard error of the subsamples with respect to a given galaxy property, with the binned parameter highlighted in bold.

(e.g., Steidel et al., 2010, 2011, 2018; Verhamme et al., 2015; Reddy et al., 2016b). Stacks of rest-UV spectra at $z \sim 2-5$ have demonstrated anti-correlations between $W_\lambda(\mathrm{Ly}\alpha)$ and both M_* and SFR (Du et al., 2018; Pahl et al., 2020), mimicking the anti-correlations between f_{esc} and these parameters that we presented in Figure 5.4. Meanwhile, surveys at this redshift have shown either no strong correlation between $W_\lambda(\mathrm{Ly}\alpha)$ and age (Du et al., 2018; Pahl et al., 2020) or only a weak correlation (Reddy et al., 2022). These analyses of Lyα escape in combination with our f_{esc} trends indicate that stellar age may not as closely linked to the configuration of neutral-phase gas in the ISM and CGM of a galaxy as much as other galaxy properties, such as M_*, L_{UV}, E(B-V), and SFR.

While the $z \sim 3-4$ universe is an excellent laboratory to test LyC escape physics in galactic environments more similar to those at $z > 6$, LyC surveys in the local universe are afforded advantages such as the ability to examine the direct ionizing signals from intrinsically fainter galaxies in the dwarf galaxy regime, which may dominate the ionizing background during the epoch of reionization (Robertson et al., 2015; Finkelstein et al., 2019). In addition, local surveys avoid the sightline variability of the IGM that necessitates binning at $z \sim 3$ (Steidel et al., 2018), enabling constraints on f_{esc} for individual objects.

The Low-redshift Lyman Continuum Survey (LzLCS) analyzed 66 galaxies at $z = 0.2 - 0.4$ observed with the *HST*/COS, and reported 35 galaxies individually detected in LyC (Flury et al., 2022a,b). The galaxies were indirectly selected to be strongly leaking using [O III]λ5007/[O II]$\lambda\lambda$3726, 3729, SFR surface density, and UV spectral slope, in contrast the LBG-selected KLCS. The correlation between a number of galaxy properties and f_{esc} were considered, where f_{esc} was inferred from stellar-population synthesis fits to COS UV spectra, similar to our determinations of f_{esc} for the KLCS SED sample. The LzLCS f_{esc} values appear to decrease as a function of increasing M_*, consistent with the negative trend we present in the upper left panel of Figure 5.4. However, the correlation coefficient between f_{esc} and M_* was determined not to be significant, mirroring other local explorations of the two

variables (Izotov et al., 2021). Augmenting this result, an examination of the LzLCS sample in tandem with archival observations (totaling 89 star-forming galaxies at $z \sim 0.3$) found that galaxies at lower M_* tend to have both bluer spectral slopes and higher f_{esc} (Chisholm et al., 2022). This analysis focused on the strong inverse correlation found between f_{esc} and the UV spectral slope at 1550Å (β_{1500}) in the expanded sample. While UV spectral slope can encapsulate both the intrinsic spectral slope of the stellar population and the degree of dust reddening in the UV, β_{1500} was strongly correlated with E(B-V) and uncorrelated with stellar age, indicating that β_{1500} for the LzLCS is primarily reflecting the degree of dust reddening (also see Saldana-Lopez et al., 2022b). Apparent anti-correlations between f_{esc} and β_{1500} are supported by our anti-correlation of f_{esc} and E(B-V) seen in the center left panel of Figure 5.4. Nonetheless, Chisholm et al. (2022) make predictions for f_{esc} vs. M_{UV} at $z \sim 3$ that are too low when compared to the $\langle f_{esc} \rangle = 0.06 \pm 0.01$ of the KLCS (at $M_{UV} \sim -21$), despite reproducing our qualitative relationship between f_{esc} and L_{UV} shown in the center right panel of Figure 5.4.

A weaker, but still significant trend of f_{esc} and sSFR was also observed in the LzLCS, contrasting with the lack of trend between f_{esc} and sSFR that we presented in the central panel of Figure 5.4. We note that the LzLCS modeling allows for arbitrarily-short stellar ages, in contrast with our SED fitting procedure, which ensures ages are greater than typical dynamical timescales (50 Myr). Finally, f_{esc} is strongly correlated with $W_\lambda(\mathrm{Ly}\alpha)$ in the LzLCS analysis, which is broadly consistent with the correlation found for both the KLCS SED sample in Figure 5.4 and the full KLCS in Steidel et al. (2018) and Pahl et al. (2021).

We also note that strong correlations are found between f_{esc} and Lyα peak separation (v_{sep}) and star-formation rate surface density (Σ_{SFR}) in the LzLCS. These potential tracers of f_{esc} remain unconfirmed at $z \sim 3$, and, in particular, the number of galaxies with *HST* imaging in the KLCS remains insufficient for testing the connection between f_{esc} and Σ_{SFR} (Pahl et al., 2022). However, future work will examine potential connections between f_{esc} and v_{sep}

at $z \sim 3$ by leveraging higher-resolution spectroscopy of the Lyα profiles of KLCS galaxies and $f_{\rm esc}$ and $\Sigma_{\rm SFR}$ in the VANDELS survey.

5.4.2 Comparison to models

Theoretical predictions for $f_{\rm esc}$ in a variety of galactic environments can help elucidate fundamental relationships between galactic physics and escaping ionizing radiation in the earliest galaxy populations, where direct LyC detections are impossible. The Feedback in Realistic Environments (FIRE-2; Hopkins et al., 2018) project was coupled with radiative transfer in post-processing to examine $f_{\rm esc}$ in cosmological zoom-in simulations of galaxies, evolved down to $z = 5$ (Ma et al., 2020). An increase of $f_{\rm esc}$ with increasing mass was found up to $\log(\mathrm{M}_*/M_\odot) \sim 8$, and a subsequent decrease in $f_{\rm esc}$ was found at $\log(\mathrm{M}_*/M_\odot) > 8$. The increasing relationship between $f_{\rm esc}$ and M_* was determined to be due to an increasing efficiency of star formation and feedback, while the decrease at the high-mass end can be explained by increasing dust attenuation. We display this trend as dark, horizontal bars in Figure 5.5, specifying the simulation resolution that extends to the stellar masses of our sample (baryonic particle mass $m_b \sim 7000 M_\odot$). The negative trend between $f_{\rm esc}$ and M_* at $\log(\mathrm{M}_*/M_\odot) > 8$ found in Ma et al. (2020) is consistent with the negative trend we find in the KLCS SED sample, which has a median $\log(\mathrm{M}_*/M_\odot) = 9.6$. We do find overall lower $f_{\rm esc}$ values than the FIRE-2 results at fixed M_*. This discrepancy may be expected considering higher $f_{\rm esc}$ values were found with increasing redshift at fixed M_* in the simulated galaxies.

Evidence of a turnover in the relationship between $f_{\rm esc}$ and M_* was also found in Kostyuk et al. (2022), which utilized the IllustrisTNG (Marinacci et al., 2018; Naiman et al., 2018; Nelson et al., 2018, 2019; Pillepich et al., 2018; Springel et al., 2018) cosmological simulations coupled with the radiative transfer code CRASH (Graziani et al., 2013). These authors also found significant scatter in the relationship between $f_{\rm esc}$ and M_*, due to both differences in ionizing photon production rates and the distribution of stars within the neutral ISM. The SPHINX simulations have also similarly reported this relationship, finding a $f_{\rm esc}$ trend that

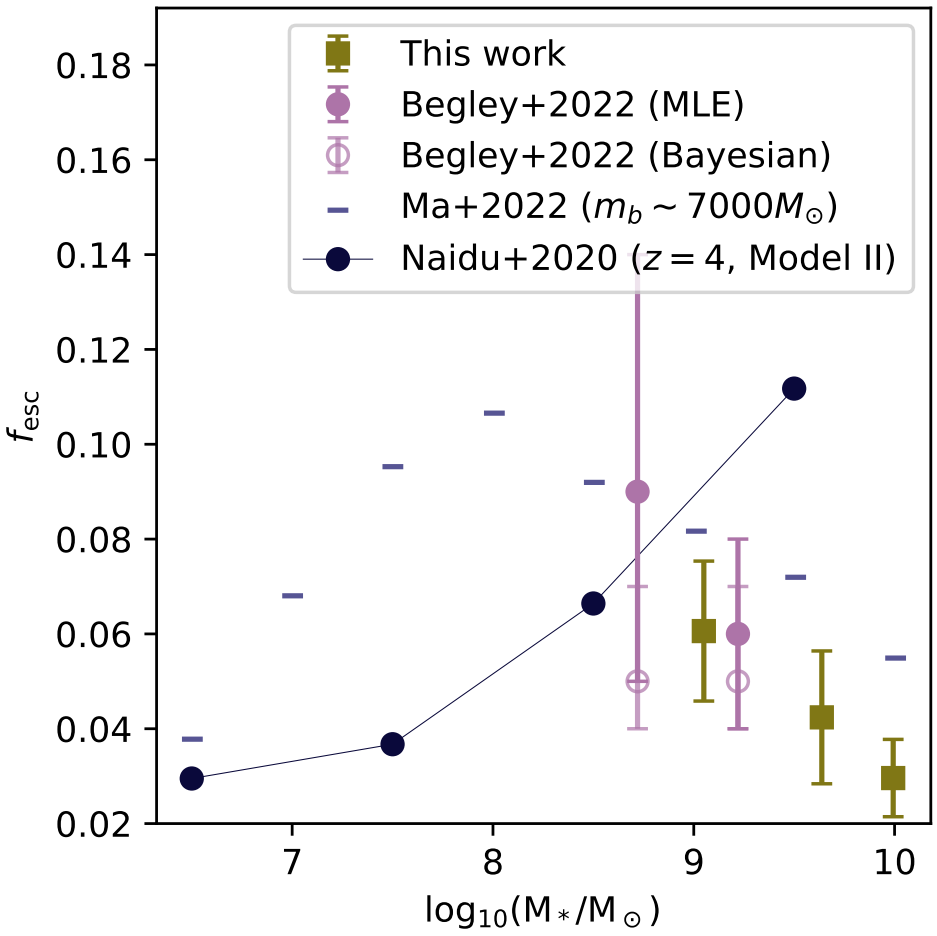

Figure 5.5: Inferred $f_{\rm esc}$ as a function of stellar mass from this work alongside trends from observation and modeling. Estimates of $f_{\rm esc}$ for three bins of increasing M_* for the KLCS SED sample are presented as yellow boxes, and are identical to values presented in Figure 5.4. The $f_{\rm esc}$ constraints from two bins of increasing M_* from 148 $z \sim 3.5$ galaxies from VANDELS (Begley et al., 2022) are displayed as purple circles. Solid, purple circles represent $f_{\rm esc}$ fit by maximum-likelihood analysis, while skeletal, purple circles represent $f_{\rm esc}$ fit by Bayesian analysis. Predictions for $f_{\rm esc}$ from the FIRE-2 cosmological simulations at a particle mass of $m_B \sim 7000 M_\odot$ are displayed as horizontal bars (Ma et al., 2020). Predicted $f_{\rm esc}$ as a function of M_* at $z \sim 4$ for the fiducial model of Naidu et al. (2020) are displayed as dark circles.

peaks at $\log(M_*/M_\odot) \sim 7$ and drops off strongly at lower and higher masses (Rosdahl et al., 2022).

Finally, Ma et al. (2020) also explored potential synchronization of periods of intense star formation and elevated f_{esc} values. They find that feedback from star formation clears sightlines in the ISM and CGM of a galaxy, creating favorable conditions for ionizing photons to escape. This process leads to a correlation between a burst of star-formation and high f_{esc}, albeit with a few Myr time delay. If true, one might expect a higher f_{esc} in galaxies with shorter stellar ages and elevated sSFR. We find no correlation between f_{esc} and these two properties in the KLCS SED analysis, as shown in Figure 5.4. The absence of an observed trend could be explained by a less bursty SFH than those found in Ma et al. (2020), which would reduce potential dependencies between f_{esc} and stellar age.

5.4.3 Implications for reionization

Models of reionization and their predicted timelines are built upon assumptions regarding f_{esc}, which are impossible to constrain directly in the reionization era. Some assume single values of f_{esc} for all galaxies for simplicity (typically $10-20\%$; Robertson et al., 2015; Ishigaki et al., 2018), others assume that f_{esc} depends on halo mass (Finkelstein et al., 2019), or that f_{esc} depends on one particular galaxy property (Naidu et al., 2020; Matthee et al., 2022).

We compare our f_{esc} vs. M_* trend to the predictions of the fiducial model of Naidu et al. (2020), which concludes that reionization is "oligarchical," such that the most luminous, massive ($M_{UV} < -18$ and $\log(M_*/M_\odot) > 8$) galaxies at $z > 6$ contribute the bulk of the ionizing photon budget. In this model, a direct relationship between f_{esc} and Σ_{SFR} is assumed such that $f_{esc} = 1.6 \times \Sigma_{SFR}^{0.4}$. As massive and UV bright galaxies tend to have high Σ_{SFR}, the assumed connection between f_{esc} and Σ_{SFR} results in a positive relationship between f_{esc} and M_*. We display this trend determined at $z = 4$ as dark circles in Figure 5.5. The trend we observe between f_{esc} and M_* at $z \sim 3$ is inconsistent with the direction and magnitude of the

model curve within the mass range where the model and observations overlap. Specifically, we show that $f_{\rm esc}$ decreases with increasing M_* at $\log(M_*/M_\odot)\sim 9.5$. We also find significantly lower $\langle f_{\rm esc}\rangle$ at fixed M_* than is predicted by the model. Some of this offset at fixed mass is likely due to the average value of $\langle f_{\rm esc}\rangle = 0.09 \pm 0.01$ from Steidel et al. (2018), used as a constraint in the model, considering $\langle f_{\rm esc}\rangle$ of the KLCS was corrected to $\langle f_{\rm esc}\rangle = 0.06 \pm 0.01$ after removal of foreground contamination (Pahl et al., 2021). Additionally, the fiducial model of Naidu et al. (2020) does not explicitly consider dust, which we find is a significant factor modulating the escape fraction of galaxies in our sample. To conservatively match observed relationships between $f_{\rm esc}$ and M_* at $z \sim 3$, assumed $f_{\rm esc}$ values of galaxies at $M_* > 10^9 M_\odot$ should be no higher than the $f_{\rm esc}$ values of $M_* \sim 10^{8.5} M_\odot$ galaxies. Specifically, the $f_{\rm esc}$ value for the most massive M_* datapoint from Naidu et al. (2020) should shift to become lower than or equal to the $f_{\rm esc}$ value for the second-most-massive datapoint. Meeting this requirement at $z \sim 4$, which is the closest point of contact between the Naidu et al. (2020) model and our observations, would require a reduction in $f_{\rm esc}$ for $M_* > 10^9$ galaxies by a factor of two in the fiducial model of Naidu et al. (2020). Satisfying this criterion at $z > 6$ during the epoch of reionization would require a similar reduction. This adjustment would significantly shift the burden of reionization to lower mass galaxies (e.g., Finkelstein et al., 2019).

The rapidity of reionization depends strongly on the population of galaxies that dominates the ionizing emissivity over cosmic time. Our results indicate that fainter, less massive galaxies with lower dust content have conditions favorable for escaping ionizing radiation, broadly consistent with other recent LyC observations at $z \sim 3$ (Begley et al., 2022) and in the local Universe (Flury et al., 2022a,b; Chisholm et al., 2022). If these trends were present within the epoch of reionization, the process of reionization may have started early and progressed gradually, such that the IGM neutral fraction is 20% at $z \sim 7$ (Finkelstein et al., 2019), in slight tension with neutral fraction constraints from Lyα damping wing measurements (Bolton et al., 2011; Greig et al., 2017; Bañados et al., 2018). Chisholm et al.

(2022) calculate the ionizing emissivity between $z \sim 4-8$ using empirical relations between $f_{\rm esc}$ and β_{1500} found in the LzLCS, which are consistent with our results connecting $f_{\rm esc}$ and E(B-V), and match constraints indicating that the ionizing emissivity flattens out at $z < 5.5$ (Becker & Bolton, 2013; Becker et al., 2021). However, as noted in Section 5.4.1, average $f_{\rm esc}$ values assumed by Chisholm et al. (2022) are too low at $z \sim 3$ when compared to the KLCS. As an alternative scenario, Matthee et al. (2022) instead directly tie $f_{\rm esc}$ to the strength of Lyα emission and build a model that produces rapid reionization and a flattened evolution of the ionizing emissivity at $z < 6$. Predicted trends between $f_{\rm esc}$ and $L_{\rm UV}$ from the model of Matthee et al. (2022) qualitatively match the KLCS anti-correlation, but underpredict the average $f_{\rm esc}$ at $z \sim 3$. Nonetheless, such a prescription is promising considering that the relationship between $f_{\rm esc}$ and W_λ(Lyα) appears be one of the most fundamental in our analysis. The Matthee et al. (2022) model assumes $f_{\rm esc} = 50\%$ for half of Lyα emitters with $L_{\rm Ly\alpha} > 10^{42}$erg s^{-1}, based on the Lyα line profile shapes of $z \sim 2$ Lyα emitters (Naidu et al., 2022). Our ongoing spectroscopic observing program to explore the connection between Lyα profile shape and LyC escape in the KLCS will test this formalism, which relies on a correlation between $f_{\rm esc}$ and Lyα peak separation that currently lacks direct observational support at high redshift.

Both the Chisholm et al. (2022) and Matthee et al. (2022) models highlight important existing relationships found between $f_{\rm esc}$ and galaxy property in our analyses, and present ionizing emissivities that both overcome recombination in the IGM at $z \sim 8$ and avoid overproducing ionizing photons at $z < 6$. The trends between $f_{\rm esc}$ and galaxy properties presented in this work are vital for anchoring assumptions of $f_{\rm esc}$ during the epoch of reionization, where direct constraints on $f_{\rm esc}$ are impossible. Future reionization models can utilize these relationships to ensure consistency between reionization-era $f_{\rm esc}$ prescriptions and our empirical results, particularly in comparable galaxy populations that have similar luminosities and masses to those of our sample. We will extend our analysis of $f_{\rm esc}$ and galaxy properties to lower $L_{\rm UV}$ in future work, which will elucidate most fundamental predictors of $f_{\rm esc}$ for a larger dynamic

range of galaxy properties.

5.5 Summary

In this work, we examine the underlying processes behind the escape of ionizing radiation by exploring trends between $f_{\rm esc}$ and galaxy properties at $z \sim 3$. We accomplish this goal by leveraging multi-band photometry of galaxies observed spectroscopically as part of KLCS. We examined a subsample of 96 KLCS galaxies with photometry suitable for SED fitting, and determined galaxy population parameters of M_*, SFR, sSFR, E(B-V), and age from these stellar-population synthesis fits. For each galaxy property, we sorted the 96 galaxies and divided them into three equal-sized bins, constructing a rest-UV composite spectrum for each bin. The main results regarding the estimated Lyman-continuum escape parameters of $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and $f_{\rm esc}$ and their relationships with galaxy properties are as follows:

1. We find significant correlations between $f_{\rm esc}$ and W_λ(Lyα) and anti-correlations between $f_{\rm esc}$ and $L_{\rm UV}$ in the KLCS SED subsample, indicating that our sample is representative of the full KLCS and appropriate for constraining $f_{\rm esc}$ as a function of other galaxy properties (Pahl et al., 2022).

2. We find significant anti-correlation between $f_{\rm esc}$ and E(B-V) across three bins of increasing E(B-V), although no correlation between $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and E(B-V). The $f_{\rm esc}$ result indicates that dust modulates escaping ionizing radiation at $z \sim 3$. Such modulation naturally arises due to the spatial coincidence of neutral-phase gas and dust (Reddy et al., 2016b; Du et al., 2018; Pahl et al., 2020), and the fact that dust directly absorbs LyC photons (Reddy et al., 2016a). These results are broadly consistent with anti-correlations found between $f_{\rm esc}$ and UV spectral slope at $z \sim 3.5$ (Begley et al., 2022) and in the local universe (Flury et al., 2022a,b; Chisholm et al., 2022).

3. Both $\langle f_{900}/f_{1500}\rangle_{\rm out}$ and $f_{\rm esc}$ are significantly correlated with M_* and SFR. Trends

between $f_{\rm esc}$ and M_* have also been suggested in other LyC surveys at high and low redshift (Begley et al., 2022; Flury et al., 2022a,b). The sense of the relationships we observe is consistent with recovered anti-correlation between $f_{\rm esc}$ and M_* for $\log(M_*/M_\odot)>8$ galaxies in cosmological simulations (Ma et al., 2020; Kostyuk et al., 2022).

4. Some cosmological zoom-in simulations of reionization-era galaxies connect stellar feedback and favorable ISM/CGM conditions for LyC escape (e.g., Ma et al., 2020), which would plausibly manifest as elevated estimates of $f_{\rm esc}$ in galaxies with shorter inferred stellar ages and higher sSFR. However, we find no correlation between $\langle f_{900}/f_{1500}\rangle_{\rm out}$ or $f_{\rm esc}$ and stellar age or sSFR, providing no direct observational support for the synchronization of recent bursts of star formation and the escape of ionizing photons at these masses. These trends are consistent with the absent or weak correlation found between W_λ(Lyα) and stellar age in earlier work (Du et al., 2018; Pahl et al., 2020; Reddy et al., 2022).

These results represent a comprehensive exploration of $f_{\rm esc}$ and SED-modeled properties at high redshift, grounding assumptions of $f_{\rm esc}$ for galaxies in the reionization era. Significant unknowns still remain for $f_{\rm esc}$ and its dependencies in galaxies less luminous than those in our sample, particularly at $z\sim3$. In future work, we will extend our examination of $f_{\rm esc}$ and galaxy properties down to lower UV luminosities. Additional indirect diagnostics of $f_{\rm esc}$ that have proven promising in the local Universe can also be tested at high redshift with the KLCS. Ongoing follow-up of the Lyα line profiles of KLCS galaxies will elucidate potential trends between $f_{\rm esc}$ and Lyα peak separation, and Keck/MOSFIRE spectra in hand for a substantial subset of the KLCS will enable an examination of the relationships between nebular emission-line properties and $f_{\rm esc}$. Using the results summarized in this section in tandem with future analyses of the KLCS, we will attempt to offer a unified picture of escaping ionizing radiation at $z\sim3$. This picture is vital for understanding the contribution

of star-forming galaxies to reionization at earlier times.

CHAPTER 6

Conclusions

This dissertation presents results that improve our understanding of reionization through an investigation of two key parameters: the escape fraction of ionizing radiation $f_{\rm esc}$, and ionizing photon production efficiency, $\xi_{\rm ion}$. The results presented herein can be used to delineate between previously-constructed reionization models, and inform future ones.

In Chapter 2, I explore the evolution of rest-UV spectral features out to $z \sim 5$ for the first time using a spectroscopic sample sample of 175 galaxies, utilizing theoretical IGM absorption curves to perform an important correction to the Lyα profile. I compared these results to similar analyses of $z \sim 2-4$ galaxies performed in Du et al. (2018). I found that Lyα equivalent width is inversely proportional to both low-ionization absorption line strength and E(B-V), indicating that the neutral gas covering fraction modulates the strength of Lyα emission from $z \sim 2-5$. This result also demonstrates that dust is likely physically coincident with neutral-phase gas in the ISM and CGM. Through the observed increase of Lyα emission line strength at fixed low-ionization line absorption strength at $z \sim 5$, I concluded that $\xi_{\rm ion}$ of star-forming galaxies is heightened at $z \sim 5$ as compared to $z \sim 2-4$. This conclusion is supported by the detection of nebular CIV emission in the galaxies with the highest Lyα equivalent widths at $z \sim 5$, which indicates significant subsolar metallicity is present in a subsample of $z \sim 5$ galaxies near reionization.

In Chapter 3, I used new *HST* data to investigate and remove the effects of foreground contamination on LyC measurements in the KLCS survey. After measuring multi-wavelength

photometry of individual subcomponents of 15 LyC detections in the KLCS, I found that two objects contained subcomponents with colors that were consistent with lower-redshift interlopers. I additionally found that two LyC non-detections with *HST* imaging had similar evidence of contamination. After removing these four contaminated objects from the KLCS, I re-measured the average $f_{\rm esc}$ of KLCS galaxies as 0.06 ± 0.01, which was significantly lower than the original $f_{\rm esc} = 0.09 \pm 0.01$ reported in Steidel et al. (2018). I also found that the trends between $f_{\rm esc}$, Lyα equivalent width, and UV Luminosity found in the KLCS were preserved after the removal of contamination.

In Chapter 4, I analyzed KLCS galaxies to explore the potential connection between $f_{\rm esc}$ and star-formation rate (SFR) surface density ($\Sigma_{\rm SFR}$), which is a promising predictor of escaping ionizing radiation in the local Universe. I examined a sample of 35 galaxies drawn from the cleaned KLCS survey which also had *HST* imaging from Chapter 3. First, I measured galaxy sizes using Sérsic profile fits to high-resolution, V_{606} images, and SFR from SED fits to ground- and space-based photometry. I split the sample into two bins of $\Sigma_{\rm SFR}$, and, based on composite spectra, I estimated the average $f_{\rm esc}$ for both $\Sigma_{\rm SFR}$ subsamples, and found no significant difference in $f_{\rm esc}$ between the two. After performing stacking simulations using the KLCS parent sample, I concluded that the size of the sample was not sufficient for robustly probing the $f_{\rm esc}$-$\Sigma_{\rm SFR}$ connection, and determined the optimal sample size to inform future *HST* LyC observing programs.

I extended these analyses in Chapter 5 by examining composite spectra in bins of stellar mass (M_*), dust reddening (E(B-V)), and stellar age, as inferred from SED fits to multi-band photometry. This analysis did not require *HST* data, thus, a the vast majority of the KLCS was included (96 galaxies). I showed that $f_{\rm esc}$ is significantly anti-correlated with M_* in the KLCS. This result is in direct tension with the prominent reionization model of Naidu et al. (2020), which predicts reionization is primarily driven by the most luminous, massive ($M_{\rm UV} < -18$, $\log(M_*/M_\odot) > 8$) galaxies. I also demonstrated that $f_{\rm esc}$ is significantly anti-

correlated with E(B-V), reflecting that LyC escapes through holes in a neutral-phase and dusty ISM/CGM at $z \sim 3$ (Reddy et al., 2016b). Finally, I found no significant trend between $f_{\rm esc}$ and stellar age, surprising considering the apparent connection between the feedback of recent star formation and the escape of LyC in cosmological zoom-in simulations (Ma et al., 2020). These simulations use bursty star-formation histories, which may be inappropriate for galaxies with stellar masses similar to those in the KLCS.

My analysis presented in this dissertation has lead to a clean estimate of $f_{\rm esc}$ at $z \sim 3$; new trends connecting $f_{\rm esc}$, M_*, and E(B-V); and evidence for a redshift evolution of $\xi_{\rm ion}$ from $z \sim 2-5$. Significant work remains for understanding which galaxy properties are most appropriate to map onto $f_{\rm esc}$ in the epoch of reionization.

A particularly important path forward is understanding the relation between $f_{\rm esc}$ and galaxy properties at lower luminosities and stellar masses than the KLCS. Lower-luminosity galaxies are both more numerous at $z > 6$ and likely have elevated $f_{\rm esc}$ values, as indicated by the trends between $f_{\rm esc}$ and M_* presented in Chapter 5. Recent reionization models, built upon assumptions that $f_{\rm esc}$ depends on dust content (Chisholm et al., 2022) or the shape of the Lyα emission line (Matthee et al., 2022), indicate that the population of galaxies that dominate the ionizing emissivity may lie at $M_* \sim 10^8 M_\odot$. The importance of this lower-mass population is underlined by the results of hydrodynamic simulations that find that the $f_{\rm esc}$—M_* relation likely also peaks at $\sim 10^8 M_\odot$ (Ma et al., 2020; Kostyuk et al., 2022), but this parameter space remains observationally-unexplored at $z \sim 3$. An extension to the KLCS is currently underway which will be used to extend the relations of Chapter 5 to lower luminosities and masses. Additionally, observational follow up has been performed for a subset of the KLCS with a higher-resolution grating on Keck/LRIS, with spectral coverage of the Lyα emission line. Analysis of $f_{\rm esc}$ as a function of Lyα profile shape and kinematics enabled by this higher-resolution data will directly test the assumptions of the Matthee et al. (2022) reionization model. In addition, we will be able to test if there is evolution in the

relation between $f_{\rm esc}$ and Lyα kinematics between $z \sim 0$ and $z \sim 3$.

With the launch of *JWST*, more information can be gleaned about the galaxies in reionization and those in the $z \sim 3$ Universe. Measurements of $\xi_{\rm ion}$ have thus-far been limited in the epoch of reionization, with only a handful of individual $\xi_{\rm ion}$ measurements of $z \sim 7$ galaxies determined from rest-UV metal lines (Stark et al., 2015, 2017). The coverage and sensitivity of the *JWST*/Near Infrared Spectrograph (NIRSpec) opens up an ideal opportunity to explore $\xi_{\rm ion}$ for an ensemble of $z > 6$ galaxies, which would significantly reduce uncertainties present in existing reionization models that arise from assumptions of this parameter. Additionally, the spectroscopic coverage of *JWST*/NIRSpec will allow the recovery of Hα at $z \gtrsim 2.7$, which can ensure both $\xi_{\rm ion}$ and metallicity measurements for reionization-era analog galaxy samples at $z \sim 3$. These additional measurements will allow a more sophisticated and comprehensive understanding of the mechanics of ionizing-photon production and escape at the highest redshifts feasible. With these discoveries on the horizon, we grow closer to a complete understanding of the contribution of star-forming galaxies to reionization.

Bibliography

Adelberger K. L., Steidel C. C., Shapley A. E., Hunt M. P., Erb D. K., Reddy N. A., Pettini M., 2004, The Astrophysical Journal, 607, 226

Bañados E., et al., 2018, Nature, 553, 473

Becker G. D., Bolton J. S., 2013, Monthly Notices of the Royal Astronomical Society, 436, 1023

Becker G. D., D'Aloisio A., Christenson H. M., Zhu Y., Worseck G., Bolton J. S., 2021, Monthly Notices of the Royal Astronomical Society, 508, 1853

Begley R., et al., 2022, Monthly Notices of the Royal Astronomical Society, 513, 3510

Berg D. A., Erb D. K., Auger M. W., Pettini M., Brammer G. B., 2018, The Astrophysical Journal, 859, 164

Berg D. A., Chisholm J., Erb D. K., Pogge R., Henry A., Olivier G. M., 2019, The Astrophysical Journal, 878, L3

Bertin E., Arnouts S., 1996, Astronomy and Astrophysics Supplement Series, 117, 393

Bezanson R., et al., 2016, The Astrophysical Journal, 822, 30

Bian F., Fan X., McGreer I., Cai Z., Jiang L., 2017, The Astrophysical Journal Letters, 837

Bielby R., et al., 2012, Astronomy and Astrophysics, 545, A23

Bolton J. S., Haehnelt M. G., Warren S. J., Hewett P. C., Mortlock D. J., Venemans B. P., McMahon R. G., Simpson C., 2011, Monthly Notices of the Royal Astronomical Society, 416, L70

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

Bouwens R. J., et al., 2014, The Astrophysical Journal, 793, 115

Bouwens R. J., Illingworth G. D., Oesch P. A., Caruana J., Holwerda B., Smit R., Wilkins S., 2015, The Astrophysical Journal, 811

Bouwens R. J., et al., 2016, The Astrophysical Journal, 833, 72

Bradley L., et al., 2020, Astropy/Photutils: 1.0.1, Zenodo, doi:10.5281/ZENODO.4049061

Brammer G. B., et al., 2012, The Astrophysical Journal, Supplement Series, 200, 13

Bruzual G., Charlot S., 2003, Monthly Notices of the Royal Astronomical Society, 344, 1000

Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, The Astrophysical Journal, 533, 682

Capak P., et al., 2007, The Astrophysical Journal Supplement Series, 172, 99

Chabrier G., 2003, Publications of the Astronomical Society of the Pacific, 115, 763

Chen Y., et al., 2021, Monthly Notices of the Royal Astronomical Society, 508, 19

Chisholm J., Tremonti Christy A., Leitherer C., Chen Y., 2016, Monthly Notices of the Royal Astronomical Society, 463, 541

Chisholm J., et al., 2018, Astronomy and Astrophysics, 616, A30

Chisholm J., Rigby J. R., Bayliss M., Berg D. A., Dahle H., Gladders M., Sharon K., 2019, The Astrophysical Journal, 882, 182

Chisholm J., et al., 2022, Monthly Notices of the Royal Astronomical Society, stac2874

Cowie L. L., Barger A. J., Trouille L., 2009, The Astrophysical Journal, 692, 1476

De Barros S., et al., 2016, Astronomy and Astrophysics, 585, 51

De Barros S., et al., 2017, Astronomy and Astrophysics, 608, A123

Domínguez A., Siana B., Brooks A. M., Christensen C. R., Bruzual G., Stark D. P., Alavi A., 2015, Monthly Notices of the Royal Astronomical Society, 451, 839

Du X., Shapley A. E., Martin C. L., Coil A. L., 2016, The Astrophysical Journal, 829, 64

Du X., et al., 2018, The Astrophysical Journal, 860, 75

Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A., Taylor G., Ng M., Greis S. M., Bray J. C., 2017, Publications of the Astronomical Society of Australia, 34, 61

Erb D. K., Pettini M., Shapley A. E., Steidel C. C., Law D. R., Reddy N. A., 2010, The Astrophysical Journal, 719, 1168

Faisst A. L., et al., 2016, The Astrophysical Journal, 822, 29

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

Finkelstein S. L., et al., 2012, The Astrophysical Journal, 758

Finkelstein S. L., et al., 2019, The Astrophysical Journal, 879, 36

Fletcher T. J., Tang M., Robertson B. E., Nakajima K., Ellis R. S., Stark D. P., Inoue A., 2019, The Astrophysical Journal, 878, 87

Flury S. R., et al., 2022a, The Astrophysical Journal Supplement Series, 260, 1

Flury S. R., et al., 2022b, The Astrophysical Journal, 930, 126

Förster Schreiber N. M., et al., 2006, The Astronomical Journal, 131, 1891

Fruchter A. S., Hack W., Dencheva N., Droettboom M., Greenfield P., 2010, 2010 Space Telescope Science Institute Calibration Workshop, pp 382–387

Garn T., Best P. N., 2010, Monthly Notices of the Royal Astronomical Society, 409, 421

Gazagnes S., Chisholm J., Schaerer D., Verhamme A., Rigby J. R., Bayliss M., 2018, Astronomy and Astrophysics, 616, 29

Gazagnes S., Chisholm J., Schaerer D., Verhamme A., Izotov Y., 2020, Astronomy and Astrophysics, 639, A85

Gillman S., et al., 2020, Monthly Notices of the Royal Astronomical Society, 492, 1492

Gordon K. D., Clayton G. C., Misselt K. A., Landolt A. U., Wolff M. J., 2003, The Astrophysical Journal, 594, 279

Grazian A., et al., 2017, Astronomy and Astrophysics, 602, A18

Graziani L., Maselli A., Ciardi B., 2013, Monthly Notices of the Royal Astronomical Society, 431, 722

Greig B., Mesinger A., Haiman Z., Simcoe R. A., 2017, Monthly Notices of the Royal Astronomical Society, 466, 4239

Hainline K. N., Shapley A. E., Greene J. E., Steidel C. C., 2011, The Astrophysical Journal, 733, 31

Hasinger G., et al., 2018, The Astrophysical Journal, 858, 77

Hayes M., et al., 2014, The Astrophysical Journal, 782, 6

Heckman T. M., Alexandroff R. M., Borthakur S., Overzier R., Leitherer C., 2015, The Astrophysical Journal, 809

Hellwing W. A., Frenk C. S., Cautun M., Bose S., Helly J., Jenkins A., Sawala T., Cytowski M., 2016, Monthly Notices of the Royal Astronomical Society, 457, 3492

Hopkins P. F., Richards G. T., Hernquist L., 2007, The Astrophysical Journal, 654, 731

Hopkins P. F., et al., 2018, Monthly Notices of the Royal Astronomical Society, 480, 800

Ishigaki M., Kawamata R., Ouchi M., Oguri M., Shimasaku K., Ono Y., 2018, The Astrophysical Journal, 854, 73

Izotov Y. I., Schaerer D., Thuan T. X., Worseck G., Guseva N. G., Orlitova I., Verhamme A., 2016, Monthly Notices of the Royal Astronomical Society, 461, 3683

Izotov Y. I., Worseck G., Schaerer D., Guseva N. G., Thuan T. X., Fricke K. J., Verhamme A., Orlitová I., 2018, Monthly Notices of the Royal Astronomical Society, 478, 4851

Izotov Y. I., Worseck G., Schaerer D., Guseva N. G., Chisholm J., Thuan T. X., Fricke K. J., Verhamme A., 2021, Monthly Notices of the Royal Astronomical Society, 503, 1734

Ji Z., et al., 2020, The Astrophysical Journal, 888, 109

Jones T., Stark D. P., Ellis R. S., 2012, The Astrophysical Journal, 751

Jones L. H., Barger A. J., Cowie L. L., 2021, The Astrophysical Journal, 908, 222

Jung I., et al., 2018, The Astrophysical Journal, 864, 103

Kakiichi K., Gronke M., 2021, The Astrophysical Journal, 908, 30

Kornei K. A., Shapley A. E., Erb D. K., Steidel C. C., Reddy N. A., Pettini M., Bogosavljević M., 2010, The Astrophysical Journal, 711, 693

Kostyuk I., et al., 2022, arXiv, arXiv:2207.11278

Kulkarni G., Worseck G., Hennawi J. F., 2019, Monthly Notices of the Royal Astronomical Society, 488, 1035

Laigle C., et al., 2016, The Astrophysical Journal Supplement Series, 224, 24

Laursen P., Sommer-Larsen J., Razoumov A. O., 2011, The Astrophysical Journal, 728

Laursen P., Sommer-Larsen J., Milvang-Jensen B., Fynbo J. P. U., Razoumov A. O., 2019, Astronomy & Astrophysics, 627, A84

Law D. R., Steidel C. C., Shapley A. E., Nagy S. R., Reddy N. A., Erb D. K., 2012, The Astrophysical Journal, 745, 85

Law D. R., et al., 2016, The Astronomical Journal, 152, 83

Leitherer C., Otálvaro P. A., Bresolin F., Kudritzki R. P., Lo Faro B., Pauldrach A. W., Pettini M., Rix S. A., 2010, Astrophysical Journal, Supplement Series, 189, 309

Leitherer C., Tremonti C. A., Heckman T. M., Calzetti D., 2011, The Astronomical Journal, 141

Ma X., Hopkins P. F., Kasen D., Quataert E., Faucher-Giguère C. A., Kereš D., Murray N., Strom A., 2016, Monthly Notices of the Royal Astronomical Society, 459, 3614

Ma X., Quataert E., Wetzel A., Hopkins P. F., Faucher-Giguère C. A., Kereš D., 2020, Monthly Notices of the Royal Astronomical Society, 498, 2001

Mack J., Olszewksi H., Pirzkal N., 2021, Technical report, WFC3/IR Filter-Dependent Sky Flats

Madau P., Dickinson M., 2014, Annual Review of Astronomy and Astrophysics, 52, 415

Marchi F., et al., 2017, Astronomy and Astrophysics, 601, 73

Marchi F., et al., 2018, Astronomy and Astrophysics, 614, 11

Marchi F., et al., 2019, Astronomy & Astrophysics

Marinacci F., et al., 2018, Monthly Notices of the Royal Astronomical Society, 480, 5113

Mason C. A., Treu T., Dijkstra M., Mesinger A., Trenti M., Pentericci L., de Barros S., Vanzella E., 2018, The Astrophysical Journal, 856, 2

Matthee J., et al., 2022, Monthly Notices of the Royal Astronomical Society, 512, 5960

McGreer I. D., Mesinger A., D'Odorico V., 2015, Monthly Notices of the Royal Astronomical Society, 447, 499

McKinney J. H., Jaskot A. E., Oey M. S., Yun M. S., Dowd T., Lowenthal J. D., 2019, The Astrophysical Journal, 874, 52

McLure R. J., et al., 2018, Monthly Notices of the Royal Astronomical Society, 476, 3991

Momcheva I. G., et al., 2016, The Astrophysical Journal Supplement Series, 225, 27

Mostardi R. E., Shapley A. E., Steidel C. C., Trainor R. F., Reddy N. A., Siana B., 2015, Astrophysical Journal, 810, 107

Naidu R. P., Tacchella S., Mason C. A., Bose S., Oesch P. A., Conroy C., 2020, The Astrophysical Journal, 892, 109

Naidu R. P., et al., 2022, Monthly Notices of the Royal Astronomical Society, 510, 4582

Naiman J. P., et al., 2018, Monthly Notices of the Royal Astronomical Society, 477, 1206

Nakajima K., Fletcher T., Ellis R. S., Robertson B. E., Iwata I., 2018, Monthly Notices of the Royal Astronomical Society, 477, 2098

Nakajima K., Ellis R. S., Robertson B. E., Tang M., Stark D. P., 2019, arXiv, 889, 161

Nelson D., et al., 2018, Monthly Notices of the Royal Astronomical Society, 475, 624

Nelson D., et al., 2019, Monthly Notices of the Royal Astronomical Society, 490, 3234

Oesch P. A., et al., 2012, The Astrophysical Journal, Volume 772, Issue 2, article id. 136, 13 pp. (2013)., 772

Oke J. B., Gunn J. E., 1983, The Astrophysical Journal, 266, 713

Oke J. B., et al., 1995, Publications of the Astronomical Society of the Pacific, 107, 375

Onodera M., et al., 2016, The Astrophysical Journal, 822, 42

Ouchi M., et al., 2009, The Astrophysical Journal, 706, 1136

Pahl A. J., Shapley A., Faisst A. L., Capak P. L., Du X., Reddy N. A., Laursen P., Topping M. W., 2020, Monthly Notices of the Royal Astronomical Society, 493, 3194

Pahl A. J., Shapley A., Steidel C. C., Chen Y., Reddy N. A., 2021, Monthly Notices of the Royal Astronomical Society, 505, 2447

Pahl A. J., Shapley A., Steidel C. C., Reddy N. A., Chen Y., 2022, Monthly Notices of the Royal Astronomical Society, 516, 2062

Pahl A. J., Shapley A., Steidel C. C., Reddy N. A., Chen Y., Rudie G. C., Strom A. L., 2023, Monthly Notices of the Royal Astronomical Society, 521, 3247

Parsa S., Dunlop J. S., McLure R. J., 2018, Monthly Notices of the Royal Astronomical Society, 474, 2904

Parzen E., 1962, The Annals of Mathematical Statistics, 33, 1065

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, The Astronomical Journal, 124, 266

Peng C. Y., Ho L. C., Impey C. D., Rix H. W., 2010, Astronomical Journal, 139, 2097

Pentericci L., Grazian A., Fontana A., Salimbeni S., Santini P., De Santis C., Gallozzi S., Giallongo E., 2007, Astronomy and Astrophysics, 471, 433

Pettini M., Rix S. A., Steidel C. C., Adelberger K. L., Hunt M. P., Shapley A. E., 2002, The Astrophysical Journal, 569, 742

Pillepich A., et al., 2018, Monthly Notices of the Royal Astronomical Society, 475, 648

Planck Collaboration et al., 2016, Astronomy and Astrophysics, 594

Reddy N. A., Steidel C. C., 2009, The Astrophysical Journal, 692, 778

Reddy N. A., Steidel C. C., Pettini M., Adelberger K. L., Shapley A. E., Erb D. K., Dickinson M., 2008, The Astrophysical Journal Supplement Series, 175, 48

Reddy N. A., Pettini M., Steidel C. C., Shapley A. E., Erb D. K., Law D. R., 2012, The Astrophysical Journal, 754

Reddy N. A., et al., 2015, The Astrophysical Journal, 806, 259

Reddy N. A., Steidel C. C., Pettini M., Bogosavljević M., 2016a, The Astrophysical Journal, 828, 107

Reddy N. A., Steidel C. C., Pettini M., Bogosavljević M., Shapley A. E., 2016b, The Astrophysical Journal, 828, 108

Reddy N. A., et al., 2018a, The Astrophysical Journal, 853, 56

Reddy N. A., et al., 2018b, The Astrophysical Journal, 869, 92

Reddy N. A., et al., 2022, The Astrophysical Journal, 926, 31

Rivera-Thorsen T. E., et al., 2015, Astrophysical Journal, 805, 14

Rix S. A., Pettini M., Leitherer C., Bresolin F., Kudritzki R.-P., Steidel C. C., 2004, The Astrophysical Journal, 615, 98

Robertson B. E., et al., 2013, The Astrophysical Journal, 768, 71

Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, The Astrophysical Journal Letters, 802, L19

Rosdahl J., et al., 2022, Monthly Notices of the Royal Astronomical Society, 515, 2386

Rosenblatt M., 1956, The Annals of Mathematical Statistics, 27, 832

Rudie G. C., et al., 2012, The Astrophysical Journal, 750, 67

Rudie G. C., Steidel C. C., Shapley A. E., Pettini M., 2013, The Astrophysical Journal, 769, 146

Saldana-Lopez A., et al., 2022a, The VANDELS Survey: The Ionizing Properties of Star-Forming Galaxies at 3<= Z<=5 Using Deep Rest-Frame Ultraviolet Spectroscopy, doi:10.48550/arXiv.2211.01351

Saldana-Lopez A., et al., 2022b, Astronomy and Astrophysics, 663, A59

Salpeter E. E., 1955, The Astrophysical Journal, 121, 161

Sanders R. L., et al., 2015, Astrophysical Journal, 799, 138

Sanders R. L., et al., 2018, The Astrophysical Journal, 858, 99

Sawala T., et al., 2016, Monthly Notices of the Royal Astronomical Society, 457, 1931

Saxena A., et al., 2021, Monthly Notices of the Royal Astronomical Society, 511, 120

Scarlata C., et al., 2009, Astrophysical Journal, 704, 98

Schenker M. A., Ellis R. S., Konidaris N. P., Stark D. P., 2014, The Astrophysical Journal, 795

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, The Astrophysical Journal, 500, 525

Scott D., 2015, Multivariate Density Estimation: Theory, Practice, and Visualization, 2nd Edition. Wiley

Scoville N., et al., 2007, The Astrophysical Journal Supplement Series, 172, 1

Senchyna P., et al., 2017, Monthly Notices of the Royal Astronomical Society, 472, 2608

Senchyna P., Stark D. P., Chevallard J., Charlot S., Jones T., Vidal-García A., 2019, Monthly Notices of the Royal Astronomical Society, 488, 3492

Shapley A. E., Steidel C. C., Adelberger K. L., Dickinson M., Giavalisco M., Pettini M., 2001, ApJ, 562, 95

Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, The Astrophysical Journal, 588, 65

Shapley A. E., Steidel C. C., Strom A. L., Bogosavljević M., Reddy N. A., Siana B., Mostardi R. E., Rudie G. C., 2016, The Astrophysical Journal, 826, L24

Sharma M., Theuns T., Frenk C., Bower R., Crain R., Schaller M., Schaye J., 2016, Monthly Notices of the Royal Astronomical Society: Letters, 458, L94

Shen X., Hopkins P. F., Faucher-Giguère C. A., Alexander D. M., Richards G. T., Ross N. P., Hickox R. C., 2020, Monthly Notices of the Royal Astronomical Society, 495, 3252

Shibuya T., Ouchi M., Harikane Y., 2015, Astrophysical Journal, Supplement Series, 219, 15

Shivaei I., et al., 2018, The Astrophysical Journal, 855, 42

Siana B., et al., 2015, The Astrophysical Journal, 804, 17

Skelton R. E., et al., 2014, Astrophysical Journal, Supplement Series, 214, 24

Springel V., et al., 2018, Monthly Notices of the Royal Astronomical Society, 475, 676

Stanway E. R., Eldridge J. J., 2018, Monthly Notices of the Royal Astronomical Society, 479, 75

Stanway E. R., Eldridge J. J., Becker G. D., 2016, Monthly Notices of the Royal Astronomical Society, 456, 485

Stark D. P., Ellis R. S., Chiu K., Ouchi M., Bunker A., 2010, Monthly Notices of the Royal Astronomical Society, 408, 1628

Stark D. P., et al., 2014, Monthly Notices of the Royal Astronomical Society, 445, 3200

Stark D. P., et al., 2015, Monthly Notices of the Royal Astronomical Society, 454, 1393

Stark D. P., et al., 2017, Monthly Notices of the Royal Astronomical Society, 464, 469

Steidel C. C., Hunt M. P., Shapley A. E., Adelberger K. L., Pettini M., Dickinson M., Giavalisco M., 2002, The Astrophysical Journal, 576, 653

Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Giavalisco M., 2003, The Astrophysical Journal, 592, 728

Steidel C. C., Shapley A. E., Pettini M., Adelberger K. L., Erb D. K., Reddy N. A., Hunt M. P., 2004, The Astrophysical Journal, 604, 534

Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, The Astrophysical Journal, 717, 289

Steidel C. C., Bogosavljević M., Shapley A. E., Kollmeier J. A., Reddy N. A., Erb D. K., Pettini M., 2011, The Astrophysical Journal, 736, 160

Steidel C. C., et al., 2014, The Astrophysical Journal, 795, 165

Steidel C. C., Strom A. L., Pettini M., Rudie G. C., Reddy N. A., Trainor R. F., 2016, The Astrophysical Journal, 826, 159

Steidel C. C., Bogosavljević M., Shapley A. E., Reddy N. A., Rudie G. C., Pettini M., Trainor R. F., Strom A. L., 2018, The Astrophysical Journal, 869, 123

Strom A. L., Steidel C. C., Rudie G. C., Trainor R. F., Pettini M., Reddy N. A., 2017, The Astrophysical Journal, 836, 164

Sunnquist B., 2018, Technical report, WFC3/IR Blob Monitoring

Tacchella S., Bose S., Conroy C., Eisenstein D. J., Johnson B. D., 2018, The Astrophysical Journal, 868, 92

Trainor R. F., Strom A. L., Steidel C. C., Rudie G. C., Chen Y., Theios R. L., 2019, The Astrophysical Journal, 887, 85

Van Der Wel A., et al., 2012, Astrophysical Journal, Supplement Series, 203, 24

Vanzella E., et al., 2005, Astronomy and Astrophysics, 434, 53

Vanzella E., et al., 2006, Astronomy and Astrophysics, 454, 423

Vanzella E., et al., 2008, Astronomy and Astrophysics, 478, 83

Vanzella E., et al., 2009, The Astrophysical Journal, 695, 1163

Vanzella E., et al., 2012, The Astrophysical Journal, 751, 70

Vanzella E., et al., 2015, Astronomy and Astrophysics, 576, 116

Vanzella E., et al., 2016a, The Astrophysical Journal, 821, L27

Vanzella E., et al., 2016b, The Astrophysical Journal, 825, 41

Vanzella E., et al., 2017, The Astrophysical Journal, 842, 47

Vanzella E., et al., 2020, Monthly Notices of the Royal Astronomical Society, 491, 1093

Verhamme A., Schaerer D., Atek H., Tapken C., 2008, Astronomy and Astrophysics, 491, 89

Verhamme A., Orlitová I., Schaerer D., Hayes M., 2015, Astronomy and Astrophysics, 578, A7

Virtanen P., et al., 2020, Nature Methods, 17, 261

Whitaker K. E., Pope A., Cybulski R., Casey C. M., Popping G., Yun M. S., 2017, The Astrophysical Journal, 850, 208

Zackrisson E., Inoue A. K., Jensen H., 2013, The Astrophysical Journal, 777, 39