



The Redshift Evolution of Rest-UV Spectroscopic Properties in Lyman-break Galaxies at $z \sim 2-4$

Xinnan Du (杜辛楠)¹, Alice E. Shapley¹, Naveen A. Reddy² , Tucker Jones³ , Daniel P. Stark⁴, Charles C. Steidel⁵ , Allison L. Strom⁶ , Gwen C. Rudie⁶ , Dawn K. Erb⁷ , Richard S. Ellis⁸ , and Max Pettini⁹

¹University of California, Los Angeles, CA 90095, USA

²University of California, Riverside, CA 92521, USA

³University of California, Davis, CA 95616, USA

⁴University of Arizona, Tucson, AZ 85719, USA

⁵California Institute of Technology, Pasadena, CA 91125, USA

⁶Carnegie Observatories, Pasadena, CA 91101, USA

⁷University of Wisconsin Milwaukee, Milwaukee, WI 53211, USA

⁸University College London, London, WC1E 6BT, UK

⁹Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

Received 2018 March 15; revised 2018 April 18; accepted 2018 April 19; published 2018 June 13

Abstract

We present the first comprehensive evolutionary analysis of the rest-frame UV spectroscopic properties of star-forming galaxies at $z \sim 2-4$. We match samples at different redshifts in UV luminosity and stellar mass, and perform systematic measurements of spectral features and stellar population modeling. By creating composite spectra grouped according to Ly α equivalent width (EW) and various galaxy properties, we study the evolutionary trends among Ly α , low- and high-ionization interstellar (LIS and HIS) absorption features, and integrated galaxy properties. We also examine the redshift evolution of Ly α and LIS absorption kinematics, and fine-structure emission EWs. The connections among the strengths of Ly α , LIS lines, and dust extinction are redshift independent, as is the decoupling of the Ly α and HIS line strengths, and the bulk outflow kinematics as traced by the LIS lines. Stronger Ly α emission is observed at higher redshift at fixed UV luminosity, stellar mass, SFR, and age. Much of this variation in the average Ly α strength with redshift, and the variation in Ly α strength at fixed redshift, can be explained in terms of variations in the neutral gas covering fraction and/or dust content in the ISM and CGM. However, based on the connection between Ly α and C III] emission strengths, we additionally find evidence for variations in the intrinsic production rate of Ly α photons at the highest Ly α EWs. The challenge now is to understand the observed evolution of the neutral gas covering fraction and dust extinction within a coherent model for galaxy formation, and make robust predictions for the escape of ionizing radiation at $z > 6$.

Key words: galaxies: evolution – galaxies: high-redshift – galaxies: ISM

1. Introduction

The rest-frame ultraviolet (UV) spectra of star-forming galaxies provide rich insights into the physical properties of not only their massive stars, but also their gas. This gas includes the multiphase interstellar medium (ISM) and circumgalactic medium (CGM), which extends to the virial radius (Tumlinson et al. 2017). Both low- and high-ionization interstellar (LIS and HIS, respectively) absorption lines are probes of the ISM and CGM, which typically reflect the kinematic signatures of galaxy-wide outflows in star-forming galaxies at these redshifts (e.g., Shapley et al. 2003; Steidel et al. 2010; Jones et al. 2012). LIS features primarily trace the neutral phase of outflows while HIS lines mainly trace the ionized phase. Fine-structure emission lines (e.g., Si II*, Fe II*) arise from the emission of a photon corresponding to a drop down to the excited ground state following resonant absorption. These fine-structure features are typically coupled with corresponding resonant LIS absorption and are thus effective tracers of the structure and spatial extent of the CGM (e.g., Erb et al. 2012; Jones et al. 2012; Kornei et al. 2013). The H I Ly α feature, on the other hand, has a more complex nature. Ly α photons are produced by recombination in H II regions and then propagate through the ISM, interacting with both neutral hydrogen and dust grains. As such, Ly α photons potentially offer insights into the properties of both their H II regions of origin and the

more extended ISM and CGM through which they propagate. Nebular emission lines, such as C III] $\lambda\lambda 1907, 1909$ and [O III] $\lambda\lambda 1661, 1666$, are produced in H II regions and serve as useful probes of the ionized ISM and radiation field produced by massive stellar populations (e.g., Erb et al. 2010; Berg et al. 2016; Du et al. 2017; Senchyna et al. 2017; Stark et al. 2017). When measured over a wide range of redshifts, the properties of rest-UV absorption and emission features provide a window into the evolution of ISM/CGM properties. Properties of particular interest include the evolving gas covering fraction, which modulates the emergent Ly α properties and escape of ionizing radiation, and the characteristics of galaxy outflows, which play a crucial role in the formation and evolution of galaxies.

As Ly α photons are produced in H II regions and are scattered and absorbed in the dusty ISM, Ly α is a key probe of the physical conditions of interstellar gas (e.g., metallicity, ionization parameter, H I column density and covering fraction, and dust distribution) given its unique detectability especially at high redshift. Numerous studies have been carried out to understand the factors modulating the strength of Ly α emission at $z \sim 2-6$. Observationally, larger Ly α equivalent widths (EWs) are typically found in galaxies with bluer UV colors, lower metallicities, lower stellar masses, lower UV luminosities, and lower SFRs (e.g., Shapley et al. 2003; Reddy et al. 2006; Kornei et al. 2010; Pentericci et al. 2010; Stark et al. 2010; Berry et al. 2012; Jones et al. 2012; Erb et al. 2016; Hathi et al.

2016; Trainor et al. 2016). Furthermore, understanding the mechanisms underlying the existence of strong Ly α emitters (LAEs; rest-frame Ly α EW > 20 Å) is critical for our interpretation of the galaxy populations during the reionization epoch, where, thus far, the majority of data are only photometric.

The connections between Ly α and other spectral features, including LIS and HIS absorption lines and nebular emission lines, have been extensively studied in star-forming galaxies. Stronger Ly α appears to be associated with weaker LIS absorption lines (Shapley et al. 2003; Berry et al. 2012; Jones et al. 2012) and stronger nebular emission lines (Shapley et al. 2003; Stark et al. 2014, 2015; Rigby et al. 2015; Erb et al. 2016; Jaskot & Ravindranath 2016; Trainor et al. 2016). In contrast, the Ly α EW does not show a significant correlation with the strength of the HIS absorption lines (Shapley et al. 2003; Berry et al. 2012). The shape of the Ly α profile (e.g., peak location, number of peaks, asymmetry) can also be used as a probe of the kinematics and density distribution of outflowing neutral gas (e.g., Verhamme et al. 2006; Steidel et al. 2010; Kulas et al. 2012).

While much of the work on Ly α emission and other rest-UV spectroscopic properties of star-forming galaxies at high redshift has focused on characterizing samples at individual redshifts (e.g., Shapley et al. 2003; Pentericci et al. 2007, 2009; Kornei et al. 2010; Steidel et al. 2010), there have been limited studies on the evolution of these spectroscopic properties at $z \gtrsim 2$. LAEs are found to be more prevalent at higher redshifts up to $z \sim 6$ (Reddy et al. 2008; Stark et al. 2010; Schenker et al. 2014), and stronger Ly α and weaker LIS absorption are observed with increasing redshift at fixed UV luminosity (Stark et al. 2010; Jones et al. 2012). Specifically, by comparing a sample of $z \sim 4$ Lyman-break galaxy (LBG) spectra to the $z \sim 3$ LBG composites from Shapley et al. (2003), Jones et al. (2012) found no evidence for the evolution of the correlation between Ly α and LIS absorption EWs. However, these authors discovered smaller characteristic radii of fine-structure emission and overall weaker LIS absorption at higher redshifts. While Jones et al. (2012) created a subsample of $z \sim 4$ galaxies to match the $z \sim 3$ galaxies in UV absolute magnitude for a fair comparison, the study was limited in that the SED-derived galaxy properties (e.g., stellar mass) were unavailable at that time for the majority of the galaxies in the $z \sim 3$ comparison sample from Shapley et al. (2003). Furthermore, the $z \sim 3$ spectroscopic measurements (e.g., Ly α and LIS absorption EWs) were taken from Shapley et al. (2003) and not necessarily performed in a manner consistent with that of Jones et al. (2012). When conducting evolutionary studies on the connection between rest-UV spectroscopic and galaxy properties, both the spectroscopic measurements and the modeling of galaxy photometry must be performed in a uniform, controlled fashion.

The redshift range $z \sim 2$ –4 spans nearly 2 billion years of cosmic time and covers the peak epoch of star formation and beyond (Madau & Dickinson 2014). The active star formation and mass assembly over this period must have been shaped by various processes (e.g., feedback, mergers, gas accretion), which will also be reflected in the physical characteristics of the ISM/CGM. Accordingly, it is of key interest to trace the evolving ISM/CGM with galaxy spectra.

Here we present the first comprehensive study of the evolution of the ISM and CGM as probed by rest-UV spectra in star-forming galaxies at $z \sim 2$ –4. In order to perform meaningful comparisons between galaxy samples at different

redshifts, we carefully construct these samples with galaxies spanning the same range in UV absolute magnitude and stellar mass. Furthermore, we model the stellar populations and measure the spectral features in all redshift samples in a uniform manner to avoid potential systematic biases. By comparing the rest-UV spectroscopic properties in the controlled samples at different redshifts, we aim to investigate the evolution of the covering fractions of gas and dust, the structure and kinematics of the multiphase ISM/CGM, and the intrinsic Ly α production, which, in combination, provide rich insights into galaxy evolution.

We provide a brief overview of the observations and data in Section 2, along with a description of the galaxy SED modeling and the construction of controlled samples in UV luminosity and stellar mass. In Section 3, we describe methods for creating composite spectra, and measuring Ly α emission and interstellar absorption profiles. We present multiple $z \sim 2$ –4 evolutionary trends in Section 4, including the connections among key spectroscopic features (Ly α , and LIS and HIS absorption features), between Ly α and galaxy stellar populations, and fine-structure emission properties. We also investigate the correlation between Ly α and C III] emission. In Section 5, we examine the evolution of outflow kinematics as probed by both Ly α and LIS absorption features over the same redshift range. In Section 6, we connect our results to the evolution of outflow kinematics over $z \sim 0$ –2 and present a physical picture for the evolving distribution of Ly α emission and interstellar absorption measurements at $z \sim 2$ –4. Finally, we summarize our results in Section 7. In Appendix B, we consider the rest-UV spectroscopic properties of galaxies with mid-IR SED excesses, another topic of interest that can be explored with our data.

Throughout this paper, we adopt a standard Λ CDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70 \text{ km s}^{-1}$. All wavelengths are measured in the vacuum frame. Magnitudes are in the AB system.

2. Observations, Data Reduction, and Samples

In this section, we describe the $z \sim 2$ –4 LBG data used in this study, along with the derivation of stellar population parameters from spectral energy distributions (SEDs), and the sample properties. We refer readers to the original papers in which these data were presented for a more in-depth discussion of the sample selection, data reduction, and photometric measurements of the samples presented here.

2.1. Samples

2.1.1. LRIS Sample

The $z \sim 2$ –3 sample was drawn from the UV-selected galaxy surveys described in Steidel et al. (2003, 2004) and Reddy et al. (2008). These galaxies were preselected photometrically according to the U_nGR color cuts to an apparent magnitude limit of $R_{AB} = 25.5$ and spanned 15 fields covering a total area of $\sim 1900 \text{ arcmin}^2$. The candidates were then followed up spectroscopically with the Low Resolution Imager and Spectrometer (LRIS; Oke et al. 1995; Steidel et al. 2004) on the Keck I telescope. We note that galaxies selected in this manner at $z \sim 2$ are not technically LBGs, as their actual Lyman break falls bluewards of the U_n band. However, as described in Steidel et al. (2004), the $z \sim 2$ U_nGR -selected galaxies have very similar properties to those of the $z \sim 3$ LBGs. Therefore, we also refer to the $z \sim 2$ U_nGR -selected galaxies as LBGs hereafter for simplicity.

The data were collected during multiple observing runs from 1997 to 2009 using $1''.2$ slits for the multislit masks. The majority of the objects in the LRIS sample were observed with the 400 lines mm^{-1} ($\sim 57\%$) and 600 lines mm^{-1} ($\sim 33\%$) grisms. A small fraction of the data was obtained using the 300 lines mm^{-1} grism ($\sim 3\%$) and the 300 lines mm^{-1} grating ($\sim 4\%$), the latter taken before the LRIS-B upgrade in 2000 (Steidel et al. 2004). Twenty-one spectra ($\sim 2\%$) were combined from multiple observations with a mixture of the 400 and 600 lines mm^{-1} grisms. The effective spectral resolutions of the 400, 600, and 300 lines mm^{-1} grisms and the 300 lines mm^{-1} grating are $R \sim 800, 1330, 530,$ and 670 , respectively. The typical integration time for each slit mask is 1.5 hr, and is significantly longer (5–8 hr) for a small subset of the slit masks observed as part of specialized investigations (e.g., searches for Lyman-continuum radiation; Reddy et al. 2016; Steidel 2018, in preparation). All two-dimensional spectra were flat fielded, cleaned of cosmic rays, background subtracted, extracted, wavelength, and flux calibrated.

Ideally, the systemic redshift can be robustly measured from strong rest-frame optical nebular emission lines (e.g., [O II] $\lambda 3727$, H β $\lambda 4861$, [O III] $\lambda \lambda 4959, 5007$, and H α $\lambda 6563$). However, nebular lines were not measured for the vast majority of the $z \sim 2$ –3 objects in our sample, so the systemic redshift of these objects was estimated instead based on the redshift of the Ly α emission and/or LIS absorption lines. Due to the presence of outflowing neutral gas in the galaxies, the Ly α emission and LIS absorption features are typically observed to be redshifted and blueshifted, respectively, relative to each other, and therefore do not trace the galaxy systemic velocity. Instead, the systemic redshift was determined following the procedure described in Rudie et al. (2012) assuming that Ly α is redshifted by $+300 \text{ km s}^{-1}$ from the systemic velocity and the LIS absorption lines are blueshifted by -160 km s^{-1} . The Ly α velocity correction was applied to all spectra with measurable Ly α , and the LIS correction was applied to those with LIS redshift measurements only. In this manner, we obtained the systemic redshift for all of the $z \sim 2$ –3 LBGs observed with LRIS, and the typical uncertainty on the systemic redshift is $\sim 125 \text{ km s}^{-1}$. The systemic redshift, as estimated above, was used to transform each spectrum to the rest frame.

The resulting LRIS sample includes 1297 LBGs spanning the redshift range $1.7 \lesssim z \lesssim 3.65$. These galaxies have secure redshift measurements, spectral coverage of Ly α , and detections in the K band and/or IRAC channel 1 or 2, enabling reliable stellar population modeling.

2.1.2. DEIMOS/FORS2 Sample

To probe $z \sim 4$ LBGs, we primarily use the spectra presented in Jones et al. (2012). This sample includes 70 objects (81 spectra accounting for duplicate observations) with $3.5 < z < 4.5$ and apparent magnitudes of $z'_{\text{AB}} < 26.0$. These galaxies were selected as B -band dropouts in the two GOODS fields (Giavalisco et al. 2004) and have a spectroscopic completeness of $\sim 100\%$ at $z'_{\text{AB}} < 24.75$ (Stark et al. 2018, in preparation). Detailed descriptions of sample selection and photometric measurements can be found in Stark et al. (2009, 2010) and Jones et al. (2012). In the $z \sim 4$ sample, 42 out of 70 objects (53 out of 81 spectra) were observed with the Deep Imaging Multi-object Spectrograph (DEIMOS; Faber et al. 2003) on the Keck II telescope. The DEIMOS spectra have a constant resolution of $\simeq 3.5 \text{ \AA}$ across the spectral range of

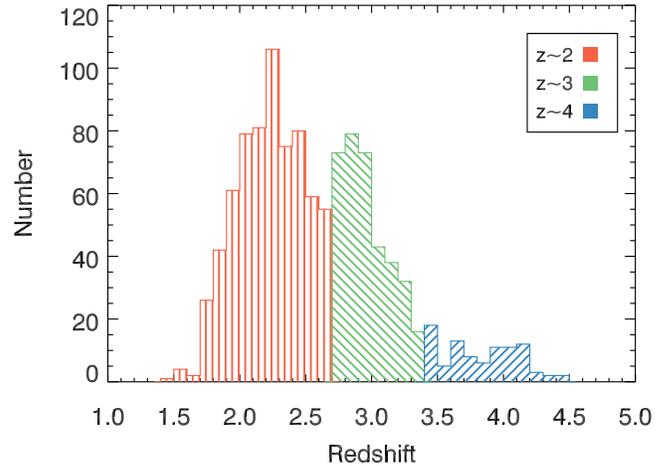


Figure 1. Redshift distribution of our sample, divided according to boundaries at $z = 2.7$ and $z = 3.4$. The $z \sim 2$ sample (671 objects) is indicated in red, the $z \sim 3$ sample (352 objects) in green, and the $z \sim 4$ sample (80 objects, 91 spectra) in blue. UV absolute magnitude and mass cutoffs have not been applied here.

4100–9300 \AA . The typical integration time for the DEIMOS spectra is 5–7 hr per mask. The other 28 objects (28 spectra) were obtained from the archival spectroscopic database of the FOcal Reducer and low dispersion Spectrograph 2 (FORS2; Vanzella et al. 2005, 2006, 2008, 2009). The FORS2 data have a spectral resolution of $R \sim 660$ over 6000–10000 \AA , corresponding to $\sim 9.0 \text{ \AA}$ near the observed wavelength of Ly α . The FORS2 spectra typically have 4–6 hr integrations for each mask. Similar to the $z \sim 2$ –3 LRIS data, standard data reduction procedures (flat fielding, cosmic-ray rejection, background subtraction, extraction, wavelength and flux calibration, and transformation to the vacuum wavelength frame) were performed on the DEIMOS and FORS2 data to obtain calibrated one-dimensional spectra.

The systemic redshift of the $z \sim 4$ galaxies was also estimated based on Ly α emission and LIS absorption features, using a procedure analogous to the one described in Section 2.1.1. To correct the Ly α and LIS absorption redshifts to the systemic value, Jones et al. (2012) assumed a velocity offset of $+330 \text{ km s}^{-1}$ for Ly α and -190 km s^{-1} for the LIS features. The systemic redshift of individual galaxies was determined from the Ly α redshift when available, given its robustness compared to that of the LIS features, and from the LIS absorption redshift when the Ly α redshift was not measurable.

2.1.3. Redshift Samples

We applied fixed redshift cuts to the LRIS, DEIMOS, and FORS2 data sets to define the $z \sim 2, 3,$ and 4 samples. We considered galaxies with $z < 2.7$ to be in the $z \sim 2$ sample, galaxies with $2.7 \leq z < 3.4$ to be in the $z \sim 3$ sample, and those with $z \geq 3.4$ to be in the $z \sim 4$ sample. The application of the redshift cuts results in 671, 352, and 80 objects (91 spectra) in the $z \sim 2, z \sim 3,$ and $z \sim 4$ samples, respectively. Figure 1 shows the redshift histograms of the $z \sim 2, 3,$ and 4 samples defined by fixed redshift cuts, with median redshifts of 2.25, 2.93, and 3.86, respectively. In practice, the $z \sim 2$ and $z \sim 3$ samples are all probed with LRIS data, while the $z \sim 4$ sample is mainly covered by DEIMOS and FORS2 spectra, with a small addition of objects observed with LRIS.

Table 1
Photometric Bands Used in SED Modeling^a

| Fields | Photometric Bands |
|--|--|
| GOODS-N ^b | <i>U</i> , <i>F435W</i> , <i>B</i> , <i>G</i> , <i>V</i> , <i>F606W</i> , <i>R</i> , <i>R_s</i> , <i>i</i> , <i>F775W</i> , <i>z</i> , <i>F850LP</i> , <i>F125W</i> , <i>F140W</i> , <i>J</i> , <i>H</i> , <i>F160W</i> , <i>K_s</i> , <i>IRAC1</i> , <i>IRAC2</i> , <i>IRAC3</i> , <i>IRAC4</i> |
| GOODS-S ^b | <i>U₃₈</i> , <i>U</i> , <i>F435W</i> , <i>B</i> , <i>V</i> , <i>F606W_{candels}</i> , <i>F606W</i> , <i>R</i> , <i>R_c</i> , <i>F775W</i> , <i>i</i> , <i>F814W_{candels}</i> , <i>F850LP</i> , <i>F850LP_{candels}</i> , <i>F125W</i> , <i>J</i> , <i>J_{tenis}</i> , <i>F140W</i> , <i>H</i> , <i>F160W</i> , <i>K_{tenis}</i> , <i>K_s</i> , <i>IRAC1</i> , <i>IRAC2</i> , <i>IRAC3</i> , <i>IRAC4</i> |
| Q0100, Q0142, Q0449, Q1009, Q1217, Q1549, Q2343 | <i>U_n^c</i> , <i>G^c</i> , <i>R_s^c</i> , <i>J1^d</i> , <i>J2^d</i> , <i>J3^d</i> , <i>J^e</i> , <i>F140 W</i> , <i>H_s^d</i> , <i>H^d</i> , <i>F160 W</i> , <i>H_l^d</i> , <i>K^e</i> , <i>IRAC1</i> , <i>IRAC2</i> , <i>IRAC3</i> , <i>IRAC4</i> |
| Q0933, Q1307, Q1422, Q1623, Q1700, Q2206, Q2346 | <i>U_n^c</i> , <i>G^c</i> , <i>R_s^c</i> , <i>i^f</i> , <i>J^g</i> , <i>H^h</i> , <i>K^g</i> , <i>F160 W</i> , <i>IRAC1</i> , <i>IRAC2</i> , <i>IRAC3</i> , <i>IRAC4</i> |

Notes.

^a While some objects in our samples miss photometric data from a subset of the bands listed here, we ensured that every object in our final, controlled samples has detections in the *K*-band and/or one of the IRAC channels.

^b Photometric bands as described in the 3D-HST catalog (Skelton et al. 2014; Brammer et al. 2012).

^c Observed with LRIS.

^d Observed with the Four Star IR camera on the Magellan Baade Telescope.

^e Observed with Multi-object Spectrometer For Infra-red Exploration (MOSFIRE) at the Keck I telescope.

^f Observed with the Kitt Peak 4 m Mayall telescope.

^g Observed with the Palomar 5.08 m telescope.

^h Observed with the Wide-field InfraRed Camera (WIRCam) at the Palomar 200-in Hale telescope.

2.2. SED Fitting

In order to make fair comparisons among LBGs at different redshifts, one important aspect of our study is the measurement of the spectral and galaxy properties in a systematic, uniform manner. In this subsection, we describe the SED-fitting procedure adopted for fitting all of the galaxies in our samples from $z \sim 2-4$.

We collected the most up-to-date photometry for all objects in the redshift samples, and the photometric bands available for SED modeling are listed in Table 1. For objects observed in the GOODS fields (145 objects in the $z \sim 2$ sample, 59 objects in the $z \sim 3$ sample, and 69 objects in the $z \sim 4$ sample), we acquired the publicly available photometric catalog from the 3D-HST survey (Skelton et al. 2014) and matched the objects in our samples with the targets observed in the GOODS-N and GOODS-S fields. The objects were matched according to their right ascension, declination, and redshift. We excluded objects from the samples if there was no unique, unambiguous match in the 3D-HST *F606W* science image. In this way, we removed seven objects from the $z \sim 2$ sample, one object from the $z \sim 3$ sample, and three objects from the $z \sim 4$ sample because we were unable to identify them in the 3D-HST photometric catalog. For galaxies in the $z \sim 2-3$ samples in fields other than GOODS-N and GOODS-S, we utilized the updated photometric measurements from Steidel et al. (2003, 2004, private communication), Reddy et al. (2012), and Strom et al. (2017).

Based on the photometric measurements of the objects in our samples, we aimed to derive key galaxy properties (e.g., stellar mass, dust extinction, age, and SFR) by fitting the SEDs of individual galaxies. In order to ensure the robustness of the stellar population modeling, we required at least one photometric measurement redward of the Balmer break. Twelve objects in the $z \sim 2$ sample and 28 objects in the $z \sim 3$ sample were removed due to a lack of near- and mid-IR photometry.

We fit the galaxy SEDs with stellar population templates from Bruzual & Charlot (2003) assuming a Chabrier (2003) initial mass function (IMF). We note that Bruzual & Charlot (2003) models do not account for emission lines or nebular continuum emission. Strong emission lines (e.g., $H\alpha$, [O III] $\lambda\lambda 4959, 5007$) could bias the shape of the SED redward of the Balmer break, resulting in an older best-fit galaxy age than the value derived from the stellar continuum alone. In the case of our $z \sim 2-4$ galaxies, [O III] and $H\alpha$ fall in the *H* and *K* bands, respectively, for the $z \sim 2$ sample; [O III] falls in the *K* band for the $z \sim 3$ sample; and $H\alpha$ falls within IRAC channel 1 for galaxies with $z \geq 3.8$ (38 out of 70 objects) in the $z \sim 4$ sample. As discussed in Section 4.2, the contamination from [O III] has a significant impact on the age estimate for the $z \sim 3$ galaxies, but nebular emission does not appear to significantly bias results from the $z \sim 2$ and $z \sim 4$ samples (bottom panel of Figure 6). Hence, we excluded *K*-band photometry from the SED fits for the $z \sim 3$ sample,¹⁰ but did not reject any photometric data in the *H* and *K* bands (IRAC channel 1) for the $z \sim 2$ ($z \sim 4$) sample.

Following Reddy et al. (2017), we adopted two combinations of metallicity and extinction curves for the SED modeling. These include 1.4 solar metallicity ($Z_{\odot} = 0.014$) with the Calzetti et al. (2000) attenuation curve (hereafter “1.4 Z_{\odot} +Calzetti”), and 0.28 Z_{\odot} with the SMC extinction curve (hereafter “0.28 Z_{\odot} +SMC”). The grid of each model includes different star formation histories (exponentially declining, constant, and rising), age ranging from 10 Myr to 5 Gyr, and $E(B - V)$ ranging from $E(B - V) = 0.00$ to 0.60.

Although our SED-fitting grid enables numerous combinations of stellar parameters, we applied several constraints on the

¹⁰ Excluding *K*-band photometry resulted in another 41 objects being removed from the $z \sim 3$ sample, as the *K* band was the reddest photometric band available for those objects.

best-fit stellar population model based on our current knowledge of the galaxies in our samples. For all the redshift samples, we considered the constant SFR model as a satisfactory description of the star formation history for the typical star-forming galaxies at $z \sim 2-4$ (Reddy et al. 2012; Steidel et al. 2014; Strom et al. 2017). Furthermore, we required the age of the $z \sim 2$ and $z \sim 3$ galaxies to be no younger than 50 Myr based on the typical dynamical timescales of these galaxies (Reddy et al. 2012). With these constraints, we calculated a set of best-fit stellar parameters for individual galaxies in our samples for both the $1.4 Z_{\odot} + \text{Calzetti}$ and the $0.28 Z_{\odot} + \text{SMC}$ models. While the former model has been traditionally used to describe the SEDs of the L^* galaxies at $z \gtrsim 2$, recent work suggests that subsolar metallicity models with an SMC curve provide a better description of the IRX– β relation for these objects (e.g., Reddy et al. 2017), especially at higher redshifts ($z \gtrsim 4$; Oesch et al. 2013; Bouwens et al. 2016). Therefore, we adopted the best-fit stellar parameters of the $0.28 Z_{\odot} + \text{SMC}$ model as the final best-fit parameters for the $z \sim 4$ galaxies, given that these objects are in general younger and likely less enriched on average, compared with the samples at lower redshift. For the $z \sim 2$ and $z \sim 3$ objects, we found that the $1.4 Z_{\odot} + \text{Calzetti}$ model gave a systemically better fit to the observed SEDs than the $0.28 Z_{\odot} + \text{SMC}$ model above a given stellar mass threshold, which is consistent with the presence of the mass–metallicity relation (Steidel et al. 2014; Sanders et al. 2015; Onodera et al. 2016). As a result, we adopted the best-fit stellar parameters yielded by the $1.4 Z_{\odot} + \text{Calzetti}$ model above $\log(M_{*,0.28Z_{\odot}+\text{SMC}}/M_{\odot}) = 10.45$ (10.65) for the $z \sim 2$ ($z \sim 3$) sample, and used those from the $0.28 Z_{\odot} + \text{SMC}$ model for galaxies below the corresponding stellar mass thresholds.

2.2.1. Controlled Samples

In order to conduct a well-controlled comparison, we need to ensure that we are comparing galaxies at different redshifts with similar galaxy properties. Two key properties that can be used to constrain the galaxy populations are luminosity and stellar mass. To determine the UV absolute magnitude for individual objects in our samples, we selected the corresponding photometric bands that cover the rest-frame wavelength $\lambda_{\text{rest}} = 1500 \text{ \AA}$. $\lambda_{\text{rest}} = 1500 \text{ \AA}$ falls in the U , G , and V bands and $F435W$ for the $z \sim 2$ sample depending on the exact galaxy redshift, in the R_i band for the $z \sim 3$ sample, and in $F775W$ for the $z \sim 4$ sample. The observed magnitude in the corresponding band was then converted into UV absolute magnitude by accounting for the monochromatic luminosity distance (since AB magnitudes correspond to flux density). The UV absolute magnitude estimated from photometry agrees well with that calculated from the best-fit galaxy SED at $\lambda_{\text{rest}} = 1500 \text{ \AA}$, and using the SED-based M_{UV} instead of the M_{UV} estimated from the photometry has a negligible effect on the sample selection as well as on the Ly α trends discussed in Section 4.2.

The $z \sim 4$ sample spans a narrower range in UV absolute magnitude and stellar mass than the $z \sim 2$ and $z \sim 3$ samples. Therefore, we selected a subsample of galaxies from each of the $z \sim 2$ and $z \sim 3$ samples that span the same range in UV absolute magnitude ($-22.62 < M_{\text{UV}} < -19.91$) and stellar mass ($8.04 < \log(M_{*}/M_{\odot}) < 11.31$) as the full $z \sim 4$ sample. In Figure 2, we plot the M_{UV} versus M_{*} diagram for the $z \sim 2, 3$, and 4 samples defined according to the fixed redshift boundaries (the same sets of objects as in Figure 1), along with the cuts in UV absolute magnitude and stellar

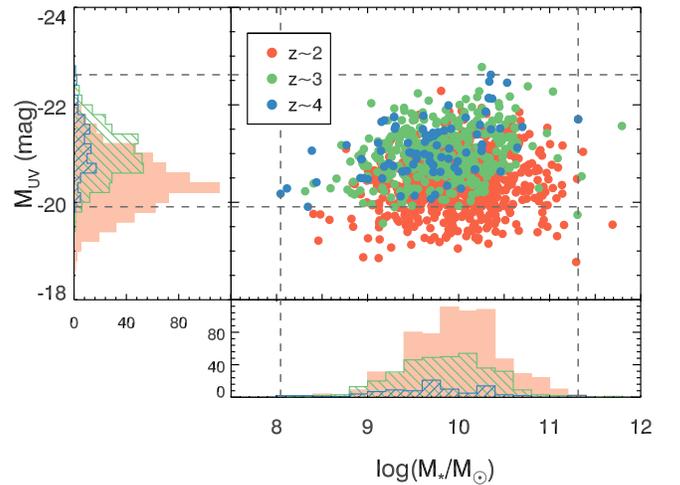


Figure 2. UV absolute magnitude vs. stellar mass for the $z \sim 2$ (red), $z \sim 3$ (green), and $z \sim 4$ (blue) samples. Histograms in M_{*} and M_{UV} are indicated, respectively, along the x- and y-axes. The vertical and horizontal gray dashed lines indicate the cuts in stellar mass and UV absolute magnitude, respectively, imposed on the $z \sim 2$ and $z \sim 3$ samples, determined by the dynamic range of these properties in the $z \sim 4$ sample.

mass that we applied to construct the final, controlled samples. In the following sections, we present results based on the analyses of the controlled samples restricted in UV absolute magnitude and stellar mass. The resulting sample sizes for the $z \sim 2, 3$, and 4 samples after being matched in UV absolute magnitude and stellar mass are 539, 309, and 91, respectively.

We list the median galaxy properties of the controlled $z \sim 2-4$ samples in Table 2. Although the final redshift samples are constrained within a range in UV luminosity and stellar mass, the median values of these properties and other stellar parameters differ slightly among the samples. The $z \sim 2$ sample median is ~ 0.5 mag fainter than that in the $z \sim 3$ and $z \sim 4$ samples, and the median stellar masses agree within a factor of 2, with $z \sim 2$ ($z \sim 4$) having the highest (lowest) mass. We note that objects falling onto the faint tail of the UV luminosity in the $z \sim 2$ sample will not affect the evolutionary trends we explore here. Specifically, we have verified that the results presented in Sections 4 and 5 do not change significantly if the $z \sim 2$ sample is further restricted to having median UV luminosity and stellar mass similar to those of the $z \sim 3$ and $z \sim 4$ samples (i.e., by removing the high-mass, faint wedge in the $z \sim 2$ sample that does not overlap with the $z \sim 3$ and $z \sim 4$ samples in Figure 2). As shown in the top-right panel of Figure 6, the strength of Ly α barely changes with M_{UV} , indicating that having a slightly larger portion of fainter objects will not affect the collective neutral ISM/CGM properties noticeably at $z \sim 2$. Similarly, given that the Ly α EW has little variation with stellar mass on the higher-mass end (middle-left panel of Figure 6), we do not consider the $z \sim 2$ sample, which has a relatively larger median stellar mass, to have a potential bias against the results presented in Sections 4 and 5. As for other galaxy properties, the lower-redshift samples on average have a higher level of dust extinction and older age. The median SFRs of the samples are comparable, with the SFR of the $z \sim 3$ sample being slightly higher than the SFRs of $z \sim 2$ and $z \sim 4$ samples.

Table 2
Median Galaxy Properties in Controlled $z \sim 2$ –4 Samples

| Sample | Redshift | M_{UV} (Magnitude) | $\log(M_*/M_\odot)$ | $E(B - V)$ | Age (Myr) | SFR ($M_\odot \text{ yr}^{-1}$) | Sample Size |
|------------|----------|-------------------------|---------------------|------------|--------------|--------------------------------------|-------------|
| $z \sim 2$ | 2.267 | −20.51 | 10.00 | 0.09 | 806 | 13 | 539 |
| $z \sim 3$ | 2.925 | −21.00 | 9.87 | 0.08 | 404 | 16 | 309 |
| $z \sim 4$ | 3.856 | −21.06 | 9.72 | 0.04 | 321 | 13 | 91 |

3. Measurements

The rest-UV spectral range covered by the $z \sim 2$ –3 LRIS spectra and the $z \sim 4$ DEIMOS and FORS2 spectra includes multiple interstellar absorption lines and fine-structure emission features in addition to Ly α . These include the low-ionization absorption lines Si II λ 1260, O I λ 1302+Si II λ 1304, C II λ 1334, and Si II λ 1527; the high-ionization features Si IV λ 1393, 1402, and C IV λ 1548, 1550; and the Si II* λ 1265, 1309, 1533 fine-structure emission lines. While Ly α and the strong absorption features are detected individually in many, but not all, cases, composite spectra enable us to study the universal correlations between the spectral and galaxy properties by utilizing all objects in the samples. Moreover, composite spectra provide significantly higher continuum signal-to-noise ratio (S/N) than individual detections, which potentially allows the measurement of weak emission and absorption lines that are not typically detected on an individual basis. In light of these advantages, we investigate the correlations among spectral and galaxy properties by studying the spectral features in the composite spectra at different redshifts. In this section, we describe our methods for creating composite spectra and measuring the rest-frame EW of both the Ly α feature and interstellar absorption lines.

3.1. Composite Spectra

To examine how Ly α correlates with other spectral lines and galaxy properties, we divided each redshift sample into four bins in Ly α EW, UV luminosity, stellar mass, $E(B - V)$, age, and SFR, with each bin containing nearly the same number of galaxies. Given the fact that galaxies in the same bin have different rest-frame spectral coverage, resulting from the difference in their redshifts and observed-frame coverage, we further required that, in every composite spectrum, the same set of objects contribute to all the wavelengths. For the measurements of Ly α , Si II λ 1260, Si II* λ 1265, O I λ 1302 + Si II λ 1304, and C II λ 1334 in the composite spectra, we required individual contributing objects to have spectral coverage from Ly α to at least 1340 Å. Similarly, we set the minimum reddest wavelength of individual contributing spectra to be 1410 Å for measuring Si IV λ 1393, 1402, and 1560 Å for measuring Si II λ 1527, Si II* λ 1533 and C IV λ 1448, 1550 from the composites. In general, the extra requirement on spectral coverage has an insignificant impact on the sample size for the measurement near Ly α (up to C II λ 1334; 0%, 6%, and 8% reduction for the $z \sim 2$, 3, and 4 samples, respectively). Objects included in the composites made for the measurement of Si IV are 99%, 60%, and 96% of the full $z \sim 2$, 3, and 4 samples, respectively, and 97%, 53%, and 81% for the spectral line measurements near C IV.

We constructed the composite spectra according to the following steps. First, we converted the individual spectra from F_ν space to L_λ space in order to capture the contribution of their intrinsic luminosity to the overall composite. The 600-line

spectra were smoothed to the resolution of the 400-line spectra for the $z \sim 2$ and $z \sim 3$ samples,¹¹ and the DEIMOS and LRIS spectra were smoothed to match the resolution of the FORS2 spectra in the $z \sim 4$ sample. Next, all of the individual spectra were interpolated onto a grid with 0.15 Å increments in wavelength and scaled to have the same median value over the wavelength range 1270–1290 Å, which itself represents the median of the median values estimated over 1270–1290 Å of all individual contributing spectra. After the scaling, the median of all input spectra was estimated at each wavelength increment to create a composite spectrum. Figure 3 shows the overall composite spectra of the $z \sim 2$, 3, and 4 samples created from the luminosity- and stellar-mass-constrained galaxies, with strong emission and absorption features marked. We describe our method of estimating the individual and composite error spectra in Appendix A.

We note that adopting a different stacking method (e.g., the sigma-clipped mean) has a negligible effect ($\lesssim 5\%$ difference) on the EW and centroid measurements of the interstellar absorption and fine-structure and nebular emission features, as the measurements were performed on the continuum-normalized spectra. As for the Ly α measurements, median stacking yields a systemically lower rest-frame Ly α EW and a more redshifted Ly α centroid. However, given that our main results about Ly α are presented in a relative sense among the redshift samples, and that the spectra were analyzed in a consistent manner, they are robust regardless of the manner in which the composite spectra are produced.

Although the establishment of the systemic redshift was not verified for individual galaxy spectra in our samples, the rest-frame composite spectra at all three redshifts indicate a close alignment with the systemic velocity. The centroid velocities of C III λ 1176, a stellar absorption feature that typically indicates the galaxy systemic velocity (i.e., $v = 0$), are measured to be $-22 \pm 20 \text{ km s}^{-1}$, $-24 \pm 27 \text{ km s}^{-1}$, and $18 \pm 123 \text{ km s}^{-1}$, respectively, at $z \sim 2$, $z \sim 3$, and $z \sim 4$ in the overall composites controlled in UV luminosity and stellar mass (Figure 3).

3.2. Ly α Measurement

To measure the Ly α rest-frame EW in both the individual and composite spectra (the former being a property for binning), we adopted the procedure described in Kornei et al. (2010). In brief, the morphology of Ly α in individual galaxy spectra is classified into four categories through visual inspection: “emission,” “absorption,” “combination,” and “noise.” The blue- and red-side continuum levels were measured over the wavelength range of 1225–1255 Å and 1120–1180 Å, respectively. The integrated Ly α flux was calculated between the blue and red wavelength “boundaries,” λ_{blue} and λ_{red} , which are defined as where the flux-density level

¹¹ Given the negligible fraction of 300-line spectra we have in our samples and the similarity of the spectral resolutions of the 400- and the 300-line spectra, we did not further smooth the spectra down to the resolution of the 300-line spectra.

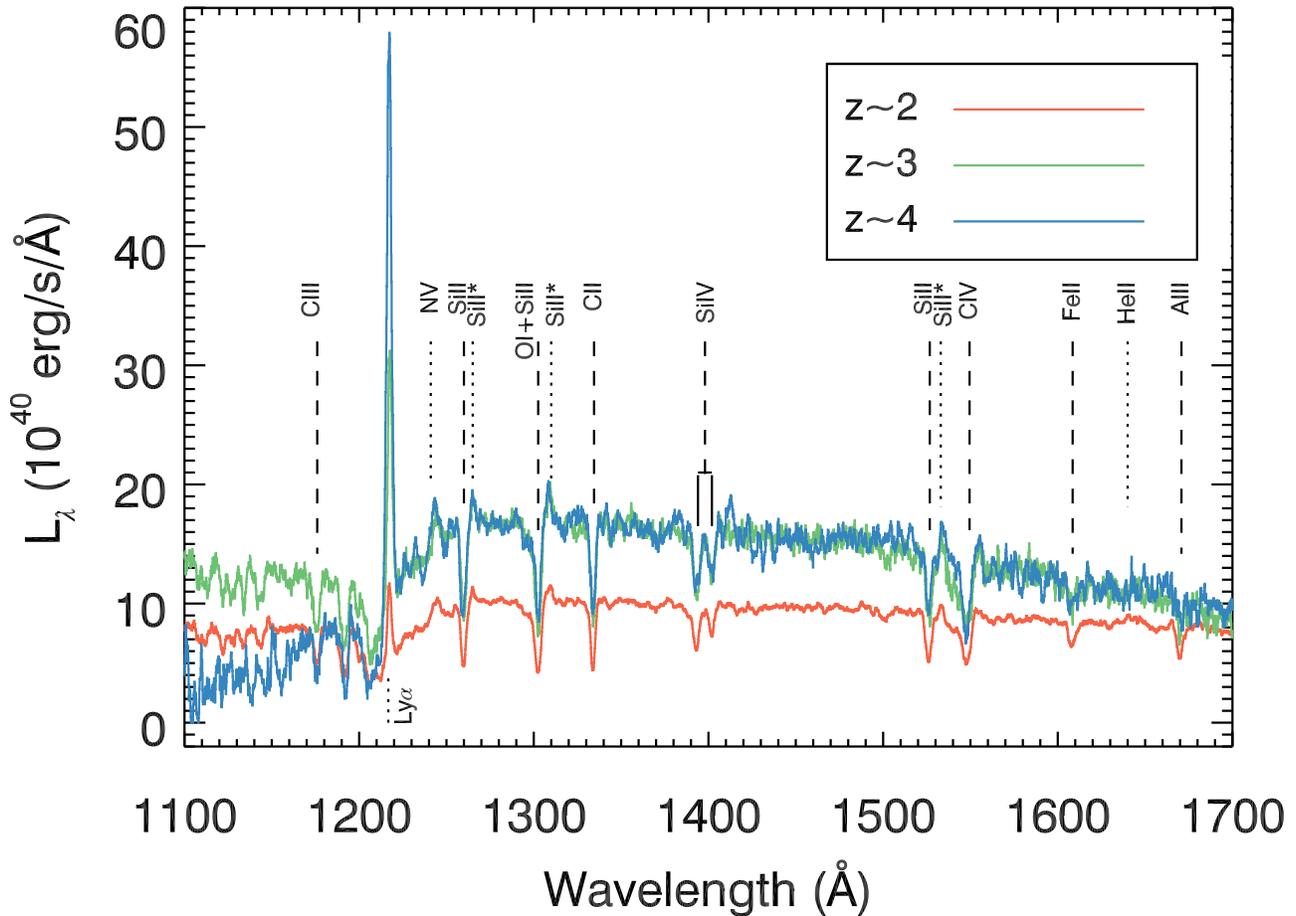


Figure 3. Composite UV spectra of the LBGs in the $z \sim 2$ (red), $z \sim 3$ (green), and $z \sim 4$ (blue) samples with UV luminosity and stellar mass constrained. The emission (dotted lines) and absorption (dashed lines) lines analyzed in this study are identified. At the resolution of the LRIS spectra ($z \sim 2$ and 3 objects) and FORS2 spectra ($z \sim 4$ objects), O I $\lambda 1302$ and Si II $\lambda 1304$ are blended, as is the C IV $\lambda \lambda 1548, 1550$ doublet.

first meets the blue- and red-side continuum, respectively, on either side of the Ly α feature. However, without additional constraints imposed on λ_{blue} , the measured Ly α EWs appear to have a bimodal distribution, with an apparent deficit near $\text{EW}_{\text{Ly}\alpha} \simeq 0$. A solution to this issue is fixing λ_{blue} at 1208 Å for the “emission” objects and requiring λ_{blue} to be no bluer than 1208 Å for the “combination” objects. The Ly α rest-frame EW was then estimated by dividing the enclosed Ly α flux by the red-side continuum level.

While the method described above worked well for measuring the Ly α EW in individual galaxy spectra at $z \sim 2$ and $z \sim 3$, it failed to handle the $z \sim 4$ spectra due to the limited spectral coverage blueward of Ly α . Five out of 91 spectra do not have Ly α coverage, and an additional 19 spectra do not have full spectral coverage of 1120–1180 Å, the default window used to determine the blue-side continuum in Kornei et al. (2010). Out of these 19 spectra, 16 have their bluest wavelength greater than 1160 Å, such that the spectral region available for estimating the blue-side continuum level is less than 20 Å, and therefore insufficient for obtaining a robust value. In order to estimate the Ly α EW for these objects, we measured the relative level of the blue- and red-side continua from objects with sufficient spectral coverage on both sides within the $z \sim 4$ sample. The median blue-to-red continuum ratio is 0.434, the value we adopted to obtain a rough estimate of the blue-side continuum for the objects without

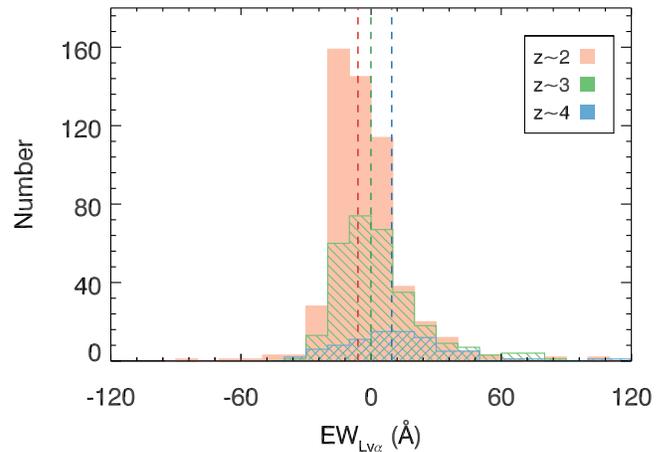


Figure 4. Ly α rest-frame EW distribution for the $z \sim 2$ (red), $z \sim 3$ (green), and $z \sim 4$ (blue) samples constrained in UV luminosity and stellar mass. The colored dashed lines mark the median Ly α EW in the corresponding samples: -6.07 Å, -0.08 Å, and 9.51 Å at $z \sim 2$, $z \sim 3$, and $z \sim 4$, respectively.

sufficient blue-side spectral coverage. The same method as described above was then applied to calculate the Ly α EW for these 16 objects. Figure 4 shows the rest-frame EW distribution of Ly α in the controlled redshift samples. While spanning the same range in UV absolute magnitude and stellar mass, objects

at higher redshifts in general have a higher Ly α EW than the lower-redshift counterparts. The median Ly α EWs in the three redshift samples are consistent with the Ly α EWs measured from the overall composites (Figure 3), which are $-4.27 \pm 0.09 \text{ \AA}$, $-0.12 \pm 0.12 \text{ \AA}$, and $8.02 \pm 1.15 \text{ \AA}$ for the $z \sim 2$, $z \sim 3$, and $z \sim 4$ samples, respectively.

The Ly α EW of the composite spectra was measured in the same manner as in the individual spectra at $z \sim 2-3$, except that the composite spectra have much higher S/N and accordingly never fall in the “noise” category. In order to obtain the uncertainty associated with the Ly α measurement in the composites, we perturbed the composite science spectra 100 times with the corresponding composite error spectra. The Ly α profiles in the 100 fake composite spectra were measured, and we adopted the sigma-clipped average and standard deviation of the measurements as the composite Ly α EW and the 1σ error bar, respectively.

3.3. Absorption Line Measurement

Many interstellar absorption lines are covered within the $z \sim 2-4$ rest-frame UV spectra described here. These include the low-ionization features of Si II $\lambda 1260$, O I $\lambda 1302 + \text{Si II } \lambda 1304$, C II $\lambda 1334$, and Si II $\lambda 1527$. As for the high-ionization lines, Si IV $\lambda\lambda 1393, 1402$ and C IV $\lambda\lambda 1548, 1550$ are covered in a majority of the spectra. Interstellar absorption lines were only measured in composite spectra, which have a fairly low spectral resolution. Hence, we adopted single-component Gaussian fits as the simplest possible functional form to describe the interstellar absorption lines in the composites. The key best-fit parameters from such fits are the line centroid and EW.

We continuum-normalized the rest-frame composite spectra using spectral regions (“windows”) that are clean of spectral features defined by Rix et al. (2004). Based on these windows, we modeled the continuum for all composite spectra with the IRAF *continuum* routine, using a *spline3* function of order = 16 in order to provide a reasonable fit near Ly α , where the continuum is fairly curved. In cases where the fitted continuum level did not provide a proper description of the observed spectrum due to the limited coverage of windows from Rix et al. (2004), additional windows customized for each object were added to provide reasonable constraints on the fit.

The absorption line profile fitting was performed on the continuum-normalized composite spectra. In general, we used the IDL program MPFIT (Markwardt 2009) with the initial values of the continuum flux level, line centroid, EW, and Gaussian FWHM estimated from the program *splot* in IRAF. The best fit was then determined where the χ^2 of the fit reached a minimum. We iterated the fitting over a narrower wavelength range for all of the interstellar absorption lines: centroid $- 4\sigma < \lambda < \text{centroid} + 4\sigma$, where the centroid and σ are, respectively, the central wavelength and standard deviation of the best-fit Gaussian profile returned by MPFIT from the initial fit to the respective lines over $\lambda_{\text{rest}} - 10 \text{ \AA}$ to $\lambda_{\text{rest}} + 10 \text{ \AA}$.

Some extra care was required for the high-ionization lines, Si IV $\lambda\lambda 1393, 1402$, and the blended C IV $\lambda\lambda 1548, 1550$ feature. As Si IV is a spectrally resolved doublet, each doublet member was individually fit with a single-component Gaussian profile, and the centroids of the doublet members were fixed at the rest-wavelength ratio in both the initial and iterated fits. As for C IV, the overall absorption profile includes a P-Cygni profile originating from the stellar wind common in the most massive stars. Without proper treatment, the presence of stellar absorption may bias the measurement of interstellar C IV.

Accordingly, we removed the stellar component following the method presented in Du et al. (2016), and the remaining interstellar absorption trough was then fit with a single-component Gaussian profile for the EW measurement.¹²

4. Line Strength

A key goal of our study is to examine the redshift evolution of the neutral and ionized ISM/CGM using luminosity- and mass-controlled samples. In particular, in this section we investigate the relations in line strength among multiple rest-frame far-UV spectral features in an evolutionary context. These include the relations between Ly α and (1) both LIS and HIS absorption features (Section 4.1), (2) multiple galaxy properties (Section 4.2), and (3) nebular C III] emission (Section 4.5). We also investigate the mutual relationships among Ly α , LIS absorption lines, and dust extinction (Section 4.3), and the fine-structure emission properties at $z \sim 2-4$ (Section 4.4).

4.1. Ly α Versus Interstellar Absorption Lines

Previous studies have shown that the strength of Ly α is tightly correlated with that of LIS absorption lines in LBGs at $z \sim 3-4$, such that stronger Ly α emission corresponds to weaker LIS absorption (e.g., Shapley et al. 2003; Quider et al. 2009; Vanzella et al. 2009; Jones et al. 2012). In contrast, no significant correlation has been found between Ly α and HIS absorption lines (e.g., Shapley et al. 2003). These observations can be explained by a physical model in which clumpy neutral clouds are embedded in a more smoothly distributed halo of ionized gas (Shapley et al. 2003). Both the EWs of LIS absorption features, which are typically saturated in the galaxy spectra in our samples, and Ly α emission are modulated by the covering fraction of neutral gas: the higher the covering fraction of the neutral ISM/CGM, the higher the fraction of Ly α photons that are resonantly scattered out of the line of sight, resulting in a weaker Ly α emission profile (e.g., Reddy et al. 2016). At the same time, an increasing covering fraction of neutral gas leads to stronger LIS absorption lines as these neutral and singly ionized metal species primarily arise in neutral gas. The lack of a direct link between HIS absorption and Ly α emission strengths is consistent with a picture in which the HIS absorption primarily arises in gas that is in a phase different from the phase responsible for the radiative transfer of Ly α . Furthermore, recent work suggests that while LIS absorption occurs spatially closer to the galaxy, HIS gas is seen with a high covering fraction out past the virial radius (Rudie 2018, in preparation). Thus, the different spatial distribution of the LIS and HIS absorbers is likely to be another factor responsible for the different behaviors of LIS and HIS with Ly α in EW.

Motivated by this physical picture, we investigated whether the correlations between Ly α and the absorption features from different ionization states evolve with time from $z \sim 4$ down to $z \sim 2$, using carefully controlled comparison samples and uniformly measured spectra. Figure 5 shows the relation between rest-frame Ly α EW ($\text{EW}_{\text{Ly}\alpha}$) and the rest-frame LIS (top) and HIS (bottom) absorption line EWs (EW_{LIS} and EW_{HIS} , respectively) at $z \sim 2-4$ from the composite spectra

¹² We did not try to model the individual C IV doublet members as done in Du et al. (2016), since in the present study we did not infer any kinematic information from the interstellar C IV profile. The single-component Gaussian profile is sufficient in characterizing the line profile for the purpose of EW measurement.

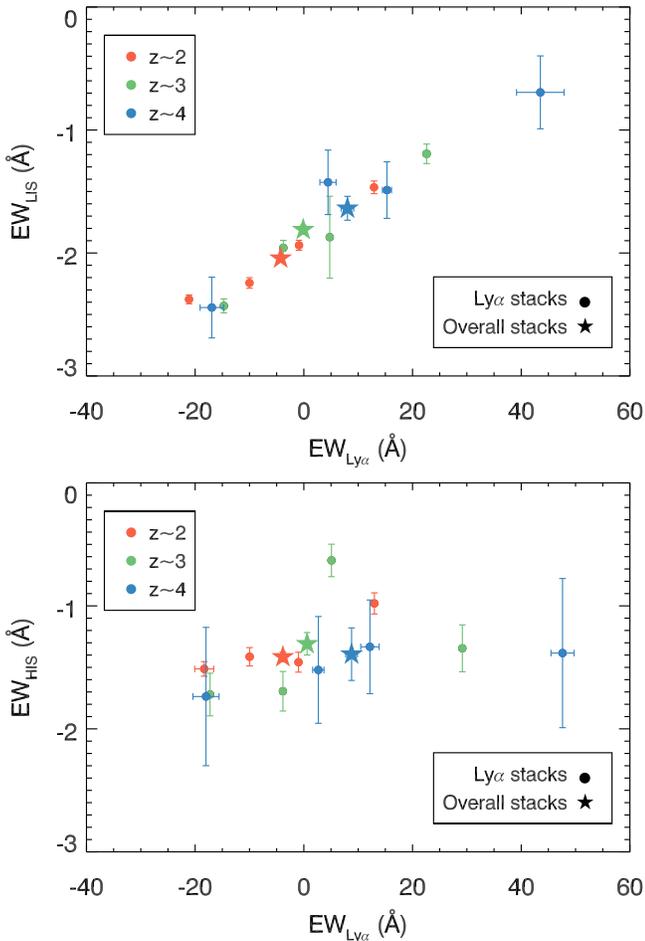


Figure 5. Top: EW_{LIS} vs. $EW_{\text{Ly}\alpha}$ in the composite spectra binned according to $EW_{\text{Ly}\alpha}$ (circles) and in the overall composite spectra (stars, as measured from the spectra shown in Figure 3). The plotted EW for LIS absorption features was calculated by averaging the EW of Si II $\lambda 1260$, O I $\lambda 1302$ + Si II $\lambda 1304$, C II $\lambda 1334$, and Si II $\lambda 1527$. Bottom: EW_{HIS} vs. $EW_{\text{Ly}\alpha}$ in the same set of composites. The color coding of the symbols is the same as in the top panel. The averaged EW of Si IV $\lambda\lambda 1393$, 1402 and C IV $\lambda\lambda 1548$, 1550 is shown.

binned according to $EW_{\text{Ly}\alpha}$. We also show the stacks of the full samples at each redshift. EW_{LIS} was calculated by averaging the EW of Si II $\lambda 1260$, O I $\lambda 1302$ + Si II $\lambda 1304$, C II $\lambda 1334$, and Si II $\lambda 1527$, and EW_{HIS} was estimated by averaging the EW of Si IV $\lambda\lambda 1393$, 1402 and interstellar C IV $\lambda\lambda 1548$, 1550.

From the EW_{LIS} versus $EW_{\text{Ly}\alpha}$ relation, we find that the $z \sim 4$ sample spans a wider dynamic range in $EW_{\text{Ly}\alpha}$, extending toward a larger emission EW than the $z \sim 2$ and $z \sim 3$ samples. In addition, the $z \sim 4$ galaxies on average show stronger Ly α emission and weaker LIS absorption in the overall stacks than the lower-redshift samples, as suggested by the measurements from the overall stacks. This result is in agreement with what Jones et al. (2012) found when comparing the strengths of Ly α and the LIS lines in the $z \sim 4$ and $z \sim 3$ LBG composites. In terms of redshift evolution, it is important to note that there is no apparent evolutionary trend associated with the EW_{LIS} versus $EW_{\text{Ly}\alpha}$ correlation (Figure 5). The absence of evolution in the trend itself suggests that Ly α and the LIS absorption features are fundamentally related, indicating that the gas giving rise to LIS absorption modulates the radiative transfer of Ly α photons. While not fixing the galaxies in UV luminosity and stellar mass would introduce a slightly

larger scatter into the EW_{LIS} versus $EW_{\text{Ly}\alpha}$ relation, this trend remains qualitatively the same in terms of the absence of redshift evolution, offering additional support to the picture of this relationship being fundamental.

As for the HIS absorption lines, we notice that they appear to be decoupled from Ly α , in that little variation in the HIS absorption strength is observed with increasing $EW_{\text{Ly}\alpha}$. This result is consistent with previous findings (e.g., Shapley et al. 2003) and provides further evidence that the HIS absorption features are produced in a different phase of gas from that controlling the radiative transfer of Ly α . In addition, we discover that the “flat” trend of HIS absorption with $EW_{\text{Ly}\alpha}$ does not seem to evolve with redshift either. Instead, the measured EW_{HIS} at all redshifts seems to scatter around a constant value of ~ -1.5 Å. We note that in contrast to the LIS absorption lines, Si IV appears to be at least partially optically thin, based on the observed $EW_{\text{Si IV } \lambda 1393} / EW_{\text{Si IV } \lambda 1402}$ doublet ratios that are greater than unity at all redshifts we study here. In the optically thin regime, the rest-frame EW of Si IV, $EW_{\text{Si IV}}$, does not depend on only the ionized gas covering fraction and velocity range, but is also determined by the column density. In the overall composite spectra, the Si IV doublet ratio decreases from 2.13 ± 0.13 Å at $z \sim 2$ to 1.67 ± 0.18 Å and 1.70 ± 0.55 Å at $z \sim 3$ and $z \sim 4$, respectively. The decreasing doublet ratio suggests that a larger fraction of gas is optically thick at $z \sim 3$ –4 than at $z \sim 2$, and that the column density of Si IV increases with increasing redshifts. In the meantime, the total $EW_{\text{Si IV}}$ and the FWHM of individual doublet members remain almost constant, suggesting that the covering fraction of the Si IV gas may decrease at higher redshifts. We will expand the discussion of the covering fraction of the ionized gas and relate it to a physical picture for the evolving ISM/CGM in Section 6.2.

4.2. Ly α Versus Galaxy Properties

Extensive studies have been carried out to examine the connection between Ly α emission and galaxy properties in an effort to understand the factors modulating the escape fraction of Ly α photons. Among all of the relations, the one for which there is the most consensus is between Ly α and the rest-frame UV color (alternatively, dust reddening, i.e., $E(B - V)$), according to which galaxies with stronger Ly α emission have bluer UV continua (e.g., Shapley et al. 2003; Vanzella et al. 2009; Kornei et al. 2010; Pentericci et al. 2010; Berry et al. 2012; Jones et al. 2012; Erb et al. 2016). In contrast, controversies remain as to how $EW_{\text{Ly}\alpha}$ relates to galaxy ages. While some studies claim that stronger Ly α is typically found in older galaxies (e.g., Shapley et al. 2001; Kornei et al. 2010), others found no apparent relation between $EW_{\text{Ly}\alpha}$ and age (e.g., Pentericci et al. 2009), or that galaxies with stronger Ly α emission are much younger on average (e.g., Pentericci et al. 2007). In terms of the trends with UV luminosity, stellar mass, and SFR, studies have shown that stronger Ly α emission is typically associated with fainter (e.g., Vanzella et al. 2009; Stark et al. 2010; Jones et al. 2012), lower-mass galaxies (e.g., Erb et al. 2006b; Pentericci et al. 2007, 2009; Jones et al. 2012) with lower SFRs (e.g., Shapley et al. 2003; Kornei et al. 2010; Hathi et al. 2016).

In most cases, the relations described above have been investigated at individual redshifts. Now, we examine whether these relations ($EW_{\text{Ly}\alpha}$ versus UV luminosity, stellar mass, dust extinction, SFR, and galaxy age) exhibit any evolutionary trends between $z \sim 2$ –4. We plot the relations between $EW_{\text{Ly}\alpha}$ and

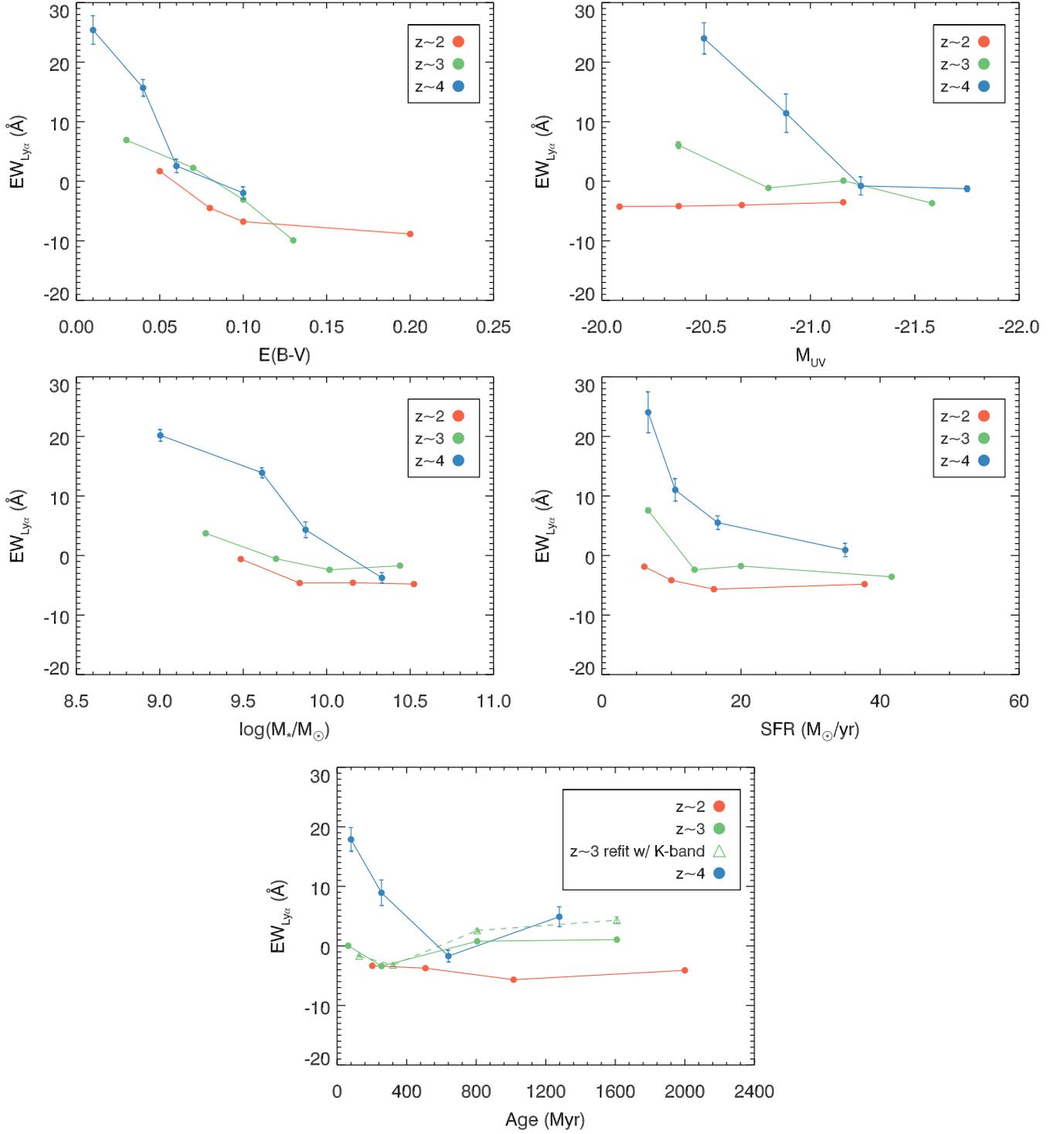


Figure 6. Top to bottom, left to right: $EW_{Ly\alpha}$ vs. UV absolute magnitude, stellar mass, dust extinction, star formation rate, and galaxy age in the composite spectra. The triangles and dashed lines in the bottom panel represent the measurements from fitting the $z \sim 3$ galaxy SED with the K -band photometry included.

$E(B - V)$, M_{UV} , M_* , SFR, and age in Figure 6. The composite spectra were created by binning individual spectra according to the respective galaxy properties, and the median galaxy property in each bin is plotted. The choice of SED models is not the dominant factor for the relations between $EW_{Ly\alpha}$ and galaxy properties at $z \sim 2-4$ (i.e., the results will remain qualitatively the same with properties derived from the $1.4 Z_\odot + \text{Calzetti}$ model for all redshift samples) except for the $EW_{Ly\alpha}$ versus age relation at $z \sim 3$, as discussed below.

Among all of the Ly α relations with galaxy properties, the most striking one is between Ly α and $E(B - V)$. In addition to the previously observed negative correlation between $EW_{Ly\alpha}$ and dust extinction at individual redshifts, we find that the $EW_{Ly\alpha}$ versus $E(B - V)$ relation lacks significant evolution from $z \sim 4$ to $z \sim 2$. This result implies that dust affects the Ly α photons in the same manner in galaxies at all redshifts studied here, and that the connection between Ly α emission and dust extinction is more direct than those observed between

$\text{Ly}\alpha$ emission and the other galaxy properties described in this subsection. Shapley et al. (2003) proposed that the patchy neutral gas lying within the more extended ionized gas halo could also be dusty, and the $\text{Ly}\alpha$ photons would be absorbed if they encounter the dust grains. We revisit this physical picture in Section 6.2.2.

For the relation between $\text{Ly}\alpha$ and UV absolute magnitude, we note that only the $z \sim 4$ galaxies display noticeably stronger $\text{Ly}\alpha$ in fainter galaxies. The same relation at $z \sim 3$ is not as strong, and at $z \sim 2$ the UV absolute magnitude has a flat correlation with $\text{EW}_{\text{Ly}\alpha}$. Furthermore, the strength of $\text{Ly}\alpha$ increases with increasing redshift at fixed UV luminosity, likely as a result of a lower covering fraction of both neutral gas and dust in the galaxies at fixed UV luminosity. This interpretation is consistent with the manner in which $E(B - V)$ evolves from $z \sim 2$ to $z \sim 4$.

Similar evolutionary trends can also be observed from the relations of $\text{Ly}\alpha$ with stellar mass and SFR: in general, galaxies with lower stellar mass and lower SFR have stronger $\text{Ly}\alpha$ emission, but these relations become weaker as redshift decreases. A higher $\text{EW}_{\text{Ly}\alpha}$ is also generally found in higher-redshift samples at fixed stellar mass and/or SFR. We note that the correlations among $\text{Ly}\alpha$ and UV luminosity, stellar mass, and SFR are not independent, since galaxies with larger stellar masses are likely to have higher UV luminosities and higher SFRs.

The relation between $\text{Ly}\alpha$ and galaxy age has also been explored previously. Without excluding the K -band photometry (where both $\text{H}\beta$ and $[\text{O III}]$ fall for the $z \sim 3$ LBGs) from the SED fits, Shapley et al. (2001) and Kornei et al. (2010) found that $\text{Ly}\alpha$ is stronger in older galaxies. These authors interpreted the results as an evolution in time toward higher $\text{EW}_{\text{Ly}\alpha}$ as the neutral gas and dust are blown out by supernovae and massive stellar winds, resulting in a reduced covering fraction of both at older galaxy ages. We revisited this correlation for our three redshift samples and found that, after excluding the photometric bands potentially contaminated by strong emission lines ($\text{H}\beta$, $[\text{O III}]$, and $\text{H}\alpha$), the $z \sim 2$ and $z \sim 3$ galaxies do not exhibit a significant positive trend between $\text{EW}_{\text{Ly}\alpha}$ and galaxy age. As for the $z \sim 4$ galaxies, $\text{Ly}\alpha$ emission is more prominent in younger galaxies than in the older ones, an opposite trend from that found by Shapley et al. (2001) and Kornei et al. (2010).

We further explored the connection between $\text{EW}_{\text{Ly}\alpha}$ and age at $z \sim 3$. The positive correlation of $\text{Ly}\alpha$ and age can be reproduced if we do not discard the contaminated K band from SED fitting at $z \sim 3$ (triangles in the bottom panel of Figure 6). It is worth noting, though, that when using the $1.4 Z_{\odot} + \text{Calzetti}$ model, which was adopted both in Shapley et al. (2001) and Kornei et al. (2010), the derived galaxy age $z \sim 3$ does reveal a positive correlation with $\text{EW}_{\text{Ly}\alpha}$. The exclusion of the K band, while making the relation flatter, cannot completely remove the positive trend. On the other hand, the $\text{EW}_{\text{Ly}\alpha}$ versus age relation at $z \sim 2$ and $z \sim 4$ is not sensitive to the choice of SED models. Although it is unclear why the adoption of the $1.4 Z_{\odot} + \text{Calzetti}$ model only affects the $\text{EW}_{\text{Ly}\alpha}$ relation with age specifically at $z \sim 3$, galaxy age is undoubtedly the least well-determined property among those we explore here. Therefore, we conclude with caution that, with the most reasonable assumptions of our $z \sim 3$ sample ($0.28 Z_{\odot} + \text{SMC}$ for the majority and $1.4 Z_{\odot} + \text{Calzetti}$ for very high-mass galaxies), there is no evidence for a positive correlation

between $\text{EW}_{\text{Ly}\alpha}$ and age at $z \sim 3$. The previously reported positive $\text{EW}_{\text{Ly}\alpha}$ trend with age at $z \sim 3$ is best explained as a combination of the choice of SED models and the contamination in the K band from strong nebular $[\text{O III}]$ emission.

As described in Section 2.2, excluding the H and K bands for the $z \sim 2$ sample and IRAC channel 1 for the $z \sim 4$ sample has little impact on the age estimates. However, there are plausible factors to explain the fact that strong emission lines affect the SED fits at $z \sim 3$. For the $z \sim 4$ sample, only objects with $z \geq 3.8$ (38 out of 70 objects) have SEDs that are potentially affected by emission-line contamination (i.e., $\text{H}\alpha$) in IRAC channel 1 while the remainder will not suffer from significant emission-line contribution to the SED. As for the $z \sim 2$ sample, the best-fit SED models are less sensitive to the emission lines because (1) 75 objects (14% of the sample) already had the H and K bands corrected for strong nebular emission, and (2) a larger fraction of objects in the $z \sim 2$ sample (474 out of 539 galaxies) than in the $z \sim 3$ sample (273 out of 349 galaxies) have IRAC coverage to constrain the overall shape of the SEDs. Therefore, the relative importance of the H and K bands in determining the best-fit stellar population parameters is lower in the $z \sim 2$ sample than at $z \sim 3$.

In summary, we do not observe an apparent correlation between $\text{EW}_{\text{Ly}\alpha}$ and galaxy age within the $z \sim 2$ and $z \sim 3$ samples, and there is no strong evidence for a physical picture in which gas and dust are systematically cleared away in galaxies as a function of time (e.g., as described in Kornei et al. 2010). It is suggestive, though, that at $z \sim 4$ $\text{Ly}\alpha$ is stronger in younger galaxies. At fixed galaxy age, the $z \sim 4$ galaxies have stronger $\text{Ly}\alpha$ emission than galaxies at $z \sim 2-3$. Furthermore, there appears to be a tendency toward smaller $\text{EW}_{\text{Ly}\alpha}$ (and thus increased dust content and HI covering fraction) at older ages within the $z \sim 4$ sample. In contrast to the physical picture proposed by Shapley et al. (2001) and Kornei et al. (2010), the trend we observe at $z \sim 4$ suggests that galaxies at the youngest ages in our sample (< 200 Myr) have a less substantial neutral ISM/CGM and are possibly experiencing the first generations of star formation. At the same time, comparably young $z \sim 2-3$ galaxies may be more chemically enriched than those at $z \sim 4$, and the relatively higher neutral gas and dust covering fractions have reduced the observed $\text{EW}_{\text{Ly}\alpha}$.

In this section, we found that galaxies with stronger $\text{Ly}\alpha$ emission tend to be fainter, younger, lower-mass galaxies with lower SFRs. The correlations between $\text{Ly}\alpha$ and these galaxy properties are the strongest at $z \sim 4$ and become much weaker at $z \sim 2-3$. The evolutionary change of the $\text{Ly}\alpha$ dependence on galaxy properties has also been reported in previous studies (e.g., Pentericci et al. 2010). However, before making any physical interpretation of the observed evolution, one needs to consider the larger spectroscopic incompleteness of the $z \sim 4$ sample relative to that of the $z \sim 2$ and $z \sim 3$ samples especially at the faint, low-mass end (i.e., the $z \sim 4$ galaxies have to show a stronger $\text{Ly}\alpha$ in order to be detected and included in the sample at faint luminosities).

To investigate how the sample incompleteness impacts our results, we conducted two tests. First, we constructed composites from a “complete” subsample at each redshift. We selected galaxies with $z' \leq 24.75$ at $z \sim 4$, where the spectroscopic completeness is nearly 100%, to be in the “complete” subsample. Galaxies in this subsample have UV absolute magnitudes extending down to $M_{\text{UV}} = -20.78$. We

then constructed subsamples at $z \sim 2$ and $z \sim 3$ with galaxies brighter than $M_{UV} \leq -20.78$, which are also spectroscopically complete (Shapley et al. 2003; Steidel et al. 2003). The progressive evolution of $EW_{Ly\alpha}$, being stronger at higher redshifts, still exists in the “complete” composites, with $EW_{Ly\alpha} = -4.0 \text{ \AA}$, -2.2 \AA , and 6.5 \AA at $z \sim 2$, $z \sim 3$, and $z \sim 4$, respectively. In our second test, we limited the $z \sim 2$ and $z \sim 3$ samples in each galaxy property bin such that they had the same fraction of galaxies with only Ly α redshift as the $z \sim 4$ sample and examined how the trends in Figure 6 would be affected. We note that intentionally selecting objects with “ $z_{Ly\alpha}$ only” in the lower-redshift samples would bias the samples to objects with stronger Ly α . As expected, the resulting $EW_{Ly\alpha}$ in the censored $z \sim 2$ and $z \sim 3$ composites is systemically larger, but the trends at $z \sim 4$ are still steeper than those at lower redshifts across the same dynamic range in galaxy properties. Additionally, we measure larger $EW_{Ly\alpha}$ at $z \sim 4$ in the overall composite than in the censored $z \sim 2$ and $z \sim 3$ stacks. Therefore, we conclude that while the difference in sample incompleteness needs to be seriously taken into account when studying galaxies at different redshifts, our key results are robust regarding the redshift evolution of Ly α emission and LIS absorption strengths, and $E(B - V)$.

The steeper Ly α trends with galaxy properties at $z \sim 4$ can be attributed to a relatively larger range in H I and dust covering fractions, which results in a larger dynamic range in $EW_{Ly\alpha}$ (see Figure 5) especially toward the stronger-Ly α end where galaxies have very little dust. In contrast, the $z \sim 2-3$ galaxies have higher $E(B - V)$ than the $z \sim 4$ galaxies even in the faintest, youngest, lowest-mass, and lowest-SFR bin. The higher minimum $E(B - V)$ limits the highest $EW_{Ly\alpha}$ we could possibly observe at $z \sim 2-3$, resulting in a much smaller range in the measured $EW_{Ly\alpha}$ in composites binned according to galaxy properties.

4.3. Ly α , LIS Absorption Features, and Dust

In Sections 4.1 and 4.2, we presented our result of the seemingly invariant relations between $EW_{Ly\alpha}$ and EW_{LIS} (Figure 5) and between $EW_{Ly\alpha}$ and dust extinction (Figure 6). These non-evolving correlations indicate an interdependence among Ly α emission, LIS absorption, and dust extinction, independent of redshift. If these three properties are inherently interrelated, we would expect strong correlations among any two of them regardless of galaxy properties. As a result, we utilize the composite spectra binned according to different galaxy properties (no longer restricted to the Ly α stacks as shown in Section 4.1) and further examine the relations among Ly α , the LIS lines, and the dust content. Figure 7 shows the measurements of the strength of Ly α emission, LIS absorption, and $E(B - V)$ plotted against each other in composites divided according to M_{UV} , M_* , $E(B - V)$, SFR, age, and $EW_{Ly\alpha}$. While the measurements from these composite spectra are not entirely independent in the way that measurements from individual galaxies would be, the relation between the most fundamentally correlated parameters should show the least scatter in the composites binned in galaxy properties. With high S/N and the inclusion of the entire samples, composite spectra provide a comprehensive view of the relative tightness and redshift dependence, if any, of the correlations among $EW_{Ly\alpha}$, EW_{LIS} , and $E(B - V)$. Therefore, although we recognize that the correlation significances associated with these composite scatter plots do not hold the

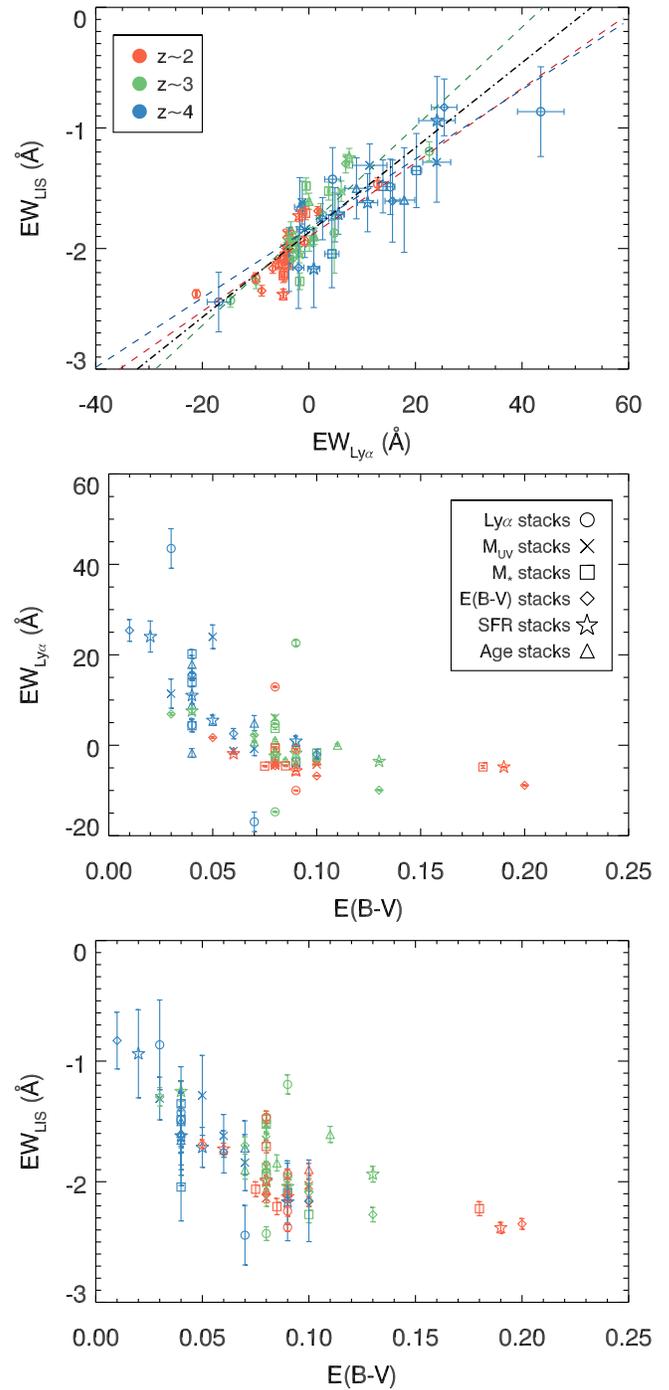


Figure 7. Top to bottom: $EW_{Ly\alpha}$ vs. EW_{LIS} , $EW_{Ly\alpha}$ vs. $E(B - V)$, and EW_{LIS} vs. $E(B - V)$ in composite spectra divided according to $EW_{Ly\alpha}$ (circles), UV absolute magnitude (crosses), stellar mass (squares), dust extinction (diamonds), star formation rate (stars), and galaxy age (triangles). The median $E(B - V)$ in each bin is plotted. The color coding of the symbols is the same as in Figure 5. In the top panel, the colored dashed lines denote the best-fit linear model for the $z \sim 2$ (red), $z \sim 3$ (green), and $z \sim 4$ (blue) samples, and the black dashed-dotted line represents the overall best-fit linear model for the $EW_{Ly\alpha}$ vs. EW_{LIS} relation accounting for all three redshift samples.

same meaning as for independent data points, we still use the Spearman correlation coefficient and significance as gauges of the relative tightness of these correlations.

To quantify the EW_{LIS} versus $EW_{Ly\alpha}$ relation, we performed an inverse-squared weighted linear regression to the data points

for each redshift sample. The best-fit linear models are parameterized as $EW_{\text{LIS}} = -1.90 + 0.031 \times EW_{\text{Ly}\alpha}$, $EW_{\text{LIS}} = -1.81 + 0.041 \times EW_{\text{Ly}\alpha}$, and $EW_{\text{LIS}} = -1.83 + 0.029 \times EW_{\text{Ly}\alpha}$ at $z \sim 2$, 3, and 4, respectively. As shown in the top panel of Figure 7, the best-fit linear model for each redshift sample predicts very similar values over the ranges of $EW_{\text{Ly}\alpha}$ and EW_{LIS} probed in our study, and shows no progressive, evolutionary trend at $z \sim 2-4$. The overall fit to all of the $z \sim 2-4$ data points yields $EW_{\text{LIS}} = -1.87 + 0.035 \times EW_{\text{Ly}\alpha}$, with a Spearman correlation coefficient $\rho = 0.87$ (7.4σ away from the null hypothesis).

Similar to Figure 6, the middle panel of Figure 7 shows that $EW_{\text{Ly}\alpha}$ decreases with increasing $E(B - V)$ in all composites binned according to galaxy properties. This correlation is less tight than the $EW_{\text{Ly}\alpha}$ versus EW_{LIS} relation, and has a Spearman correlation coefficient $\rho = -0.71$ (5.9σ away from the null hypothesis). Given the common dependence of $EW_{\text{Ly}\alpha}$ on both EW_{LIS} and dust extinction, a correlation between EW_{LIS} and $E(B - V)$ is also expected. The tightness of the EW_{LIS} versus $E(B - V)$ relation is comparable to that of the $EW_{\text{Ly}\alpha}$ versus $E(B - V)$ relation, with a Spearman $\rho = -0.70$ (5.9σ away from the null hypothesis). A comparison of the correlation coefficients and significances suggests that the dependence between $\text{Ly}\alpha$ and the LIS absorption lines is the strongest, whereas their correlation with $E(B - V)$ may be secondary.

While the grid of discrete $E(B - V)$ values in the stellar population templates may potentially introduce larger scatter to the relations with $E(B - V)$, we argue that this discretization has a negligible impact on the resulting correlation coefficient and significance. To mimic the discrete $E(B - V)$ values and match the number of $E(B - V)$ bins, we discretized $EW_{\text{Ly}\alpha}$ into 21 values from -20 \AA to 40 \AA with a 3 \AA increment. Each $EW_{\text{Ly}\alpha}$ was assigned to the closest fixed value of the grid. Discretizing $EW_{\text{Ly}\alpha}$ results in a Spearman $\rho = 0.86$ (7.2σ away from the null hypothesis), which is almost unchanged relative to the original correlation strength. We therefore conclude, as in Shapley et al. (2003), that the correlation between $\text{Ly}\alpha$ and the LIS lines is likely the primary one among the three. We note, however, that the larger scatter in the EW_{LIS} and $E(B - V)$ and $EW_{\text{Ly}\alpha}$ versus $E(B - V)$ relations may partially result from the uncertainty in $E(B - V)$ due to photometric errors and systematic uncertainty from the assumption of stellar population models and attenuation curves, and may not reflect the intrinsic scatter with $EW_{\text{Ly}\alpha}$ and EW_{LIS} . Individual measurements of $EW_{\text{Ly}\alpha}$, EW_{LIS} , and $E(B - V)$ from deep spectra will provide additional confirmation on which relationship is most fundamental.

These relations, which appear to be roughly independent of redshift and other galaxy properties, have significant implications for understanding galaxy evolution. For example, the strength of both $\text{Ly}\alpha$ emission and LIS absorption can be predicted based on the measurement of the other, and the evolutionary trends shown in the $EW_{\text{Ly}\alpha}$ versus galaxy property relations (Figure 6) can be directly linked to the evolution of the range of LIS absorption and thus the neutral gas covering fraction at different redshifts.

4.4. Fine-structure Emission

Aside from $\text{Ly}\alpha$ and LIS absorption lines, multiple fine-structure emission features, $\text{Si II}^* \lambda 1265$, 1309 , 1533 , are also detected in the LBG composite spectra at $z \sim 2-4$. Fine-structure emission comes from the re-emission of photons to the excited ground state following the absorption of continuum photons, as traced by the resonant interstellar absorption lines. Accordingly, the optically thin fine-structure emission lines are considered to

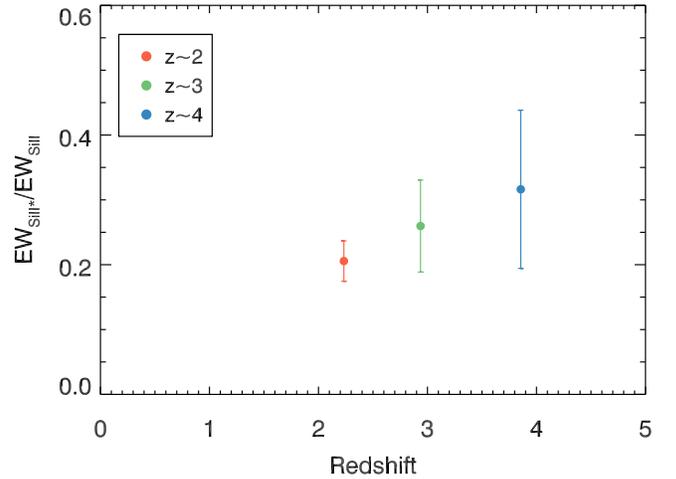


Figure 8. The fine-structure-to-LIS-absorption ratio vs. redshift in the composite spectra with fixed median $E(B - V) = 0.05$. The ratio is calculated by dividing the total EW of $\text{Si II}^* \lambda 1265 + \text{Si II}^* \lambda 1533$ by that of $\text{Si II} \lambda 1260 + \text{Si II} \lambda 1527$.

originate from the outflowing ISM/CGM (e.g., Erb et al. 2012; Kornei et al. 2013) and offer rich information on the spatial extent and geometry of galactic-scale outflows (e.g., Prochaska et al. 2011; Finley et al. 2017). In the absence of dust, the total strength of the fine-structure emission is expected to be equal to that of the associated interstellar absorption for optically thick gas probed by the saturated LIS lines, as the resonant photons cannot escape until they emerge as fine-structure photons. However, observations have shown that the fine-structure emission always appears to be weaker than the LIS absorption for the Si II (Shapley et al. 2003; Jones et al. 2012) and Fe II transitions (Erb et al. 2012). Several factors may contribute to this discrepancy: the presence of dust, geometry of the outflows, and spectroscopic slit loss (Prochaska et al. 2011; Erb et al. 2012). By comparing the EW ratio of the Si II^* fine-structure emission and Si II absorption (i.e., $R_{\text{FS}} = EW_{\text{Si II}^*}/EW_{\text{Si II}}$) in the overall LBG composites at $z \sim 3$ and $z \sim 4$, Jones et al. (2012) found that this ratio is higher for the $z \sim 4$ galaxies than for those at $z \sim 3$. These authors interpreted such results as evidence of a smaller characteristic size for the fine-structure emitting regions in $z \sim 4$ LBGs, such that a larger fraction of emission gets captured in the slit at $z \sim 4$ than at $z \sim 3$.

In order to investigate the effect of slit loss alone and whether the size of the fine-structure emission region evolves with redshift, we utilized the full $z \sim 4$ sample and constructed composites at $z \sim 2$ and $z \sim 3$ with a fixed median $E(B - V) = 0.05$ matched with that of the $z \sim 4$ sample. We only sampled galaxies on the blue tail of the $E(B - V)$ distribution of the $z \sim 2$ and $z \sim 3$ samples, selecting objects from smaller to larger $E(B - V)$ until the median $E(B - V)$ is matched. This approach results in 179, 129, and 84 spectra in the $E(B - V)$ -controlled subsamples at $z \sim 2$, 3, and 4, respectively. We measured R_{FS} using $\text{Si II}^* \lambda 1265, 1533$ and $\text{Si II} \lambda 1260, 1527$. The pair of $\text{Si II}^* \lambda 1309$ and $\text{Si II} \lambda 1304$ was excluded from the measurement because the latter is blended with $\text{O I} \lambda 1302$ in the composite spectra. We define $R_{\text{FS}} = -(EW_{1265} + EW_{1533})/(EW_{1260} + EW_{1527})$, as in Jones et al. (2012). We measure a ratio of 0.21 ± 0.03 , 0.26 ± 0.07 , and 0.32 ± 0.12 for the $z \sim 2$, $z \sim 3$, and $z \sim 4$ composites, respectively. Figure 8 shows R_{FS} plotted against redshift and does not indicate a significant evolutionary trend. Given that the composites are fixed in $E(B - V)$, our result suggests that there may not be a redshift evolution in the size of the fine-structure emission region in galaxies with similar dust extinction. In fact,

if the median $E(B - V)$ is not fixed, R_{FS} is measured to be 0.16 ± 0.02 and 0.18 ± 0.04 for the $z \sim 2$ and $z \sim 3$ samples, respectively (R_{FS} at $z \sim 4$ remains the same), from the overall composites. This trend is consistent with that reported by Jones et al. (2012), that galaxies at higher redshifts on average have stronger fine-structure emission at a given resonant absorption strength. However, our results suggest that the larger R_{FS} in the $z \sim 4$ galaxies measured by Jones et al. (2012) reflects the lower dust content in higher-redshift galaxies on average (Table 2 and Figure 6).

We note that while we reported $R_{\text{FS}} = 0.32 \pm 0.12$ for the $z \sim 4$ composite, Jones et al. (2012) measured a higher value, $R_{\text{FS}} = 0.53 \pm 0.17$, using the $z \sim 4$ LBG spectra from the same set of Si II and Si II* lines. Several factors may contribute to this discrepancy. First, the composites were constructed from different samples. By adding 14 objects with $z > 3.4$ from the LRIS sample, removing one galaxy from the DEIMOS/FORS2 sample with $z < 3.4$, and excluding three galaxies without a unique match in photometry, our $z \sim 4$ sample includes 81 objects (91 spectra)¹³ while that in Jones et al. (2012) includes 70 objects (81 spectra). Second, the individual galaxy spectra were combined in different ways. In Jones et al. (2012), all of the individual spectra were normalized to $F_{\nu} = 1$ over 1250–1500 Å, and the mean value at each wavelength was taken with 1σ clipping. In comparison, we constructed the composites accounting for the intrinsic shape of the galaxy spectra in the L_{λ} space. We scaled all galaxy spectra in the same bin such that they have the same median L_{λ} over 1270–1290 Å, a spectral window that all individual spectra have coverage for, and performed median stacking to create the composite spectra. Additionally, we required that the same set of objects contribute to all wavelengths in a single composite. For example, to measure the features of Si II $\lambda 1527$ and Si II* $\lambda 1533$, we only included objects with spectral coverage up to at least 1540 Å when creating the composite. While the requirement on wavelength coverage reduced the $z \sim 2, 3,$ and 4 sample sizes to 173, 71, and 74, respectively, the composites represent the average of the same set of objects at all wavelengths, and the S/N of the composite does not vary significantly across the spectrum due to the different numbers of objects contributing at each wavelength.

In conclusion, we created the $z \sim 4$ composite using a different method with a slightly different sample from the one in Jones et al. (2012), and a different R_{FS} value is therefore not unexpected. In fact, the R_{FS} value quoted in Jones et al. (2012) can be reproduced if we use the same sample to create the composite following the same stacking approach described in Jones et al. (2012). Since we have performed uniform stacking schemes and spectral feature measurements across the redshift samples, the results presented earlier in this subsection are free from any systematic bias and reflect the true (lack of) redshift evolution of the size of the fine-structure emitting region with fixed dust extinction. We note, however, that although the redshift rules used to determine the systemic redshift (Section 2.1) yield a well-established rest frame for the composite spectra, as evidenced by the lack of significant

velocity offset for stellar features, the average velocity shift may not be accurate in determining the redshift of each individual galaxy. As a result, redshift errors for individual objects can potentially broaden the fine-structure emission profiles in the composites, as these features were not included when determining the redshift rules. Some fine-structure emission flux can therefore be lost during our measurements of these features due to the fixed wavelength range we used for fitting the profiles, leading to smaller measured fine-structure emission EWs than the intrinsic values. Individual spectra with high S/N and spectral resolution and precise redshift measurements will provide key information on the intrinsic R_{FS} of star-forming galaxies at $z \sim 2-4$.

4.5. Ly α versus C III] at $z \sim 2$

Typically being the second strongest emission feature in the rest-frame far-UV (after Ly α), the nebular C III] $\lambda\lambda 1907, 1909$ emission doublet has received great attention as a potential redshift probe at $z > 6$ (Stark et al. 2014, 2015), where Ly α is significantly attenuated by the neutral intergalactic medium (IGM; e.g., Treu et al. 2012; Konno et al. 2014; Pentericci et al. 2014; Schenker et al. 2014). More importantly, C III] is a useful probe of the physical conditions of the H II regions in star-forming galaxies when combined with other nebular lines such as [O III] and C IV. Photoionization models predict that large rest-frame C III] EWs, $\text{EW}_{\text{C III]}}$, result from lower gas-phase metallicity, higher ionization parameters, and harder radiation fields (e.g., Erb et al. 2010; Berg et al. 2016; Steidel et al. 2016; Senchyna et al. 2017). The common dependence of C III] and other emission-line properties (both nebular and recombination features) on the hardness of the ionizing spectrum leads to positive predicted correlations of $\text{EW}_{\text{C III]}}$ with $\text{EW}_{\text{Ly}\alpha}$ (Jaskot & Ravindranath 2016), [O III] (Stark et al. 2014; Jaskot & Ravindranath 2016; Maseda et al. 2017; Senchyna et al. 2017), and the escape fraction of Lyman-continuum (LyC) photons (de Barros et al. 2016; Jaskot & Ravindranath 2016; Vanzella et al. 2016). Previous works have suggested a connection between $\text{EW}_{\text{Ly}\alpha}$ and $\text{EW}_{\text{C III]}}$ (e.g., Shapley et al. 2003; Stark et al. 2014, 2015; Rigby et al. 2015; Guaita et al. 2017). The evidence for a positive correlation between Ly α and C III] emission EWs suggests that the observed $\text{EW}_{\text{Ly}\alpha}$, aside from being modulated by the covering fraction of neutral gas and dust, may also show differences in intrinsic Ly α production. For example, the $z \sim 4$ composite made of galaxies in the quartile with the strongest Ly α emission (the rightmost blue circle in the top panel of Figure 7) falls off the mean EW_{LIS} versus $\text{EW}_{\text{Ly}\alpha}$ relation, showing stronger Ly α emission at fixed LIS absorption than all other composites. This result suggests variations in the properties of H II regions and massive stars even among galaxies in our samples. Since C III] is covered in some of the individual $z \sim 2-4$ LBG spectra, it is of great interest to examine how the strength of C III] in general relates to that of Ly α , and whether such correlation evolves with time using the large data sets at hand.

We selected a subset of objects with C III] coverage from each redshift sample controlled in UV luminosity and stellar mass. The selection results in 291, 37, and 34 objects in the $z \sim 2, z \sim 3,$ and $z \sim 4$ samples, respectively. Given the small number of objects with C III] coverage at $z \sim 3$ and $z \sim 4$, we performed binary stacks at each redshift according to $\text{EW}_{\text{Ly}\alpha}$, with the low and high bins containing nearly equal numbers of objects. We measured the C III] emission profile on the

¹³ As we discuss below, the constraint on wavelength coverage results in 66 objects (74 spectra) contributing to the $z \sim 4$ composite from which we measured R_{FS} . Among these individual spectra, 60 are in the DEIMOS/FORS2 sample Jones et al. (2012) used to measure R_{FS} , and the remaining 14 spectra are from LRIS.

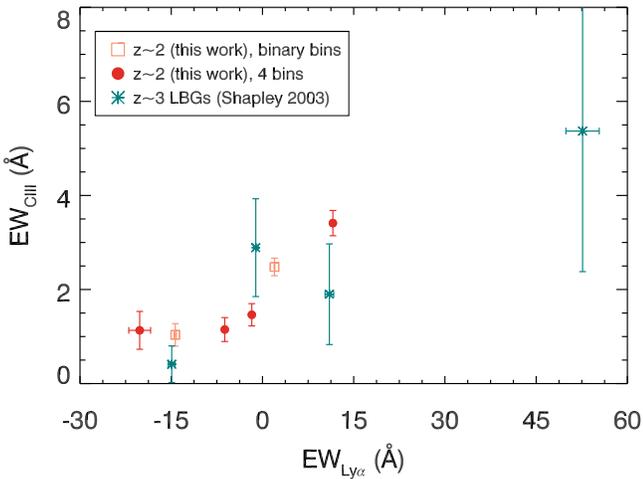


Figure 9. Rest-frame $EW_{C\text{ III]}}$ vs. $EW_{Ly\alpha}$ in the $z \sim 2$ composite spectra with C III] coverage divided into two bins (light red squares) and four bins (dark red circles). The C III] emission is not significant in any of the $z \sim 3$ and $z \sim 4$ composites and is therefore not shown in the figure.

continuum-normalized spectra in the same manner as described in Section 3.3.

Unfortunately, the C III] feature is not significantly detected in any of the $z \sim 3$ and $z \sim 4$ composites due to relatively low S/N, making it impossible to study the redshift evolution of the relation between $EW_{Ly\alpha}$ and $EW_{C\text{ III]}}$. The $z \sim 2$ composites, on the other hand, have much higher S/Ns thanks to the significantly larger sample size and therefore enable the first measurement of C III] from a statistical sample at this redshift.

As shown in Figure 9, the $z \sim 2$ binary stacks (squares) follow a clear trend of increasing $EW_{C\text{ III]}}$ at higher $EW_{Ly\alpha}$. In order to achieve a larger dynamic range in both C III] and $EW_{Ly\alpha}$, we further divided each $z \sim 2$ Ly α bin into two (four bins in total), and investigated if any underlying trends are obscured by simple binary. From the four-bin composites (circles in Figure 9), we observe a fairly flat behavior between $EW_{C\text{ III]}}$ and $EW_{Ly\alpha}$, with the exception of the strongest Ly α bin. The strongest Ly α bin has a measured $EW_{C\text{ III]}} = 3.41 \pm 0.27 \text{ \AA}$ while that in the rest of the Ly α stacks is $\sim 1.25 \text{ \AA}$. This trend is consistent with what has been found in previous studies (e.g., Shapley et al. 2003; Stark et al. 2014, 2015; Rigby et al. 2015; Jaskot & Ravindranath 2016), suggesting an intrinsic difference in the Ly α photon production between galaxies showing strong and weak Ly α emission. Although we were unable to detect the C III] feature at $z \sim 3$ –4, measurements were made using the $z \sim 3$ LBG composites by Shapley et al. (2003).¹⁴ These authors show a suggestive positive correlation between $EW_{Ly\alpha}$ and $EW_{C\text{ III]}}$, which is consistent with our much cleaner trend at $z \sim 2$.

The positive correlation between the C III] and Ly α strengths suggests that environments favorable for C III] production are also conducive to higher Ly α production and large Ly α escape fractions. As C III] is a collisionally excited transition, the emission is enhanced in H II regions with strong radiation fields and low metallicities (Jaskot &

Ravindranath 2016; Senchyna et al. 2017). In the meantime, the harder ionizing spectrum given by young, metal-poor stars leads to higher $EW_{Ly\alpha}$ by (1) boosting the intrinsic production rate of Ly α photons and (2) reducing the neutral covering fraction due to the higher ionization state of the ISM (Trainor et al. 2015), allowing more Ly α photons to escape the galaxy. Therefore, the hard ionizing spectrum and the low gas-phase metallicity modulate both the C III] and Ly α emission in the same direction, resulting in the positive correlation we observe between these two transitions. Similarly, $EW_{Ly\alpha}$ has also been demonstrated to correlate with the rest-optical nebular emission properties, such as [O III] $\lambda 5008/H\beta$ and [N II] $\lambda 6585/H\alpha$ (Erb et al. 2016; Trainor et al. 2016), which, like C III], can be used to probe the physical conditions of H II regions (e.g., gas-phase metallicity, electron temperature and density, and ionization parameter).

5. Kinematics

The measurement of the Ly α and LIS absorption profiles in the composite spectra enables us to study not only their line strengths, but also kinematics. The line centroids measured from LIS absorption features offer at least a crude measure of the bulk velocities of outflowing ISM, and the observed Ly α centroid may further constrain the kinematics of the neutral outflows (e.g., velocity distribution). In this section, we focus on the evolution of Ly α (Section 5.1) and LIS (Section 5.2) kinematics across $z \sim 2$ –4.

5.1. Ly α Kinematics

As described in Section 2, the centroids of Ly α emission (or absorption) and LIS absorption lines typically do not trace the galaxy systemic velocity because of the presence of galactic-scale outflows. Furthermore, it has been shown empirically how ISM kinematics affect the emergent Ly α profile. By studying a composite spectrum of $z \sim 3$ LBGs, Shapley et al. (2003) discovered that weaker Ly α emission corresponds to a larger velocity offset between Ly α and the LIS absorption lines, Δv_{em-abs} . A similar trend has been measured in the $z \sim 4$ LBGs by Jones et al. (2012), and these authors found no redshift evolution of the $EW_{Ly\alpha}$ versus Δv_{em-abs} relation from $z \sim 4$ to $z \sim 3$. Motivated by these previous studies, we revisit the correlation between Ly α kinematics and strength, and investigate if the covering fraction of the neutral ISM modulates the observed Ly α profile differently at different redshifts.

We characterized the kinematics of Ly α by measuring the centroid of the Ly α emission feature in the composite spectra binned according to $EW_{Ly\alpha}$. Even in the composite with the weakest Ly α where the line profile is classified as “absorption,” a small Ly α emission peak is still discernible and therefore its peak wavelength can be evaluated. We used MPFIT to fit the Ly α emission profile in a similar manner to the measurement of the LIS absorption lines, except that the fitting was performed on composites in L_λ space (i.e., not continuum-normalized). The initial estimate of the line parameters (centroid, EW, and Gaussian FWHM) was obtained from the IRAF routine *splot* by fitting the Ly α profile from the blue-side base to the red-side base of the line. We did not iterate the wavelength range for the Ly α emission profile in all Ly α morphological categories, given that the wavelength range bracketed by the blue to red bases of the Ly α profile already well defines the feature and

¹⁴ The $z \sim 3$ LRIS spectra presented in Shapley et al. (2003) were mainly collected before the LRIS-B upgrade and typically have significantly redder wavelength coverage than the LRIS blue-side dichroic spectra presented here. As such, the sample of $z \sim 3$ spectra with C III] coverage is significantly larger in Shapley et al. (2003).

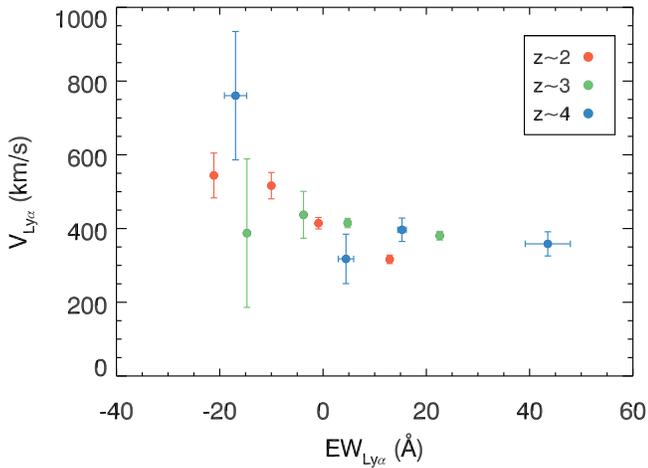


Figure 10. Centroid velocity of Ly α emission and the rest-frame Ly α EW in the composite spectra binned according to EW $_{\text{Ly}\alpha}$.

that iterations would not work for the “absorption” composites as a result of the extremely weak emission line.

We plot the centroid velocity of the Ly α emission and the rest-frame EW $_{\text{Ly}\alpha}$ in Figure 10. The three redshift samples span different dynamic ranges in EW $_{\text{Ly}\alpha}$, with the $z \sim 2$ galaxies having on average the weakest Ly α and the $z \sim 4$ galaxies showing the strongest. The $z \sim 2$ and $z \sim 4$ measurements show a clear trend of less redshifted Ly α centroid with stronger emission, similar to what has been found for LAEs in both individual and composite spectra at $z \sim 2-4$ (Erb et al. 2014; Guaita et al. 2017), while for the $z \sim 3$ sample there does not appear to be a dependence between these two parameters. Although the correlation between Ly α kinematics and strength is not conclusive at $z \sim 3$ based only on our measurements, given previous kinematics studies of Ly α at this redshift (e.g., Shapley et al. 2003), we believe that this relation holds at $z \sim 3$ as well. Generally speaking, the data points from different redshift samples seem to follow the same curve, and we do not see any evolutionary trend of the peak velocity versus EW relation for Ly α from $z \sim 4$ to $z \sim 2$.

In explaining the trend of more redshifted Ly α with weaker Ly α emission across $z \sim 2-4$, we consider two physical scenarios. The first possibility is that the larger (positive) offset of Ly α to the systemic velocity results from the faster-moving outflows on the far side of the galaxies, which at the same time have a larger covering fraction, leading to weaker Ly α emission. In light of the fact that the outflow kinematics can be characterized by the centroid velocity of LIS absorption features, we can test this hypothesis by examining whether the composites with more redshifted (and therefore weaker) Ly α also show more blueshifted LIS lines. In fact, the measured LIS centroid velocities are fairly similar between composites with the strongest and weakest Ly α , counter to what the “faster outflowing gas” hypothesis predicts.

Alternative models are therefore needed to explain the observed correlation between EW $_{\text{Ly}\alpha}$ and its peak velocity. Several studies have attempted to generate the emergent Ly α profile with assumed neutral gas kinematics (e.g., Verhamme et al. 2006, 2008; Steidel et al. 2010). Specifically, in Steidel et al. (2010), where the observed profiles of Ly α and the LIS absorption lines are simultaneously recovered by a kinematic model considering the velocity distribution of the neutral gas, these authors found that the apparent velocity shift of Ly α

primarily results from the absorption of Ly α photons by neutral gas near the systemic velocity. As the covering fraction of neutral gas increases around $v = 0$, more Ly α photons get resonantly scattered out of the line of sight, pushing the observed Ly α centroid to a redder wavelength, resulting in a weaker Ly α emission profile at the same time. The model from Steidel et al. (2010) successfully explains our results, and thus the largest EW $_{\text{Ly}\alpha}$ seen at $z \sim 4$ is likely caused by the smallest covering fraction near the systemic velocity. Furthermore, the redshift independence of the Ly α velocity versus EW relation suggests a direct impact of the covering fraction and kinematics of the neutral ISM on the observed Ly α emission profile across $z \sim 2-4$. Due to the low resolution of the spectra, we were unable to study the detailed structure of the Ly α profile (e.g., the presence and properties of multiple emission peaks). Future spectroscopic data with high spectral resolution will provide further information on the kinematics and covering fraction of the H I gas.

5.2. LIS Absorption Kinematics

While the centroid velocity of Ly α can provide valuable information on neutral gas kinematics, LIS absorption features are typically used to probe the bulk velocities of the neutral phase of the ISM/CGM. Ideally, we would like to decompose the interstellar absorption profiles into a systemic and an outflow component, with the latter representing the absorption truly occurring in the outflowing gas. Unfortunately, the relatively low resolution of our data prevents us from obtaining robust measurements from a two-component fit. Therefore, we fit the absorption profiles with a single Gaussian and used the centroid velocity of the best-fit Gaussian model as a crude estimator of the bulk kinematic properties.

In addition, it is most ideal to compare galaxies at different redshifts not only spanning the same UV luminosity and stellar mass ranges, but also having the same median values of these properties. However, due to the different distributions of the redshift samples in the $M_{\text{UV}}-M_*$ plane (Figure 2), it is difficult to construct such subsamples at each redshift with the same median M_{UV} and stellar mass at the same time. Considering that these two properties are correlated, we chose stellar mass as the “control” and created composite spectra for the $z \sim 2$ and $z \sim 3$ samples such that they have the same median stellar mass, $\log(M_*/M_\odot) = 9.98$ as the full $z \sim 4$ sample. Specifically, we constructed the matching M_* subsamples by including objects starting from the lower-mass end of the $z \sim 2$ and $z \sim 3$ samples until the median values of these subsamples become closest to that of the $z \sim 4$ sample. The resulting subsamples include 290, 235, and 84 spectra at $z \sim 2$, 3, and 4, respectively. The higher-redshift subsamples have slightly brighter UV luminosities, with the median M_{UV} for the $z \sim 2$, $z \sim 3$, and $z \sim 4$ samples being -20.48 , -20.96 , and -21.06 , respectively.

In principle, Si II $\lambda 1260$, C II $\lambda 1334$, and Si II $\lambda 1527$ are all good tracers of the ISM kinematics. However, we chose C II $\lambda 1334$ as the probe of the low-ion kinematics, mainly to ensure a fair comparison to the evolutionary study at $z \sim 0-2$ by Sugahara et al. (2017; Section 6.1), in which the authors used C II $\lambda 1334$ to trace neutral gas kinematics. Using the averaged centroid velocity of Si II $\lambda 1260$, C II $\lambda 1334$, and Si II $\lambda 1527$ yields qualitatively the same answer regarding the evolution of interstellar kinematics at $z \sim 2-4$ as using C II $\lambda 1334$ alone.

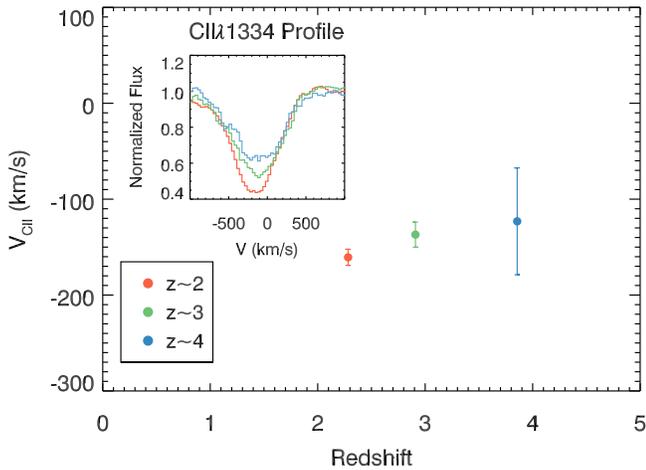


Figure 11. Centroid velocity of the C II $\lambda 1334$ absorption feature vs. redshift in the composite spectra with fixed median stellar mass, $\log(M_*/M_\odot) = 9.98$. The 1σ error bar on the centroid velocity was estimated based on the composite error spectra and does not include the uncertainty on the systemic velocity. The C II $\lambda 1334$ line profiles in the $z \sim 2$, 3, and 4 fixed-mass composites are shown in the inset panel.

Figure 11 shows the averaged centroid velocity plotted against redshift. The centroid velocities of C II are $-161 \pm 8 \text{ km s}^{-1}$, $-137 \pm 13 \text{ km s}^{-1}$, and $-123 \pm 56 \text{ km s}^{-1}$ for the $z \sim 2$, $z \sim 3$, and $z \sim 4$ samples, respectively. We note that the error bars quoted here on the velocity measurements do not include the formal uncertainty on the systemic velocity of the composites, Δv_{sys} , which is $\sim 25 \text{ km s}^{-1}$ for $z \sim 2$ –3 galaxies and $\sim 120 \text{ km s}^{-1}$ for $z \sim 4$ galaxies as estimated from MPFIT.¹⁵ Taking into account the Δv_{sys} of all redshift samples, we do not observe a noticeable evolutionary trend of the neutral gas kinematics from $z \sim 2$ to $z \sim 4$ for galaxies at fixed stellar mass. These results, as derived from the centroid velocities, suggest that the kinematics of the neutral ISM/CGM are mostly indistinguishable in galaxies with similar UV luminosity and stellar mass despite being at different redshifts. Spectroscopic data with higher S/N and spectral resolution will be of great help for examining the evolution of neutral gas kinematics in greater detail by enabling precise measurements of the outflow component from two-component profile fits. Additionally, high-quality imaging data will provide unique insights into the evolution of outflows through the study of the factors that modulate the strength of the outflows (e.g., size, SFR surface density).

6. Discussion

Although we have primarily investigated how the physical properties of the cool ISM/CGM evolve with redshift, the wide spectral coverage in the rest-UV and the unprecedentedly large size of our galaxy samples at $z \gtrsim 2$ enable various analyses that may potentially shed light on multiple key questions in the study of galaxy evolution. Here we extend our results from Sections 4 and 5, examining the redshift evolution of outflow velocities all the way from $z \sim 0$ to $z \sim 4$ (Section 6.1), and present a physical picture that accounts for the observed spectroscopic trends (Section 6.2).

¹⁵ As described in Section 3.1, these formal error bars were estimated by measuring the stellar feature C III $\lambda 1176$ in the overall composites, and therefore do not reflect the systematic uncertainty associated with the redshift rules.

6.1. Outflow Velocity at $z \sim 0$ –4

In this paper, we have explored how the kinematics of neutral gas evolve from $z \sim 2$ to $z \sim 4$, as probed by the centroid velocity of the LIS absorption features. While we did not find a noticeable evolution of the outflow velocities of the neutral gas within the $z \sim 2$ –4 redshift interval, Sugahara et al. (2017) report evidence of increased outflow velocity from $z \sim 0$ to $z \sim 2$ based on the measurement of LIS absorption features. These authors studied galaxy spectra at $z \sim 0$, $z \sim 1$, and $z \sim 2$, drawn, respectively, from the Sloan Digital Sky Survey (SDSS; York et al. 2000), DEEP2, and Erb et al. (2006b, 2006a), with the last comprising a subset of the $z \sim 2$ –3 LRIS sample presented in this paper. Considering the different rest-frame spectral coverage of the three data sets, Sugahara et al. (2017) used ions with similar ionization energies for the comparison at different redshifts: Na I D $\lambda\lambda 5891, 5897$ and Mg I $\lambda 2853$ from $z \sim 0$ to $z \sim 1$, and Mg II $\lambda\lambda 2796, 2803$ and C II $\lambda 1334$ from $z \sim 1$ to $z \sim 2$. Given that we also cover the C II feature in our $z \sim 2$ –4 spectra, it is possible to extend the $z \sim 0$ –2 evolutionary study of neutral gas kinematics by Sugahara et al. (2017) up to $z \sim 4$, connecting the lower- and higher-redshift universe.

Sugahara et al. (2017) constructed the $z \sim 0$ –2 samples such that they have similar normalized stellar mass distributions. The $z \sim 0$ and $z \sim 1$ composites were binned according to SFRs, while the $z \sim 2$ stacks included all objects in the $z \sim 2$ sample. To characterize the interstellar absorption profiles in the composite spectra, Sugahara et al. (2017) adopted a two-component model, which includes an outflow component and a systemic component. The profile of each component in the model is described by $I(\lambda) = 1 - C_f + C_f e^{-\tau(\lambda)}$ and $\tau(\lambda) = \tau_0 e^{-(\lambda - \lambda_0)^2 / (\lambda_0 b_D / c)^2}$, where C_f is the covering fraction, τ is the optical depth as a function of wavelength, τ_0 is the optical depth at the rest wavelength λ_0 , and b_D is the Doppler parameter. Since the absorption lines were fit with a single Gaussian profile in our study, we cannot directly compare our results with those presented in Sugahara et al. (2017). Instead, we conducted a crude two-component Gaussian fit to the C II $\lambda 1334$ absorption line profile in the continuum-normalized, overall composites at $z \sim 2$, 3, and 4. According to Sugahara et al. (2017), the observed flux of the interstellar absorption in the continuum-normalized spectra is $I_{\text{obs}}(\lambda) = I_{\text{out}}(\lambda)I_{\text{sys}}(\lambda)$, where $I_{\text{out}}(\lambda)$ and $I_{\text{sys}}(\lambda)$ are the outflow and systemic components with the continuum normalized to unity, respectively. Accordingly, we fit the C II profile using MPFIT to simultaneously model the systemic component at the rest wavelength (1334.5 Å) and the blueshifted outflow component. The functional form of the model we adopted is

$$I_{\text{obs}} = (1.0 - P_{\text{sys}}) \times (1.0 - P_{\text{out}}), \quad (1)$$

where P_{sys} and P_{out} are the Gaussian profiles parameterized by the line centroid, Gaussian FWHM, and enclosed area for the systemic and outflow components in the continuum-normalized composites, respectively. We list the best-fit parameters of the two-component models for the $z \sim 2$, 3, and 4 overall composites in Table 3. In contrast to Sugahara et al. (2017), we did not apply additional constraints on the covering fraction, optical depth, and impact parameter.

We show an example of the two-component best fit we obtained with the $z \sim 2$ composite in the top panel of Figure 12. The outflow velocity, V_{out} , can therefore be derived from the shift in the centroid of the outflow component. The

Table 3
Best-fit Parameters of Two-component Models for $z \sim 2-4$ Composites

| Sample | λ_{sys} (Å) | FWHM_{sys} (Å) | Area_{sys} (Å) | λ_{out} (Å) | FWHM_{out} (Å) | Area_{out} (Å) |
|------------|----------------------------|--------------------------------|--------------------------------|----------------------------|--------------------------------|--------------------------------|
| $z \sim 2$ | 1334.532 | 0.79 ± 0.29 | 0.31 ± 0.41 | 1333.647 ± 0.216 | 1.25 ± 0.09 | 1.61 ± 0.32 |
| $z \sim 3$ | 1334.532 | 0.96 ± 0.26 | 0.70 ± 0.73 | 1333.379 ± 0.881 | 1.37 ± 0.36 | 1.06 ± 0.76 |
| $z \sim 4$ | 1334.532 | 1.04 ± 0.31 | 0.87 ± 0.33 | 1332.934 ± 1.003 | 1.03 ± 0.77 | 0.53 ± 0.43 |

Note. Columns 2–4 and 5–7 represent the line centroid, Gaussian FWHM, and enclosed area for the systemic and outflow components in the continuum-normalized composites, respectively. The systemic component was fixed at the rest-frame wavelength of C II, 1334.532 Å.

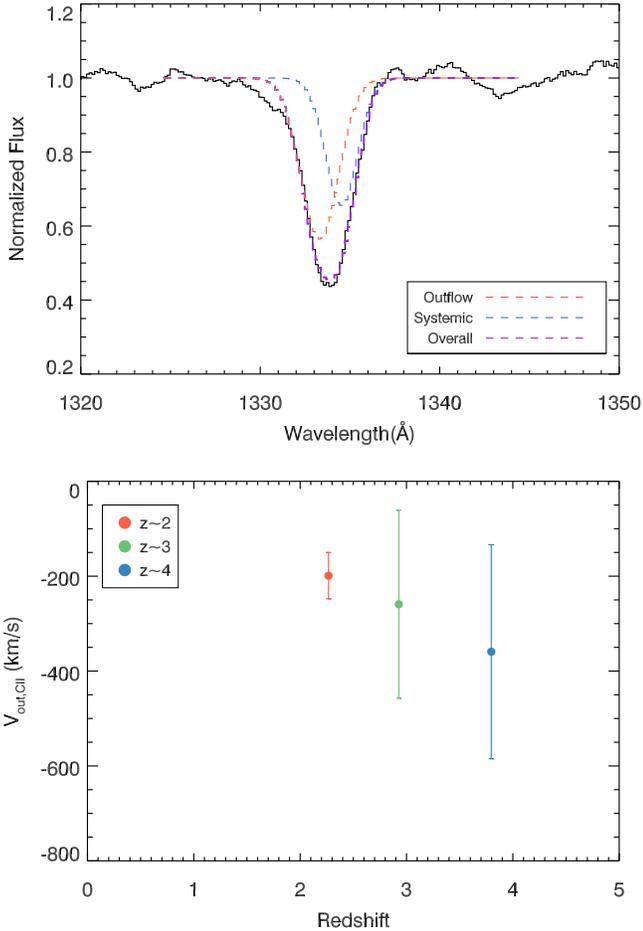


Figure 12. Top: best-fit outflow (orange), systemic (blue), and overall (purple) absorption profiles in the two-component fit to the $z \sim 2$ overall composite. Bottom: outflow velocity (determined by the centroid velocity of the outflow component) measured from the $z \sim 2$, $z \sim 3$, and $z \sim 4$ composite spectra including all objects within the UV luminosity and stellar-mass constraints.

best-fit values of $V_{\text{out,C II}}$ over $z \sim 2-4$ are plotted in the bottom panel of Figure 12. We measure an outflow velocity of $-199 \pm 49 \text{ km s}^{-1}$, $-259 \pm 198 \text{ km s}^{-1}$, and $-359 \pm 226 \text{ km s}^{-1}$ for the $z \sim 2$, $z \sim 3$, and $z \sim 4$ samples, respectively. Within the error bars,¹⁶ no clear evolution is shown for the C II kinematics. Specifically, our measurement of $V_{\text{out,C II}}$ at $z \sim 2$ is in good agreement with the value quoted in Sugahara et al. (2017), $-208 \pm 30 \text{ km s}^{-1}$, assuming $C_{f, \text{sys}} = 0.1$ without any additional constraints. We also note that this result is consistent with what we found in Section 5.2, that the neutral gas kinematics do not appear to evolve

¹⁶ We note that the uncertainty on the outflow velocities is fairly large, which is likely due to the lack of constraints on both the outflow and systemic components, leading to numerous possible combinations near the minimum χ^2 .

significantly when being probed by the Gaussian centroid velocity of multiple LIS absorption features. While the lack of redshift evolution of the outflow velocity may not be definitive based only on Figure 12, we can conclude that we do not see an apparent evolution of the neutral gas kinematics from $z \sim 2-4$ when combining our result here with that in Section 5.2.

If we consider the overall evolution of the neutral gas kinematics from $z \sim 0$ to $z \sim 4$, combining both our results and those from Sugahara et al. (2017), the outflow velocity seems to stop increasing significantly beyond $z \sim 2$. A couple of factors may contribute to this “turning point.” First, given the positive relation between outflow velocities and the SFR surface density observed in star-forming galaxies (Heckman et al. 2000; Chen et al. 2010; Kornei et al. 2012), the flattening of the outflow velocity at higher redshifts may result from a smaller change in the SFR surface density at fixed mass by a factor of ~ 2.5 at $z \sim 2$ to $z \sim 4$ relative to that at $z \sim 0$ to $z \sim 2$ (e.g., Shibuya et al. 2015). Second, galaxies at $z \leq 1.5$ tend to show a bipolar outflow geometry (e.g., Rupke et al. 2005; Martin et al. 2012; Rubin et al. 2014), as inferred by the outflow detection rate among galaxies with similar star-forming properties as well as the detection of extraplanar gas perpendicular to the plane (e.g., Bordoloi et al. 2011). Galactic-scale outflows at $z \gtrsim 2$, on the other hand, are much more ubiquitous (e.g., Pettini et al. 2002; Shapley et al. 2003; Steidel et al. 2010), suggesting a more spherical geometry. Therefore, the smaller detection fraction of blueshifted interstellar absorption lines may make those line profiles less blueshifted in the composites at $z \sim 0-2$ than at $z \sim 2-4$, resulting in an apparent plateau in outflow speed at $z \sim 2$ and higher. Since both factors can collectively affect the observed interstellar absorption profiles in the composites, individual spectra with significant outflow detections or composite spectra constructed by only stacking those with individual detections of outflows would help determine if the SFR surface density or the outflow geometry plays a more important role in this “turning point” at $z \sim 2$.

6.2. A Physical Picture

We now interpret our results in a physical picture for the evolution of Ly α emission and the properties of the absorbing ISM. We also consider the implications of our results for estimates of the LyC escape fraction during the epoch of reionization.

6.2.1. Ly α Emission

Ly α emission is commonly the strongest feature in the UV spectra of distant star-forming galaxies and has been used to confirm the spectroscopic redshifts of galaxies out to $z > 8$ (Zitrin et al. 2015). Furthermore, large samples of LAEs have

been assembled at $z \sim 6-7$ in order to probe the neutral fraction of the IGM (e.g., Konno et al. 2018). Given the prevalence of Ly α measurements, it is important to describe how $EW_{Ly\alpha}$ relates to other galaxy properties, and, more fundamentally, to understand the factors controlling the observed strength of the Ly α emission. This knowledge will enable us to relate LAEs to the star-forming galaxy population in general over a wide range of redshifts and to use the measurement of Ly α as a probe of other key galaxy properties that cannot be directly measured (e.g., escaping LyC radiation during the epoch of reionization).

There are three key factors to highlight when attempting to explain the observed Ly α emission properties of high-redshift star-forming galaxies: (1) the intrinsic production of Ly α photons through nebular recombination emission, (2) the radiative transfer of Ly α photons through the ISM and CGM, and (3) the radiative transfer of Ly α photons through the increasingly neutral IGM at higher redshifts. Here we consider the importance of these factors, as suggested by our results.

We have established the constancy of the relations among Ly α emission, LIS absorption, and dust extinction (as parametrized by $E(B - V)$), extending over the widest redshift baseline ($z \sim 2-4$), and using the most systematically controlled samples to date. In fact, a similar relationship between $EW_{Ly\alpha}$ and EW_{LIS} has even been observed in $z \sim 0$ star-forming galaxies (e.g., Henry et al. 2015; Chisholm et al. 2017). The direct and non-evolving connection between $EW_{Ly\alpha}$, EW_{LIS} , and $E(B - V)$ suggests that the evolving relations between Ly α and other galaxy properties (e.g., M_{UV} , age, M_* , SFR) arise as the LIS-absorbing gas and dust reddening (at fixed galaxy properties) evolve with redshift. Furthermore, according to one simple interpretation of the trends among $EW_{Ly\alpha}$, EW_{LIS} , and $E(B - V)$, galaxies in our samples all have similar intrinsic Ly α EWs, as determined by the ionizing photon production efficiency for a given mass of stars formed and a relatively constant (and small) escape fraction of LyC photons. Together, these two factors determine the ratio of Ly α recombination luminosity to UV continuum luminosity density (i.e., the rest-frame Ly α EW; Reddy et al. 2016). Accordingly, in this simple picture, the observed range of $EW_{Ly\alpha}$ at $z \sim 2-4$ is entirely determined by the range of LIS-absorbing gas covering fractions and dust reddenings—i.e., the radiative transfer of Ly α through the ISM and CGM.

However, our results on the correlation between Ly α and C III] emission suggest that, at the highest Ly α EWs, the H II region properties and the massive stars that ionize them are intrinsically different. Specifically, in our $z \sim 2$ sample, galaxies in the quartile with the strongest Ly α emission (median $EW_{Ly\alpha} \sim 10 \text{ \AA}$) also have significantly stronger C III] emission than those in the remainder of the sample. As shown in theoretical and observational works (e.g., Rigby et al. 2015; Jaskot & Ravindranath 2016; Senchyna et al. 2017), stronger C III] EWs are typically produced in lower-metallicity (i.e., significantly subsolar) H II regions where the electron temperature is higher and the stellar ionizing radiation is stronger. Accordingly, this result indicates a connection between Ly α emission strength and nebular metallicity. This result is consistent with the work of Trainor et al. (2016), who show that $EW_{Ly\alpha}$ is strongly correlated with nebular metallicity at $z \sim 2$, as probed by rest-optical emission-line ratios indicating the degree of excitation and ionization. Therefore, the strength of Ly α emission, at least at the highest values of $EW_{Ly\alpha}$, reflects

not only the transfer Ly α photons through the ISM but also the ionizing spectra, ionization parameters, and nebular metallicities of the H II regions where Ly α photons are originally produced.

The increasing neutral hydrogen opacity in the IGM is evident in the $z \sim 2, 3$, and 4 composite spectra shown in Figure 3. Specifically, the ratio of the continuum flux density bluewards of Ly α relative to that redwards of Ly α decreases significantly from $z \sim 2$ to $z \sim 4$. Simulations (e.g., Laursen et al. 2011) suggest that increasing IGM opacity will lead to a measurably reduced Ly α EW, by $\sim 25\%$ at $z = 3.5$ and by $\sim 75\%$ at $z = 5.8$. Jones et al. (2012) in fact find evidence for this effect when comparing the Ly α and LIS properties of composite spectra constructed from LBG subsamples at $\langle z \rangle = 3.8$ and $\langle z \rangle = 4.7$. For roughly the same $EW_{Ly\alpha}$, the $\langle z \rangle = 4.7$ sample has a significantly weaker EW_{LIS} ($\sim 1 \text{ \AA}$ at $\langle z \rangle = 4.7$ versus $\sim 1.7 \text{ \AA}$ at $\langle z \rangle = 3.8$). This difference can be explained if the $\langle z \rangle = 3.8$ and $\langle z \rangle = 4.7$ galaxies follow the same intrinsic relationship between $EW_{Ly\alpha}$ and EW_{LIS} , but the average $EW_{Ly\alpha}$ in the subset of LBGs at $\langle z \rangle = 4.7$ is suppressed by roughly a factor of 2 by IGM absorption—consistent with the predictions by Laursen et al. (2011). In our work, we find no significant difference among the EW_{LIS} versus $EW_{Ly\alpha}$ relations from $z \sim 2$ to $z \sim 4$. This result suggests that either, even in our $z \sim 4$ sample, the IGM absorption of Ly α photons produced in LBGs is not significant, or we are not measuring the part of the Ly α profile that is affected by IGM absorption. Alternatively, $z \sim 4$ galaxies may follow a different intrinsic relation between $EW_{Ly\alpha}$ and EW_{LIS} relative to the $z \sim 2-3$ samples, characterized by higher $EW_{Ly\alpha}$ for a given EW_{LIS} . The increased IGM absorption at $z \sim 4$ then attenuates the $EW_{Ly\alpha}$ to the point that the $z \sim 4$ sample follows the same patterns as the lower-redshift galaxies, mimicking a non-evolving $EW_{Ly\alpha}$ versus EW_{LIS} relation. For now, we favor the simpler explanation that the EW_{LIS} versus $EW_{Ly\alpha}$ relation does not evolve between $z \sim 2$ and $z \sim 4$ and that the effects of IGM absorption are not significant within our $z \sim 2-4$ samples. We also note that the lack of evolution between Ly α kinematics ($v_{Ly\alpha}$) and $EW_{Ly\alpha}$ (Section 5.1) is consistent with the subdominance of evolving IGM absorption in our $z \sim 2-4$ measurements. However, it will be extremely valuable to investigate the combined possible effects of IGM and ISM evolution out to $z \sim 5-6$ using existing rest-UV spectroscopic samples (e.g., Faisst et al. 2016).

6.2.2. The Absorbing ISM

Although $z \sim 2-4$ galaxies follow the same relations among $EW_{Ly\alpha}$, EW_{LIS} , and $E(B - V)$, the average $EW_{Ly\alpha}$ increases from $z \sim 2-4$, while the average EW_{LIS} and $E(B - V)$ decrease. This evolution in average properties suggests that (at fixed UV luminosity), the covering fraction of LIS-absorbing gas and dust decreases from $z \sim 2-4$. As demonstrated by Reddy et al. (2016), the covering fractions of various low ions are correlated with, but systematically lower than, the covering fraction of H I ($f_{cov}(H I)$) by a factor of 2–3. Furthermore, the ratio between the covering fractions of low ions and H I increases as $f_{cov}(H I)$ increases. These trends, as described in Reddy et al. (2016), are consistent with a scenario in which discrete metal-enriched regions of interstellar gas are contained within the outflowing interstellar H I, whose dust-to-gas (i.e., metal-to-gas) ratio increases with increasing covering fraction of H I. Accordingly, the decrease in average EW_{LIS}

from $z \sim 2$ to $z \sim 4$ corresponds to a decrease in the average covering fraction of interstellar H I.

Two factors that may contribute to the decrease in the covering fraction of interstellar H I (at fixed galaxy properties such as UV luminosity) are (1) an evolving average stellar radiation field corresponding to a harder ionizing spectrum at $z \sim 4$ than at $z \sim 2$, and (2) a decrease in the overall covering fraction of both neutral and ionized interstellar gas. We can test for the presence of both of these effects in our data.

To investigate the nature of the ionizing radiation field, we used the relative column densities of Si II and Si IV as a proxy of the ISM ionization state. As discussed in Section 4.1, we found that the column density of Si IV appears to increase from $z \sim 2$ to $z \sim 3-4$. At the same time, the variation in the EW ratio of Si II $\lambda 1260$ and Si II $\lambda 1527$ can be used to infer a change in the column density of Si II. In the optically thin regime, the EW of Si II $\lambda 1260$ should be ~ 5.3 times larger than that of Si II $\lambda 1527$. We observe $\text{EW}_{\text{Si II } \lambda 1260} / \text{EW}_{\text{Si II } \lambda 1527}$ ratios of $0.93 \pm 0.03 \text{ \AA}$, $0.84 \pm 0.07 \text{ \AA}$, and $1.30 \pm 0.25 \text{ \AA}$ at $z \sim 2$, 3, and 4, respectively. This progression suggests a slight decrease in Si II column density from $z \sim 2$ to $z \sim 4$. The combined increase and decrease in the respective column densities of Si IV and Si II suggest a more highly ionized ISM at $z \sim 4$ than at $z \sim 2$. We also note that, in simple, empirical terms, the ratio of the EWs of the Si IV and Si II lines increases from $z \sim 2-4$, as seen in the composite spectra of Figure 3: compare the relative depths of the Si IV and Si II features in the $z \sim 2$ (red) and the $z \sim 4$ (blue) composites.

While the balance between high and low ions in the ISM of star-forming galaxies at $z \sim 2-4$ may evolve, it also appears that the covering fraction of both neutral and ionized phases decreases. This decrease is traced by the decreasing EW of saturated LIS lines and by the roughly constant EW of Si IV absorption as the column density increases (Section 4.1). The evolution in the covering fraction of interstellar gas must reflect an evolving balance among gas accretion, star formation, and feedback. For example, star-forming galaxies at fixed UV luminosity are smaller at $z \sim 4$ than at $z \sim 2-3$, with approximately doubled SFR surface densities (Shibuya et al. 2015). Given the connection between outflow properties and SFR surface density (Heckman et al. 2000; Chen et al. 2010; Kornei et al. 2012), the higher typical SFR surface densities at $z \sim 4$ may correspond to outflows that more efficiently clear out channels through the ISM, reducing the covering fraction of gas, and allowing for the escape of Ly α photons. Alternatively, if outflows have similar properties at $z \sim 2-4$ (at fixed UV luminosity) in terms of energy and momentum input and a similar ability to evacuate holes in the ISM, we may be observing the accumulation of gas in the ISM and CGM as cosmic time advances. At least in our results, we do not detect significant evolution in the bulk outflow kinematics from $z \sim 2-4$ (Sections 5.2, 6.1), so this second alternative is also plausible. A comparison with gas in galaxies modeled in cosmological zoom-in simulations (e.g., Feldmann et al. 2017), as well as observations at both lower and higher redshifts, will be crucial for understanding the evolution of the covering fraction of metals and H I in the neutral ISM of star-forming galaxies.

We now consider the connection between the covering fraction of H I and dust reddening. As shown in this work and previously (e.g., Shapley et al. 2003; Jones et al. 2012; Reddy et al. 2016), there is a strong correlation between EW_{LIS} and

$E(B - V)$, such that stronger LIS absorption is associated with increased dust reddening. At the same time, there is no significant connection between the high-ionization absorption strength, EW_{HIS} and $E(B - V)$. These results suggest that dust responsible for reddening the stellar continuum in the outflowing ISM/CGM is confined to the neutral regions probed by the LIS features. We would also like to understand what fraction of the reddening and attenuation of the UV continuum takes place in the outflowing ISM/CGM as opposed to gas that is much more local to H II regions and not participating in large-scale outflows. For example, does the evolution at $z \sim 2-4$ toward lower $E(B - V)$ at fixed galaxy properties such as M_{UV} and SFR reflect not only simply the evolution toward a lower H I gas covering fraction, but also the overall lower degree of H II region chemical enrichment (and dust content) at higher redshift (at fixed galaxy properties)? Spatially resolved high-S/N measurements of EW_{LIS} and $E(B - V)$ indicating the degree of correlation between small-scale EW_{LIS} and $E(B - V)$ variations may help address this important question of where significant dust attenuation occurs.

6.2.3. Lyman-continuum Emission

One of the frontiers for galaxy formation studies is to quantify the contribution of star-forming galaxies to reionization (for a review, see Stark et al. 2017). Measurements of the non-ionizing rest-frame UV luminosity function are extending to extremely faint magnitudes based on Hubble Frontier Fields observations (e.g., Bouwens et al. 2017; Livermore et al. 2017). However, since neither LyC radiation nor any of the Lyman series absorption features bluewards of Ly α can be measured directly during the epoch of reionization (Vanzella et al. 2012; Shapley et al. 2016), we require another proxy for the LyC radiation properties of $z > 6$ galaxies. Recent work at both low redshift and $z \sim 3$ has highlighted the connection between the strength of Ly α emission, the H I covering fraction, and the escape of LyC radiation (Shapley et al. 2003; Jones et al. 2013; Trainor et al. 2015; Reddy et al. 2016; Chisholm et al. 2017; Marchi et al. 2017, Steidel et al., 2018, in preparation). Galaxies with higher ratios of LyC to non-ionizing UV radiation tend to show stronger Ly α emission and weaker LIS absorption, and models of their UV spectra suggest lower H I covering fractions.

Here we suggest that measuring Ly α and LIS absorption properties of $z > 6$ star-forming galaxies holds great promise for estimating how LyC escape varies with UV luminosity during the Epoch of Reionization. In particular, one promising path forward consists of using the Ly α escape fraction as a proxy for the H I covering fraction (and, accordingly the LyC escape fraction, which is $1 - f_{\text{cov}}(\text{H I})$; Jones et al. 2013; Leethochawalit et al. 2016) in bins of UV luminosity. For this method to be effective, there are key uncertainties that will need to be addressed: (1) we will need to estimate and correct for the typical IGM absorption of the Ly α emission profile at $z \sim 6$, and (2) we will require constraints on the intrinsic production rate of Ly α photons in $z \sim 6$ star-forming galaxies in order to convert the observed $\text{EW}_{\text{Ly}\alpha}$ (corrected for IGM absorption) to a Ly α escape fraction. It is also tempting to use measurements of LIS absorption at $z \sim 6$ to predict $f_{\text{cov}}(\text{H I})$. However, given the hints of significant metallicity evolution out to $z \sim 6$ in star-forming galaxies (e.g., Stark et al. 2017), typical LIS absorption features at $z \sim 6$ may not be saturated,

and therefore may not offer a direct probe of the H I covering fraction. More generally, the H I gas may be optically thinner at $z \sim 6$ than at $z \sim 3$, which indicates that the absolute escape fraction of LyC photons would no longer be $1 - f_{\text{cov}}(\text{H I})$ as stated earlier. In order to understand the utility of EW_{LIS} for probing the H I covering fraction at $z \sim 6$, we should first trace the $\text{EW}_{\text{Ly}\alpha}$ versus EW_{LIS} relation out to $z \sim 6$ using high-quality $z \sim 4\text{--}6$ ground-based galaxy spectra. At even higher redshifts, rest-UV spectroscopy with the *James Webb Space Telescope (JWST)* will prove a powerful tool for constraining the LyC emission properties of star-forming galaxies.

7. Summary

Rest-UV spectroscopic features provide rich insights into the key properties (structure, kinematics, covering fraction) of the ISM/CGM in star-forming galaxies. By constructing LBG samples controlled in UV absolute magnitude and stellar mass at $z \sim 2\text{--}4$ and performing systematic measurements of the spectral and galaxy properties in a uniform manner, we primarily studied the evolution in line strengths and kinematics of both emission and absorption features in the rest-UV spectra. We have found the following key results:

1. No noticeable redshift evolution is found for the relations among $\text{EW}_{\text{Ly}\alpha}$, EW_{LIS} , and $E(B - V)$. In other words, Ly α , LIS lines, and dust extinction are fundamentally intercorrelated. Specifically, the redshift independence of the $\text{EW}_{\text{Ly}\alpha}$ versus EW_{LIS} relation suggests either that the covering fraction of neutral gas is the dominant factor for modulating both $\text{EW}_{\text{Ly}\alpha}$ and EW_{LIS} in the same manner at all redshifts, or that the intrinsically higher $\text{EW}_{\text{Ly}\alpha}$ due to lower metallicity and stronger radiation fields at higher redshifts at fixed galaxy property is compensated by the increased absorption from the IGM. In contrast, the strength of Ly α exhibits no correlation with the H I absorption strength, suggesting that Ly α and the H I absorption features trace different phases of gas. The approximately constant $\text{EW}_{\text{H I}}$ over $z \sim 2\text{--}4$ suggests a lower covering fraction of the ionized gas at higher redshifts, given the apparent increasing Si IV column density.

2. We observe a redshift-invariant correlation between EW_{LIS} and $E(B - V)$, which is consistent with a physical picture in which dust and metal ions are embedded in the outflowing H I gas. Both the covering fraction of neutral gas, as probed by EW_{LIS} , and dust extinction seem to decrease with increasing redshift. We interpret this redshift evolution with two possible scenarios: either the $z \sim 4$ galaxies are more effective in reducing the H I covering fraction and dust content by increasing the ionization state of the ISM and driving powerful outflows, or the $z \sim 2\text{--}3$ galaxies are more efficient in accumulating gas and dust as they evolve. Individual spectra with high S/N and spectral and spatial resolution will shed light on the physical location of the dust with respect to the neutral gas.

3. The strength of Ly α emission decreases with decreasing redshift at fixed stellar mass, UV luminosity, and SFR, which likely results from a larger covering fraction of the neutral gas and higher $E(B - V)$ at lower redshifts. In the $z \sim 2$ and $z \sim 3$ samples, $\text{EW}_{\text{Ly}\alpha}$ shows little to no correlation with stellar mass, UV luminosity, and SFR. However, a larger $\text{EW}_{\text{Ly}\alpha}$ is observed at fainter galaxies with lower stellar mass and SFR at $z \sim 4$. The weakening of these trends from $z \sim 4$ to $z \sim 2$ is likely caused by a larger dynamic range of the neutral gas covering fraction at $z \sim 4$ than at $z \sim 2$ and $z \sim 3$. Although the sample

incompleteness enhances the observed Ly α trends, the evolution of these relations (i.e., stronger Ly α emission at higher redshifts at fixed galaxy properties, steeper Ly α trends at $z \sim 4$ than at $z \sim 2\text{--}3$) still exists after we have accounted for differential sample incompleteness. We thus conclude that the trends are real and robust, and not simply an effect of increasing sample incompleteness to galaxies lacking Ly α emission at higher redshift and fainter magnitude.

4. Younger galaxies at $z \sim 4$ show stronger Ly α , which can be attributed to their lower gas-phase metallicities and harder ionizing spectra. On the other hand, no visible trends are seen at $z \sim 2$ and $z \sim 3$ between $\text{EW}_{\text{Ly}\alpha}$ and galaxy age. The different behavior of the $\text{EW}_{\text{Ly}\alpha}$ versus age relation can be explained by the $z \sim 2\text{--}3$ galaxies being more chemically enriched at the youngest ages than the $z \sim 4$ galaxies, which are possibly experiencing their first generation of star formation. Our results here do not support the positive correlation between Ly α and age previously reported in some studies of $z \sim 3$ LBGs (Shapley et al. 2001; Kornei et al. 2010). In this earlier work, the authors modeled galaxy SEDs assuming 1.4 solar metallicity and the Calzetti extinction curve, and did not account for the contamination of strong nebular emission lines in the photometric bands. We found that with the most reasonable description of the stellar populations in the $z \sim 3$ galaxies in our sample (0.28 solar metallicity and the SMC extinction curve except for the highest-mass objects), the positive correlation between $\text{EW}_{\text{Ly}\alpha}$ and age disappears after removing the contaminated K_s -band from SED modeling.

5. The ratio of the fine-structure emission to the corresponding resonant absorption does not evolve significantly with redshift at fixed dust extinction. While on average a larger $\text{EW}_{\text{Si II}^*}/\text{EW}_{\text{Si II}}$ ratio is observed at $z \sim 4$, this trend is a result of the lower dust extinction in higher-redshift galaxies, rather than an evolution with cosmic time toward a larger size of the fine-structure emission region.

6. We find a flat trend between $\text{EW}_{\text{C III]}}$ and $\text{EW}_{\text{Ly}\alpha}$ at $z \sim 2$ except in the quartile with the strongest Ly α emission, where the C III] emission is significantly higher than in the remaining Ly α stacks. Given that $\text{EW}_{\text{C III]}}$ depends on the physical properties of the H II regions (e.g., gas-phase metallicity, ionization parameter), this result suggests variation in the intrinsic Ly α production among galaxies in our samples. Galaxies with larger observed $\text{EW}_{\text{Ly}\alpha}$ may not only have lower H I gas covering fractions, but also intrinsically produce more ionizing (and Ly α) photons per unit mass of stars formed.

7. We measure no strong evolution in the key rest-UV spectroscopic trends tracing interstellar kinematics. The trend of larger $\text{EW}_{\text{Ly}\alpha}$ corresponding to a less redshifted Ly α emission profile seems to be universal across $z \sim 2\text{--}4$, which suggests that the covering fraction of the neutral gas near $v = 0$ modifies the Ly α profile in the same manner across $z \sim 2\text{--}4$. On the other hand, the outflow velocities traced by the centroids of the LIS absorption features do not exhibit any evolutionary trend with redshift. We therefore speculate that the kinematics of the neutral gas are similar in LBGs in fixed UV luminosity and stellar mass ranges at $z \sim 2\text{--}4$.

Answering the key outstanding questions in the study of the evolving ISM/CGM at high redshift will require spectroscopic data with high S/N and spectral resolution, along with Integral Field Unit spectroscopic maps. For example, rest-UV and optical spectral maps of lensed, spatially resolved galaxies will provide key information for dissecting the relative physical

distribution of neutral gas and dust. Furthermore, the near-IR capabilities of the *JWST* will enable rest-UV and rest-optical spectroscopic studies of star-forming galaxies out to $z > 6$. With these high-quality individual, deep spectra, detailed, quantitative conclusions can be drawn on topics including (1) the nature of the relation between dust and neutral gas, (2) the intrinsic production and escape of Ly α and LyC photons, and (3) the detailed kinematics and spatial extent of the neutral and ionized phases of outflows. Making progress on these questions from an empirical standpoint is essential for our understanding of where the ISM/CGM absorption arises in galaxies and for a complete model of feedback in galaxy formation.

We acknowledge support from the David & Lucile Packard Foundation (A.E.S). We are grateful to the 3D-HST team for providing ancillary data on galaxy properties. This paper also includes data gathered with the 6.5 m Magellan Telescopes located at Las Campanas Observatory, Chile. We 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.

Appendix A Individual and Composite Error Spectra

Individual error spectra are essential for creating the composite error spectra, which ideally account for both sample variance and measurement uncertainty. However, the individual error spectra were not available for all the $z \sim 2$ and $z \sim 3$ galaxies and for the FORS2 objects in the $z \sim 4$ sample. In this section, we describe how we constructed the composite error spectra for the $z \sim 2-3$ and the $z \sim 4$ samples, respectively, despite the lack of some individual error spectra.

To create the composite error spectra for the $z \sim 2$ and $z \sim 3$ samples, we first reconstructed individual error spectra for all $z \sim 2-3$ galaxies by utilizing the vast library of the individual LRIS error spectra attained in the $z \sim 1$ outflow kinematic study presented in Martin et al. (2012). In this study, a sample of 208 objects were drawn from the Deep Extragalactic Evolutionary Probe 2 (DEEP2; Newman et al. 2013) galaxy redshift survey spanning the range $0.4 \leq z \leq 1.4$ and apparent *B*-band magnitude $B < 24.0$ and were observed with LRIS. The 400 lines mm^{-1} and 600 lines mm^{-1} grisms were used to obtain the blue side of the spectra. A total of 145 objects were observed with the 400 lines mm^{-1} grism and 63 objects were observed with the 600 lines mm^{-1} grism. The individual one-dimensional error spectra for the $z \sim 1$ galaxies were determined by the standard deviation at each wavelength of the corresponding science spectra of the same object from multiple exposures.

To reconstruct the individual error spectra of the $z \sim 2-3$ LRIS spectra, the main goal is to recover the shape of the error spectra, which is set by the sensitivity of the instrument in the observed frame. We first made the 400-line and 600-line error spectrum templates by stacking individual error spectra with their respective spectral resolutions. When being combined, these individual error spectra were scaled to a common median over 4000–5000 Å in the observed frame. We then used the IRAF *continuum* routine to fit the general shape of the two error spectrum templates with an order of 5. After obtaining the smoothed error spectrum templates, for each object in the $z \sim 2-3$ LBG sample that was observed with the 400-line (600-line) grism, we shifted the 400-line (600-line) template to

the rest frame, and scaled it such that the pixel-to-pixel noise over 1250–1400 Å in the rest-frame science spectrum matched the average flux level of the rest-frame error template across the same wavelength range.

Unfortunately, we were unable to precisely simulate the overall shape of the 300-line grating and grism error spectra, given that the $z \sim 1$ DEEP2/LRIS data were not obtained with these configurations on LRIS. For this small fraction of our sample ($\sim 7\%$ combined at $z \sim 2-3$), we adopted the shape of the 400-line grism error template for the 300-line grating and grism objects as a crude estimate. For individual 300-line galaxies, the 400-line error template was also transformed into their respective rest frames and scaled to match the rms of corresponding science spectra over 1250–1400 Å. We note that approximating the individual 300-line error spectra with the 400-line error template does not affect the line measurements in the composite spectra significantly. Removing the 300-line grating and grism objects from the composites (and therefore having “clean” $z \sim 2$ and $z \sim 3$ samples of only 400- and 600-line objects with well-reconstructed individual error spectra) results in a change of only $\lesssim 8\%$ or within the 1σ uncertainty for the $\text{EW}_{\text{Ly}\alpha}$ measurements and $\lesssim 5\%$ of change in the EW_{LRIS} .

With estimates of the individual error spectra of the $z \sim 2$ and $z \sim 3$ galaxies in hand, we bootstrap-resampled the objects in each bin and perturbed each spectrum (in the L_λ space) in the bootstrap sample according to its own error spectrum. The perturbed spectra in the bootstrap sample were then scaled and combined (following the same procedure of constructing the science composites) to create a new composite spectrum. The process was repeated 100 times and the standard deviation of these 100 fake composites at each wavelength was adopted as the composite error spectrum for each bin.

The construction of the composite error spectra for the $z \sim 4$ sample is a bit different. While the DEIMOS objects have available individual error spectra, the FORS2 objects do not. Considering that the individual error spectra were only needed for creating the composite error spectra, we evaluated the relative contribution of bootstrap resampling and individual error spectra to the overall uncertainty for both the DEIMOS and FORS2 objects, based on their proportion in the $z \sim 4$ sample and typical S/N.¹⁷ We first calculated the ratio of the uncertainty estimated from “bootstrap only” to that from both bootstrap and individual error spectra for the DEIMOS data. The median value of the ratio is 0.404, which was estimated over 1250–1400 Å, a spectral region within which DEIMOS has a decent sensitivity. Given that the DEIMOS spectra on average have lower S/Ns at these wavelengths than the FORS2 spectra (median S/N = 2.76 and 4.07, respectively), we scaled the individual DEIMOS error spectra down by a factor of $4.07/2.76 = 1.47$ to match the S/N of the FORS2 spectra, assuming that the FORS2 error spectra have the same shape as the DEIMOS error spectra. We repeated the same process, estimating the ratio between the “bootstrap only” and “bootstrap + individual error spectra” for the DEIMOS spectra scaled to the S/N of the FORS2 spectra, and the resulting median ratio is 0.564. These two ratios were combined with weights that correspond to the proportion of the DEIMOS and

¹⁷ As the $z \sim 4$ sample mainly consists of DEIMOS and FORS2 objects (80 out of 91 spectra), we did not take into account the $z > 3.4$ LRIS galaxies and their reconstructed individual error spectra when estimating the overall uncertainty.

FORS spectra in the entire $z \sim 4$ sample, yielding an overall ratio of 0.460. As a result, the overall error composite spectra of the $z \sim 4$ sample was determined by dividing the error spectra evaluated from “bootstrap only” by a factor of 0.460 to account for the contribution of individual error spectra with different S/Ns.

Appendix B IR-excess Galaxies

The process of SED modeling reveals a small fraction of objects that stand out because of their flux excess in the IRAC channels relative to the best-fit galaxy SED template. Studies have shown that monotonically rising IRAC SEDs can be an effective tool for selecting AGNs (Reddy et al. 2006; Donley et al. 2012; Hainline et al. 2012), complementary to other AGN identification methods based on X-ray properties or rest-frame optical and UV spectra. By studying a sample of 33 AGNs identified on the basis of narrow, high-ionization emission lines in the rest-frame UV, Hainline et al. (2012) found that 11 out of 16 objects with IRAC photometry show a monotonically increasing flux in the IRAC channels toward longer wavelength. These authors discovered that the “excess” in the IRAC bands could be fit by an additional power law on top of the best-fit galaxy SED to represent the emission from hot dust associated with the AGN. Motivated by the idea that the flux excess in the IRAC bands can possibly be used as an indicator of AGN emission, we select galaxies with IR excesses in our samples and compare their rest-UV spectra with those showing no flux excess in the mid-IR.

Fortunately, the $z \sim 2-4$ LBG spectra cover multiple high-ionization emission lines characteristic of AGN spectra, including N V $\lambda 1240$, C IV $\lambda 1549$, and He II $\lambda 1640$. These emission features are clear indicators of photoionization by a nonstellar source and are therefore typically weak and undetected individually in our redshift samples of star-forming galaxies. However, the use of composite spectra makes it

possible to study these lines in greater detail and examine potential AGN activity in the LBGs in our samples.

Although objects with high-ionization UV emission features detected on an individual basis were classified as AGNs and therefore already removed from the samples presented in this work prior to the construction of composite spectra, there may exist an underlying, low-level AGN contribution in the remaining sample that can only be identified within composite spectra. In order to select objects with potential low-level AGN activity, we searched for objects in our redshift samples based on a $\geq 2\sigma$ flux excess threshold relative to the best-fit SED in (1) either IRAC channel 3 or channel 4 if only one band is detected, or (2) both channel 3 and channel 4 if both bands are detected. In the latter case, we further required that channel 4 has a higher flux density (lower value in magnitude) than channel 3 to ensure the shape of a power law. The flux-density excess was calculated by comparing the observed magnitude in corresponding bands with the theoretical magnitude derived from the best-fit galaxy SED fit excluding IRAC channels 3 and 4. Figure 13 shows the SED of an IR-excess galaxy (GOODS-BX1100 in the $z \sim 2$ sample) selected based on these criteria.

A total of 55, 28, and 4 objects were selected based on the flux-density excess for the $z \sim 2$, $z \sim 3$, and $z \sim 4$ samples, respectively. We did not stack the $z \sim 4$ “excess” spectra given that this sample is too small to be considered a statistical one. Figure 14 shows the comparison of the composite spectra with IRAC excess (red) and without (black) for the $z \sim 2$ (top) and $z \sim 3$ samples (bottom) near the wavelengths of N V, C IV, and He II. For a better visual comparison, we have smoothed the “excess” spectrum to match the pixel-to-pixel rms of its “non-excess” counterpart in order to account for the difference in continuum S/N. We found that among the three high-ionization emission features, N V is the strongest in all composites. The C IV emission is weak, and He II is almost not detected at all. It is worth noting that the “excess” and

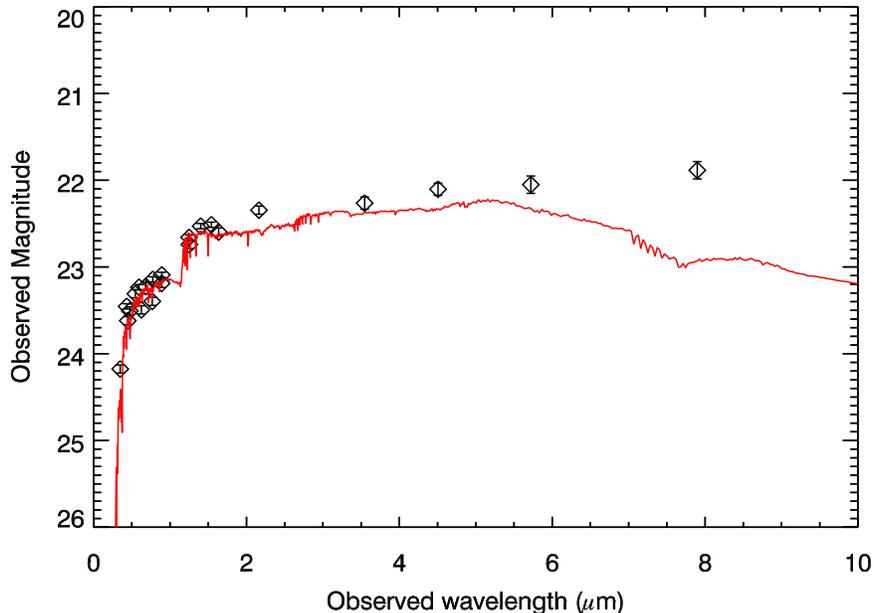


Figure 13. Example SED of one of the IR-excess galaxies, GOODS-BX1100 at $z = 2.08$, selected based on the $\geq 2\sigma$ flux excess criteria in IRAC channels 3 and 4, as described in Appendix B. The diamonds represent the observed magnitudes in the photometric bands used for SED modeling (Table 1), and the red curve shows the best-fit galaxy SED excluding photometric data from IRAC channels 3 and 4.

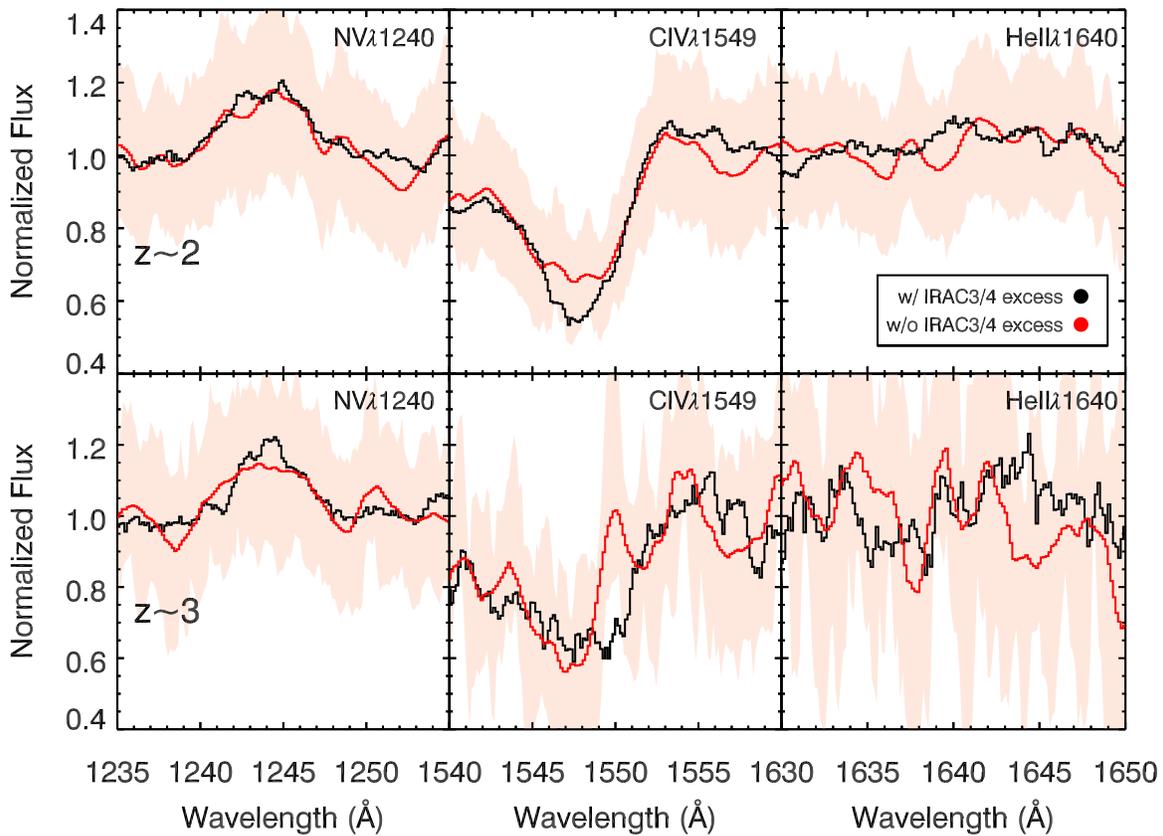


Figure 14. Comparison of the AGN signature profiles (N V, C IV, and He II) between the stacks of objects showing a $\geq 2\sigma$ excess in either IRAC channel 3 or 4 (red) and those without (black) for the $z \sim 2$ (top) and $z \sim 3$ (bottom) samples. The light red shaded area represents the 1σ uncertainty level of the “excess” composites. The $z \sim 4$ composites are not plotted because of the small number of objects meeting the flux excess threshold. Given the much smaller size of the “excess” sample compared to that of the “non-excess” sample, the “excess” composites have been smoothed to match the continuum S/N of the counterpart “non-excess” composites. This smoothing minimizes the visual difference caused by the pixel-to-pixel rms.

“non-excess” composites look fairly similar in that the former do not show specifically stronger N V, C IV, or He II emission. This result suggests that the IR-excess galaxies in our star-forming galaxy samples—which are already cleaned of IR-excess galaxies identified individually as AGNs based on strong rest-UV high-ionization lines—may not be AGNs. The flux-density excess in the mid-IR may suggest evidence for hot dust due to active star formation, as observed in some local dwarf galaxies (Hainline et al. 2016). Alternatively, the minor AGN activity in these galaxies, if any, is highly obscured and does not result in distinct features in the rest-frame UV.

ORCID iDs

Naveen A. Reddy  <https://orcid.org/0000-0001-9687-4973>
 Tucker Jones  <https://orcid.org/0000-0001-5860-3419>
 Charles C. Steidel  <https://orcid.org/0000-0002-4834-7260>
 Allison L. Strom  <https://orcid.org/0000-0001-6369-1636>
 Gwen C. Rudie  <https://orcid.org/0000-0002-8459-5413>
 Dawn K. Erb  <https://orcid.org/0000-0001-9714-2758>
 Richard S. Ellis  <https://orcid.org/0000-0001-7782-7071>

References

Berg, D. A., Skillman, E. D., Henry, R. B. C., Erb, D. K., & Carigi, L. 2016, *ApJ*, 827, 126
 Berry, M., Gawiser, E., Guaita, L., et al. 2012, *ApJ*, 749, 4
 Bordoloi, R., Lilly, S. J., Knobel, C., et al. 2011, *ApJ*, 743, 10
 Bouwens, R. J., Aravena, M., Decarli, R., et al. 2016, *ApJ*, 833, 72

Bouwens, R. J., van Dokkum, P. G., Illingworth, G. D., et al. 2017, arXiv:1711.02090
 Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, *ApJS*, 200, 13
 Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000
 Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, 533, 682
 Chabrier, G. 2003, *PASP*, 115, 763
 Chen, Y.-M., Tremonti, C. A., Heckman, T. M., et al. 2010, *AJ*, 140, 445
 Chisholm, J., Orlitová, I., Schaerer, D., et al. 2017, *A&A*, 605, A67
 de Barros, S., Vanzella, E., Amorín, R., et al. 2016, *A&A*, 585, A51
 Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, *ApJ*, 748, 142
 Du, X., Shapley, A. E., Martin, C. L., & Coil, A. L. 2016, *ApJ*, 829, 64
 Du, X., Shapley, A. E., Martin, C. L., & Coil, A. L. 2017, *ApJ*, 838, 63
 Erb, D. K., Pettini, M., Shapley, A. E., et al. 2010, *ApJ*, 719, 1168
 Erb, D. K., Pettini, M., Steidel, C. C., et al. 2016, *ApJ*, 830, 52
 Erb, D. K., Quider, A. M., Henry, A. L., & Martin, C. L. 2012, *ApJ*, 759, 26
 Erb, D. K., Steidel, C. C., Shapley, A. E., et al. 2006a, *ApJ*, 647, 128
 Erb, D. K., Steidel, C. C., Shapley, A. E., et al. 2006b, *ApJ*, 646, 107
 Erb, D. K., Steidel, C. C., Trainor, R. F., et al. 2014, *ApJ*, 795, 33
 Faber, S. M., Phillips, A. C., Kibrick, R. I., et al. 2003, *Proc. SPIE*, 4841, 1657
 Faisst, A. L., Capak, P. L., Davidzon, I., et al. 2016, *ApJ*, 822, 29
 Feldmann, R., Quataert, E., Hopkins, P. F., Faucher-Giguère, C.-A., & Kereš, D. 2017, *MNRAS*, 470, 1050
 Finley, H., Bouché, N., Contini, T., et al. 2017, *A&A*, 605, A118
 Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, *ApJL*, 600, L93
 Guaita, L., Talia, M., Pentericci, L., et al. 2017, *A&A*, 606, A19
 Hainline, K. N., Reines, A. E., Greene, J. E., & Stern, D. 2016, *ApJ*, 832, 119
 Hainline, K. N., Shapley, A. E., Greene, J. E., et al. 2012, *ApJ*, 760, 74
 Hathi, N. P., Le Fèvre, O., Ilbert, O., et al. 2016, *A&A*, 588, A26
 Heckman, T. M., Lehnert, M. D., Strickland, D. K., & Armus, L. 2000, *ApJS*, 129, 493
 Henry, A., Scarlata, C., Martin, C. L., & Erb, D. 2015, *ApJ*, 809, 19
 Jaskot, A. E., & Ravindranath, S. 2016, *ApJ*, 833, 136
 Jones, T., Stark, D. P., & Ellis, R. S. 2012, *ApJ*, 751, 51

- Jones, T. A., Ellis, R. S., Schenker, M. A., & Stark, D. P. 2013, *ApJ*, 779, 52
- Konno, A., Ouchi, M., Ono, Y., et al. 2014, *ApJ*, 797, 16
- Konno, A., Ouchi, M., Shibuya, T., et al. 2018, *PASJ*, 70, S16
- Kornei, K. A., Shapley, A. E., Erb, D. K., et al. 2010, *ApJ*, 711, 693
- Kornei, K. A., Shapley, A. E., Martin, C. L., et al. 2012, *ApJ*, 758, 135
- Kornei, K. A., Shapley, A. E., Martin, C. L., et al. 2013, *ApJ*, 774, 50
- Kulas, K. R., Shapley, A. E., Kollmeier, J. A., et al. 2012, *ApJ*, 745, 33
- Laursen, P., Sommer-Larsen, J., & Razoumov, A. O. 2011, *ApJ*, 728, 52
- Leethochawalit, N., Jones, T. A., Ellis, R. S., Stark, D. P., & Zitrin, A. 2016, *ApJ*, 831, 152
- Livermore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, *ApJ*, 835, 113
- Madau, P., & Dickinson, M. 2014, *ARA&A*, 52, 415
- Marchi, F., Pentericci, L., Guaita, L., et al. 2017, *A&A*, 601, A73
- Markwardt, C. B. 2009, in ASP Conf. Ser. 411, *Astronomical Data Analysis Software and Systems XVIII*, ed. D. A. Bohlender, D. Durand, & P. Dowler (San Francisco, CA: ASP), 251
- Martin, C. L., Shapley, A. E., Coil, A. L., et al. 2012, *ApJ*, 760, 127
- Maseda, M. V., Brinchmann, J., Franx, M., et al. 2017, *A&A*, 608, A4
- Newman, J. A., Cooper, M. C., Davis, M., et al. 2013, *ApJS*, 208, 5
- Oesch, P. A., Labbé, I., Bouwens, R. J., et al. 2013, *ApJ*, 772, 136
- Oke, J. B., Cohen, J. G., Carr, M., et al. 1995, *PASP*, 107, 375
- Onodera, M., Carollo, C. M., Lilly, S., et al. 2016, *ApJ*, 822, 42
- Pentericci, L., Grazian, A., Fontana, A., et al. 2007, *A&A*, 471, 433
- Pentericci, L., Grazian, A., Fontana, A., et al. 2009, *A&A*, 494, 553
- Pentericci, L., Grazian, A., Scarlata, C., et al. 2010, *A&A*, 514, A64
- Pentericci, L., Vanzella, E., Fontana, A., et al. 2014, *ApJ*, 793, 113
- Pettini, M., Rix, S. A., Steidel, C. C., et al. 2002, *ApJ*, 569, 742
- Prochaska, J. X., Kasen, D., & Rubin, K. 2011, *ApJ*, 734, 24
- Quider, A. M., Pettini, M., Shapley, A. E., & Steidel, C. C. 2009, *MNRAS*, 398, 1263
- Reddy, N. A., Oesch, P. A., Bouwens, R. J., et al. 2018, *ApJ*, 853, 56
- Reddy, N. A., Pettini, M., Steidel, C. C., et al. 2012, *ApJ*, 754, 25
- Reddy, N. A., Steidel, C. C., Erb, D. K., Shapley, A. E., & Pettini, M. 2006, *ApJ*, 653, 1004
- Reddy, N. A., Steidel, C. C., Pettini, M., et al. 2008, *ApJS*, 175, 48
- Reddy, N. A., Steidel, C. C., Pettini, M., Bogosavljević, M., & Shapley, A. E. 2016, *ApJ*, 828, 108
- Rigby, J. R., Bayliss, M. B., Gladders, M. D., et al. 2015, *ApJL*, 814, L6
- Rix, S. A., Pettini, M., Leitherer, C., et al. 2004, *ApJ*, 615, 98
- Rubin, K. H. R., Prochaska, J. X., Koo, D. C., et al. 2014, *ApJ*, 794, 156
- Rudie, G. C., Steidel, C. C., Trainor, R. F., et al. 2012, *ApJ*, 750, 67
- Rupke, D. S., Veilleux, S., & Sanders, D. B. 2005, *ApJS*, 160, 115
- Sanders, R. L., Shapley, A. E., Kriek, M., et al. 2015, *ApJ*, 799, 138
- Schenker, M. A., Ellis, R. S., Konidaris, N. P., & Stark, D. P. 2014, *ApJ*, 795, 20
- Senchyna, P., Stark, D. P., Vidal-García, A., et al. 2017, *MNRAS*, 472, 2608
- Shapley, A. E., Steidel, C. C., Adelberger, K. L., et al. 2001, *ApJ*, 562, 95
- Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, *ApJ*, 588, 65
- Shapley, A. E., Steidel, C. C., Strom, A. L., et al. 2016, *ApJL*, 826, L24
- Shibuya, T., Ouchi, M., & Harikane, Y. 2015, *ApJS*, 219, 15
- Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, *ApJS*, 214, 24
- Stark, D. P., Ellis, R. S., Bunker, A., et al. 2009, *ApJ*, 697, 1493
- Stark, D. P., Ellis, R. S., Charlot, S., et al. 2017, *MNRAS*, 464, 469
- Stark, D. P., Ellis, R. S., Chiu, K., Ouchi, M., & Bunker, A. 2010, *MNRAS*, 408, 1628
- Stark, D. P., Richard, J., Charlot, S., et al. 2015, *MNRAS*, 450, 1846
- Stark, D. P., Richard, J., Siana, B., et al. 2014, *MNRAS*, 445, 3200
- Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2003, *ApJ*, 592, 728
- Steidel, C. C., Erb, D. K., Shapley, A. E., et al. 2010, *ApJ*, 717, 289
- Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, *ApJ*, 795, 165
- Steidel, C. C., Shapley, A. E., Pettini, M., et al. 2004, *ApJ*, 604, 534
- Steidel, C. C., Strom, A. L., Pettini, M., et al. 2016, *ApJ*, 826, 159
- Strom, A. L., Steidel, C. C., Rudie, G. C., et al. 2017, *ApJ*, 836, 164
- Sugahara, Y., Ouchi, M., Lin, L., et al. 2017, *ApJ*, 850, 51
- Trainor, R. F., Steidel, C. C., Strom, A. L., & Rudie, G. C. 2015, *ApJ*, 809, 89
- Trainor, R. F., Strom, A. L., Steidel, C. C., & Rudie, G. C. 2016, *ApJ*, 832, 171
- Treu, T., Trenti, M., Stiavelli, M., Auger, M. W., & Bradley, L. D. 2012, *ApJ*, 747, 27
- Tumlinson, J., Peebles, M. S., & Werk, J. K. 2017, *ARA&A*, 55, 389
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2005, *A&A*, 434, 53
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2006, *A&A*, 454, 423
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2008, *A&A*, 478, 83
- Vanzella, E., De Barros, S., Cupani, G., et al. 2016, *ApJL*, 821, L27
- Vanzella, E., Giavalisco, M., Dickinson, M., et al. 2009, *ApJ*, 695, 1163
- Vanzella, E., Guo, Y., Giavalisco, M., et al. 2012, *ApJ*, 751, 70
- Verhamme, A., Schaerer, D., Atek, H., & Tapken, C. 2008, *A&A*, 491, 89
- Verhamme, A., Schaerer, D., & Maselli, A. 2006, *A&A*, 460, 397
- York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, *AJ*, 120, 1579
- Zitrin, A., Ellis, R. S., Belli, S., & Stark, D. P. 2015, *ApJL*, 805, L7