RECONCILING THE STELLAR AND NEBULAR SPECTRA OF HIGH-REDSHIFT GALAXIES[*]

CHARLES C. STEIDEL[1], ALLISON L. STROM[1], MAX PETTINI[2], GWEN C. RUDIE[3], NAVEEN A. REDDY[4,6], AND RYAN F. TRAINOR[5,7]

[1] Cahill Center for Astronomy and Astrophysics, California Institute of Technology, MS 249-17, Pasadena, CA 91105, USA
[2] Institute of Astronomy, Madingley Road, Cambridge CB3 OHA, UK
[3] Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101, USA
[4] Department of Physics and Astronomy, University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA
[5] Department of Astronomy, University of California, Berkeley, Campbell Hall, Berkeley, CA 94720 USA

ABSTRACT

We present a combined analysis of rest-frame far-UV (FUV; 1000–2000 Å) and rest-frame optical (3600–7000 Å) composite spectra formed from very deep Keck/LRIS and Keck/MOSFIRE observations of a sample of 30 star-forming galaxies with $z = 2.40 \pm 0.11$, selected to be broadly representative of the full KBSS-MOSFIRE spectroscopic survey. Since the same massive stars are responsible for the observed FUV continuum *and* for the excitation of the observed nebular emission, a self-consistent stellar population synthesis model should simultaneously match the details of the FUV stellar+nebular continuum and—when inserted as the excitation source in photoionization models—predict all observed nebular emission line ratios. We find that only models including massive star binaries, having low stellar metallicity ($Z_*/Z_\odot \simeq 0.1$) but relatively high nebular (ionized gas-phase) abundances ($Z_{neb}/Z_\odot \simeq 0.5$), can successfully match all of the observational constraints. We show that this apparent discrepancy is naturally explained by highly super-solar O/Fe ($\simeq$4–5 (O/Fe)$_\odot$), expected for a gas whose enrichment is dominated by the products of core-collapse supernovae. While O dominates the physics of the ionized gas (and thus the nebular emission lines), Fe dominates the extreme-UV (EUV) and FUV opacity and controls the mass-loss rate from massive stars, resulting in particularly dramatic effects for massive stars in binary systems. This high nebular excitation—caused by the hard EUV spectra of Fe-poor massive stars—is much more common at high redshift ($z \gtrsim 2$) than low redshift due to systematic differences in the star formation history of typical galaxies.

Key words: cosmology: observations – galaxies: evolution – galaxies: high-redshift – H II regions – ISM: abundances – stars: massive

1. INTRODUCTION

The formation and evolution of massive stars affect most—if not all—observable properties of the most distant galaxies. Their winds, radiation pressure, and supernova explosions are the principal sources of feedback that ultimately control the galaxy formation process (e.g., Murray et al. 2005; Faucher-Giguère et al. 2013; Hopkins et al 2014; Muratov et al. 2015) and they dominate the bolometric luminosity of rapidly forming galaxies, whether most of the luminosity emerges in the UV or is re-radiated in the far-IR. Massive stars are responsible for producing—and, arguably, dispersing—most of the metals that have enriched galaxies and the circumgalactic and intergalactic media (CGM and IGM) over cosmic time, and were also the most likely sources responsible for producing the photons that reionized the universe at $z \gtrsim 7$ (e.g., Robertson et al. 2015).

In spite of the crucial role played by massive stars in all of these baryonic processes, we know surprisingly little about them. Contributing to our ignorance is the fact that massive stars are naturally short-lived, and therefore any accounting must be done in situ at the epoch of their formation. Short lifetimes also mean that massive stars die close to their birthplace and these sites are often heavily obscured by dense

ISM and dust. When not obscured, massive stars emit most of their energy in the UV, requiring observations from space for nearby examples. Although they produce most of the H-ionizing photons within galaxies and, at $z \gtrsim 2$, in the IGM medium as well, no direct observation of the ionizing spectrum of an O-star has ever been made, at any redshift or metallicity.

On top of observational challenges, massive stars are also more difficult to model than their less massive counterparts. The ionizing spectra, main-sequence lifetimes, evolutionary stages, end states, and chemical yields from massive stars all depend on the details of mass loss (e.g., Smith 2014), which itself depends on the initial chemical composition (e.g., Kudritzki & Puls 2000; Cantiello et al. 2007; Puls et al. 2008; Brott et al. 2011; Eldridge et al. 2011; Langer 2012). The last decade has witnessed a rapidly changing paradigm for massive star winds due to the appreciation that the wind material is clumpy, reducing the mass-loss rates by factors of $\sim$2–3 compared to what was believed previously (Smith 2014, and references therein). In parallel, it has also been established that *most* massive stars do not evolve as single stars, but as part of binary systems (Crowther 2007; Sana et al. 2012). Multiplicity affects nearly all aspects of the expected evolutionary path of massive stars, and can also lead to phenomena that could not occur for single massive stars, particularly in low-metallicity environments. For example, binary massive stars with metallicities $Z_* < 0.004$ ($Z_*/Z_\odot < 0.35$) can lead to "quasi-homogeneous evolution" (QHE; Yoon & Langer 2005; Cantiello et al. 2007; Brott et al. 2011; Eldridge et al. 2011; Eldridge & Stanway 2012) in which the more massive star over-flows its Roche lobe and dumps mass onto the secondary,

which (because of weak winds inherent to lower stellar metallicity) accretes rapidly and spins up to high rotational velocity, causing it to become fully mixed. Such stars can burn all of their H to He, last up to 3 times longer on the main sequence than a single star of the same mass, would not produce stellar winds, and become progressively *hotter* as they evolve away from the zero-age main sequence. QHE could produce sources of extreme-UV (EUV) radiation that would mimic classical Wolf–Rayet stars, but would be much longer-lived, would require less massive progenitors, and may be not have the emission features of C or N.

In the context of ionizing photon production, since the more massive donor star reaches its end state first, the hot Wolf–Rayet-like companion may receive a kick capable of perturbing it well out of its birth cloud—potentially resolving the puzzle of how stellar ionizing photons manage to escape their immediate surroundings to contribute to the ionization of the IGM. As we discuss below, there may be significant observational evidence that QHE is important, particularly at high redshift.

Most importantly, the uncertainties associated with massive star populations have profound effects on the interpretation of observations of the high-redshift universe. The recent sea-changes in our understanding of massive stars—lower mass loss, the dominance of binary evolution—have been slow to percolate to the galaxy formation community, and most investigations that make use of massive star population synthesis use models based on key assumptions now known to be incorrect, or at least incomplete. Quantities that may be subject to revision include estimates of star formation rate (SFR), stellar mass (M_*), stellar population ages from spectral energy distribution (SED) fitting, the ratio of ionizing to non-ionizing-UV photons produced by young stellar populations, the rate of core-collapse supernovae and associated chemical yields, and the measurement of gas-phase metallicities from observations of nebular emission lines. These uncertainties propagate through to affect our understanding of stellar and supernova feedback, gas accretion, and outflows, and the chemical evolution of galaxies—in short, all of the details relevant to baryonic processes in forming galaxies.

The goal of this paper is to seek a deeper understanding of these processes by combining two complementary observational signatures of massive stars in forming galaxies—the rest-frame far-UV (FUV) continuum ($1000 \lesssim \lambda_0 \lesssim 2000$ Å) and nebular emission lines from H II regions—in the same galaxies. Clearly the two are complementary as well as causally connected: the FUV continuum provides direct access to the light produced by OB stars, encoding information on stellar metallicity, initial mass function (IMF), population age, SFR, continuum attenuation, etc. The nebular spectrum is the response of the galaxy ISM to the ionizing-UV continuum ($200 \lesssim \lambda_0 \lesssim 900$ Å)—which cannot be observed directly—from the same OB stars, and is also sensitive to the physical conditions (metallicity, electron density n_e) in the ionized gas.

Because many diagnostic nebular emission lines for H II regions are found in the rest-wavelength range $3500 \lesssim \lambda_0 \lesssim 9000$, easily accessible from the ground for nearby galaxies, there is a long history of nebular spectroscopy and its application to measuring abundances and other gas-phase physical parameters. Large optical spectroscopic data-bases such as the Sloan Digital Sky Survey (SDSS) are now playing a major role in establishing a statistical baseline in the $z \sim 0$ universe. However, complementary FUV spectroscopy of low-redshift galaxies requires observations from space, where the absence of multiplexing capability, limitations in sensitivity and wavelength coverage, and competition for scarce *Hubble Space Telescope* (*HST*) resources has kept the samples small.[8]

The situation is nearly the reverse in the high-redshift universe, where a large fraction of what we know about star-forming galaxies at $z \gtrsim 1.5$ comes from rest-UV spectra obtained using multiplexed optical spectrometers on large ground-based telescopes. Rest-frame optical spectroscopy has been slower to develop, awaiting sensitive multiplexed near-IR spectroscopic capabilities. Nevertheless, it has long been appreciated that, once such capabilities arrive, there are specific redshift ranges which maximize access to diagnostic nebular emission lines due to fortuitous placement with respect to the atmospheric transmission windows: $1.4 \lesssim z \lesssim 1.7$, $2.0 \lesssim z \lesssim 2.6$, and $3 \lesssim z \lesssim 3.8$. Of these intervals, the first is too low to provide full access (from the ground) to the rest-frame FUV, and the last is too high to permit observations of Hα due to prohibitively bright thermal backgrounds.

The redshift range $z = 2$–2.6, however, is optimal for ground-based observations: it provides access in the J, H, and K band atmospheric windows to a suite of nebular emission lines collectively sensitive to electron density (n_e), SFR, ionization parameter (U), metallicity [$12 + \log(O/H)$], extinction, and the stellar ionizing radiation field in the 1–4 Ryd (13.6–54.4 eV) range. At the same time, $2 \lesssim z \lesssim 2.6$ brings the rest-frame FUV (1000–2000 Å) portion of galaxy spectra above the atmospheric cutoff near 3100 Å, and into a region which enjoys unparalleled sensitivity for direct observation of the integrated FUV light from massive OB stars in the same galaxies (see, e.g., Pettini et al. [2002](#); Shapley et al. [2003](#); Steidel et al. [2003](#), [2004](#), [2010](#); Rix et al. [2004](#)). The information content of the FUV spectra is extremely high: in addition to the integrated stellar UV continuum, with thousands of stellar photospheric absorption features and several strong P-Cygni lines from massive star winds, the $\lambda_0 = 1000$–2000 Å wavelength range includes nebular emission lines that complement those available in the rest-frame optical. These include lines of O III, Si III, N III, and C III—in each case the dominant ionization stage in high-excitation H II regions—as well as the He II $\lambda1640$ emission feature, which can be observed as a nebular recombination line and/or as a broad stellar emission feature in the spectra of very hot stars.

It was clear from the initial KBSS-MOSFIRE sample presented by Steidel et al. ([2014](#), hereafter S14) that standard population synthesis models of massive stars could not easily account for the high-excitation nebular emission observed in most of the high-redshift galaxies. Recognizing the large uncertainties in massive star population synthesis models, S14 used single-temperature blackbody spectra to approximate the integrated stellar ionizing spectrum in the 1–4 Ryd energy range. It was found that the observed ratios of strong nebular

---

emission lines could be reproduced if the energy distribution of the 1–4 Ryd radiation field resembles that of a 55,000–60,000 K blackbody in the *steady state*. Such high ionizing radiation field temperatures can be produced only over very short durations in single-star population synthesis models, over timescales much shorter than the (central) dynamical times of $t_d \gtrsim 20$–30 Myr typical of the $z \sim 2.3$ KBSS-MOSFIRE galaxies (e.g., Erb et al. 2006; Reddy et al. 2008). The fact that these galaxies are both relatively massive ($\langle \log(M_*/M_\odot) \rangle \sim 10$) and very common at high redshift suggests a high duty cycle for the prevailing ionization conditions. S14 speculated further that massive star population synthesis models that include rapidly rotating massive stars (e.g., Brott et al. 2011) and/or that account for massive star binary evolution (Eldridge & Stanway 2012) might mitigate the problem, for the reasons discussed above.

In any case, it seems likely that the use of strong nebular lines to infer gas-phase metallicity—a technique now employed en masse at all redshifts—could be strongly affected by the degeneracy between gas-phase physical conditions and the nature of the ionizing stars. It appears that uncertainties in the EUV stellar ionizing spectrum are at least as significant—if not more so—than uncertainties in gas-phase abundances and physical conditions when it comes to modeling the nebular spectra of high-redshift galaxies. Fortunately, it is now feasible to make considerable headway in constraining the *combination* of massive star properties and nebular physical conditions that can explain the rapidly improving observations.

In this paper, we present the results of a pilot program intended to test our ability to construct a self-consistent model that accounts for all observational constraints provided by joint rest-frame FUV and rest-optical spectra of a representative sample of star-forming galaxies at $z \sim 2.4$. In our view, a successful population synthesis+photoionization model must *simultaneously* match the overall FUV stellar continuum shape, the observed nebular line intensity ratios, and the detailed stellar spectral features (including photospheric line blanketing and stellar wind features) of the UV stellar spectrum. The foundation of the project is a set of high-quality near-IR spectra obtained as part of KBSS-MOSFIRE (S14; A. L. Strom et al. 2016, in preparation), to which we have added new, very deep optical spectra, providing complementary rest-frame FUV spectra of comparable quality, using Keck/LRIS.

The paper is organized as follows: Section 2 details the observations and reductions of the optical and near-IR spectra used for this study; Section 3 describes the stellar population synthesis models, and the procedure used to compare them to the observed rest-frame FUV spectrum, while Section 4 makes more detailed comparisons of the observed and model spectra. Section 5 describes the measurement of nebular emission lines in the rest-frame FUV and rest-frame optical, and Section 6 derives physical parameters from the nebular line ratios that are useful as input to photoionization models of the nebulae. The photoionization models—using the population synthesis models as input—and the detailed comparison of the model grids with the observed nebular emission measurements are covered in Section 7, and ionized gas-phase abundances and abundance ratios are derived in Section 8. Finally, Section 9 discusses the results and their significance in interpreting FUV and nebular spectra of high-redshift galaxies.

Throughout this paper, where required we have assumed a cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $h = 0.7$, and a Chabrier

Table 1
KBSS-LM1 Sample Statistics

Property	Median[a]	Range
Redshift	2.396 ± 0.111[b]	[2.113, 2.572]
$\log(M_*/M_\odot)$	9.8 ± 0.3	[9.0, 10.8]
$SFR_{UV}/(M_\odot \ yr^{-1})$	32.9 ± 26.7	[8, 102]
$SFR_{H\alpha}/(M_\odot \ yr^{-1})$	29.2 ± 17.6	[2, 330]
$\log(sSFR/Gyr^{-1})$	0.54 ± 0.33	[−0.30, +1.45]
$E(B-V)_{sed}$	0.16 ± 0.07	[0.00, 0.32]
$\mathcal{R}$	24.44 ± 0.46	[23.49, 25.43]
K_s	23.35 ± 0.68	[22.12, 24.55]
$U_n - G$	0.90 ± 0.38	[0.37, 2.49]
$G - \mathcal{R}$	0.25 ± 0.09	[0.07, 0.18]
$\mathcal{R} - K_s$	0.69 ± 0.45	[−0.21, +1.99]

Notes.
[a] Error bars on median values are the inter-quartile range.
[b] Mean and rms, rather than median.

(2003) stellar IMF. All wavelengths of spectral lines are referred to using their vacuum values. For definiteness, we generally refer to metallicity using the metal fraction by mass, Z; when conversion between Z and abundances relative to solar are needed, we have assumed that $Z_\odot = 0.0142$, with oxygen abundance $12 + \log(O/H) = 8.69$ (Asplund et al. 2009). Unless indicated otherwise, solar abundance ratios are assumed.

2. OBSERVATIONS AND DATA

We refer to the sample of galaxies analyzed in this paper as KBSS-LM1: Table 1 summarizes its most important properties. The data presented in this paper are drawn from the first two Keck/LRIS (Oke et al. 1995; Steidel et al. 2004) slitmasks observed as part of a larger program to obtain high-quality rest-UV spectra for a sample of $\simeq100$ galaxies with $\langle z \rangle \simeq 2.4$ drawn from the deepest subsample of the KBSS-MOSFIRE survey (S14; Strom et al. 2016). The main criteria used to select targets for LRIS observations are: (1) the redshift lies in the range $2.10 \leqslant z_{gal} \leqslant 2.58$ so that a full set of strong rest-frame optical nebular emission lines is accessible in the near-IR atmospheric windows (see e.g., Steidel et al. 2014) and (2) the galaxy observations with MOSFIRE have already been completed in the J, H, and K bands, and have yielded emission line measurements or sensitive limits on the fluxes of the following emission lines: [O II]$\lambda\lambda$3727, 3729, [Ne III]λ3870 (J band); Hβ and [O III]$\lambda\lambda$4960, 5008 (H band); and Hα, [N II]$\lambda\lambda$6549, 6585, and [S II]$\lambda\lambda$6718, 6732 (K band). Subsequent to applying these criteria, we attempted to ensure that the observed sample spans the full range of properties in the KBSS-MOSFIRE sample, in terms of inferred SFR, M_*, and metallicity. Any remaining space on the slitmasks was assigned to galaxies drawn from the parent KBSS photometric sample with no prior rest-UV spectra. The choice of KBSS field was then dictated by the number of high-priority galaxies accommodated on a single slitmask, and by the time of scheduled observations. One mask each was observed in the KBSS1442+295 and KBSS2343+125 fields, as summarized in Table 2; Table 2 also summarizes the available MOSFIRE observations of the high-priority targets. Each of the two slitmasks discussed here includes 15 high-priority KBSS-MOSFIRE galaxies, for a total of 30.

Table 2
Summary of Observations

Field	Telescope	Instrument	Wavelength Range (Å)	Integration Time (s)	Dates
KBSS2343+125	Keck I 10 m	LRIS-B 600/4000	3400–5100	37,800	2014 Sep 19,20,21
	Keck I 10 m	LRIS-R 600/5000	4900–7650	35,880	2014 Sep 19,20,21
	Keck I 10 m	MOSFIRE J	11,500–13,500	14,420[a]	2012–2014 Sep/Oct
	Keck I 10 m	MOSFIRE H	14,300–18,100	17,500[a]	2012–2014 Sep/Oct
	Keck I 10 m	MOSFIRE K	19,400–24,060	22,150[a]	2012–2014 Sep/Oct
KBSS1442+295	Keck I 10 m	LRIS-B 600/4000	3400–5700	36,000	2015 May 17,18,19
	Keck I 10 m	LRIS-R 600/5000	5500–8000	37,100	2015 May 17,18,19
	Keck I 10 m	MOSFIRE J	11,500–13,500	9260[a]	2013–2014 May/Jun
	Keck I 10 m	MOSFIRE H	14,300–18,100	18,900[a]	2013–2014 May/Jun
	Keck I 10 m	MOSFIRE K	19,400–24,060	9260[a]	2013–2014 May/Jun

Note.
[a] Average total integration time for the galaxies included in this study.

2.1. MOSFIRE Observations

Observations and data reduction for the KBSS-MOSFIRE survey are described in detail elsewhere (Steidel et al. 2014; Strom et al. 2016). Here, we briefly address the issue of "cross-band" calibration, i.e., ensuring that emission line fluxes measured in each atmospheric band are corrected for slit losses so that ratios of emission lines falling in different bands are accurate. Considerable care was used to establish the cross-band calibration for KBSS-MOSFIRE observations, described in detail by Strom et al. (2016). We use a combination of "slit star" observations—the placement of a star with good broad-band photometric measurements (bright enough to ensure a continuum spectrum of high signal-to-noise ratio; S/N) on each observed mask—and multiple observations of a given target on independent masks, to correct for slit losses on an object-by-object and band-by-band basis. Each slit-loss-corrected spectrum was assigned a quality flag based on the internal consistency of measurements on different masks after applying the nominal correction from the slit star measurements. All spectra in the current KBSS-LM1 sample were assigned the highest quality flag, meaning that the relative flux calibrations have estimated systematic errors of <10%.

The slit-loss-corrected J, H, and K band spectra of the 30 galaxies comprising the KBSS-LM1 sample were shifted into the rest-frame according to the measured nebular redshift z_{neb};[9] the flux scale of each spectrum was adjusted to account for the small differences in redshift among the sample so that a line of the same luminosity and rest wavelength would map to the same line flux in each spectrum after resampling to a common rest-frame wavelength scale. In practice, the relative scale factors are small ($\lesssim$10%) and make little difference to the final results as long as they are applied in the same way to the spectra in each band. The final stacked spectrum in each band was formed by averaging all unmasked pixels at each dispersion point; unless otherwise noted, the composite spectra are used for measurements only where the wavelength of a spectral feature was observed for all 30 objects. The 1σ error spectra for each object were propagated to produce a statistical 1σ error estimate at each dispersion point in the composite spectrum within each atmospheric band (J, H, and K). Portions

of the resulting KBSS-LM1 composite spectra, together with their 1σ error spectra, are shown in Figure 1; all 30 individual objects contributed to the composite spectra over the entire wavelength regions shown.

2.2. LRIS-B+R Observations

The LRIS spectrograph configuration was chosen to optimize the sensitivity and spectral resolution over the observed wavelength range $3400 \lesssim \lambda \lesssim 7300$ Å, corresponding to a rest-wavelength range of $1000 \lesssim \lambda_0 \lesssim 2150$ Å at the mean redshift of the sample, $\langle z \rangle = 2.396 \pm 0.111$. We used a 600 line/mm grism blazed at 4000 Å in the blue channel (hereinafter LRIS-B), and a 600 line/mm grating blazed at 5000 Å in the red channel (hereinafter LRIS-R); see Table 2. For the observations of mask KBSS2343-LM1, we used the d500 dichroic beamsplitter, which sends wavelengths shorter than $\simeq$5000Å to LRIS-B and longer wavelengths to LRIS-R; a configuration using the d560 (with crossover wavelength near 5600 Å) was used for mask KBSS1442-LM1. Both configurations achieve a spectral resolving power of $R \sim 1300–1500$ over the full wavelength range when combined with $1\farcs2$ entrance slits and typical seeing of $0\farcs7$ (see, e.g., Steidel et al. 2010).

The LRIS instrument is located at the Cassegrain focus of the Keck I 10 m telescope, behind the Cassegrain atmospheric dispersion corrector (ADC). The ADC allows slitmask position angles to be chosen to optimize the targets on the mask without concern about operating far from the parallactic angle during the course of the observations. The LRIS-B data were obtained as a series of typically 1800s individual integrations, with the $4k \times 4k$ E2V CCD mosaic binned 1×2 (binning in the dispersion direction). The LRIS-R exposures, obtained contemporaneously, were generally 1200s, with 1×2 binning of the $4k \times 4k$ deep depletion detector as on the blue side. The shorter individual exposure times in the red were used to mitigate the much higher rate of cosmic-ray events owing to the thick CCD substrate of the LBNL deep depletion CCDs (Rockosi et al. 2010). Wavelength calibration was achieved for both LRIS-B and LRIS-R using a combination of Hg, Cd, Zn, and Ne arc lamps obtained during afternoon and morning calibrations; the lamp solutions were adjusted to account for any small illumination or flexure-induced shifts using night sky emission lines observed in the individual science frames.

---

[9] The redshifts were measured independently in each band, with rms difference $\sigma(\Delta v) = 18$ km s^{-1}. The average of the H band and K band redshifts was adopted as z_{neb}.
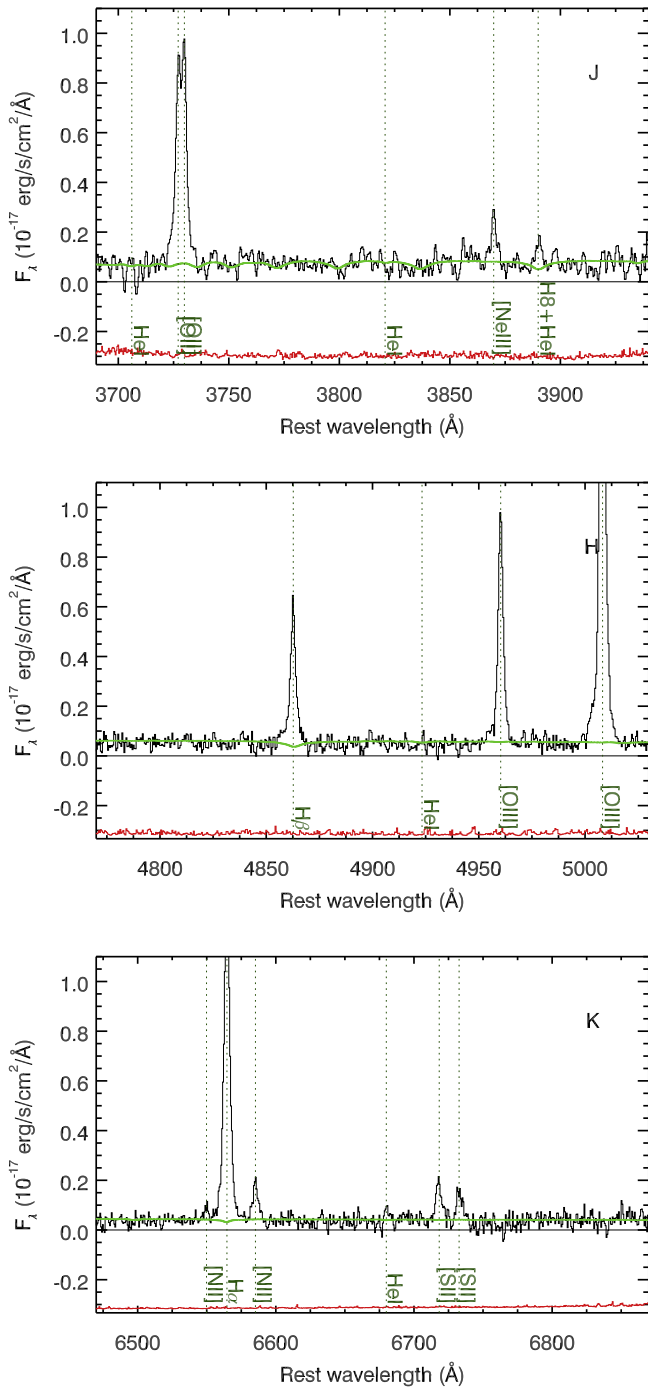
Figure 1. Portions of the MOSFIRE KBSS-LM1 composite spectra (comprised of the same 30 galaxies included in the composite LRIS spectrum in Figure 2) in the J, H, and K bands (top, middle, and bottom panels, respectively). Each panel also shows (red) the 1σ error spectrum for the composite. The locations of various strong emission lines are indicated in each panel; the green curve is the adopted stellar continuum fit. The ordinate in each panel is the average flux density that would be *observed* if all objects were at the mean redshift, $z = 2.396$, while the wavelength scale is shifted to the rest frame.

Typical residuals to the wavelength solutions were $\lesssim 0.10$ Å for both LRIS-B and LRIS-R spectra.

The data were reduced using a suite of custom IRAF and IDL routines, described in detail elsewhere (e.g., Steidel et al. 2003, 2010). In brief, each slit is edge-traced and rectified to remove curvature in the spatial direction (i.e., to produce rectilinear regions for which the traces of dispersed object spectra follow straight lines). The two-dimensional spectrograms for each observation are flat-fielded, cosmic-ray rejected, and background subtracted. A stacked average two-dimensional spectrogram is then formed after shifting each exposure into spatial and spectral registration and averaging all unmasked pixels. The final one-dimensional spectra were extracted from the two-dimensional spectrograms based on the observed profiles of the two-dimensional traces of each galaxy. Night sky spectra were extracted using the same profile used for the object spectra, from frames processed identically to the science spectra except for the background subtraction. These were used to determine small shifts to the wavelength solutions relative to the arc lamp spectra, and to construct error spectra based on a noise model that includes the contributions of Poisson uncertainties associated with the object+sky background counting statistics plus that of the detector read noise. The final one-dimensional spectra were corrected to vacuum, heliocentric wavelengths and resampled to a log-linear dispersion of 65 km s^{-1} pix^{-1} ($\simeq 1/3$ of a spectral resolution element, and close to the original pixel sampling) for both LRIS-B and LRIS-R.

Flux calibration was achieved using observations of spectrophotometric standard stars from the list of Massey et al. (1988); for some purposes, including those addressed in this paper, it is important to tie the flux calibration of LRIS-B and LRIS-R together, and to tie both to line flux measurements observed in the near-IR using MOSFIRE. The former is made possible by the contemporaneous measurement of the standard star (thus both red and blue channels are subject to identical slit losses and atmospheric conditions), and using as a cross-check the wavelength region of overlap between blue and red, within $\simeq \pm 100$ Å of the nominal dichroic cutoff. In general, it was not necessary to apply any scaling between the flux-calibrated spectra in the two different wavelength channels to produce spectra with good relative spectrophotometry. Absolute spectrophotometry was then achieved by comparing the flux-calibrated LRIS-B+R optical spectra to existing broadband photometry of the same objects; such corrections were typically $\lesssim 0.3$ mag.

The final, stacked, $z = 2.4$ KBSS-LM1 rest-UV spectrum was produced in a manner identical to that used for the MOSFIRE spectra discussed above: the data were shifted to the rest frame using $z_{\rm neb}$ measured from the MOSFIRE spectra, the flux scales were adjusted to the values they would have at $z = 2.396$, and averaged without additional weighting. Outlier pixels due to sky subtraction residuals near bright night sky emission lines were excluded, but otherwise all 30 spectra contributed to the stack over the full range in rest wavelength 1000–2150 Å. The final stacked spectrum is shown in Figure 2, together with the 1σ error spectrum propagated from the error spectra of the individual objects. The S/N per pixel in the composite spectrum varies, with $5 \lesssim$ S/N $\lesssim 25$ over the wavelength range shown in Figure 2, with median S/N $\simeq 15$; the S/N per spectral resolution element ($\simeq 3$ pixels) is higher by a factor of ~ 1.7. For the purpose of measuring observed emission line intensities, the stacked rest-frame composite was shifted back to the observed frame using the mean redshift of the sample, $\langle z \rangle = 2.396$, and converted to an f_λ flux density scale.
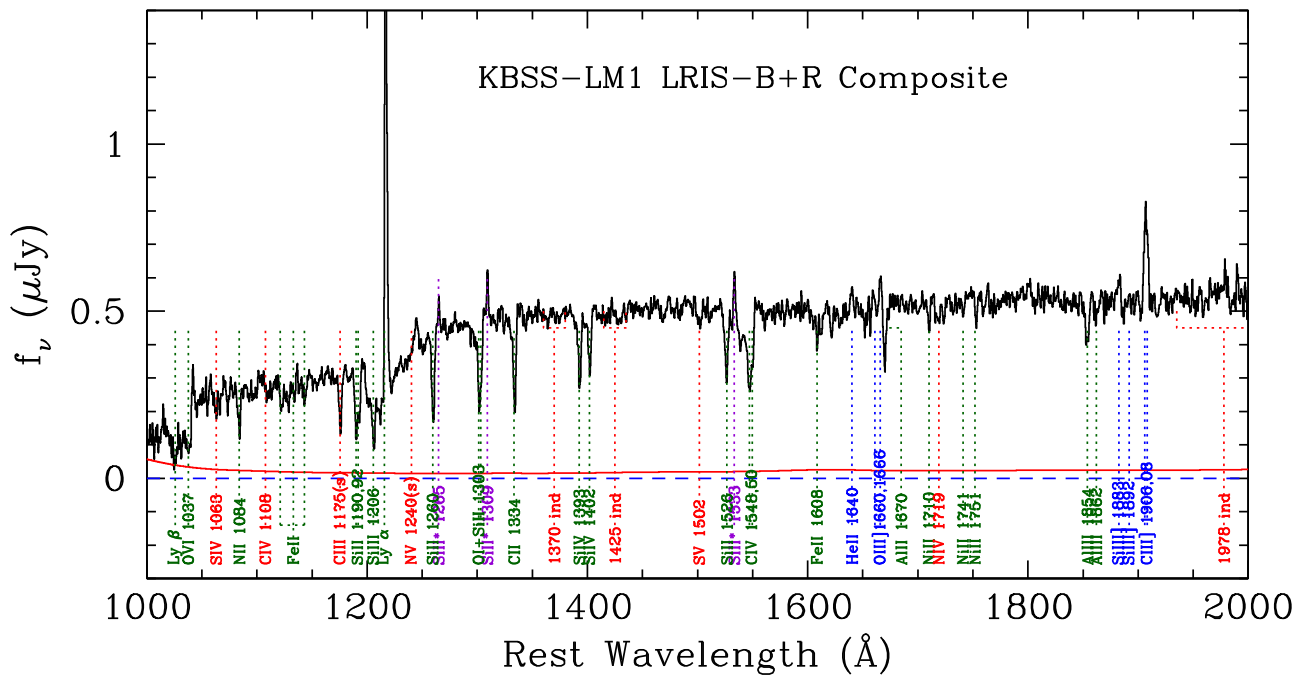
Figure 2. Stacked composite rest-frame UV spectrum of 30 galaxies in the initial KBSS-LM1 sample, with $\langle z \rangle = 2.396 \pm 0.111$ (black histogram). Some prominent emission and absorption features are identified, with color-coded labels: stellar absorption features (red), interstellar absorption features (dark green), nebular emission lines (blue), and excited fine structure emission lines (dark violet). The emission line spectrum is discussed in Section 5.

3. STELLAR POPULATION SYNTHESIS MODELS

In Steidel et al. (2014) we were concerned primarily with constraining the *shape* of the ionizing radiation field responsible for producing the observed nebular emission lines in the KBSS-MOSFIRE $z \sim 2.3$ sample; as a first approximation, we parametrized the ionizing spectrum using a single-temperature blackbody with temperature $T_{\rm eff}$.[10] In this paper, we are interested in a more detailed comparison of the observed FUV spectrum—in both its overall spectral shape and its photospheric and stellar wind features–with state-of-the-art population synthesis models having a range of stellar metallicities and including additional physics deemed potentially important for massive star populations.

As discussed in Section 1, advances in understanding massive stars suggest that a realistic population synthesis of the FUV spectrum of star-forming galaxies should include the effects of lower mass-loss rates, the preponderance of binary evolution, and the related phenomena of rapid stellar rotation and QHE (e.g., Brott et al. 2011; Leitherer et al. 2014), all of which may be essential to correctly predicting the stellar UV spectrum and its evolution with time. As in Steidel et al. (2014), we deliberately refrain from assuming that the stellar metallicity of a population synthesis model need necessarily match the ionized gas-phase abundance in the nebulae; in part this is motivated by the desire to include the widest range of plausible stellar energy distributions, but also because the *stellar* spectra are primarily dependent on the total opacity, which is dominated by Fe, and are relatively insensitive to O/H (see, e.g., Rix et al. 2004). To date, essentially all population synthesis models of the UV spectra of star-forming galaxies

assume solar abundance ratios.[11] However, it seems possible— or perhaps even likely, given the young inferred ages for most high-redshift star-forming galaxies—that O/Fe may be enhanced relative to solar as expected for an ISM that is enriched primarily by core-collapse supernovae. If this is the case, one might find that the best-matching stellar model is one having a metallicity several times *lower* than the nebular oxygen abundance. We return to a discussion of this issue in Sections 8 and 9 below.

3.1. Model Details

The models we considered include the most recent revision of Starburst99 (Leitherer et al. 2014), using the new weaker-wind Geneva tracks without stellar rotation[12] (referred to as "S99-v00–[Z_*]," where Z_* is the assumed mass fraction of metals). These are all "single-star" models, and thus do not include binary evolution. We considered models with the default Kroupa (2001) stellar IMF, with a power law index of -2.30 over the mass range $0.5 \leqslant M_*/M_\odot \leqslant 100$, as well as models with "flatter" IMF slopes of -2.00 and -1.70 (referred to as S00-v00-[Z_*]-IMF2.0 and S99-v00-[Z_*]-IMF1.7). We considered each of these for metallicities $Z_* = (0.001, 0.002, 0.008, 0.014)$, dictated by the available Geneva tracks.

We have also included a suite of models from the new release of "Binary Population and Spectral Synthesis" (BPASSv2[13]; Eldridge & Stanway 2016; Stanway

---

[10] $T_{\rm eff}$ should not be confused with a stellar effective temperature that reproduces the bolometric luminosity of a star given its physical size–in our usage, $T_{\rm eff}$ refers only to the overall spectral shape of the radiation field over the range 1–4 Ryd that is most relevant to the observed emission lines.

[11] An exception is Eldridge & Stanway (2012), who considered models with depletion of C relative to O; C/O is discussed in more detail in 8.2.

[12] We also considered the corresponding models with rotation as implemented in Starburst99 using the Geneva tracks with rotation, but found that they were indistinguishable (in terms of the observational constraints) from the non-rotating models.

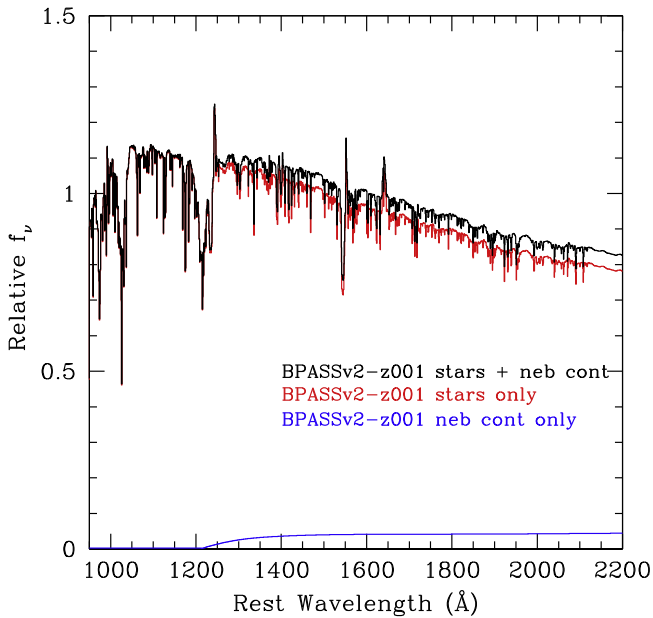[13] http://bpass.auckland.ac.nz

Figure 3. An example illustrating the calculated contribution of nebular continuum emission (blue) to a BPASSv2-z001-100bin stellar population synthesis model assuming continuous star formation over 10^8 years. The red spectrum is that of the stars only, while the black spectrum is the sum of the stellar and nebular continuum that would be used for comparison to the observed spectrum. The spectra are scaled such that the purely stellar spectrum is normalized to $f_\nu = 1$ at 1500 Å. For this particular model, the contribution of the nebular continuum emission to the total ranges from $\simeq 3.5\%$ at 1300 Å to $\simeq 6.5\%$ at 2200 Å.

et al. 2016) all with an IMF index of -2.35. We refer to three sets of models according to the stellar metallicity Z_*, the upper mass cutoff of the assumed IMF (in units of $M_\odot$), and whether or not binary evolution is included: BPASSv2-[Z_*]-100bin models include binary evolution and an IMF with upper mass cutoff of $100\ M_\odot$, BPASSv2-[Z_*]-300bin include binary evolution with upper mass cutoff of $300\ M_\odot$, and BPASSv2-[Z_*]-300 are "single-star" models with cutoff mass $300\ M_\odot$. The models considered have stellar metallicities $Z_* = (0.001, 0.002, 0.003, 0.004, 0.006, 0.008, 0.010, 0.014)$.

All models (both S99 and BPASSv2) assume continuous star formation with a duration of 10^8 years. [14]

For direct comparison to the KBSS-LM1 FUV spectrum, we used the high-resolution WM-Basic spectra covering the wavelength range 900–3000 Å spectral resolution of 0.4 Å provided by Starburst99, which we smoothed and resampled to match the wavelength scale and spectral resolution of the LM1 composite, 65 km s^{-1} pix^{-1} and FWHM $\simeq 210$ km s^{-1}, respectively. The BPASSv2 model spectra are provided with 1 Å pix^{-1} sampling in the UV; these were spline-interpolated onto the observed pixel scale.

The high spectral resolution models for both S99 and BPASSv2 do not include the contribution of the nebular continuum to the FUV spectra; to include this contribution self-consistently, we used the photoionization models (Section 7) that produced the best matches to the observed nebular emission spectrum to calculate the emergent nebular continuum

contribution relative to the stellar continuum. This contribution was then added to the stellar spectra provided by S99 and BPASSv2 models prior to performing the comparison to the data. For the assumed star formation ages, the nebular continuum contributes $\simeq 3\%$–7% to the total FUV continuum flux density for rest wavelengths in the range $1200 \lesssim \lambda_0 \lesssim 2150$ Å (see the example in Figure 3.) Including the nebular continuum has the effect of making the total continuum slightly redder than the purely stellar continuum over this range–fits using the stellar continuum only for the same population synthesis models require slightly larger values of $E(B - V)_{\rm cont}$ ($\simeq 0.21$) to match the observed spectrum.

In preparation for fitting models to the observations, the KBSS-LM1 spectrum shown in Figure 4 has been corrected for the mean IGM and CGM opacity due to neutral hydrogen (H I) along the line of sight, for a source with $z = 2.40$. This was accomplished using Monte Carlo sampling of the statistical distribution function of neutral hydrogen column density $f(N_{\rm HI}, z)$ measured by Rudie et al. (2013) from high-resolution spectra of the KBSS QSOs. Including the CGM component accounts for the average additional opacity experienced by a galaxy in the KBSS spectroscopic sample, over and above a source located at a random position within the IGM. The IGM/CGM correction, which affects the spectrum only below $\lambda_0 = 1216$ Å and amounts to an average upward correction to the observed continuum of $\sim 20\%$, was applied to the *observations* so that the resulting spectrum represents, as closely as possible, the intrinsic FUV spectrum of the stellar populations.

We then created masks which exclude wavelength pixels where the IGM-corrected LM1 spectrum (Figure 4) is affected by features unrelated to the intrinsic stellar spectrum–i.e., the positions of nebular emission lines and interstellar absorption lines, as illustrated in Figure 4. The parameters of two different masks are given in Table 3—Mask 1 is intended for global FUV fitting, and includes pixels over the full wavelength range 1040–2000 Å as well as the vicinity of strong P-Cygni features from stellar winds. Mask 2 is intended to emphasize spectral regions that are more sensitive to stellar photospheric line blanketing, for reasons discussed in Section 4.2 below.

The spectrum generated by each model was compared directly to the observed spectrum, varying only the degree of continuum reddening by dust and the overall normalization of the reddened model spectrum relative to the observations. The model spectra were reddened using the Calzetti et al. (2000) continuum attenuation relation[15] parametrized by the continuum color excess, $E(B - V)_{\rm cont}$, which was varied over the range $0.000 \leqslant E(B - V)_{\rm cont} \leqslant 0.600$, in steps of 0.001. Using Mask 1 (Table 3), the pixels in each reddened model spectrum were compared to the same pixels in the LM1 composite after multiplying the model by a normalization constant that forced the median intensity of included pixels to be equal to that in the data. The total χ^2 was then computed,

$$\chi^2 = \sum_i \left(\frac{\text{model}(i) - \text{data}(i)}{\sigma(i)} \right)^2 \qquad (1)$$

---

[14] The models considered reach equilibrium in the FUV after a few $\times 10^7$ years, after which the UV ionizing SEDs are time-independent. The typical ages inferred from SED fits over the full UV–mid-IR range are $\simeq 300$ Myr (e.g., Reddy et al. 2012).

[15] We also evaluated a grid of models using an SMC extinction curve as parametrized by Gordon et al. (2003), which can reproduce the observed spectrum as well as Calzetti et al. (2000) attenuation, albeit with smaller values of the best-fit E_{B-V}. We adopt Calzetti et al. (2000) to maintain consistency with the attenuation relation assumed in the SED fitting.
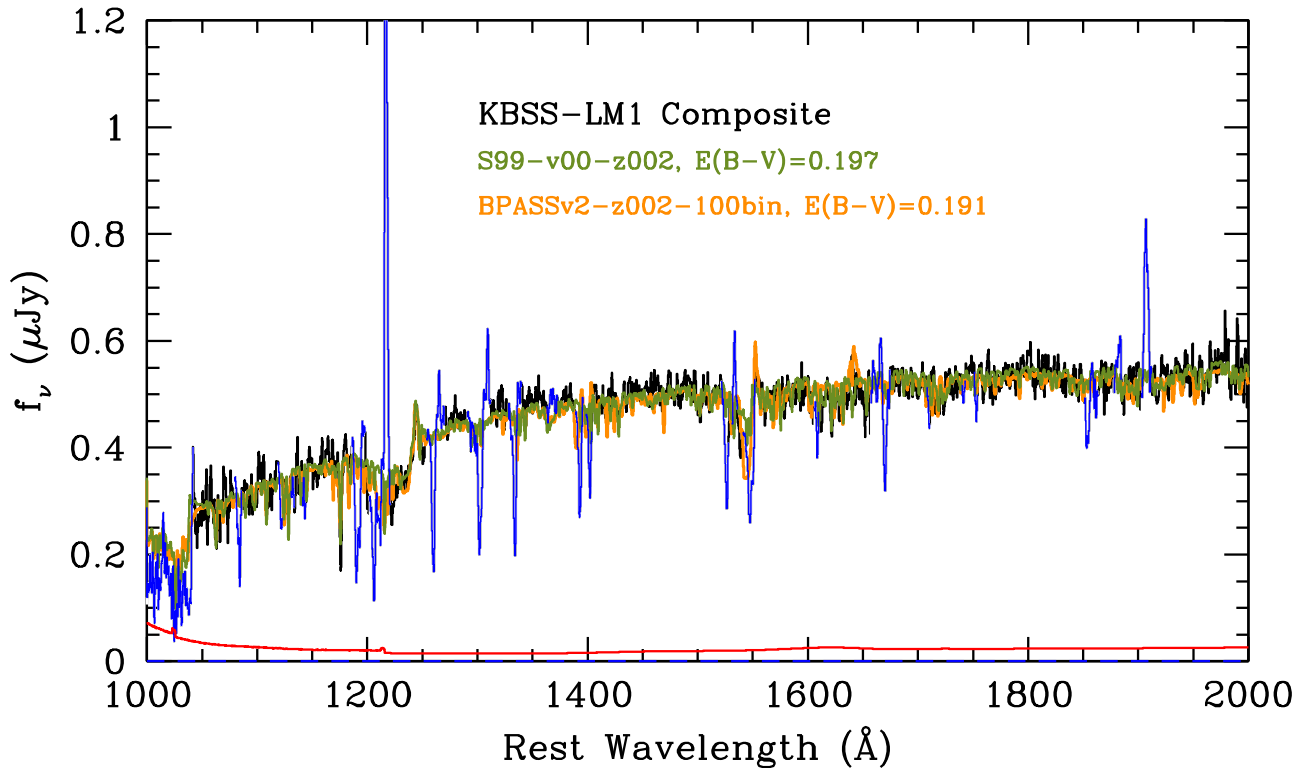
Figure 4. As in Figure 2, the KBSS-LM1 composite FUV spectrum is plotted as a black histogram. The observed spectrum has been corrected for the mean intergalactic+circumgalactic attenuation appropriate at $z \simeq 2.40$ (Rudie et al. 2013). Superposed are two of the best-fitting population synthesis models (see Table 4 and Section 3.1). Both model spectra include the predicted contribution of the nebular continuum emission, calculated using the photoionization models described in Section 4. The model spectra were reddened assuming the Calzetti et al. (2000) starburst attenuation law, where $E(B - V)$ was adjusted to minimize χ^2 with respect to the observed spectrum after masking spectral regions containing strong interstellar absorption or nebular emission lines (violet).

Table 3
Rest-wavelength Regions Used for χ^2 Fits[a]

λ_{min} (Å)	λ_{max} (Å)	Mask 1 ($N_{pix} = 2771$)	Mask 2 ($N_{pix} = 1710$)
1042	1080	1	0
1086	1119	1	0
1123	1131	1	0
1134	1141	1	0
1144	1186	1	0
1198	1202	1	0
1221	1254	1	0
1270	1291	1	0
1312	1328	1	0
1340	1363	1	0
1373	1389	1	0
1396	1398	1	0
1404	1521	1	1
1528	1531	1	0
1536	1541	1	0
1552	1606	1	1
1610	1657	1	1
1675	1708	1	1
1711	1740	1	1
1743	1751	1	1
1754	1845	1	1
1864	1878	1	0
1885	1903	1	0
1920	2000	1	1

Note.
[a] "1" indicates that the wavelength region is included in the mask; "0" indicates that the region is not included.

where $\sigma(i)$ is the value of the error spectrum at pixel i and the summation is over all included pixels. The value of $E(B - V)_{cont}$ that minimized χ^2 for each population synthesis model, and the corresponding value of χ^2/ν, where ν is the number of spectral pixels in the applied mask, are recorded in the top portion of Table 4. For reasons of space, we have omitted from Table 4 the results for metallicities $Z_* > 0.008$ since the fits were increasingly poor with increasing Z_*.

Also listed in Table 4 are two additional parameters: $\Delta\chi^2/\nu$ is the difference in χ^2/ν between that of each model and the best-fitting model in the ensemble. The last column, $\Delta\sigma$, is a measure of the significance of the difference between the goodness of fit of each model and the best-fitting model of the same type (i.e., BPASSv2 or S99), in terms of the number of σ (assuming normally distributed errors) calculated for a χ^2 distribution with ν degrees of freedom. With these assumptions, $\Delta\sigma = 1$ corresponds to $\Delta\chi^2/\nu \approx 0.031$ for $\nu = 2771$ and $\Delta\chi^2/\nu \approx 0.039$ for $\nu = 1710$.

As can be seen in Table 4, the best-fit values of $E(B - V)_{cont}$ are similar across all of the model spectra (they vary between 0.172–0.200), reflecting the fact that the unreddened FUV model spectral shape depends only modestly on the stellar abundance and IMF. However, the fits are quite sensitive to the degree of photospheric line blanketing, which is a relatively strong function of stellar metallicity. For the global fits using Mask 1, the best-fitting population synthesis models are BPASSv2-z001-300 and BPASSv2-z002-100bin (which differ by $\Delta\sigma < 1$) and S99-v00-z002. Figure 4 shows both of these models (BPASSv2-z002-100bin and S99-v00-z002) after

Table 4
Results of χ^2 Minimization for Stellar Models

Model	$Z_*/Z_\odot$ [a]	$E(B-V)$ [b]	χ^2/ν	$\Delta\chi^2/\nu$ [c]	$\Delta\sigma$ [d]
Mask 1 (Global) Fit ($\nu = 2771$)					
BPASSv2-z001-100bin	0.07	0.185	1.098	0.026	0.84
BPASSv2-z002-100bin	0.14	0.191	1.090	0.018	0.58
BPASSv2-z003-100bin	0.21	0.192	1.134	0.062	2.00
BPASSv2-z004-100bin	0.28	0.196	1.236	0.164	5.29
BPASSv2-z006-100bin	0.42	0.187	1.385	0.313	10.09
BPASSv2-z008-100bin	0.56	0.181	1.582	0.510	16.45
BPASSv2-z001-300bin	0.07	0.187	1.186	0.114	3.68
BPASSv2-z002-300bin	0.14	0.190	1.159	0.087	2.81
BPASSv2-z003-300bin	0.21	0.192	1.201	0.129	4.16
BPASSv2-z004-300bin	0.28	0.192	1.284	0.212	6.84
BPASSv2-z006-300bin	0.42	0.186	1.369	0.297	9.58
BPASSv2-z008-300bin	0.56	0.183	1.560	0.488	15.74
BPASSv2-z001-300	0.07	0.173	1.072	0.000	0.00
BPASSv2-z002-300	0.14	0.181	1.075	0.003	0.10
BPASSv2-z003-300	0.21	0.185	1.137	0.065	2.10
BPASSv2-z004-300	0.28	0.188	1.253	0.181	5.84
BPASSv2-z006-300	0.42	0.181	1.384	0.312	10.06
BPASSv2-z008-300	0.56	0.176	1.585	0.513	16.54
S99-v00-z001	0.07	0.193	1.111	0.096	3.15
S99-v00-z002	0.14	0.197	1.015	0.000	0.00
S99-v00-z008	0.56	0.200	1.530	0.515	16.90
S99-v00-z001-IMF2.0	0.07	0.188	1.144	0.129	4.24
S99-v00-z002-IMF2.0	0.14	0.181	1.061	0.046	1.52
S99-v00-z008-IMF2.0	0.56	0.197	1.467	0.452	14.83
S99-v00-z001-IMF1.7	0.07	0.174	1.185	0.170	5.57
S99-v00-z002-IMF1.7	0.14	0.174	1.150	0.135	4.42
S99-v00-z008-IMF1.7	0.56	0.187	1.456	0.441	14.45
Mask 2 Fit ($\nu = 1710$)					
BPASSv2-z001-100bin	0.07	0.203	1.089	0.038	0.97
BPASSv2-z002-100bin	0.14	0.201	1.126	0.075	1.92
BPASSv2-z003-100bin	0.21	0.203	1.181	0.130	3.33
BPASSv2-z004-100bin	0.28	0.202	1.270	0.219	5.62
BPASSv2-z006-100bin	0.42	0.183	1.393	0.342	8.77
BPASSv2-z008-100bin	0.56	0.177	1.533	0.482	12.36
BPASSv2-z001-300bin	0.07	0.205	1.107	0.056	1.43
BPASSv2-z002-300bin	0.14	0.205	1.140	0.089	2.28
BPASSv2-z003-300bin	0.21	0.204	1.187	0.136	3.49
BPASSv2-z004-300bin	0.28	0.200	1.261	0.210	5.39
BPASSv2-z006-300bin	0.42	0.187	1.367	0.316	8.10
BPASSv2-z008-300bin	0.56	0.183	1.508	0.457	11.72
BPASSv2-z001-300	0.07	0.186	1.051	0.000	0.00
BPASSv2-z002-300	0.14	0.187	1.093	0.042	1.08
BPASSv2-z003-300	0.21	0.190	1.166	0.115	2.95
BPASSv2-z004-300	0.28	0.189	1.264	0.213	5.46
BPASSv2-z006-300	0.42	0.177	1.383	0.332	8.51
BPASSv2-z008-300	0.56	0.176	1.529	0.478	12.26
S99-v00-z001	0.07	0.204	0.986	0.000	0.00
S99-v00-z002	0.14	0.193	1.024	0.038	0.98
S99-v00-z008	0.56	0.196	1.502	0.516	13.41
S99-v00-z001-IMF2.0	0.07	0.200	0.994	0.008	0.20
S99-v00-z002-IMF2.0	0.14	0.188	1.004	0.018	0.46
S99-v00-z008-IMF2.0	0.56	0.187	1.439	0.453	11.77
S99-v00-z001-IMF1.7	0.07	0.192	1.005	0.019	0.50
S99-v00-z002-IMF1.7	0.14	0.180	1.004	0.018	0.47
S99-v00-z008-IMF1.7	0.56	0.182	1.360	0.374	9.71

Notes.
[a] Stellar metallicity relative to solar, where $Z_\odot = 0.0142$ (Asplund et al. 2009).
[b] Value of $E(B-V)_{\rm cont}$ that minimizes χ^2.
[c] Difference in χ^2/ν compared to best-fitting model within group (BPASSv2 or S99).
[d] Number of σ deviation from best-fitting models within group (BPASSv2 or S99).

applying the reddening values that minimized χ^2, $E(B-V)_{\rm cont} = 0.191$ and 0.197, respectively. Table 4 shows that all of the models within $\Delta\sigma < 2$ of the best-fitting model have $Z_* \leqslant 0.002$, and that all models with $Z_* > 0.003$ ($Z_*/Z_\odot > 0.21$) are strongly dis-favored in comparison. We return to a more detailed comparison of the model versus observed FUV spectrum in Section 4 below.

The values of $E(B-V)_{\rm cont}$ for the two best-fitting models are slightly higher than the median value from the SED fits of the individual galaxies in the LM1 sample, $E(B-V)_{\rm sed} = 0.16$ (Table 1). The difference is likely due to the fact that the SED fits used a different grid of stellar models (the solar metallicity models of Bruzual & Charlot 2003, which are systematically somewhat redder in the FUV, and thus require smaller values of $E(B-V)$ to match the same observed spectrum[16]). We compare the inferred continuum attenuation, which implies extinction $A_\lambda \approx 2.0$ mag at $\lambda = 1500$ Å, to that inferred from the nebular emission lines in Section 5.3 below.

4. DETAILED COMPARISON OF THE FUV SPECTRA

Whereas the global FUV spectral shape over the fitted range 1000 Å $\leqslant \lambda_0 \leqslant 2000$ Å is easily matched by the two families of population synthesis models after applying a standard starburst attenuation relation (the only free parameters are $E(B-V)$ and the overall normalization), the KBSS-LM1 composite spectrum is of high enough spectral resolution and S/N to merit a more detailed comparison of the observed and predicted spectral features.

To facilitate such comparisons, we normalized the observed and model spectra by the unabsorbed stellar spectrum (i.e., the stellar continuum, in the absence of photospheric absorption lines). In the case of the S99 models, the WM-Basic high-resolution UV spectra include such a normalized version. For the BPASSv2 models and the observed LM1 composite spectrum, we used the continuum shape of the best-fitting S99 model as a first guess for the continuum; the initial continuum fit was then modified interactively by adding spline points constraining the fit in wavelength regions deemed relatively free of photospheric absorption (including the regions recommended for this purpose by Rix et al. 2004). For the observed spectrum, this procedure was repeated until upward fluctuations relative to the continuum fit (excluding regions containing nebular, stellar, or excited fine structure emission lines) had a distribution consistent with those expected from the 1σ error spectrum in the same spectral region.[17] In practice, since the reddened model spectra described in the previous section resulted in excellent fits to the shape of the observed spectrum (e.g., Figure 4), only small adjustments to the theoretical unabsorbed stellar continuum of the S99 models were needed to fit the continua of the LM1 (observed) and BPASSv2 model spectra. Figures 5 and 6 show zoom-in comparisons of continuum-normalized data and models in selected spectral regions.
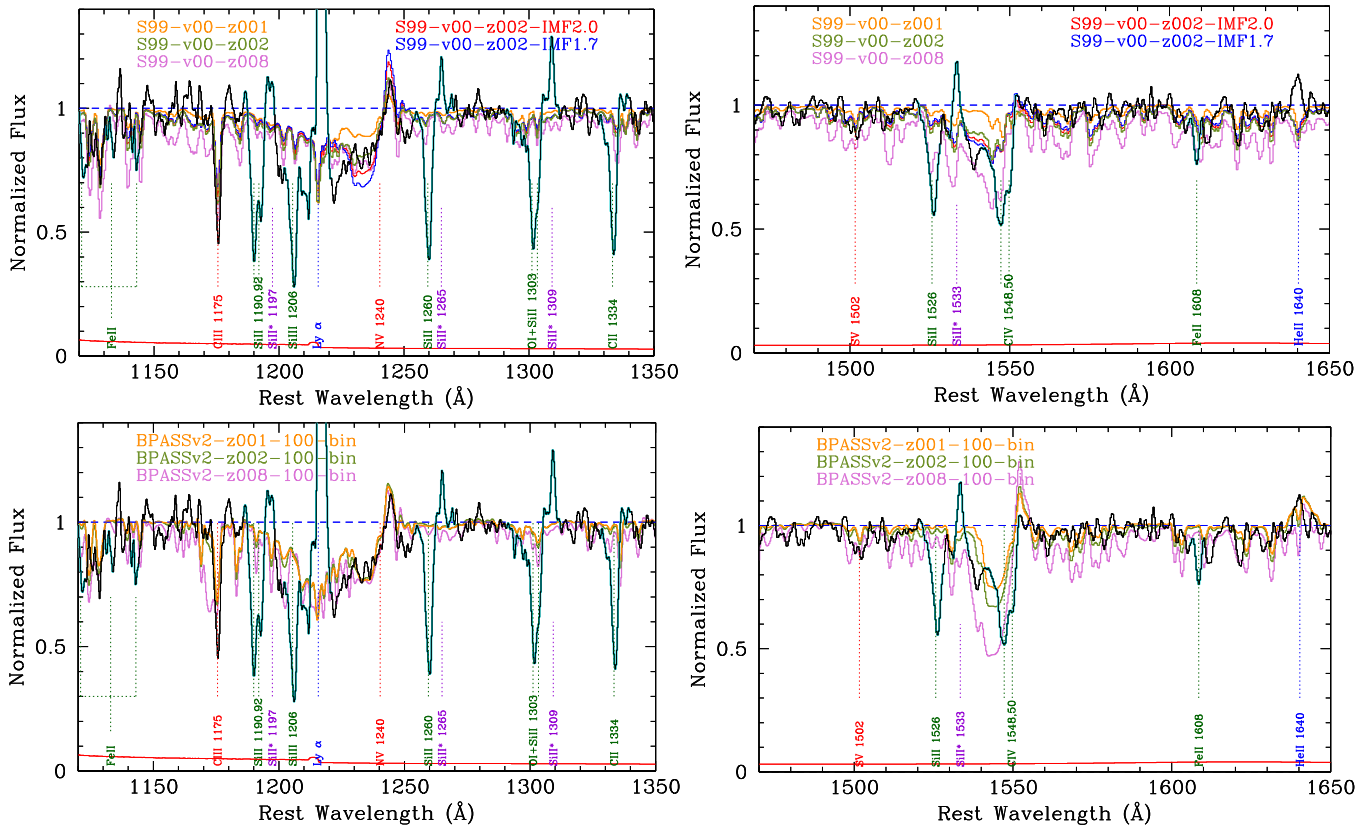

Figure 5. Zoom-in of the spectral regions near the N V (left panels) and C IV (right panels) wind lines, comparing various S99 (top) and BPASSv2 (bottom) models with the KBSS-LM1 spectrum. The top panels include three different IMFs for the S99-z002 models to illustrate the IMF dependence of the P-Cygni profiles. As in Figure 4, regions of the spectrum that were excluded from the global fits (see Section 4) are shaded in cyan. The color coding of the line identifications is the same as in Figure 2.

4.1. Stellar Wind Features

Both sets of model spectra shown in Figure 5 do reasonably well reproducing the N V $\lambda 1240$ wind feature, though there are systematic uncertainties in the data in this region because the models include no interstellar Lyα absorption, and the IGM/CGM correction to the observed spectrum is not expected to fully capture this contribution.[18] The N V feature does not appear to be particularly sensitive to stellar metallicity in the BPASSv2 models, as the three metallicities shown in Figure 5 are nearly indistinguishable. However, for the S99 models, there is a clear preference for the S99-v00-z002 model, which offers a nearly perfect match to the observed N V profile, in both emission and absorption. The effect of varying the stellar IMF in the context of the S99 models with fixed $Z_* = 0.002$ is shown in the upper left-hand panel of Figure 5, where the flatter IMF models produce N V profiles that are slightly too strong in both emission and absorption.

The C IV stellar wind feature has been discussed by many authors as a potential diagnostic because of its strength and its dependence on stellar metallicity (e.g., Pettini et al. 2000; Leitherer et al. 2001, 2010; Shapley et al. 2003; Rix et al. 2004; Steidel et al. 2004; Quider et al. 2009, 2010). The only disadvantages are the substantial contribution of interstellar C IV absorption, and the presence of narrow Si II*$\lambda 1533.4$ emission, both of which can be difficult to separate from the stellar wind P-Cygni profile at low spectral resolution and/or

S/N (see Crowther et al. 2006.) Nevertheless, the LM1 composite spectrum is of high enough S/N and spectral resolution that the broad P-Cygni absorption profile of the C IV feature is easily distinguishable in spite of the superposition of these narrower spectral features, as shown in the right-hand panels of Figure 5. The top-right panel of Figure 5 shows a clear preference among the S99 models for metallicity $Z_* = 0.002$, and a slight preference for a stellar IMF with slope > -2.3. The BPASSv2 models are less able to match the details of the C IV absorption profile, and only the lowest-metallicity models (BPASSv2-z001) are consistent with the observed depth of the P-Cygni absorption in the LM1 composite. The same BPASSv2-z001 models appear to over-produce the emission component of the observed spectrum, but it is important to note that the apparent emission component of the P-Cygni profile is dependent on the kinematics, optical depth, and covering fraction of gas giving rise to the interstellar C IV *absorption* in the vicinity of the galaxy systemic redshift; such absorption would tend to diminish the apparent strength of the P-Cygni emission component relative to its intrinsic strength.[19]

Over the rest-wavelength range $\lambda_0 \simeq 1000\text{--}2000$ Å, the most significant difference between the S99 and BPASSv2 models including binary evolution is that the latter predict broad He II $\lambda 1640$ stellar emission that is absent from any of the continuous star formation S99 models (compare the right-hand

---

[18] The IGM/CGM correction is based on the statistics of $N_{\mathrm{H\,I}}$ for galactocentric distances $d > 50$ kpc (see Rudie et al. 2013).

[19] A similar ambiguity is much less important for the N V feature because there is little or no contribution to the observed line profile from interstellar N V absorption.
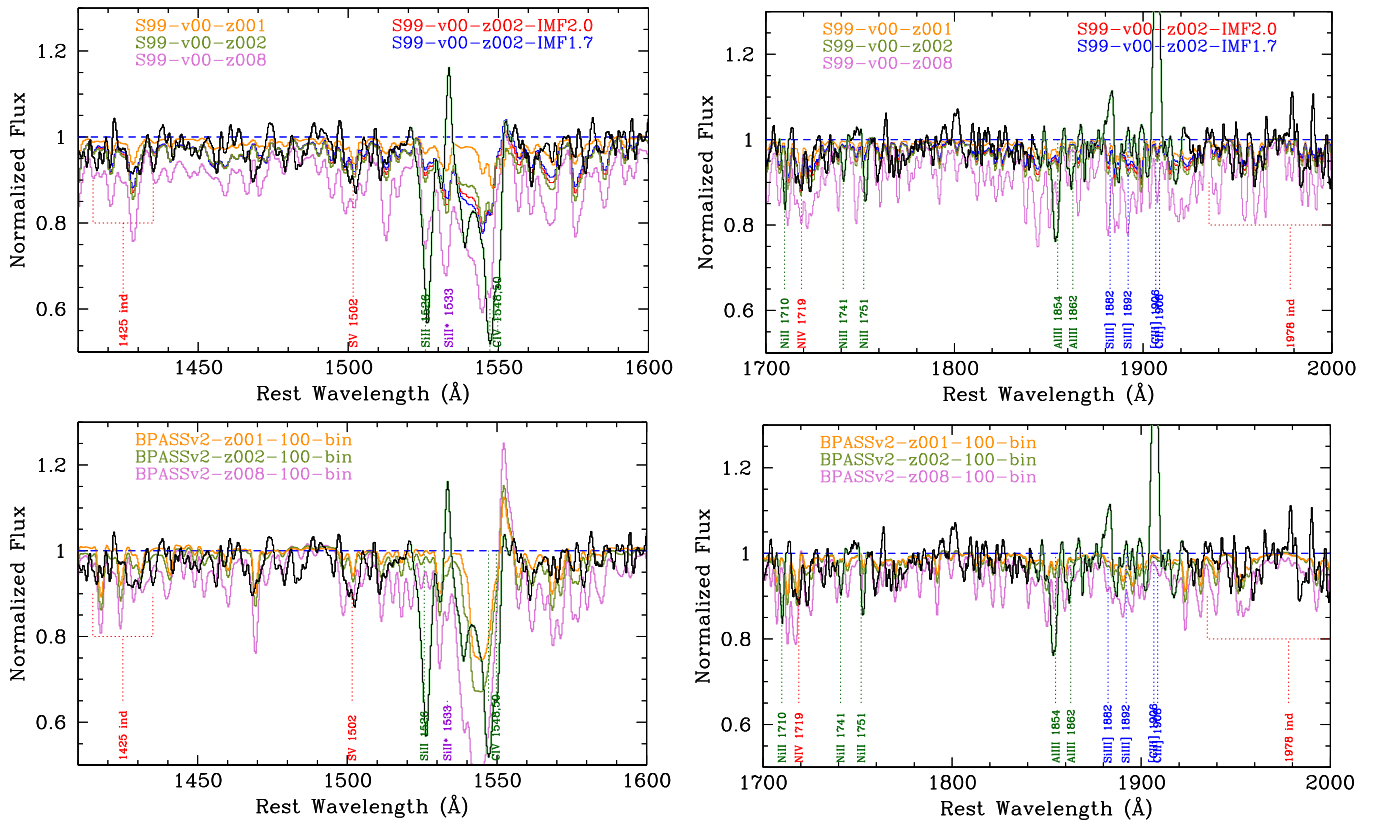
Figure 6. Comparison of the KBSS-LM1 composite spectrum (black histogram) with the predicted stellar spectra at three values of stellar metallicity in the S99 family of models (top panels) and BPASSv2 (bottom panels): $Z_* = 0.001$ ($Z_*/Z_\odot = 0.07$), $Z_* = 0.002$ ($Z_*/Z_\odot = 0.14$), and $Z_* = 0.008$ ($Z_*/Z_\odot = 0.56$). Selected stellar absorption features (including the 1425 and 1978 photospheric absorption indices suggested by Rix et al. (2004) as stellar metallicity indicators) are labeled in red. Spectral pixels excluded from the Mask 2 χ^2 minimization are shaded light green, and individual feature identifications are as in Figure 2.

panels of Figure 5). Note that this feature is also absent from the single-star BPASSv2-z001-300 model (see Figure 7). He II emission can be present as either a stellar feature, broadened by processes associated with winds from massive stars, or as a nebular feature, from recombination in H II regions; of course, it is possible that both mechanisms may contribute in any single case (see, e.g., Erb et al. 2010). As illustrated in Figure 5 (and discussed in Section 5 below) the observed He II feature in the LM1 LRIS spectrum may be attributed almost entirely to stars if the BPASSv2 models are used for the stellar continuum, whereas adopting the S99 models would imply that the line must be entirely nebular in origin. We will return to a discussion of the importance of this distinction in Sections 5 and 7.4 below.

4.2. Stellar Photospheric Absorption

We showed above that global χ^2 minimization of the stellar population synthesis models relative to the LM1 composite FUV (1000–2000 Å) spectrum favors a stellar metallicity of $Z_* = 0.001$–0.002 ($Z_*/Z_\odot = 0.07$–0.14) for both BPASSv2 and S99 models (see Table 4, Figure 4) using the Mask 1 wavelength intervals (Table 3), which includes all wavelength pixels that are unaffected by the strongest interstellar absorption and nebular emission features. Mask 2 includes a subset of Mask 1, confined to a more restricted wavelength interval 1400 Å $\leqslant \lambda_0 \leqslant$ 2000 Å and which also excludes strong P-Cygni stellar wind features. The smaller wavelength range gives less weight to matching the overall continuum shape, in order to increase the weight given to the details of photospheric line blanketing. We re-fit the same suite of models using Mask 2,

again parametrizing the overall shape of the SED using $E(B - V)$ in the context of the Calzetti et al. (2000) attenuation relation and determining the combination of $E(B - V)$ and overall normalization factor that minimizes χ^2 as defined in Equation (1). Note that $E(B - V)$ was used only as a parametrization of spectral shape, and that it was not fixed at the same value that minimized χ^2 over the wider wavelength baseline of Mask 1. The results of this exercise are presented in the bottom half of Table 4. As expected, the results are similar to those obtained using Mask 1, with the exception that the best-fitting models using Mask 2 have $Z_* = 0.001$ ($Z_*/Z_\odot = 0.07$) for both S99 and BPASSv2 models; while the $Z_*/Z_\odot = 0.14$ models produce χ^2 only marginally higher for each, there is a systematic trend among all of the models that using Mask 2 moves the minimum χ^2 toward slightly lower-metallicity models.

Figure 6 compares the KBSS-LM1 spectrum with a subset of the models over wavelength regions included on Mask 2 to illustrate the results summarized in Table 4, chosen to highlight regions with the strongest variations in line blanketing as a function of assumed stellar metallicity Z_*. Figures 5 and 6 show that in some spectral regions the lower-metallicity $Z_* = 0.001$ models provide better matches to individual photospheric absorption features. An example can be seen within the stellar absorption blend at 1415-1435 Å (left-hand panels of Figure 6), which according to Leitherer et al. (2001) is a blend of Si III $\lambda 1417$, C III $\lambda 1427$, and Fe V $\lambda 1430$. The deepest absorption in the S99 model spectra (probably associated with the Fe V feature) is over-predicted relative to the observations even for the $Z_* = 0.002$ model, whereas the
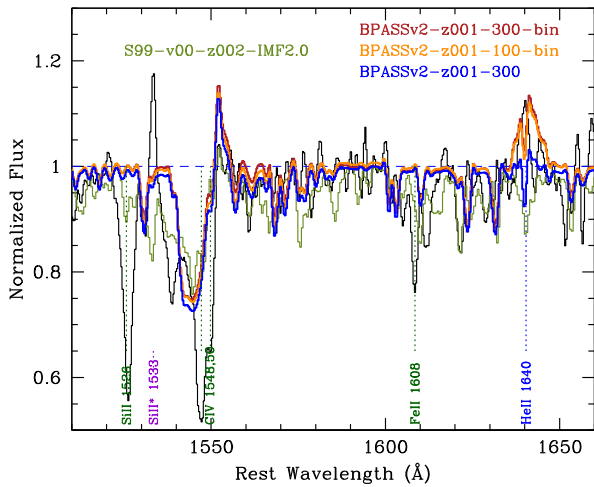
Figure 7. Illustration of the differences in the predictions for stellar He II $\lambda1640$ emission for models which otherwise all produce good fits to the FUV spectra. Both BPASSv2 models that include binary evolution (orange and dark red) predict nearly identical stellar C IV and He II emission; the single-star models (S99 with upper mass cutoff of 100 $M_{\odot}$, green, and BPASSv2 with upper mass cutoff of 300 $M_{\odot}$, blue) predict no detectable stellar He II emission.

same model is more consistent with the observations elsewhere within the same complex.

A number of authors have used the strength of the broad complex of stellar absorption features spanning the wavelength range 1935–2050 Å (see the right-hand panels of Figure 6) originally proposed as a metallicity indicator by Rix et al. (2004). [20] The total equivalent width of the blend, which is believed to be dominated by many photospheric lines of Fe III in the spectra of early B-stars, was identified as one of the more promising indicators of stellar photospheric metallicity. However, as pointed out by Rix et al. (2004), measurement of the equivalent width of the feature can be noisy due to its sensitivity to continuum placement for such a broad and relatively shallow feature. Figure 6 shows that the so-called "1978 index" is certainly dependent on assumed stellar metallicity (for both the S99 and BPASSv2 families of models), but there are additional spectral regions regions that appear to have similar metallicity sensitivity: e.g., 1560–1590 Å (left-hand panels of Figure 6), and 1770–1850 Å (right-hand panels of Figure 6). Given the relatively high resolution of the KBSS-LM1 FUV spectrum, and the sensitivity to stellar metallicity evident in many spectral regions shown in Figures 5 and 6, we believe that a χ^2 minimization similar to that described above makes optimal use of the available spectral information for estimating photospheric abundances.

In general, there are clear differences in detail between the photospheric absorption predicted by the S99 and BPASSv2 models, but the quantitative results in Table 4 are borne out by the visual impression of Figure 6: the data are most closely matched by stellar metallicity between $Z_* = 0.001$ and $Z_* = 0.002$ ($0.07 \lesssim Z_*/Z_{\odot} \lesssim 0.14$). Stellar metallicities as high as $Z_* = 0.008$ ($Z_*/Z_{\odot} = 0.56$) are ruled out with a very high significance ($>10\sigma$) for both families of models. Although stellar metallicity models with $0.002 < Z_* < 0.008$ are not yet available within S99 using the most recent stellar tracks, the similarity of the results for the values of Z_* in common to the

BPASSv2 and S99 models, and the quantitative trend of $\Delta\sigma$ versus Z_* for BPASSv2 models in the $Z_* = 0.003$–0.006 range (Mask 2 in Table 4), strongly indicate that the most likely stellar metallicity of the KBSS-LM1 composite is in the range $Z_* = 0.001$–0.002, or $0.07 \lesssim Z_*/Z_{\odot} \lesssim 0.14$.

4.3. Stellar IMF Dependence

As mentioned in Section 4.1 (see Figure 5), for the S99 models at fixed stellar metallicity of $Z_* = 0.002$, the N V and C IV stellar wind profiles are modulated by changing the slope of the stellar IMF, with the stronger dependence exhibited by the N V feature. The Mask 2 results section of Table 4 shows that varying the stellar IMF slope over the range between -2.30 and -1.70 has a barely perceptible effect on the best-fitting stellar photospheric metallicity, though there is a trend toward slightly increasing the best-matching metallicity as the IMF slope increases (i.e., as it becomes flatter). Taken together, in the context of S99 models and the non-ionizing FUV spectrum, neither the stellar wind lines nor the spectral regions dominated by photospheric stellar absorption lines offer strong evidence in favor of flatter IMF models over the default slope (-2.3) models.

The upper mass cutoff of the IMF can have a strong effect on the net EUV spectrum of massive stars. The spectra predicted by models with a high IMF cutoff mass look similar in many ways to those that incorporate massive binaries; the most dramatic way in which they differ is in the predicted stellar He II $\lambda1640$ emission. Within the first few Myr of an instantaneous burst, when the most massive stars are evolving off the main sequence to become Wolf–Rayet stars, the hardness of the EUV spectrum and the strength of stellar He II reach a maximum relative to the FUV continuum. While observation of substantial stellar He II emission in the integrated spectrum of massive star clusters *can* be explained by the presence of evolved stars with $M > 100\,M_{\odot}$ (e.g., Crowther et al. 2016), such a short-lived phenomenon cannot explain He II emission observed in the integrated FUV spectrum of entire galaxies (e.g., Shapley et al. 2003; Eldridge & Stanway 2012). As discussed in Section 1, binary evolution of massive stars can produce stars with effective temperatures similar to those of the most massive Wolf–Rayet stars from much lower initial stellar masses even in low-metallicity environments.

Indeed, Figure 7 shows that the BPASSv2-z001 *binary* model with upper mass cutoff of 100 $M_{\odot}$ predicts much stronger stellar He II $\lambda1640$ emission than the BPASSv2-z001 *single-star* model with the same IMF slope and an upper mass cutoff of 300 $M_{\odot}$. [21] In fact, we will show in Section 7 that, among the models providing a reasonable fit to the observed FUV spectrum, only the BPASSv2 *binary* models have a sufficiently hard spectral shape in the ionizing EUV to successfully predict all of the observed nebular line intensity ratios *and* predict the observed stellar He II emission for continuous star formation histories.

5. EMISSION LINES IN THE COMPOSITE SPECTRA

Because of the way in which the rest-optical and rest-UV composite spectra have been constructed (Section 2), they are

---

[20] The original calibration was subsequently updated by James et al. (2014).

---

[21] We also show in Section 7 (see in particular Sections 7.3–7.5) that the EUV continuum spectrum of the $M_{\rm up} = 100\,M_{\odot}$ binary model is harder than that of the $M_{\rm up} = 300\,M_{\odot}$ single-star model.

Table 5
Nebular Emission Line Measurements[a]

Ion	λ_0 (Å)[b]	$F(\lambda)$[c]	$I(\lambda)$[d]
		LRIS-B+R	
Lyα	1215.67	2.591 ± 0.019	6.064 ± 0.044
He II	1640.42	0.096 ± 0.009[e]	0.100 ± 0.010[e]
	1640.42	0.075 ± 0.009[f]	0.078 ± 0.010[f]
	1640.42	0.016 ± 0.009[g]	0.017 ± 0.010[g]
O III]	1660.81	0.019 ± 0.006	0.020 ± 0.006
O III]	1666.15	0.054 ± 0.006	0.056 ± 0.007
[Si III]	1882.47	0.072 ± 0.015	0.078 ± 0.016
[Si III]	1892.03	0.025 ± 0.015	0.028 ± 0.016
[C III]	1906.68	0.173 ± 0.006	0.190 ± 0.007
C III]	1908.73	0.119 ± 0.006	0.131 ± 0.007
		MOSFIRE	
[O II]	3727.09	2.25 ± 0.05	1.36 ± 0.03
[O II]	3729.88	2.38 ± 0.06	1.45 ± 0.03
[Ne III]	3869.81	0.66 ± 0.05	0.43 ± 0.03
[O III][d]	4364.44	⩽0.12	⩽0.06
Hβ	4862.69	2.15 ± 0.04	1.00 ± 0.02
[O III]	4960.30	3.14 ± 0.03	1.42 ± 0.02
[O III]	5008.24	9.39 ± 0.04	4.25 ± 0.02
Hα	6564.61	7.76 ± 0.04	2.89 ± 0.02
[N II]	6549.84	0.26 ± 0.04	0.10 ± 0.02
[N II]	6585.23	0.75 ± 0.04	0.28 ± 0.02
[S II]	6718.32	0.82 ± 0.05	0.29 ± 0.02
[S II]	6732.71	0.66 ± 0.04	0.23 ± 0.02

Notes.

[a] Only emission lines whose observed wavelengths were included in the spectra of all 30 galaxies have been measured.
[b] All wavelengths are in vacuum.
[c] All fluxes have units of 10^{-17} ergs s^{-1} cm^{-2}.
[d] Extinction-corrected line intensities relative to I(Hβ); extinction corrections assume nominal ratio I(Hα)/I(Hβ) = 2.89, which implies $E(B - V) = 0.21$ with the Galactic extinction curve of Cardelli et al. (1989).
[e] For Starburst99 continua; assumes that the feature is entirely nebular.
[f] For BPASSv2-z001-300 single-star model, after subtraction of the stellar continuum.
[g] For BPASSv2 binary stellar models; after subtraction of the model, the residual emission is assumed to be the nebular component.

on the same flux density scale, in the sense that all of the composites have been cross-calibrated so that they can be used together across the full range of observed wavelengths (3400–24000 Å, or ~1000–7000 Å in the rest-frame). We checked the continuum fluxes (well-detected for the ensemble in all three near-IR bands as well as in the LRIS spectrum) against broadband optical and near-IR photometry in hand for the full ensemble to verify that the flux density scale is consistent with that expected from photometric measurements across all wavelengths. This procedure showed that any residual systematic errors in flux scale between the LRIS spectra and the MOSFIRE J, H, and K band composites are <2%–3% (rms).

5.1. Rest-frame Optical

Emission line fluxes were measured directly from composite spectra shifted to the mean redshift of the sample, $\langle z \rangle = 2.396$. For MOSFIRE spectra, the fluxes were measured using techniques described in detail by Steidel et al. (2014), with

the following exception: because emission lines in the stacked spectra are not necessarily well-described by a single Gaussian component (usually adequate to describe the lines in the spectrum of an individual galaxy), we used direct integration of the line fluxes after subtraction of a locally normalized stellar continuum model (so that it includes appropriate Balmer line absorption, see Figure 1). The measured fluxes are compiled in Table 5, together with uncertainties estimated from a combination of the statistical errors and possible systematic errors due to ambiguity in the normalization of the local continuum; for the weakest lines, the continuum uncertainty accounts for ~50% of the estimated error bars.

5.2. Rest-frame FUV

For the rest-frame UV spectrum, observed emission line fluxes were measured using the *splot* task in IRAF, using both Gaussian fits and direct integration of the line fluxes. Some of the rest-UV emission lines are statistically well-detected but are very weak relative to the continuum, again causing the continuum determination to be one of the largest sources of uncertainty. In Section 3, we converged on a set of stellar population synthesis models providing the best fits to the observed LM1 FUV continuum spectrum. These models provide important templates to be used in measuring the relatively weak nebular emission lines present in the rest-frame FUV (Figure 2). As discussed in Section 4 above, the FUV stellar continuum contains a large number of stellar photo-spheric absorption lines, most of which are unresolved and thus produce variable line blanketing that depends on the detailed properties of the stellar population. We used the best-fitting population synthesis models (described in Sections 3 and 4) to model and subtract the stellar continuum near weak nebular emission lines; the uncertainties in Table 5 include both statistical uncertainties and the dispersion in flux measurements after subtraction of the stellar continua produced by the acceptable range of models. Figure 8 shows zoomed-in regions near weak nebular emission lines for the KBSS-LM1 rest-UV spectrum, where the bottom panels show the emission spectrum after subtraction of the various stellar continuum models; as in previous figures showing portions of the rest-UV spectrum, the notable nebular emission lines are indicated with blue labels.

The UV intercombination doublet O III]λλ1661, 1666 is expected to have a fixed intensity ratio of ≃1 : 2.5, but its apparent ratio may be affected by the nearby interstellar Al IIλ 1670.79 absorption feature; however, even without fixing the ratio of the two lines, the measured intensities are consistent within the errors with the expected ratio. As discussed in Section 8.1, this feature is important to the measurement of the nebular electron temperature (T_e) for for direct oxygen abundance determination.

The detection of nebular He II λ1640 is a potentially important constraint on the ionizing sources in the galaxies; as discussed in Section 3 above, this feature can be either nebular or stellar, with possible contribution from both (see, e.g., Erb et al. 2010). It is notoriously difficult to produce measurable nebular He II emission assuming photoionization by normal main-sequence O-stars. Table 5 includes three measurements of the residual He II λ1640 after subtraction of three different stellar models with large variations in the predicted strength of *stellar* He II, and thus very different predictions for the residual He II emission (Figure 8(a)). We discuss the implications of the
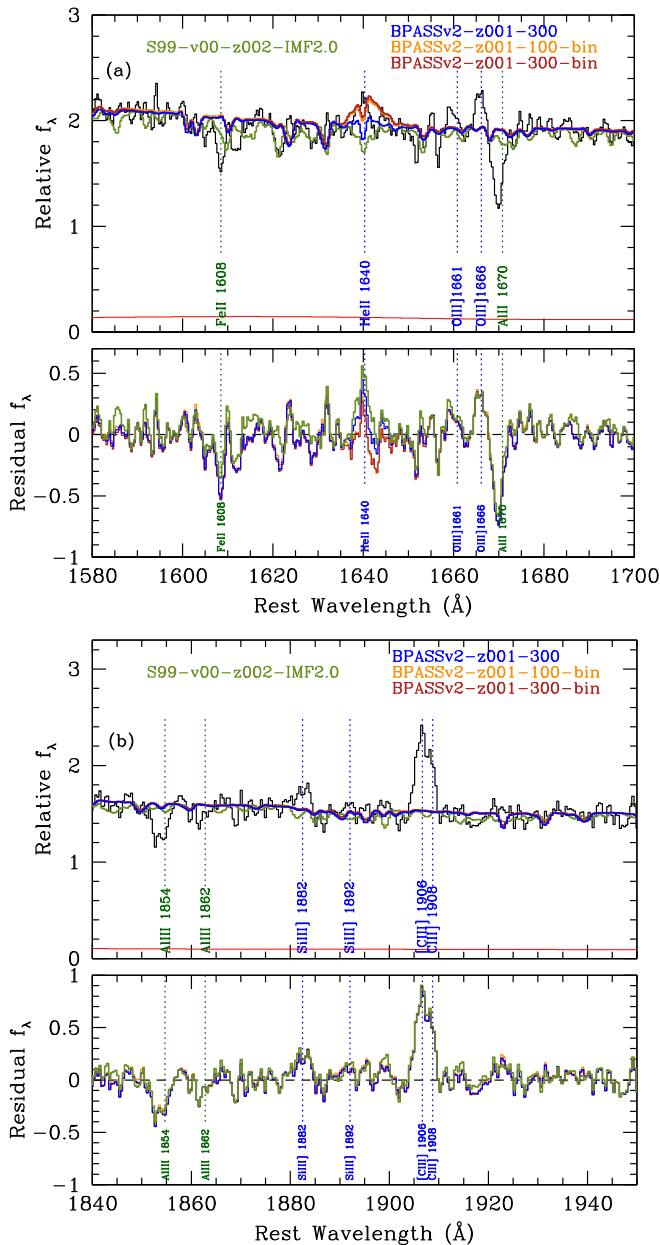
Figure 8. Portions of the LM1 FUV composite near weak nebular emission lines. The top panel in each case shows the observed spectrum (black histogram) together various stellar continuum models. The lower panel shows the residual spectrum after subtraction of the stellar continua; the residuals are plotted in the same color as the models in the upper panels. (a) The region near the He II λ1640.42 and O III] $\lambda\lambda$1660.81, 1666.15 features. Note that for the BPASSv2 binary models most of the He II feature is consistent with being stellar, with weak residual emission possibly of nebular origin. For the S99 and BPASSv2 single-star models, which do not exhibit obvious stellar He II emission, the residuals are much stronger (see Table 5). (b) The spectral region near the [C III]/C III] $\lambda\lambda$1906.68, 1908.73 and [Si III] $\lambda\lambda$1882.47, 1892.03 features, with color coding as in (a). In this case, all of the stellar models yield similar residual line profiles, with only small variations.

observed He II residuals in the context of photoionization models in Section 7.4.

In the case of the [C III]/C III] forbidden/semi-forbidden line pair, which has a rest-frame separation of only 2.05 Å and is therefore only marginally resolved at the resolution of the LRIS composite ($R = 1500$ corresponds to FWHM ≈ 1.25 Å in the rest frame at 1908 Å), the lines were measured using a simultaneous fit of two Gaussian components constrained to

Table 6
Excited Fine Structure Emission Lines[a]

Ion	λ_0 (Å)	$\lambda_{\rm obs}$ (Å)	W_λ (Å)	Δv[a] (km s^{-1})
SiII*	1197.39	1197.37	⋯	− 6.0
SiII*	1265.00	1265.06	0.23	+14.2
SiII*	1309.28	1309.20	0.42	−18.3
SiII*	1533.43	1533.45	0.43	+ 3.9
				$\langle \Delta v \rangle = 1.3 \pm 12.1$ km s^{-1}

Note.
[a] Redshift of observed feature relative to $z_{\rm neb}$ defined by strong rest-optical nebular lines.

have a fixed separation and equal width, again after subtraction of several different continuum representations as illustrated in Figure 8(b).

In addition to nebular emission lines from H II regions, the FUV also contains numerous emission features from excited fine structure transitions (purple labels in Figures 2, 5, and 6). These features are listed separately in Table 6, since they are not "nebular," but instead are produced by resonant scattering and re-emission of strong ground-state interstellar absorption features (rather than by collisional excitation or recombination). We note that in the LM1 composite, the excited fine structure emission lines agree to within $\sim$12 km s^{-1} (rms) with the nebular redshift $z_{\rm neb}$ measured from strong *nebular* emission lines in the KBSS-LM1 rest-optical composite spectra.

5.3. Nebular Extinction

Since nebular emission lines can be measured accurately from the composite spectra over a relatively wide wavelength baseline, full use of the diagnostic power of the line ratios depends on accurate estimation of the wavelength-dependent nebular extinction due to dust. The standard means for achieving this involves measurement of the "Balmer decrement," i.e., comparing the observed ratio $I({\rm H}\alpha)/I({\rm H}\beta)$ to that expected in the case of zero nebular reddening. The expected intensity ratio is usually taken to be that predicted by "Case B" recombination, in the low-density limit and assuming electron temperature $T_{\rm e} = 10,000$ K, $I({\rm H}\alpha)/I({\rm H}\beta) = 2.86$ (Osterbrock & Ferland 2006). However, for consistency with the analysis in Section 7 below, we adopted the intrinsic Balmer line ratio predicted by the photoionization models that best-reproduce the observed intensity ratios of lines that have little or no dependence on nebular extinction corrections; these models tend to predict slightly higher values of the intrinsic Balmer decrement, $I({\rm H}\alpha)/I({\rm H}\beta) \simeq 2.89$. Adopting this intrinsic ratio and the Cardelli et al. (1989) Galactic extinction curve, the observed $I({\rm H}\alpha)/I({\rm H}\beta) = 3.61 \pm 0.07$ maps to $E(B-V)_{\rm neb} = 0.21 \pm 0.02$. The fourth column of Table 5 shows the resulting extinction-corrected line strengths relative to Hβ. The implications of the observed line intensities and their ratios are discussed in Section 6.

We showed in Section 3 above that the global fits to the FUV LM1 composite yielded *continuum* reddening $E(B-V)_{\rm cont}$= 0.190–0.197 for the best-fitting population synthesis models (cf. Table 4 and Figure 4). There is a rapidly developing literature comparing dust attenuation factors estimated from nebular line ratios to those estimated from continuum

Table 7
Measured and Inferred Properties of the KBSS-LM1 Composite Spectra

Quantity	Value	Units/Notes
Extinction and Star Formation Rate (UV and Nebular)		
$F(H\alpha)/F(H\beta)$	3.61 ± 0.07	observed
$E(B-V)_{neb}$	0.21 ± 0.02	a
$E(B-V)_{UV}$	0.192 ± 0.004	b
SFR_{UV}	28.6 ± 0.8	$M_\odot \, yr^{-1}$
$SFR_{H\alpha}$	26.8 ± 1.2	$M_\odot \, yr^{-1}$
$\log(sSFR)$	0.54 ± 0.09	Gyr^{-1}
Electron-density-sensitive Line Ratios		
[C III] 1906/[C III 1909	1.469 ± 0.028	$n_e = 1470 \pm 660 \, cm^{-3}$
[O II] 3727/3729	0.945 ± 0.033	$n_e = 360 \pm 45 \, cm^{-3}$
[S II] 6718/6732	1.242 ± 0.090	$n_e = 160 \pm 90 \, cm^{-3}$
Ionization-sensitive Line Ratios		
O3	0.629 ± 0.008	$\log([O\,III]5008/H\beta)$
O3$_{tot}$	0.753 ± 0.008	$\log([O\,III]4960 + 5008/H\beta)$
O2	0.449 ± 0.007	$\log([O\,II]3727 + 3729/H\beta)$
O32	0.303 ± 0.007	O3$_{tot}$ − O2
Ne3O2	-0.812 ± 0.027	$\log([Ne\,III]3870/[O\,II]3727 + 3729)$
Abundance-sensitive Line Ratios		
O3uv	0.019 ± 0.003	O III](1661 + 1666)/[O III]5008
O3opt[c]	<0.014	O III] 4364/[O III] 5008
N2	-1.02 ± 0.02	$\log([N\,II]6585/H\alpha)$
S2	-0.74 ± 0.02	$\log([S\,II](6718 + 6732)/H\alpha)$
O3N2	1.64 ± 0.02	O3 − N2
N2O2	-1.00 ± 0.02	$\log([N\,II]6585/[O\,II](3727 + 3729))$
N2S2	-0.28 ± 0.03	$\log([N\,II]6585/[S\,II](6718 + 6732))$
R23	0.93 ± 0.01	$\log([O\,III](4960 + 5008) + [O\,II](3727 + 3729)]/H\beta)$
C3O3	0.62 ± 0.07	$\log(C\,III](1907 + 1909)/O\,III](1661 + 1666))$
Si3O3	0.00 ± 0.08	$\log([Si\,III]1882/O\,III](1661 + 1666))$
Inferred Stellar Metallicity		
Photospheric Absorption	$Z_* \simeq 0.001$–0.002	$Z_*/Z_\odot \simeq 0.10$
C IV, N V Wind	$Z_* \simeq 0.002$	$Z_*/Z_\odot \simeq 0.14$
Nebular Oxygen Abundances (Direct and Modeled)		
$12 + \log(O/H)_{CEL}$	8.14 ± 0.04	$Z_{neb}/Z_\odot = 0.29 \pm 0.03$; $T_e([O\,III]) = 12250 \pm 600$ K[d]
$12 + \log(O/H)_{REL}$	8.38 ± 0.05	$Z_{neb}/Z_\odot = 0.49 \pm 0.05$; Equation (6)[e]
$12 + \log(O/H)_{mod}$	8.39 ± 0.09	$Z_{neb}/Z_\odot = 0.50 \pm 0.10$; from photoionization modeling (this work)
Nebular Oxygen Abundances (Strong-line Methods)		
$12 + \log(O/H)_{R23}$	8.20 ± 0.03	$Z_{neb}/Z_\odot = 0.32 \pm 0.03$; lower branch McGaugh (1991)
$12 + \log(O/H)_{O3N2}$	8.23 ± 0.02	$Z_{neb}/Z_\odot = 0.35 \pm 0.02$; Equation (10)
$12 + \log(O/H)_{N2}$	8.29 ± 0.02	$Z_{neb}/Z_\odot = 0.40 \pm 0.02$; Equation (11)
$12 + \log(O/H)_{N2O2}$	8.23 ± 0.02	$Z_{neb}/Z_\odot = 0.35 \pm 0.02$; Equation (12)
$12 + \log(O/H)_{N/O}$	8.13 ± 0.02	$Z_{neb}/Z_\odot = 0.28 \pm 0.02$; "N/O analogs"; Section 8.1.4
Inferred Gas-phase Abundance Ratios		
$\log(N/O)$	-1.24 ± 0.04	[N/O] $= -0.38 \pm 0.04$; Equations (1), (2), Section 8.1.2
$\log(C/O)$	-0.60 ± 0.09	[C/O] $= -0.34 \pm 0.09$; Equation (17)
$\log(Si/O)$	-1.81 ± 0.08	[Si/O] $= -0.63 \pm 0.08$; Equation (18)

Notes.
[a] From the Balmer decrement, assuming the Cardelli et al. (1989) extinction curve (see Table 5).
[b] From the UV continuum shape, assuming the Calzetti et al. (2000) attenuation relation.
[c] [O III]λ4364 was observed in only ~50% of the galaxies.
[d] Total oxygen abundance calculated assuming that $T_e([O\,II]) = 0.7\,T_e([O\,III]) + 3000$ K, and that $(O/H) = (O^{++}/H^+) + (O^+/H^+)$.
[e] Total oxygen abundance calculated assuming $\log(O/H)_{REL} - \log(O/H)_{CEL} = 0.24 \pm 0.02$

reddening (e.g., Calzetti et al. 2000; Kashino et al. 2013; Price et al. 2014; Reddy et al. 2015; Shivaei et al. 2016). Since we are primarily concerned with correcting nebular emission line fluxes, this comparison is not essential, but because of the high quality of the composite spectra in both the rest-UV and rest-optical, it may be of some interest. Taken at face value, the continuum reddening inferred from the FUV spectrum is slightly smaller than, but consistent within the uncertainties, with the inferred nebular reddening $E(B - V)_{\rm neb} = 0.21 \pm 0.02$ (Table 7), bearing in mind that the values refer to different assumed extinction/attenuation relations (Calzetti et al. 2000 for the continuum and Cardelli et al. 1989 for the nebulae). If we had used purely stellar continuum models (rather than stellar plus nebular continuum), the two values of $E(B - V)$ would have been indistinguishable.

The similarity in the inferred $E(B - V)_{\rm cont}$ and $E(B - V)_{\rm neb}$ supports the underlying assumption that the nebular emission lines are responding to the same massive stars that dominate the FUV continuum—i.e., that there is not likely to be a significant contribution of very dust-obscured massive stars (with potentially different stellar metallicity) contributing to nebular line luminosities but remaining obscured in the FUV.

6. INFERENCES FROM THE NEBULAR MEASUREMENTS

In this section, we describe several relatively model-independent measurements that provide useful constraints on the range of parameters to be explored in the photoionization models detailed in Section 7.

6.1. Electron Density

Three sets of nebular emission line doublets detected in the KBSS-LM1 composites are among those used for estimating electron density (n_e) in H II regions: [O II]$\lambda\lambda$3727, 3729 and [S II]$\lambda\lambda$6718, 6732 in the J and K band MOSFIRE spectra, respectively, and [C III]λ1907/C III]λ1909 in the LRIS UV spectrum. Conversion of a measured intensity ratio into a constraint on n_e depends weakly on electron temperature T_e, and strongly on the actual n_e compared to the critical densities of the two transitions, which in turn depend on knowledge of the correct rate coefficients for collisional excitation/de-excitation. We used the measured line ratios (Tables 5 and 7) to calculate the constraints on n_e provided by each pair of density-sensitive lines. For [O II] and [Si II] we used the fitting formula recently calculated by Sanders et al. (2016), which incorporates all of the recent updates to the atomic data for these species, and assumes $T_e = 10^4$ K. In the case of the C III calculation, we used a grid of Cloudy models (v13.02; Ferland et al. 2013) with assumed $T_e \approx 10,000$, with the BPASSv2-z001-100bin model used as the input ionizing spectrum (see Section 4). The results for all three measurements are summarized in Table 7.

Clearly, the values of n_e obtained from [O II] ($n_e = 360 \pm 45$ cm^{-3}) and [S II] ($n_e = 160 \pm 90$ cm^{-3}) are marginally consistent with one another, with the value of n_e from [O II] being much better constrained because of the higher S/N of the line ratio measurement and the larger difference in critical density for the two transitions of [O II] compared to those of [S II].

The implied $n_e = 1470 \pm 660$ cm^{-3} from the ratio of the C III lines is substantially higher than implied by [O II] or [S II]; however, the observed ratio ($I(1907)/I(1909) = 1.469 \pm 0.028$; see Table 7) is only $\sim 2\sigma$ from the the theoretical low-density limit ($I(1907)/I(1909) = 1.529$), and $\sim 1.5\sigma$ from the ratio expected if n_e(C III) $\simeq 360$ cm^{-3}. While it is possible that the physical conditions in the H II regions producing the bulk of the [C III] emission are different from those in which most of the [O II] is produced, the high critical density for both [C III]λ1907 ($\simeq 7.4 \times 10^4$ cm^{-3}) and [C III]λ1909 ($\simeq 9.7 \times 10^8$ cm^{-3}) make their intensity ratio not particularly well-suited for measuring n_e of a few hundred cm^{-3}.

6.2. N/O Abundance Ratio

The ratio of gas-phase nitrogen to oxygen (N/O) is well known to depend sensitively on (O/H) in H II regions and star-forming galaxies, but (as discussed by Steidel et al. 2014) the precise dependence of log(N/O) versus log(O/H) is a subject about which there is little consensus in the literature. It is straightforward to measure N/O because of the similar ionization potential of N and O, so that (N/O) $\simeq$ (N$^+$/O$^+$). Differences in the shape or intensity of the ionizing radiation field thus make only very small differences to the mapping of the N2O2 line index (see definitions in Table 7) to the N/O abundance ratio. The disagreements in the behavior of N/O versus O/H in the literature are dominated by systematic differences in O/H among the methods used to estimate it.

We can therefore use a calibration sample drawn from the local universe for an initial estimate of N/O in the KBSS-LM1 composite. For reasons discussed in more detail by Strom et al. (2016), we use a sample of 412 extragalactic H II regions compiled by Pilyugin et al. (2012), which spans approximately the same range in N2O2 as the KBSS-MOSFIRE sample and includes T_e-based abundances of N, O, and S. A linear fit to the Pilyugin et al. (2012) data set (see Strom et al. 2016) yields

$$\log(\mathrm{N/O}) = 0.65 * \mathrm{N2O2} - 0.57; \quad \sigma = 0.05 \,\mathrm{dex} \qquad (2)$$

where $\sigma = 0.05$ dex refers to the scatter of individual measurements relative to the fit.

A similar fit using the ratio N2S2 (see Table 7), sensitive to N/S and useful as a proxy for N/O (since S and O are believed to be produced by the same nucleosynthetic processes), is given by

$$\log(\mathrm{N/O}) = 0.68 * \mathrm{N2S2} - 1.08; \quad \sigma = 0.12 \,\mathrm{dex}. \qquad (3)$$

As expected, the scatter is larger using N2S2 (see also Pérez-Montero & Contini 2009), but for application to the KBSS-LM1 composite, Equation (3) has the advantage of being independent of cross-band calibrations (both [N II] and [S II] are measured in the K band for the KBSS-LM1 sample) and nearly independent of reddening/extinction.

Substituting the measured values of N2O2 and N2S2 (see Table 7) into Equations (2) and (3) results in values of log(N/O) in very good agreement with one another: (log(N/O) = -1.22 ± 0.02 from N2O2 and log(N/O) = -1.26 ± 0.02 from N2S2), suggesting that the values are reasonable and that our cross-band calibration and reddening corrections are accurate. Hereafter, we adopt the average log(N/O) = -1.24 ± 0.03 as a fiducial value, but we re-assess the N/O abundance ratio for the KBSS-LM1 composite
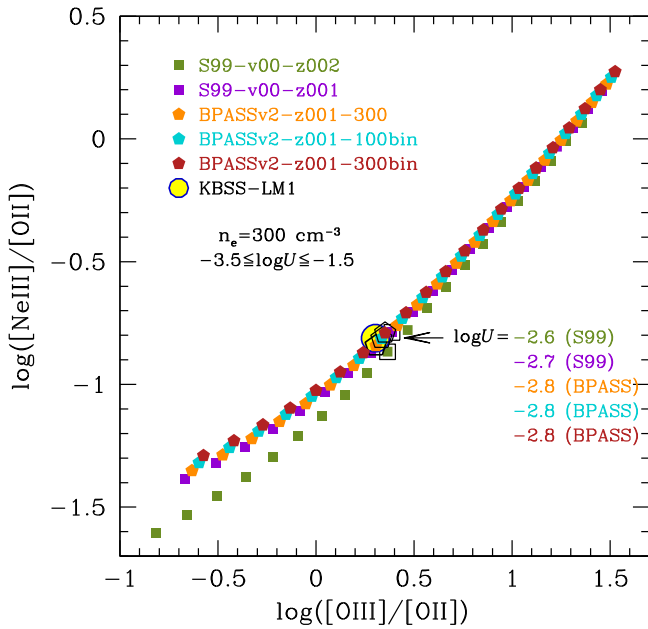
Figure 9. The predictions of photoionization models using the stellar population synthesis models that best-fit the observed FUV spectrum, for two nebular line ratios sensitive to ionization parameter U. The yellow point is measured from the KBSS-LM1 composite (Table 7); its error bars are smaller than the size of the point. The models shown assume $Z_{neb}/Z_\odot = 0.5$, but there is only very weak dependence on gas-phase metallicity of the line ratios shown. The value of $\log U$ that most closely matches the observed line indices (symbols enclosed by black outlines) is listed for each model shown.

using a method entirely independent of the empirical N2O2 $\rightarrow$ N/O calibration (Equations (2) and (3)) in Section 7.6 below.

7. PHOTOIONIZATION MODELS

Since we are interested in identifying stellar population synthesis models that can simultaneously account for both the FUV OB-star continuum and the excitation of nebular emission in the same galaxies, we use the same population synthesis models as described in Sections 3 and 4 as the source of the ionizing radiation field in photoionization models. For this purpose, we used Cloudy (v13.02; Ferland et al. 2013); in brief, the models assume a constant density $n_e = 300 \text{ cm}^{-3}$ (see Section 6.1) plane-parallel geometry where the radiation field intensity is characterized by the dimensionless ionization parameter U,

$$U \equiv \frac{n_\gamma}{n_H} \approx \frac{n_\gamma}{n_e} \qquad (4)$$

where n_H is the number density of hydrogen atoms and n_γ is the equivalent density of photons capable of ionizing hydrogen impinging on the face of the gas layer.

We varied the nebular abundances from $Z_{neb}/Z_\odot = 0.1$ to $Z_{neb}/Z_\odot = 1.0$ in steps of $0.1 Z_\odot$, using the solar abundance set from Asplund et al. (2009) for all relevant elements. Dust grains were included using the "Orion" mix provided by Cloudy, assuming that the grain abundance (and thus the dust-to-gas ratio) scales linearly with metallicity relative to solar. We allowed $\log U$ to vary over the range $-3.5 \leqslant \log U \leqslant -1.5$ in steps of $\Delta \log U = 0.1$ for each grid.

As mentioned above, we have explicitly decoupled the *stellar metallicity* (Z_*) of the population synthesis models from the ionized gas-phase *nebular abundances* (Z_{neb}), so that each

population synthesis model was tested over the full range of assumed gas-phase metallicity. As output, the final predictions for the nebular spectrum at the edge of the fully ionized region were saved, along with the associated nebular continuum emission. As discussed in Section 3, the nebular continuum (generated self-consistently along with the line emission) was then added to the high-resolution, purely stellar models for detailed comparison with the FUV spectrum. Having generated many such grids of photoionization models (210 grid points in U and Z_{neb} space for each of the population synthesis models described in Section 3) we sought those that could most closely reproduce the observed emission line intensity ratios. In what follows, we focus primarily on the models that produced the best fits to the global FUV spectra in Section 3 (Table 4), and discuss inferences based on combining the observational constraints with the photoionization model results. As we show below, the stellar models capable of reproducing the observed nebular emission line indices are a subset of those which provided reasonable matches to the FUV stellar spectra.

7.1. Ionization Parameter

The most useful line ratios for constraining the ionization parameter U are those with different ionization potentials but fixed elemental abundance ratios- the best such lines observed in the KBSS-LM1 composite spectra are $\log([O\,\text{III}](4960 + 5008)/\log[O\,\text{II}](3727 + 3729)) \equiv$ O32 and $\log([Ne\,\text{III}]3870/[O\,\text{II}](3727 + 3729)) \equiv$ Ne3O2 (see Table 7). The advantage of the former is that both lines derive from the same element; the advantage of the latter is the proximity of the two features in wavelength—both appearing in the J band over the redshift range of the LM1 sample—meaning that uncertainties in relative slit losses are absent, and relative extinction corrections are small. The abundance ratio (Ne/O) is also not expected to vary significantly with (O/H). A successful photoionization model should correctly predict the observed values of both O32 and Ne3O2 with a single value of $\log U$.

Figure 9 compares the predictions of the photoionization models using the population synthesis models that provide the closest match to the observed FUV spectrum as described in Section 3. The observed point (labeled KBSS-LM1) can be reproduced by both S99 and BPASSv2 models, but the value of $\log U$ at which the models produce their closest match is model-dependent: higher values of $\log U$ (by $\simeq 0.1$–0.2 dex) are needed for the S99 models, as indicated in Figure 9. We find that the predicted locus of O32 versus Ne3O2 for a given population synthesis model is insensitive to assumed gas-phase metallicity; Figure 9 shows grid points for $Z_{neb}/Z_\odot = 0.5$ ($12 + \log(O/H) = 8.39$) for illustration purposes and for consistency with the discussion in Sections 7.2 and 7.3.

7.2. N2 and S2 BPT Planes

Figure 10 shows the predictions of the Cloudy models in the N2 and S2 Baldwin et al. 1981 (BPT) planes. In the top panel, the model grid assumes $\log(N/O) = -1.24$ as obtained in Section 6.2 using the N2O2 calibration in Equation (2). Note first that the KBSS-LM1 measurement falls directly on the fit to the BPT locus of the full $z \sim 2.3$ KBSS-MOSFIRE sample (Steidel et al. 2014; Strom et al. 2016); i.e., the composite is representative of the full KBSS sample. The grid point corresponding to the best value of U for each model determined
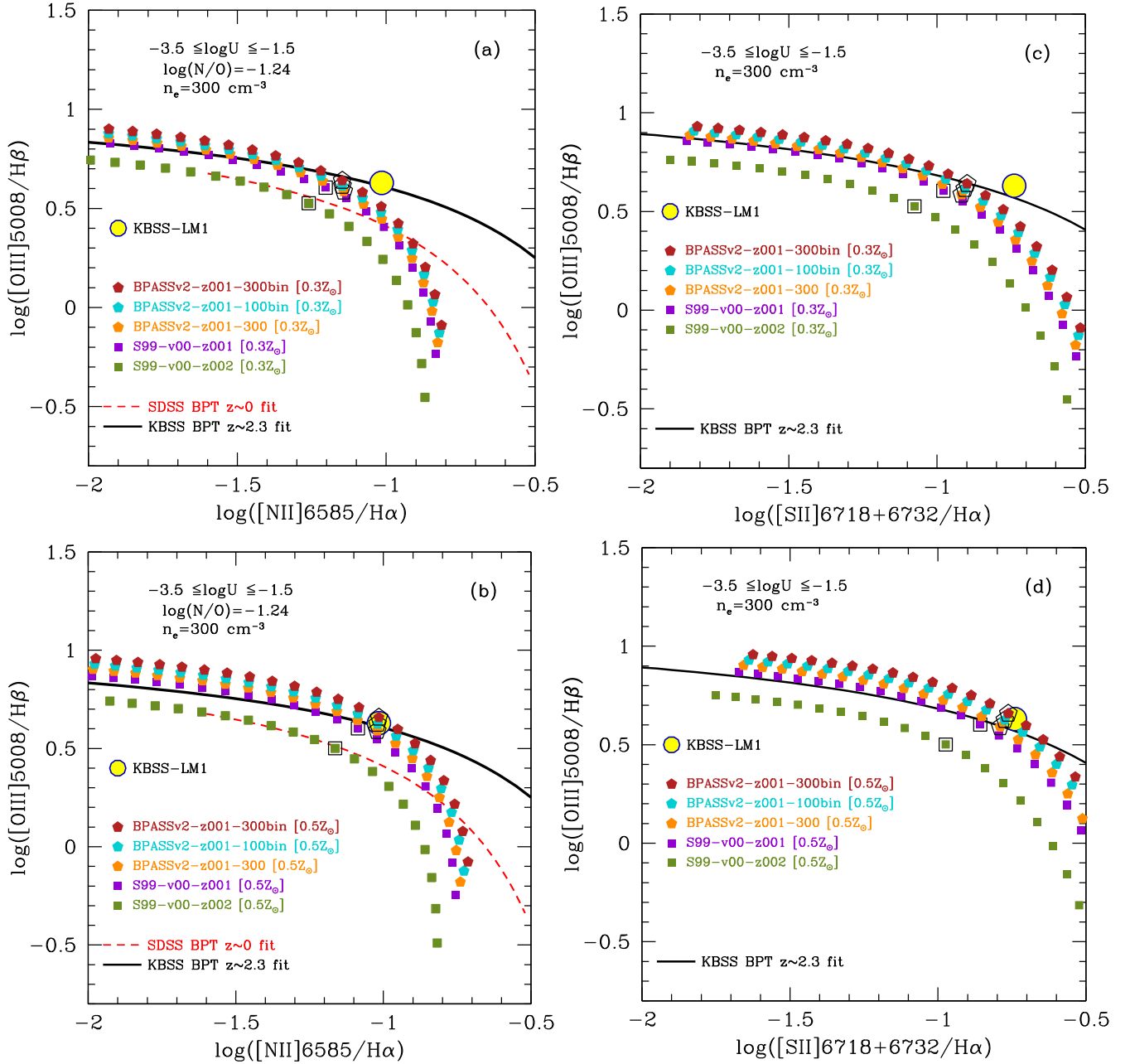
Figure 10. Predictions of the same set of models as in Figure 9 in the N2 (left-hand panels) and S2 (right-hand panels) BPT plane, compared to the KBSS-LM1 measurement. The same $\log U$ grid points that most closely matched the observed line ratios in Figure 9 are indicated with black outline symbols. As in Figure 9, the error bars for the LM1 measurements are smaller than the size of the point. The upper panels show the predictions assuming $Z_{neb}/Z_{\odot} = 0.3$, while the lower panels assume $Z_{neb}/Z_{\odot} = 0.5$; note that the higher gas-phase metallicity is strongly preferred.

from Ne3O2 and O32 (Figure 9) is indicated with a black outline symbol in each panel. The BPASSv2-z001-100bin and BPASSv2-z001-300bin population synthesis models with $\log U = -2.8$ both predict values in excellent agreement with the observations for $Z_{neb}/Z_{\odot} = 0.5$; the predictions for the BPASSv2-z001-300 single-star model come close to the observations in both BPT planes for $Z_{neb}/Z_{\odot} = 0.5$. In contrast, the S99 models cannot simultaneously match the O3 and N2, or O3 and S2 constraints at any gas-phase metallicity or value of $\log U$; the S99 models, when $\log U$ is fixed by the O32 versus Ne3O2 constraint (Figure 9), underpredict both N2 and O3 ($\simeq 0.15$ dex in each). The S99 models are capable of producing O3 as high as observed, but only if $\log U$ were $\simeq 0.5$ dex higher than the value constrained by Ne3O2 and O32;

such high values of $\log U$ would predict much *lower* N2 (or S2) than observed.

7.3. R23 versus O32

Figure 11 examines the efficacy of the photoionization models in the O32 versus R23 plane. A combination of R23 and O32 has been proposed as an indicator of gas-phase oxygen abundance by a number of authors (e.g., McGaugh 1991; Kobulnicky et al. 1999; Kewley & Dopita 2002), where O32 is used to set the ionization parameter whereupon R23 is sensitive to overall O/H. The same models illustrated in Figures 9 and 10 are shown in Figure 11; the S99 models cannot produce values of R23 as large as those observed at fixed O32, for *any* assumed gas
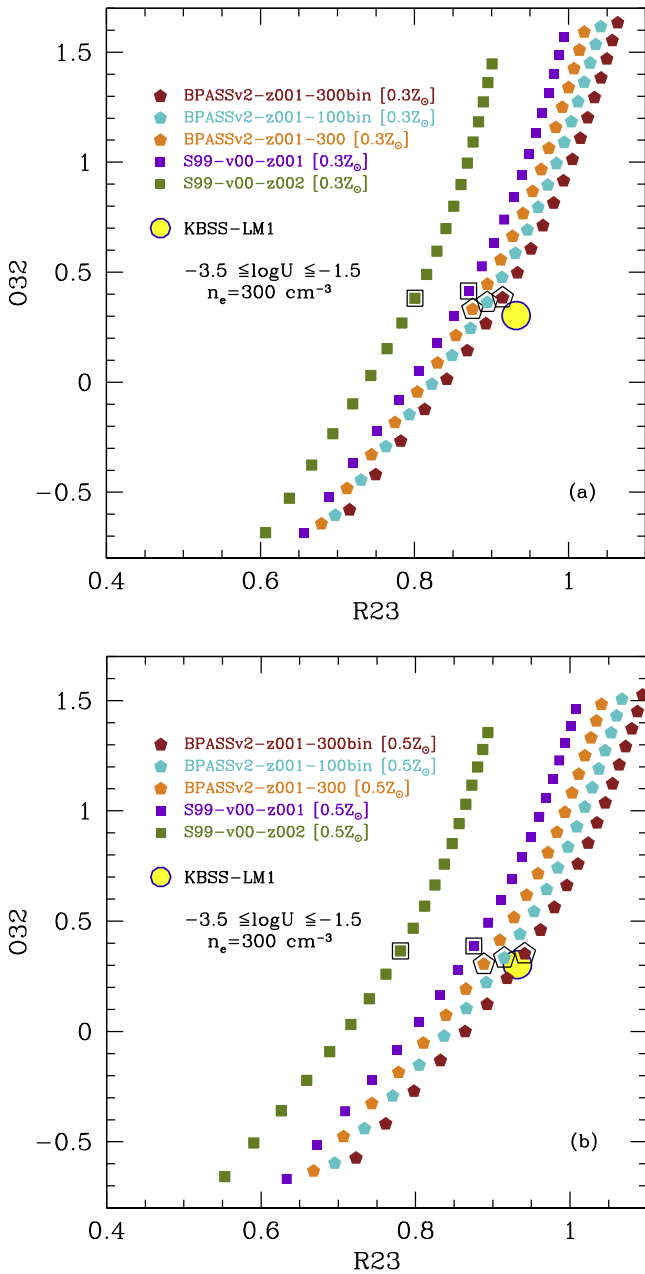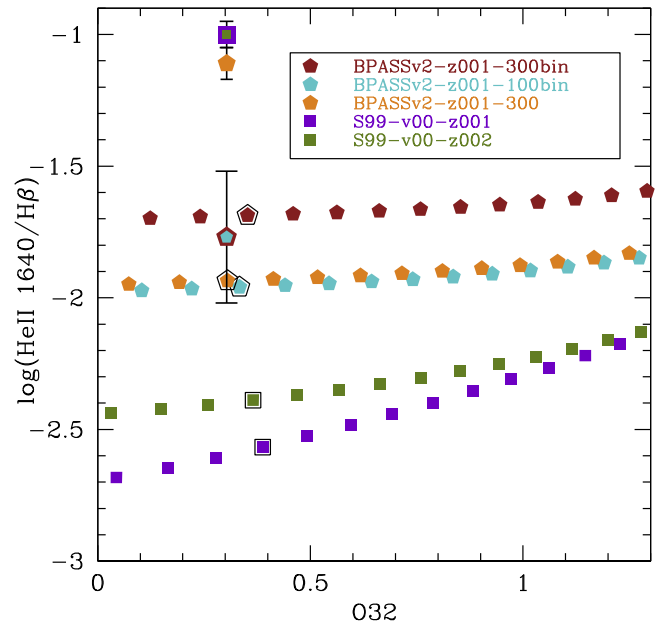
Figure 12. Illustration of the substantially different predictions for *nebular* He II $\lambda1640$ emission relative to Hβ. The colored points with error bars show the residual He II $\lambda1640$/Hβ measurements after subtraction of the corresponding stellar model. Note that the S99 models and the BPASSv2-z001-300 single-star model have much larger He II residual emission than the BPASSv2 binary models because of much weaker predicted *stellar* He II. Only the BPASSv2-z001-100bin and BPASSv2-z001-300bin models predict internally consistent stellar and nebular contributions to the observed He II $\lambda1640$ spectral feature. The colored symbols with black outlines surrounding them have the same meaning as in previous figures.

the O32 versus R23 plane is most sensitive to differences in spectral hardness among the models shown: note that at fixed O32, R23 increases monotonically with increasing spectral hardness—and that only the BPASSv2-z001 *binary* models produce predicted line indices consistent with the observations, and then only when $Z_{neb}/Z_\odot = 0.5$.

7.4. He II Emission

Figure 12 shows that the largest differences between BPASSv2 binary and S99 population synthesis model predictions are for the ratio of He II$\lambda1640$/Hβ. As discussed in Section 4.3 above, the BPASSv2 binary models with continuous star formation predict a stellar He II feature that is absent from the single-star (S99 or BPASSv2) models–this can be understood qualitatively as a direct result of the evolution of massive stars in binaries, which produce an extended evolutionary phase during which stellar effective temperatures can be as high as 10^5 K over a much longer period than classical Wolf–Rayet stars (see, e.g., Eldridge & Stanway 2012). Massive star binaries naturally produce a harder ionizing-UV spectrum with a higher duty cycle than single-star models, which can produce He II-ionizing photons only during brief intervals when very high mass Wolf–Rayet stars are present (see, e.g., Shirazi & Brinchmann 2012; Crowther et al. 2016).

In any case, Figure 12 compares the predictions of the various models for the ratio of *nebular* He II $\lambda1640$/Hβ with the observed residuals after subtraction of the corresponding predicted *stellar* He II emission (Table 5, Figure 8). The points with error bars are the measured residuals from Table 5, color-
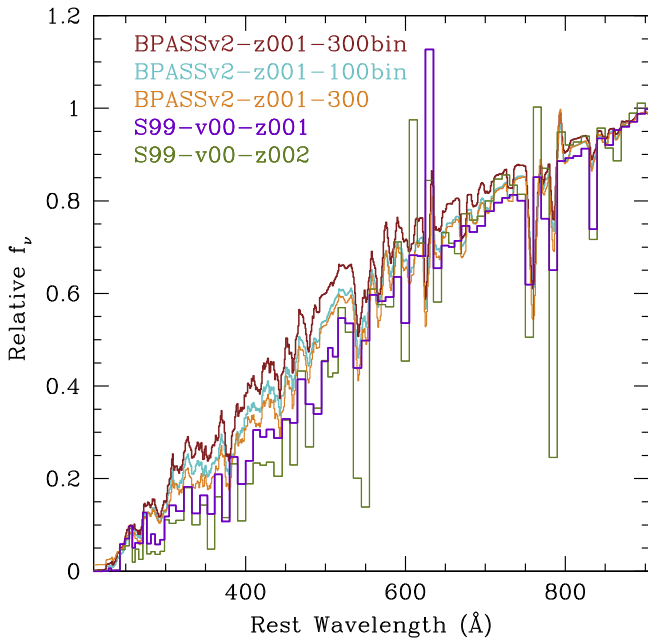

Figure 11. (a) As in Figures 9 and 10 (the LM1 error bars are again smaller than the size of the point), a comparison of the Cloudy model predictions for various assumed input stellar population models in the O32 vs. R23 plane. All of the models shown can reproduce the observed value of O32 with $-2.8 \leqslant \log U \leqslant -2.6$ (as in Figure 9, where the point corresponding to the value of U that most closely reproduces the observed Ne3O2 and O32 ratios is indicated with a black open box/pentagon), but R23 is matched only by the BPASSv2 binary models with $Z_{neb}/Z_\odot = 0.5$. Note that the O32 vs. R23 plane provides the greatest sensitivity to the hardness of the stellar EUV continuum, and that (unlike the BPT diagrams in Figure 10) both BPASSv2 binary models are clearly favored over the BPASSv2-z001-300 single-star model.

phase (O/H). Note that the BPASSv2 models are more sensitive to gas-phase metallicity for values of $\log U$ fixed by O32 and Ne3O2 ratios—the emission in the [O II] and [O III] nebular lines continues to increase until reaching a peak at $Z_{neb}/Z_\odot \simeq 0.5$ ($12 + \log(O/H) = 8.39$), while the S99 models produce their maximum R23 closer to $Z_{neb}/Z_\odot \simeq 0.3$. Comparison of Figures 11 with 10 shows that

Figure 13. Comparison of the EUV 1–4 Rydberg spectra for the same models shown in Figures 9 through 12. All of the models assume constant SFR over 10^8 years, and are normalized so that $f_\nu = 1.0$ at $\lambda_0 = 910$ Å. The models that come closest to reproducing the observed nebular emission line ratios are BPASSv2-z001-300bin and BPASSv2-z001-100bin, the models with the hardest EUV spectra. The other models do not produce enough photons at $\simeq 2$–4 Ryd (225–450 Å) relative to 1 Ryd (912 Å) to match the observed nebular excitation.

Figure 14. The predictions of a subset of the population synthesis models (each assuming the value of log U constrained by the observed O32 and Ne3O2 values shown in Figure 9) for the N2O2 index with respect to N2 $\equiv$ log([N II]$\lambda 6585$/Hα). The large yellow point is the LM1 measurement, where the radius of the circle represents the uncertainties on the measurements. Each model is plotted for a range of assumed $Z_{\rm neb}/Z_\odot = 0.3$–0.7, in linear steps of 0.1, where the symbol size scales with the value. Outside the range log(N/O) $= -1.24 \pm 0.04$ with $Z_{\rm neb}/Z_\odot = 0.5 \pm 0.1$, no combination of N/O and O/H ($Z_{\rm neb}/Z_\odot$) can reproduce the observed point. The BPASSv2-z001 binary models match the observations, with log(N/O) $= -1.24$ when $Z_{\rm neb}/Z_\odot = 0.5$; the marginal values of log(N/O) $= (-1.28, -1.20)$ match when $Z_{\rm neb}/Z_\odot = (0.6, 0.4)$. The allowed range, log(N/O) $= -1.24 \pm 0.04$, is in excellent agreement with that obtained from the local calibration of N2O2 $\rightarrow$ N/O (Section 6.2).

coded by the stellar model subtracted from the LM1 FUV spectrum. A disagreement between the predicted ratio He II $\lambda 1640$/Hβ (locus of color-coded points without error bars) and the residual He II/Hβ indicates an inconsistency between the model and the observations. Only the BPASSv2-z001 binary models (dark red and turquoise pentagons) suggest consistency between the predicted stellar He II and the strength of the residual emission when interpreted as the nebular component of He II $\lambda 1640$. When the BPASSv2-z001 binary population synthesis models are subtracted from the observed KBSS-LM1 spectrum (Figure 8), the *residual* He II feature is narrow and is reduced in intensity by a factor of ~ 6 (turquoise and dark red pentagon with error bar); though its significance is marginal, the residual He II emission is consistent with that predicted by the photoionization models using the same BPASSv2 spectra. The BPASSv2-z001-300 single-star model predicts nebular He II/Hβ comparable to that of the BPASSv2-z001-100bin model shown, but, because it predicts negligible stellar He II, the residual emission (orange pentagon with error bar) is much higher than predicted for nebular He II by the same model.

One issue that becomes relevant in the current context is that photospheric absorption lines in the stellar spectrum are present at wavelengths coincident with the broad stellar He II feature (see Figure 8(a))–this "blanketing" might conceivably mask the presence of a broad He II line in spectra with low S/N and/or low spectral resolution, and it also introduces some degree of model dependence to the measurement of He II emission in galaxy spectra (since the blanketing depends on the details of the photospheric absorption lines).
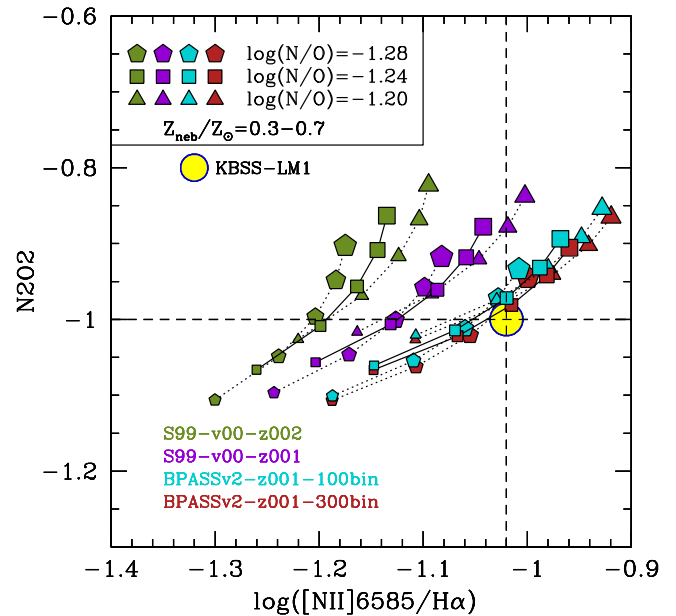
7.5. Spectral Shape in the FUV and EUV

The preceding analysis of the nebular emission line ratios essentially amounts to constraining the spectral shape of the ionizing EUV stellar continuum. Using O32 and Ne3O2 to measure U fixes the intensity of the ionizing radiation field in the 1–3 Ryd range, so that if the spectral shape of a model is off, the predictions for the relative intensities of other lines (e.g., O3, N2, S2, He II/Hβ) will not match the observations. To summarize, of the population synthesis models that provide the best matches to both the global FUV spectrum and to the details of the stellar wind and photospheric absorption lines, the BPASSv2 *binary* models appear to be capable of reproducing simultaneously all of the line ratios observed, at a fixed value of U. Evidently, the stellar ionizing spectrum of this particular sample of (typical) star-forming galaxies at $z \sim 2.4$ needs to satisfy the following:

1. It must be modestly reddened by dust, with $E(B-V)_{\rm cont} \simeq 0.19$, or $A_V \simeq 0.76$ under the assumption of the Calzetti et al. (2000) starburst attenuation relation.
2. It must have low photospheric abundances ($Z_*/Z_\odot \sim 0.1$), to reproduce the limited blanketing of the stellar FUV continuum.
3. The strongest stellar wind lines (most notably, that of C IV) suggest $Z_*/Z_\odot \simeq 0.14$.
4. It must be capable of producing nebular emission lines characterized by R23 ~ 0.93 when O32 $\simeq 0.3$.

5. It must have a sufficiently hard ionizing-UV spectrum to produce substantial stellar He II $\lambda1640$ emission and (probably) nebular He II $\lambda1640/H\beta \simeq 0.02$ (Section 7.4).

The principal reason that the BPASSv2 binary models are more successful overall is the harder ionizing spectrum under conditions of continuous star formation (i.e., star formation timescales of $t_d \gtrsim 20\text{–}30$ Myr; see Section 1). Figure 13 shows a direct comparison of the EUV ionizing spectra for the models discussed in this section. As discussed in Section 7.3, the sequence of increasing EUV hardness is reflected most directly by the R23 parameter at fixed O32 (Figure 11(b)); in both Figures 11(b) and 13, the EUV spectral hardness decreases from top to bottom in the figure legends.

7.6. N/O Revisited

The constraints on the spectral shape of the EUV ionizing radiation field provided by the combination of population synthesis models and ionization parameter as described in Section 7.1 allow us to remove the last remaining dependency on calibrations established at low redshift: the mapping of the N2O2 index (which is observed) to N/O. We showed in section Section 7.2 that the model predictions for the N2, S2, and O3 indices are very well-matched under the assumption that $\log(\text{N/O}) = -1.24$ as implied by Equations (2) and (3). We now examine the constraints on $\log(\text{N/O})$ based only on measured quantities and the population synthesis+photoionization models.

To accomplish this, we allowed N/O to vary independently of the overall scaling of the nebular abundances $Z_{neb}/Z_{\odot}$ in the models, fixed the ionization parameter as described in Section 7.1, and required that the predicted N2O2 and N2 line indices both agree with the observed point within the errors (see Table 7); the results are shown in Figure 14. We find that only the BPASSv2-z001 binary models can reproduce the observed line indices simultaneously, and then only if $\log(\text{N/O}) = -1.24 \pm 0.04$ and $Z_{neb}/Z_{\odot} = 0.5 \pm 0.1$. Figure 14 shows the model predictions for the N2O2 index versus the N2 index (see Table 7 for assumed gas-phase oxygen abundances in the range $Z_{neb}/Z_{\odot} = 0.3\text{–}0.7$.

In Figure 14, the points connected by the solid curve correspond to the $\log(\text{N/O}) = -1.24$ for each color-coded model, for which the $Z_{neb}/Z_{\odot} = 0.5$ metallicity point matches the KBSS-LM1 observation. Also shown are curves assuming the $\pm1\sigma$ marginal values of $\log(\text{N/O})$ that are still consistent with the observation, i.e., $\log(\text{N/O}) = (-1.20, -1.28)$; assuming the marginal values yields best-matching oxygen abundance $Z_{neb}/Z_{\odot} = (0.4, 0.6)$, respectively. Thus, the most successful population synthesis+photoionization models predict that $Z_{neb}/Z_{\odot} = 0.5 \pm 0.1$, or $12 + \log(\text{O/H}) = 8.39^{+0.08}_{-0.10}$; we return to a detailed discussion of the nebular oxygen abundance in Section 8.1.

The abundance ratio N/O, constrained only by the $z \sim 2.4$ KBSS-LM1 observations and the population synthesis+photoionization models that successfully reproduce all other strong-line ratios is identical to N/O obtained from local H II region calibrations. Since both methods yield consistency using either the N2O2 or N2S2 indices, it also implies that nebular S/O = $(S/O)_{\odot}$ to within $\leqslant 0.1$ dex for the KBSS-LM1 ensemble. The N/O abundance ratio for the full $z \sim 2.3$ KBSS-MOSFIRE sample is investigated in a companion paper (Strom et al. 2016).

8. IONIZED GAS-PHASE ABUNDANCES

We showed in Section 7 that the strong-line ratios observed in the nebular spectra of the KBSS-LM1 ensemble are reproduced by a model assuming nebular abundances for O (and S) of $Z_{neb}/Z_{\odot} = 0.5 \pm 0.1$ and $\log(\text{N/O}) = -1.24 \pm 0.04$ ([N/O] = -0.38 ± 0.04, or $\simeq 0.42$ (N/O)$_{\odot}$). In this section, we compare the model results for O/H with various other commonly used methods, and present additional abundance ratio determinations made possible by the combination of the KBSS-LM1 spectra and the successful population synthesis+photoionization models.

8.1. Nebular Oxygen Abundance

Table 7 summarizes the oxygen abundance measurements based on the nebular emission lines in the KBSS-LM1 composite spectra. In this section, we report a direct T_e measurement of O/H, and compare it to the results of the stellar population synthesis+photoionization modeling presented in Section 7 above. Comparisons of the direct and modeled oxygen abundances to several estimates based on strong-line indices are presented in Section 8.1.3.

8.1.1. Direct Method

The KBSS-LM1 nebular spectra are of high enough quality to yield a T_e measurement from the observed ratio O3uv $\equiv$ O III]$(1661 + 1666)/$[O III]$\lambda5008$ (see Table 7), which gives $T_e(\text{O III}) = 12250 \pm 600$ K. Assuming that $T_e(\text{O II}) = 0.7 \, T_e(\text{O III}) + 3000$ (e.g., Campbell et al. 1986; Garnett 1992; Izotov et al. 2006; Pilyugin et al. 2009; hereinafter referred to as the "T_2–T_3" relation) and that $(\text{O/H}) = (\text{O}^{++}/\text{H}^+) + (\text{O}^+/\text{H}^+)$, the corresponding oxygen abundance is

$$12 + \log(\text{O/H})_{dir} = 8.14 \pm 0.04;$$
$$Z_{neb}/Z_{\odot} = 0.29 \pm 0.03. \tag{5}$$

The error bars account both for the uncertainty in the line fluxes (dominated by the uncertainty in the UV O III] doublet) and in the extinction correction from the Balmer decrement measurement. However, the quoted error may under-estimate the true uncertainty given that we have assumed an extinction curve that may not be the correct one for nebular extinction in the high-redshift objects. [22]A smaller extrapolation would be required to use the [O III]$\lambda4364$ auroral line instead of the UV O III] intercombination doublet, but the former was covered in the observed wavelength range for only $\sim50\%$ of the KBSS-LM1 galaxies; we have therefore treated a tentative detection as a 3σ upper limit on the electron temperature, $T_e < 13370$ K (see Table 7), consistent with the T_e measurement based on the UV O III] feature.

Another potential source of systematic error is that we have not *measured* the electron temperature associated with the [O II]-emitting gas; the same problem has affected essentially all other high-redshift estimates, and most estimates at low redshift, in the literature. In the absence of a measurement of T_e for a missing ionization stage, it is common to use the T_2–T_3

---

relation mentioned above, which was originally established using photoionization models but has been generally supported by observations of local galaxy H II regions (e.g., Brown et al. 2014; Berg et al. 2015). Nevertheless, the applicability of this relation is a source of some controversy; for example, Andrews & Martini (2013) (AM13) have shown that for composite SDSS galaxy spectra in bins of M_* and SFR, the standard T_2–T_3 relation systematically over-predicts T_e(O II) at a given T_e(O III) for the vast majority of the bins in M_* and M_* − SFR. The exceptions are for objects with log SFR > 1.0 (rare in the SDSS sample), which are generally in good agreement with the T_2–T_3 formulation. The implication for the KBSS-LM1 sample is unclear, since most of the $z \sim 2.4$ galaxies have log(SFR) > 1.0.

However, it is interesting to ask how the possible offset of T_2 as found by AM13 would affect the direct method oxygen abundance for the KBSS-LM1 composite. For what AM13 called their "M_* − SFR" stacks, they found an average offset of $\Delta T_2 = -1300$ K relative to the T_2–T_3 relation. Applying this offset to the KBSS-LM1 composite would change the direct method oxygen abundance $12 + log(O/H)_{\rm dir}$ from 8.14 to 8.25 (i.e., from $Z_{\rm neb}/Z_\odot = 0.28$ to $Z_{\rm neb}/Z_\odot = 0.36$).

8.1.2. Comparison to Population Synthesis+Photoionization Model Results

There is a well-known tendency for measurements of T_e based on collisionally excited lines (CELs) to systematically under-estimate nebular O/H (see e.g., Peimbert & Peimbert 2002 and references therein); this effect is generally ascribed to the fact that, for nebulae having zones of differing temperature, the high temperature regions will be over-represented due to the temperature sensitivity of the emissivity of CELs, leading to an over-estimate of T_e. Nebular recombination lines (RELs), on the other hand, are much less temperature sensitive, and when both CELs and RELs can be measured for the same systems, the inferred oxygen abundances based on RELs are systematically higher, and generally in better agreement with the *stellar* O/H measured from nearby regions in the same galaxy—including the well-studied Orion and 30 Doradus regions in the Galaxy and LMC, respectively. Recently, Esteban et al. (2014) used high-quality echelle spectra of star-forming "knots" in nearby galaxies to measure nebular abundances from both CELs and RELs, finding a systematic offset

$$\log(O/H)_{\rm REL} - \log(O/H)_{\rm CEL} = 0.24 \pm 0.02 {\rm dex} \quad (6)$$

(median and inter-quartile range), consistent with other similar measurements for individual H II regions in the literature (e.g., Blanc et al. 2015).

For some purposes, it is sufficient simply to be consistent in the method one uses to estimate nebular O/H, as the *relative* values of O/H will be preserved within a sample. Because the metallic RELs are extremely weak, it is far more difficult to measure O/H from RELs—such measurements may remain out of reach for high-redshift galaxies for the foreseeable future. It therefore makes sense to use CEL-based direct measurements (or their strong-line proxies) for wholesale measurements of gas-phase oxygen abundances. However, for sensible comparisons between stars and ionized gas in the same galaxies (of paramount interest in this paper), it is important to acknowledge that CEL-based T_e *measurements* are likely to yield numerical values of log(O/H) lower than the *actual* gas-phase O/H (or the stars in the same regions) by $\simeq 0.24$ dex. In

this context, applying the offset from Equation (6) to the KBSS-LM1 (CEL-based) T_e measurement (Equation (5)) gives

$$12 + \log(O/H)_{\rm REL} = 8.38 \pm 0.04$$
$$Z_{\rm neb}/Z_\odot = 0.49 \pm 0.05. \quad (7)$$

This is the same value for the nebular oxygen abundance favored by the the population synthesis+photoionization modeling described in Section 7; thus, we adopt $Z_{\rm neb}/Z_\odot = 0.5$ as the most probable value of the "true" nebular oxygen abundance of the KBSS-LM1 ensemble.

Recall that the most likely *stellar* metallicity obtained from matching the observed spectrum of KBSS-LM1 to population synthesis models (Section 4) is $Z_*/Z_\odot = 0.07$–0.14, lower by a factor of $\simeq 4$–5. This apparent discrepancy between the metallicity of the *stars* and that of the *ionized gas* required to simultaneously match the FUV stellar spectra and the FUV/optical nebular emission in the same galaxy ensemble is discussed in Section 9 below.

8.1.3. Strong-Line Methods

For completeness, we have evaluated the nebular oxygen abundance based on various strong-line indices commonly used for that purpose, summarized in Section 7. By construction, the strong-line methods are intended to reproduce the direct T_e metallicity scale, for use when spectra are of insufficient quality to allow an actual measurement of T_e-sensitive line ratios.

Among the strong-line abundance measurements listed in Table 7, the only one that is not calibrated using the Pilyugin et al. (2012) H II region data set is that of R23, which instead applies the low-metallicity branch R23 calibration from McGaugh (1991). This calibration used a combination of photoionization models and T_e measurements, and accounts for the ionization parameter through the measured value of O32,

$$\begin{aligned} 12 + \log(O/H)_{\rm R23} = {} & 12 - 4.944 + 0.767 * {\rm R23} \\ & + 0.602 * ({\rm R23})^2 \\ & - {\rm O32} * [0.290 + 0.332 * {\rm R23} \\ & - 0.331 * ({\rm R23})^2] \end{aligned} \quad (8)$$

where R23 and O32 are as defined in Table 7. From this, we obtain for the KBSS-LM1 composite:

$$12 + \log(O/H)_{\rm R23} = 8.20 \pm 0.03;$$
$$Z_{\rm neb}/Z_\odot = 0.32 \pm 0.03, \quad (9)$$

where the quoted error includes uncertainties in the R23 and O32 line indices and their nebular extinction corrections. Systematic errors were estimated by McGaugh (1991) to be $\sim \pm 0.1$ dex.

For the other strong-line estimates of oxygen abundance, we use calibrations based on the same Pilyugin et al. (2012) H II region sample discussed in Section 6.2. Strom et al. (2016) present linear fits of the strong-line O3N2 and N2 indices to T_e-based oxygen abundances:[23]

$$12 + \log(O/H)_{O3N2} = 8.56 - 0.20 * {\rm O3N2}$$
$$\sigma = 0.08 {\rm \ dex} \quad (10)$$

---

[23] Only the H II regions with $12 + \log(O/H)_{\rm dir} \geqslant 8.0$ were included in the fit; objects with lower metallicity would not be compatible with the observed locations of the KBSS-MOSFIRE sample on the BPT diagram.

and

$$12 + \log(O/H)_{N2} = 8.68 + 0.38 * N2$$
$$\sigma = 0.09 \text{ dex}, \quad (11)$$

where the uncertainties are the (global) rms scatter between the fit and individual measurements in the calibration set. One can also use the Pilyugin et al. (2012) data set to calibrate the N2O2 index (used in Section 6.2 for an initial estimate of N/O) directly onto O/H, for $12 + \log(O/H) \geqslant 8.0$:

$$12 + \log(O/H)_{N2O2} = 8.63 + 0.40 * N2O2$$
$$\sigma = 0.09 \text{ dex}. \quad (12)$$

The calibrations given in Equations (10) and (11) use somewhat different parameters compared to the commonly used calibrations of Pettini & Pagel (2004; PP04); however, for the line indices measured in the KBSS-LM1 composite the results are very similar, as we show below. Error bars assigned to all strong-line metallicity estimates below (see also Table 7) reflect uncertainties in the measurement of the line indices (including extinction corrections, where relevant) as well as the rms scatter in the calibration itself, accounting for the fact that the KBSS-LM1 measurements are based on an ensemble, rather than a single galaxy. They do not attempt to account for possible systematic errors caused by differences between the calibration sample and the KBSS-LM1 ensemble, and all are based on the simplest linear relationship between line index and direct T_e (CEL) oxygen abundance in the calibration set.

For O3N2, we find

$$12 + \log(O/H)_{O3N2} = 8.23 \pm 0.02$$
$$Z_{neb}/Z_\odot = 0.35 \pm 0.02. \quad (13)$$

An identical value would be obtained using the PP04 O3N2 metallicity calibration.

For N2, we obtain

$$12 + \log(O/H)_{N2} = 8.29 \pm 0.02$$
$$Z_{neb}/Z_\odot = 0.40 \pm 0.02. \quad (14)$$

Applying the (linear) PP04 N2 calibration instead would yield an oxygen abundance *higher* by 0.03 dex ($\sim$10%).

The results for O3N2 and N2 differ in the same sense for both the PP04 calibration and those in Equations (10) and (11)—with N2 yielding higher log(O/H) than O3N2—but the amplitude of the difference is reduced: $\log(O/H)_{N2} - \log(O/H)_{O3N2} = 0.06$ dex for the new calibration, compared to $\log(O/H)_{N2} - \log(O/H)_{O3N2} = 0.12$ dex using PP04 calibrations. As discussed by Steidel et al. (2014), the smaller difference is due in large part to a closer match between the range of line indices represented in the calibration sample and those in the sample being calibrated.

Using the measured N2O2 index for the KBSS-LM1 spectrum, Equation (12) yields

$$12 + \log(O/H)_{N2O2} = 8.23 \pm 0.03$$
$$Z_{neb}/Z_\odot = 0.35 \pm 0.03, \quad (15)$$

identical to the estimate based on the O3N2 index.

8.1.4. N/O "Analogs"

Alternatively, within the Pilyugin et al. (2012) calibration data set there are 42 H II regions ($\simeq$10% of the full sample) with $-1.28 \leqslant \log(N/O) \leqslant -1.20$, thus matching the range of

N/O inferred for the KBSS-LM1 composite discussed in Section 7.6; hereafter, we refer to this subsample as "N/O Analogs" (cf. Section 6.2) for KBSS-LM1. We have already pointed out that using the N2S2 and N2O2 indices with the local calibration sample predicts the same value of N/O obtained using our photoionization modeling; here we can test whether the modeled N/O successfully predicts the direct T_e abundances in the calibration sample. The median direct T_e (CEL) oxygen abundance of the N/O analog sample is

$$12 + \log(O/H)_{N/O} = 8.13 \pm 0.02$$
$$Z_{neb}/Z_\odot = 0.28 \pm 0.02 \quad (16)$$

where the errors represent the uncertainty in the median of the calibration sample of 42.

In other words, matching the *modeled* N/O for the KBSS-LM1 composite to analogs in the H II region calibration sample predicts a T_e (CEL) oxygen abundance within 0.01 dex of the value *measured* in KBSS-LM1 (Equation (5)). We have already shown that the oxygen abundance predicted by the modeling, without reference to the local calibration sample, is the same value obtained from the direct (CEL) measurement after converting to the REL abundance scale using Equation (6). By extension, there appears to be full consistency in the mapping between N/O and O/H of the calibration sample and the measurements (both direct and modeled) of the KBSS-LM1 composite. The level of consistency implies that N/O at fixed O/H (or O/H at fixed N/O), is nearly identical in $z \simeq 0$ extragalactic H II regions and $z \sim 2.4$ star-forming galaxies.

Since the measurement of N/O via N2O2 is not strongly dependent on ionization parameter and/or the shape of the EUV ionizing radiation field (unlike O32, R23, N2, and O3N2), and *is* strongly dependent on O/H, the most reliable strong-line measurement of O/H (when information is limited) may be obtained through N/O.

8.1.5. Implications for Strong-line Measurements

To summarize, estimates of the oxygen abundances based on the measurement of various strong-line ratios—anchored by the same H II region calibration data set—span the range $8.20 \leqslant 12 + \log(O/H) \leqslant 8.29$ depending on the method used. All of the strong-line methods appear to over-estimate, by varying degrees, the T_e (CEL) abundance measured in the high-redshift LM1 composite. Use of the N2 index results in the largest over-estimate ($\simeq$0.15 dex). Nevertheless, we have shown that a higher N2 index at fixed N/O and O/H is expected when the EUV spectrum of ionizing stars is systematically harder compared to the calibration sample at the same gas-phase O/H. Once this has been accounted for, there is no evidence for redshift-dependent behavior of N/O versus O/H.

The comparatively poor performance of N2 $\equiv$ [N II]$\lambda6585$/Hα for predicting the T_e (CEL) oxygen abundance of the high-redshift composite can be understood by referring to Figure 14: for fixed N/O, increasing the hardness of the radiation field (e.g., between the S99-v00-z002 model and the BPASSv2-z001 binary models shown) increases the N2 index by $\sim$0.2 dex while leaving N2O2 nearly unchanged at fixed O/H. In contrast, the similarity (in both slope and normalization) of Equations (11) and (12) suggests that the N2 and N2O2 indices are nearly indistinguishable in the way they map to direct oxygen abundances within the calibration sample. Harder ionizing spectra at fixed O/H and N/O at high redshift
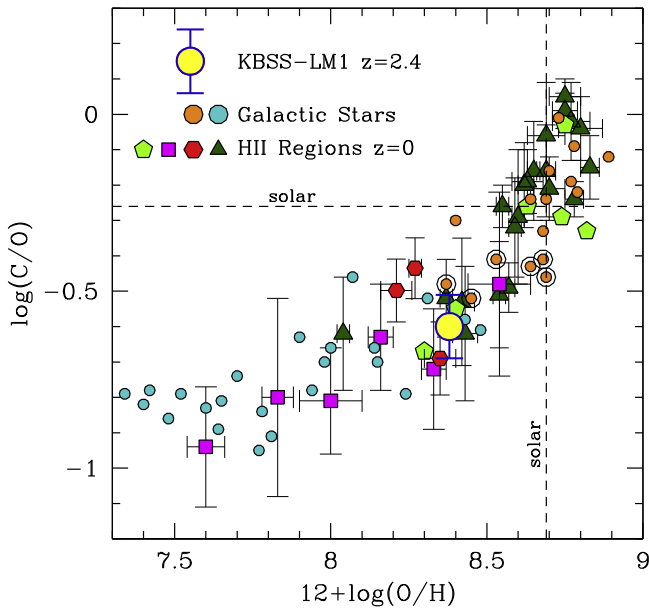
Figure 15. Comparison of log(C/O) vs. 12+log(O/H) for Galactic stars, low-metallicity dwarf galaxies, and extragalactic H II regions with the measurement from the $z \sim 2.4$ KBSS-LM1 composite (large yellow point). The low-redshift data include metal-poor halo (Akerman et al. 2004; turquoise points) and "thick-disk" (Bensby & Feltzing 2006; orange points) stars, dwarf galaxies (Garnett et al. 1995b; Kobulnicky & Skillman 1998; magenta squares and red hexagons, respectively), local spiral galaxy H II regions (Garnett et al. 1999; light green pentagons), and measurements of bright emission line knots in local star-forming galaxies (Esteban et al. 2014; dark green triangles). All nebular values of O/H are based on T_e measurements from collisionally excited emission lines, and have been adjusted to the REL oxygen abundance scale (presumed to match the *stellar* abundance scale) by adding $\Delta \log(O/H) = 0.24$ dex (Equation (6); see text for discussion). Stars with [O/Fe] > +0.4 from the Bensby & Feltzing (2006) sample are surrounded by an additional black circle. The solar values of log(O/H) and log(C/O) are indicated with dashed lines.

cause asymmetric shifts in the N2 and N2O2 line indices relative to the calibration sample, causing a systematic shift in the mapping between strong-line index and abundance. The O3N2 index comes closer than N2 to predicting the T_e oxygen abundance because both O3 and N2 *increase* with increasing spectral hardness (at fixed O/H and N/O)—so that the amplitude of the systematic shift is reduced simply because the relevant line index is the difference, O3–N2.

It is important to note that the amplitude of the systematic errors (relative to the T_e/CEL abundances) introduced by the strong-line methods will depend on the details of the differences in excitation at fixed O/H between the calibration sample and the sample to which the calibration is applied. For reasons discussed in detail in Section 9, the amplitude of systematic errors caused by such "calibration sample mismatch" is likely to increase with $Z_{neb}/Z_\odot$.

8.2. C/O

The use of the rest-UV line ratio C3O3 (see Table 7) as an indicator of the C/O abundance ratio in low-metallicity H II regions was originally proposed by Garnett et al. (1995b), and has been employed in a number of studies of high-redshift galaxies (e.g., Shapley et al. 2003; Erb et al. 2010) thanks to the relative ease with which FUV spectra of such objects can be obtained. In the rest-frame FUV, the C III] and O III] features are
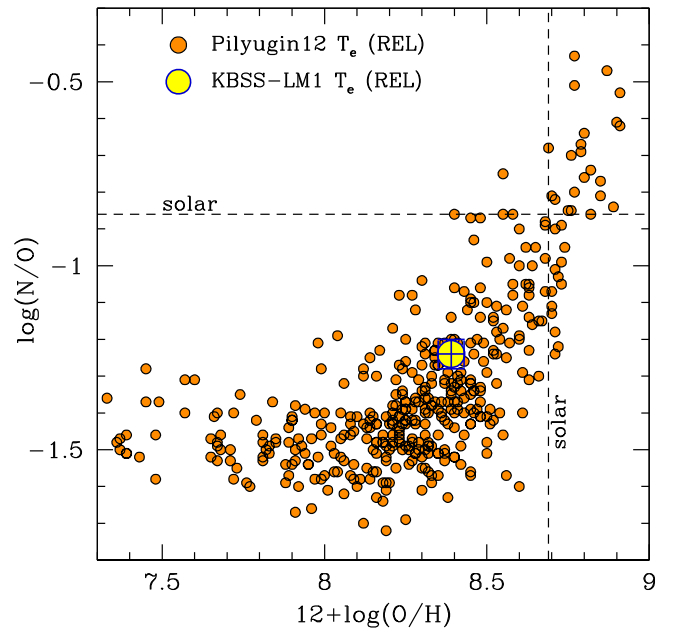

Figure 16. Similar to Figure 15, but for N/O vs. 12+log(O/H). As in Figure 15, both the local measurements (in this case, the Pilyugin et al. 2012 compilation of extragalactic H II regions) and the KBSS-LM1 T_e measurements have been shifted by +0.24 dex in log(O/H) as in Equation (6). Recall that the oxygen abundance favored by the best-fitting photoionization model (Section 7.6) is identical to the T_e oxygen abundance after applying this shift from the CEL to REL scale. The KBSS-LM1 value of N/O is as determined in Section 8.3.

close enough in wavelength that they are relatively immune to uncertainties in extinction, and at the same time C^{++} and O^{++} are expected to be by far the dominant ions for both C and O, so that ionization corrections needed to infer C/O from C^{++}/O^{++} are small and easily modeled.

Rather than using the approximations given by Garnett et al. (1995b), we make use of the photoionization grid point that combines the BPASSv2-z001 binary population synthesis model with $\log U = -2.8$ (Figure 9) and oxygen abundance in the range discussed in the previous section ($Z_{neb}/Z_\odot = 0.5 \pm 0.1$, or $12 + \log(O/H) \simeq 8.4 \pm 0.1$); this allows us to apply internally consistent ionization corrections and to include the expected dependence of C3O3 on O/H in estimating the uncertainties. We re-ran the photoionization models, varying C/O while holding O/H fixed within the range $Z_{neb}/Z_\odot = 0.5 \pm 0.1$, as discussed in Section 8.1.2, and $\log U$ fixed at the same value shown in Figure 9. The observed ratio (Table 7) is C3O3$_{obs}$ = 0.62 ± 0.07, which we find is reproduced by the BPASSv2-z001 binary models when $\log(C/O) = -0.60 \pm 0.09$, where the error bars include the uncertainties in inferred $Z_{neb}/Z_\odot$ as well as the observational uncertainties in the C3O3 index. Comparing the inferred C/O with the solar ratio $\log(C/O)_\odot = -0.26$ (Asplund et al. 2009),

$$[C/O] = C3O3_{obs} - C3O3_\odot = -0.34 \pm 0.09. \quad (17)$$

Figure 15 shows that this value is very much in line with the trend of C/O versus O/H observed in Galactic stars and H II regions in nearby dwarf and spiral galaxies. Note that the value of log(O/H) for the KBSS-LM1 point has systematic uncertainty for the reasons discussed in Section 8.1; the same is true for local H II region measurements. For internal consistency, all of the nebular values in Figure 15 are based

on T_e (CEL) measurements from collisionally excited emission lines, but have been corrected to the REL abundance scale according to Equation (6) by adding 0.24 dex (see the discussion in Section 8.1.2). [24] As discussed by Esteban et al. (2014), a similar offset to that given in Equation (6) is likely to be present between CEL-based direct determinations and the *stellar* abundance scale; thus, we include samples of Galactic halo and thick-disk stars with accurate photospheric abundance measurements of C/O and O/H in Figure 15 as well. Clearly, correcting the nebular measurements to the REL oxygen abundance scale brings them into excellent agreement with the locus occupied by Galactic stars over the full range in common; in addition, the locus of both stars and nebulae crosses $(C/O)_\odot$ at $(O/H)_\odot$.

There have been several previous estimates of C/O at high redshifts: the most directly comparable to the KBSS-LM1 composite is that of Shapley et al. (2003), who measured C3O3 from composite spectra of $z \simeq 3$ LBGs, finding that $\log(C/O) = -0.68 \pm 0.13$ for the full galaxy sample and $\log(C/O) = -0.74 \pm 0.14$ for the subset with the strongest Lyα emission lines. Both values assumed that C/O = $C^{++}/O^{++} \times$ ICF with assumed ICF = 1.0. [25] Erb et al. (2010) derived a value of $\log(C/O) = -0.62 \pm 0.09$ with ICF = 1.7 ± 0.2 for a single galaxy, Q2343-BX418 (see also Steidel et al. 2014), for which a much higher ionization parameter was assumed ($\log U = -1.0 \pm 0.15$). [26] These previous measurements are all consistent with KBSS-LM1 within their observational and modeling uncertainties; the new KBSS-LM1 measurement has considerably smaller systematic uncertainties, however, because it is derived based on a stellar +nebular model that simultaneously reproduces the FUV and optical line intensity ratios as well as the FUV stellar spectrum.

8.3. N/O

We have already discussed the determination of N/O for the KBSS-LM1 composite in Sections 6.2 and 7.6, and have shown that, in spite of the fact that the N2 index over-estimates the T_e oxygen abundance by ~0.15 dex, the measured N/O and the T_e (CEL) measurement of O/H are in excellent agreement with the locus occupied by local H II regions in the calibration sample (Figure 16 and Section 8.1.1).

In view of the results for C/O given in Section 8.2 above, it is instructive to cast the N/O measurements in similar terms. Figure 16 shows the the Pilyugin et al. (2012) calibration sample and the KBSS-LM1 point after applying the offset given in Equation (6) to the measured T_e (CEL) oxygen abundances, as was done for C/O and local comparison samples in Figure 15. As was the case for C/O versus O/H, the shift applied to $\log(O/H)$ from Equation (6) causes the locus of H II regions to cross $(O/H)_\odot$ when $(N/O) \simeq (N/O)_\odot$.

---

[24] The values of C/O should require no systematic correction, since nebular measurements of the ratio have only a weak dependence on the assumed oxygen abundance–see also the discussion in Esteban et al. (2014).

[25] The ionization correction factor implied by the best-fitting model for KBSS-LM1 is ICF = 0.78, which if applied to the Shapley et al. (2003) results would *lower* the inferred value of $\log(C/O)$ by ~0.11 dex.

[26] The emission line ratios of Q2343-BX418 from more recent KBSS-MOSFIRE observations can be reproduced with $\log U \approx -2.3$ using the same BPASSv2 population synthesis model as in the present work, which predicts ICF $\simeq 0.97$; a new LRIS observation gives C3O3 = 0.48 ± 0.03, and inferred $\log(C/O) = -0.62 \pm 0.05$, in very good agreement with that of the LM1 composite (Table 7).

The axis scales in Figure 16 have been deliberately adjusted to show the same range in N/O as for C/O in Figure 15, to emphasize the similarity of the behavior of C/O and N/O versus O/H. Note that the KBSS-LM1 composite exhibits nearly identical under-abundance of C and N with respect to O: $[C/O] = -0.34 \pm 0.09$ compared to $[N/O] = -0.37 \pm 0.04$ (see Table 7). The overall patterns of C/O and N/O are remarkably similar for the whole range of O/H in common; while the nucleosynthetic origins of C and N are beyond the scope of this work, comparison of Figures 15 and 16 does suggest a common origin of N and C in which both are closely tied to O/H.

8.4. Si/O

As for the C/O abundance determination, we make use of the observed ratio Si3O3 (Table 7) in conjunction with the most successful photoionization model. The solar ratio of silicon to oxygen according to Asplund et al. (2009) is $\log(Si/O)_\odot = -1.18$, for which the BPASSv2-z002 model ($\log U = -2.8$) predicts $0.61 \leqslant Si3O3_\odot \leqslant 0.65$ for $0.4 \leqslant Z_{neb}/Z_\odot \leqslant 0.6$. The measurement of Si3O3$_{obs}$= 0.00 ± 0.08 (Table 7) thus implies that

$$\log(Si/O) = Si3O3_{obs} - Si3O3_\odot = -1.81 \pm 0.10 \quad (18)$$

or $[Si/O] = -0.63 \pm 0.10$ in the ionized gas phase of the $z \simeq 2.4$ galaxies. Since Si and O are generally believed to be produced by the same processes, the observed [Si/O] under-abundance relative to solar is likely dominated by depletion of Si onto dust grains. Adopting the definition

$$[X/Y]_{gas} \equiv \log(X/Y)_{neb} - \log(X/Y)_{tot} \quad (19)$$

where X and Y are elements, $\log(X/Y)_{neb}$ is the nebular (gas-phase) abundance ratio, and $\log(X/Y)_{tot}$ is the element ISM ratio (gas+dust), we find $[Si/H]_{gas} \simeq -0.63$. This value represents an intermediate depletion level relative to the full range seen in Galactic sightlines ($-1.36 \leqslant [Si/H]_{gas} \leqslant -0.22$; Jenkins 2009; see Section 8.5).

Nebular measurements of Si/O at low redshift are few, owing to the need for FUV spectroscopic observations. In an *HST*/UV spectroscopic study of H II regions in dwarf galaxies, Garnett et al. (1995a) found a weighted average value $\langle \log(Si/O) \rangle = -1.59 \pm 0.07$, and saw no obvious trend with (O/H) over the range probed. At the time of that work, the solar ratio $\log(Si/O)_\odot$ was believed to be ~0.2 dex *lower* than the best current values, which led the authors to speculate that depletion of Si onto dust grains was significantly lower than expected by then-current grain formation models; however, with the currently accepted value of $(Si/O)_\odot$, the dwarf galaxy measurements ($[Si/H]_{gas} \simeq -0.4$) fall well within the range inferred from interstellar absorption line measurements, and have slightly lower depletion than inferred for the KBSS-LM1 composite.

8.5. Effects of Depletion in the Ionized Gas

The inferred depletion of Si onto dust grains (Section 8.4) can be used to estimate the degree to which other elements–in this case, we are interested primarily in C, N, O, and S–are likely to be depleted as well. Assuming gas-phase depletion $[Si/H]_{gas} = -0.63$, the depletion expected for the other

elements can be calculated from the fits given by Jenkins (2009) with the parameter $F_* = 0.36 \pm 0.09$ (see his Equation (10) and Table 4) determined from Si: $[\mathrm{O/H}]_{\mathrm{gas}} \simeq -0.09$, $[\mathrm{C/H}]_{\mathrm{gas}} \simeq -0.15$, $[\mathrm{N/H}]_{\mathrm{gas}} \simeq -0.11$, and $[\mathrm{S/H}]_{\mathrm{gas}} \simeq -0.15$, with uncertainties of ± 0.05 dex dominated by the scatter in the calibration. If these values are approximately correct, then the total oxygen abundance would be systematically higher by ~ 0.09 dex ($\sim 20\%$), but the inferred nebular abundance *ratios* C/O, N/O, S/O would be essentially unaffected by depletion ($\leqslant 0.05$ dex).

Note that increasing the nebular oxygen abundance by 0.1 dex would *increase* the apparent difference between stellar and nebular metallicity for the KBSS-LM1 composite spectra.

8.6. Abundances: Summary

1. The stellar abundances are traced directly by photospheric and wind lines in the FUV spectrum, and indirectly by the inferred EUV spectrum constrained by the nebular line ratios. Both suggest that $Z_*/Z_\odot \simeq 0.10 \pm 0.03$.

2. The nebular oxygen abundance lies in the range $0.3 \leqslant Z_{\mathrm{neb}}/Z_\odot \leqslant 0.6$ according to the best-fitting photoionization models, a direct T_e measurement of O III, and several locally calibrated strong-line estimates. However, an ionized gas-phase metallicity of $Z_{\mathrm{neb}}/Z_\odot = 0.5$ is favored by the photoionization modeling, which does not rely on the applicability of local calibrations and has been subjected to the strongest constraints imposed by the requirement of internal consistency. $Z_{\mathrm{neb}}/Z_\odot = 0.5$ is also inferred from the direct T_e measurement assuming that collisionally excited emission lines (CELs) underestimate the true gas-phase oxygen abundance by $\Delta \log(\mathrm{O/H}) = 0.24$ dex as in the local universe.

3. Using FUV emission lines of C III] and O III], whose ratios are insensitive to extinction corrections and only mildly dependent on O/H, we have measured $\log(\mathrm{C/O}) = -0.60 \pm 0.09$. The measurement is consistent with the observed trend of C/O versus O/H in Galactic stars and the H II regions of nearby galaxies, and provides external support for the inferred O/H. Together, the C/O and O/H measurements suggest that $[\mathrm{C/H}] \simeq -0.7$, i.e., $(\mathrm{C/H}) \simeq 0.2 \, (\mathrm{C/H})_\odot$.

4. The nebular abundance ratio $[\mathrm{Si/O}] = -0.63$ is presumably due to depletion onto dust grains, suggesting that $[\mathrm{Si/H}]_{\mathrm{gas}} \simeq -0.63$ if $[\mathrm{Si/O}]_{\mathrm{tot}} \simeq 0.0$. The level of Si depletion implies gas-phase depletions of C, N, O, and S of $[\mathrm{X/H}]_{\mathrm{gas}} \simeq -0.1$. Thus, inferred metallic abundance ratios N/O, C/O, and S/O should be largely unaffected by dust grains.

9. SUMMARY AND DISCUSSION

In this paper, we have demonstrated that spectroscopic measurements of the rest-frame FUV stellar continuum, *and* nebular emission lines in the FUV ($1000 \lesssim \lambda_0 \lesssim 2000$ Å) and optical ($3700 \lesssim \lambda_0 \lesssim 7000$ Å) can be used to extract detailed physical and chemical properties of stars and ionized gas in high-redshift star-forming galaxies. Perhaps the most significant development is that the observations can be understood in the context of a self-consistent model of the massive star population and ionized gas-phase physical conditions.
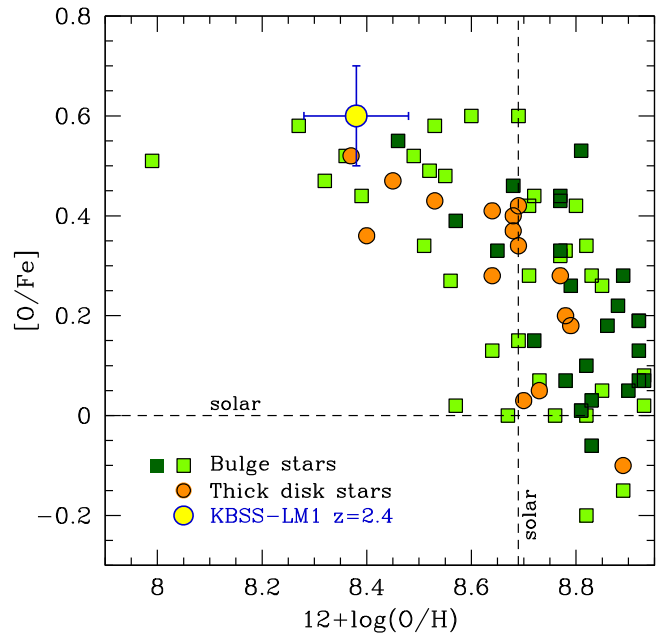

Figure 17. The enhancement in O/Fe relative to solar as a function of *stellar* O/H for the same Galactic thick-disk stars as in Figure 15 (orange points; Bensby & Feltzing 2006) and for two samples of Galactic bulge stars: microlensed dwarfs (light green squares; Bensby et al. 2013) and K-giants (dark green squares; Lecureur et al. 2007). The point representing KBSS-LM1 (large yellow point) is shown for comparison, assuming $(\mathrm{O/Fe}) \approx (4 \pm 1) \times (\mathrm{O/Fe})_\odot$ and the nebular oxygen abundance (Section 8.1.2) from direct T_e measurements and photoionization modeling, $12 + \log(\mathrm{O/H}) = 8.38 \pm 0.10$ ($Z_{\mathrm{neb}}/Z_\odot \approx 0.5 \pm 0.1$).

The results for the KBSS-LM1 composite spectra offer a global cross-check on the physical interpretation of the most readily accessible observational signatures of high-redshift galaxies, and are almost entirely independent of calibrations anchored at low redshift. Below, we briefly summarize and discuss the most significant results and their implications.

9.1. Which Metallicity?

We saw above that the most successful model in simultaneously predicting the observed stellar FUV and nebular FUV+optical spectrum of the $z = 2.4$ KBSS-LM1 composite has $Z_*/Z_\odot \simeq 0.1$ and $Z_{\mathrm{neb}}/Z_\odot \simeq 0.4$–$0.5$; that is, a factor of $\simeq 4$–5 difference between stellar and nebular "metallicity"!

However, such behavior is actually *expected* for galaxies whose ISM has been enriched primarily by Type II (corecollapse) supernovae, by the following argument: first, the FUV and EUV spectra of massive stars are most sensitive to the elements dominating the opacity in stellar interiors, atmospheres, and winds—Fe being especially important—whereas elements acting as important coolants in ionized gas with $T \simeq 10^4$ K, and thus producing strong emission lines—notably O—play only a minor role within stars. Similarly, in spite of its importance in stars, Fe is depleted onto dust grains in the ISM and therefore relatively unimportant as a coolant in most H II regions.

It is interesting to consider the limiting case that stellar opacity and mass-loss rates depend only on stellar Fe/H, and nebular emission lines only on gas-phase O/H. Assuming constant SFRs, the typical inferred ages of high-redshift starforming galaxies is a few hundred Myr; since the rate of Fe

enrichment of the ISM by SNe Ia depends on stars that formed some ~1 Gyr earlier than the epoch at which the massive stars are observed, the metals deposited in the ISM over the star-forming lifetime of the galaxy (and thus, the abundances of massive stars at the epoch of observation) would be dominated by the nucleosynthetic products of core-collapse supernovae. Moreover, it has been argued (e.g., Papovich et al. 2011; Reddy et al. 2012) that typical high-redshift galaxies have SFRs that *increase* with time at $z \gtrsim 2$—rising star formation histories could easily extend the timescale over which core-collapse supernovae dominate the gas-phase enrichment to well beyond a Gyr, since once again the production rate of Fe by SNe Ia depends on the galaxy's SFR as it was ~1 Gyr in the past.

According to Nomoto et al. (2006), the Salpeter IMF-averaged yields of O and Fe from core-collapse supernovae are [O/Fe] = (0.74, 0.67, 0.60) dex for initial metallicity $Z_* = (0.001, 0.004, 0.020)$. That is, the predicted enhancement in abundance of O relative to Fe is (O/Fe) = (5.5, 4.7, 4.0) × (O/Fe)$_\odot$—remarkably close to the estimated factor of $\simeq 4$–5 difference between Z_{neb} and Z_* inferred from the observations of KBSS-LM1.

Figure 17 shows three samples of Galactic stars with measurements of stellar O and Fe abundance, where we show the O/Fe abundance ratio relative to solar, [O/Fe], as a function of log(O/H), rather than the more commonly plotted [X/Fe] versus [Fe/H]. The sample represented by orange points is the same "thick-disk" sample of Bensby & Feltzing (2006) as in Figure 15; the squares are Galactic bulge stars: micro-lensed dwarfs from Bensby et al. (2013) (light green), and K-giants from Lecureur et al. (2007) (dark green). Also shown is our estimate for KBSS-LM1 assuming that the inferred Z_* tracks Fe/H and that Z_{neb} tracks O/H. The oxygen abundance values for all points in Figure 17 are given using the nebular convention $12 + \log(O/H)$ to emphasize the following point: essentially all of the Galactic stars in the thick disk and bulge that formed out of material with similar O/H as inferred for the high-redshift galaxies are also significantly enhanced in O/Fe. In Figure 15, the thick-disk stars with C/O similar to those of KBSS-LM1 are those with the highest values of O/Fe. Also noteworthy in Figure 17 is that Galactic thick-disk and bulge stars remain relatively Fe-poor even as O/H approaches (O/H)$_\odot$, suggesting that massive stars forming out of the same gas—which of course disappeared long ago—would have produced H II regions of significantly higher excitation than those in present-day disks with the same nebular oxygen abundance (see Section 9.4).

One important implication, practically speaking, is that in attempting to reproduce the stellar and nebular spectra of high-redshift galaxies using massive star population synthesis and photoionization models, one should *not* assume that the "stellar abundance" and the "gas-phase abundance" are the same! While it is likely that O/H in the gas phase is very close to O/H in the nearby massive stars, the most easily observed properties of the massive stars—the FUV stellar spectra and excitation of the nebular emission—will appear more like those with much lower overall metallicity. Until models of massive stars with non-solar initial abundance ratios become available in the future, successful models of high-redshift systems may require decoupling the nebular and stellar metallicity. To a first approximation, the stellar population synthesis model should match the expected abundance of Fe, while the gas-phase abundances should reflect the abundance of O.

9.2. Abundance Ratios

We have shown that the KBSS-LM1 composite spectra suggest modest under-abundances of N and C relative to O in the ionized gas—by $\simeq 0.38$ and $\simeq 0.34$ dex compare to solar, respectively (see Table 7). Both values are consistent with those of H II regions in the local universe with similar values of O/H (Figures 15 and 16). The C abundance inferred from nebular C/O (Section 8.2) and O/H (Section 8.1) is in broad agreement with that inferred from the strength of the C IV P-Cygni stellar wind feature, and together they imply C/H $\simeq 0.20 \pm 0.08$ (C/H)$_\odot$. This relatively low abundance of C, which is apparently typical of high-redshift galaxies with bolometric luminosities 10^{11}–$10^{12} L_\odot$ from which the KBSS-LM1 sample is drawn, may have significant implications for their detectability in atomic and molecular lines by sub-mm/mm observations. In particular, if they were to mimic local galaxies with similar gas-phase C/H, the appropriate value of the L(CO)-to-M_{mol} conversion factor α_{CO} could be 1–2 orders of magnitude higher (i.e., lower L(CO) for a given M$_{mol}$) than applies to more chemically evolved galaxies (see, e.g., Bolatto et al. 2013; Hunt et al. 2015).

The behavior of N/O among the high-redshift galaxy samples has been somewhat controversial in the recent literature. Anomalously high N/O at a given O/H has been advocated as the primary driver of the N2 BPT "offset" (Masters et al. 2014; Jones et al. 2015; Sanders et al. 2015; Shapley et al. 2015). Part of the argument in favor of this interpretation is the apparent absence of a similar offset in the S2 BPT diagram. However, we have shown that the location of the LM1 composite in both diagrams is as expected for normal N/O and N/S (given O/H): the same (normal) N/O is obtained by applying the best local calibration (based on direct T_e measurements) of N2O2 → (N/O) (Section 6.2) and by the photoionization model that reproduces the observed O32 and R23, which do not depend on N/O (Section 7.3). A more thorough treatment of the behavior of N/O versus O/H for the full KBSS-MOSFIRE sample is presented in a companion paper (Strom et al. 2016).

9.3. Population Synthesis Models

Population synthesis models without treatment of binary evolution cannot currently produce nebulae that resemble a large fraction of those observed at high redshift. Single-star models cannot produce steady-state ionizing radiation fields sufficiently hard to produce high [O III]/Hβ and R23 at fixed O32 (Section 7.3), nor can they produce the observed stellar He II $\lambda 1640$ emission, unless the galaxy is an extremely young starburst caught at a particular time post-burst when Wolf–Rayet stars briefly shine. This post-burst timing argument may be adequate to explain rare dwarf galaxies in the local universe (see, e.g., Shirazi & Brinchmann 2012), but the high-redshift examples have ~10–100 times larger stellar masses and thus much longer dynamical times that make the required burst timescale uncomfortably short. Perhaps most convincingly, the high-redshift galaxies we have discussed are *typical*, not rare, indicating long duty cycles for the periods of high nebular excitation. Moreover, we *know* that binary evolution of massive stars is the rule rather than the exception, thus models that do not include it cannot be correct, unless by accident.

Although the BPASSv2 models highlighted above work very well to explain the observed nebular emission and

reasonably well for the stellar FUV continuum, it does not necessarily follow that the details of the model—many of which remain uncertain—are correct. However, we can say that the ionizing spectra of the stars that dominate the nebular excitation in typical star-forming galaxies at $z \sim 2.4$ must have EUV spectra similar to those of the best-matching BPASSv2 models (see Figure 13).

9.4. What (if Anything) is Different at High Redshift?

Although analogs of galaxies that comprise the KBSS-LM1 composite can be found at all redshifts (e.g., Brown et al. 2014; Shirazi et al. 2014; Steidel et al. 2014; Jones et al. 2015; Bian et al. 2016), galaxies with very high specific SFRs and high excitation are common at high redshift but rare in the relatively local universe. Rarer still are nearby galaxies exhibiting these properties with M_* ($\sim 10^{10} M_\odot$) and SFR ($\sim 30 M_\odot \, \mathrm{yr}^{-1}$) comparable to the galaxies comprising the LM1 sample (Table 1). From the standpoint of nebular astrophysics, it is likely that the low-redshift analogs can be explained in a manner similar to those we have advocated for the high-redshift systems: hard EUV ionizing spectra and moderately sub-solar oxygen abundances. But the presence of large numbers of relatively high mass galaxies with high-excitation nebulae requires more explanation, since one cannot rely on the same duty cycle arguments that might account for a rare population of objects as at low redshift.

We have argued above that galaxies whose ISM metallicity is dominated by metals produced within the past $\sim$Gyr by core-collapse supernovae can attain near-solar O/H while still manifesting massive star population properties normally associated with much lower metallicities. The time period over which ISM enrichment remains dominated by SNe II yields can easily extend over longer periods for a given galaxy if its SFR continues to increase as a function of time as expected for galaxies during their most rapid period of growth. During this phase, newly formed OB stars will continue to appear "chemically young," since the production of O always outpaces that of Fe. The low inferred Fe/H in the active O-star population results in weak stellar winds, lower mass-loss rates, and less photospheric line blanketing, while massive stars in binaries are much more likely to undergo QHE for similar reasons (see Section 1)—all of which favor harder EUV ionizing spectra.

As mentioned in Section 9.1 above, Figure 17 suggests that these effects might well have been important at the time of formation of the Galactic bulge and thick disk—enhanced [O/Fe] is observed in all bulge and thick-disk stars with $(\mathrm{O/H}) \lesssim (\mathrm{O/H})_\odot$. Contemporaneous massive stars forming out of the same ISM material with $(\mathrm{O/H}) \lesssim (\mathrm{O/H})_\odot$ would have had $(\mathrm{Fe/H}) \lesssim 0.25 \, (\mathrm{Fe/H})_\odot$; in terms of the parametrization of stellar metallicity in the population synthesis models discussed above, this is equivalent to $Z_* \lesssim 0.0035$—low enough to allow for the most dramatic effects of massive binary evolution (including QHE). For this reason, even at solar gas-phase oxygen abundance, the H II regions excited by such massive stars would be observed to have much higher excitation than a typical H II region in a present-day Galactic disk, where $[\mathrm{O/Fe}] \simeq 0$. Similarly, galaxies with declining star formation histories and/or low sSFRs (i.e., where the current star formation is a small perturbation on the past integral) are the rule at low redshift; these would be much less likely to have super-solar O/Fe, and, at a given O/H, would have softer EUV

spectra due to increased photospheric line blanketing. Because of much higher mass-loss rates, the probability of QHE in massive binaries would be dramatically reduced (e.g., Brott et al. 2011; Eldridge et al. 2011).

It seems likely that a similar qualitative argument might explain the otherwise puzzling observation that, at $z \sim 2.3$, the correlation between strong-line indices sensitive to nebular excitation (e.g., O3, O3N2, R32) and M_* is *tighter* than could reasonably be expected even from a "stellar mass–metallicity" relation (MZR) with zero intrinsic scatter, given the uncertainties in mapping the line indices to oxygen abundance (e.g., Equation (10))—see the discussion in Section 8 of Steidel et al. (2014). Excitation is controlled primarily by the spectrum of the ionizing stars, and we have argued that the EUV is modulated by the abundance of Fe, and not O. The enrichment pattern of O/Fe in the ISM of star-forming galaxies depends very strongly on the star formation history. Galaxies with increasingly high M_* have lower sSFR and longer star formation histories where an increasing fraction of the stellar mass was accumulated in the distant past. O-stars forming in an increasingly Fe-rich ISM will be less capable of producing high-excitation nebular emission lines; thus, the oldest (and most massive) galaxies have the lowest-excitation nebulae almost independently of the nebular oxygen abundance. An obvious corollary is that strong-line indices measured in the local universe may be systematically different from those measured at high redshift for the same gas-phase O/H; conversely, the same excitation may be observed at systematically different O/H if the star formation histories (or, equivalently, the O/Fe abundance pattern) of any two samples are significantly different.

9.5. Suggestions for Future Work

In this paper we have shown that a consistent interpretation of the most easily observed physical properties of high-redshift galaxies can be obtained by simultaneously modeling the EUV–FUV spectra of the stars and the nebular emission lines from the same galaxies. Among the areas where future work would lead to measurable improvements in our understanding of stellar populations, chemistry, ISM physics, and other physical properties of high-redshift star-forming galaxies include:

1. Construction of stellar population synthesis models that include all of the following (a) binary evolution of massive stars (b) a range of non-solar stellar abundance ratios (particularly varying Fe/O, C/O, and N/O) and (c) detailed, moderate to high-resolution FUV spectra for direct comparison with observations.

2. Creation of a streamlined method for simultaneous "fitting" of FUV stellar spectra and FUV–optical nebular spectra of the same objects or samples. This would involve a more comprehensive grid of stellar and nebular abundances and abundance patterns than currently exists, and will eventually enable direct constraints on the physics and chemistry of the ISM, forming massive stars, and ionized nebulae at high redshift—independent of any assumptions imposed by our understanding at low redshift.

3. Assembly of very deep rest-frame FUV spectra for large samples of galaxies which also have excellent nebular measurements from the new generation of near-IR multi-

object spectrographs. The results of our pilot program have been a first step in that direction.

It is worth emphasizing that progress made in these areas will be immediately relevant to the planning and interpretation of future observations of distant galaxies by the *James Webb Space Telescope*, which will be observing the same FUV continuum and FUV–optical nebular spectra of galaxies beyond $z \sim 4$ and into the reionization era.

REFERENCES

Akerman, C. J., Carigi, L., Nissen, P. E., Pettini, M., & Asplund, M. 2004, A&A, 414, 931
Andrews, B. H., & Martini, P. 2013, ApJ, 765, 140
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5 (BPT)
Bensby, T., & Feltzing, S. 2006, MNRAS, 367, 1181
Bensby, T., Yee, J. C., Feltzing, S., et al. 2013, A&A, 549, A147
Berg, D. A., Skillman, E. D., Croxall, K. V., et al. 2015, ApJ, 806, 16
Bian, F., Kewley, L., Dopita, M., & Juneau, S. 2016, arXiv:1603.05275
Blanc, G. A., Kewley, L., Vogt, F. P. A., & Dopita, M. A. 2015, ApJ, 798, 99
Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, ARA&A, 51, 207
Brott, I., de Mink, S. E., Cantiello, M., et al. 2011, A&A, 530, A115
Brown, J. S., Croxall, K. V., & Pogge, R. W. 2014, ApJ, 792, 140
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Campbell, A., Terlevich, R., & Melnick, J. 1986, MNRAS, 223, 811
Cantiello, M., Yoon, S.-C., Langer, N., & Livio, M. 2007, A&A, 465, L29
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Chabrier, G. 2003, PASP, 115, 763
Crowther, P. A. 2007, ARA&A, 45, 177
Crowther, P. A., Caballero-Nieves, S. M., Bostroem, K. A., et al. 2016, MNRAS, 458, 624
Crowther, P. A., Prinja, R. K., Pettini, M., & Steidel, C. C. 2006, MNRAS, 368, 895
Eldridge, J. J., Langer, N., & Tout, C. A. 2011, MNRAS, 414, 3501
Eldridge, J. J., & Stanway, E. R. 2012, MNRAS, 419, 479
Eldridge, J. J., & Stanway, E. R. 2016, arXiv:1602.03790
Erb, D. K., Pettini, M., Shapley, A. E., et al. 2010, ApJ, 719, 1168
Erb, D. K., Steidel, C. C., Shapley, A. E., et al. 2006, ApJ, 646, 107
Esteban, C., García-Rojas, J., Carigi, L., et al. 2014, MNRAS, 443, 624
Faucher-Giguère, C.-A., Quataert, E., & Hopkins, P. F. 2013, MNRAS, 433, 1970
Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, RMxAA, 49, 137
Garnett, D. R. 1992, AJ, 103, 1330
Garnett, D. R., Dufour, R. J., Peimbert, M., et al. 1995a, ApJL, 449, L77
Garnett, D. R., Shields, G. A., Peimbert, M., et al. 1999, ApJ, 513, 168
Garnett, D. R., Skillman, E. D., Dufour, R. J., et al. 1995b, ApJ, 443, 64
González Delgado, R. M., Leitherer, C., Stasińska, G., & Heckman, T. M. 2002, ApJ, 580, 824
González Delgado, R. M., & Pérez, E. 2000, MNRAS, 317, 64
Gordon, K. D., Clayton, G. C., Misselt, K. A., Landolt, A. U., & Wolff, M. J. 2003, ApJ, 594, 279

Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, MNRAS, 445, 581
Hunt, L. K., García-Burillo, S., Casasola, V., et al. 2015, A&A, 583, A114
Izotov, Y. I., Stasińska, G., Meynet, G., Guseva, N. G., & Thuan, T. X. 2006, A&A, 448, 955
James, B. L., Pettini, M., Christensen, L., et al. 2014, MNRAS, 440, 1794
Jenkins, E. B. 2009, ApJ, 700, 1299
Jones, T., Martin, C., & Cooper, M. C. 2015, ApJ, 813, 126
Kashino, D., Silverman, J. D., Rodighiero, G., et al. 2013, ApJL, 777, L8
Kewley, L. J., & Dopita, M. A. 2002, ApJS, 142, 35
Kobulnicky, H. A., Kennicutt, R. C., Jr., & Pizagno, J. L. 1999, ApJ, 514, 544
Kobulnicky, H. A., & Skillman, E. D. 1998, ApJ, 497, 601
Kroupa, P. 2001, MNRAS, 322, 231
Kudritzki, R.-P., & Puls, J. 2000, ARA&A, 38, 613
Langer, N. 2012, ARA&A, 50, 107
Lecureur, A., Hill, V., Zoccali, M., et al. 2007, A&A, 465, 799
Leitherer, C., Ekström, S., Meynet, G., et al. 2014, ApJS, 212, 14
Leitherer, C., Leão, J. R. S., Heckman, T. M., et al. 2001, ApJ, 550, 724
Leitherer, C., Ortiz Otálvaro, P. A., Bresolin, F., et al. 2010, ApJS, 189, 309
Massey, P., Strobel, K., Barnes, J. V., & Anderson, E. 1988, ApJ, 328, 315
Masters, D., McCarthy, P., Siana, B., et al. 2014, ApJ, 785, 153
McGaugh, S. S. 1991, ApJ, 380, 140
Muratov, A. L., Kereš, D., Faucher-Giguère, C.-A., et al. 2015, MNRAS, 454, 2691
Murray, N., Quataert, E., & Thompson, T. A. 2005, ApJ, 618, 569
Nomoto, K., Tominaga, N., Umeda, H., Kobayashi, C., & Maeda, K. 2006, NuPhA, 777, 424
Oke, J. B., Cohen, J. G., Carr, M., et al. 1995, PASP, 107, 375
Osterbrock, D. E., & Ferland, G. J. 2006, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei (Mill Valley, CA: Univ. Science Books)
Papovich, C., Finkelstein, S. L., Ferguson, H. C., Lotz, J. M., & Giavalisco, M. 2011, MNRAS, 412, 1123
Peimbert, M., & Peimbert, A. 2002, in X Reunión Regional Latinoamericana de Astronomía, Vol. 14, ed. J. J. Claria, D. Garcia Lambas, & H. Levato (Cordoba, Argentina: IAR), 47
Pérez-Montero, E., & Contini, T. 2009, MNRAS, 398, 949
Pettini, M., & Pagel, B. E. J. 2004, MNRAS, 348, L59
Pettini, M., Rix, S. A., Steidel, C. C., et al. 2002, ApJ, 569, 742
Pettini, M., Steidel, C. C., Adelberger, K. L., Dickinson, M., & Giavalisco, M. 2000, ApJ, 528, 96
Pilyugin, L. S., Grebel, E. K., & Mattsson, L. 2012, MNRAS, 424, 2316
Pilyugin, L. S., Mattsson, L., Vílchez, J. M., & Cedrés, B. 2009, MNRAS, 398, 485
Price, S. H., Kriek, M., Brammer, G. B., et al. 2014, ApJ, 788, 86
Puls, J., Vink, J. S., & Najarro, F. 2008, A&AR, 16, 209
Quider, A. M., Pettini, M., Shapley, A. E., & Steidel, C. C. 2009, MNRAS, 398, 1263
Quider, A. M., Shapley, A. E., Pettini, M., Steidel, C. C., & Stark, D. P. 2010, MNRAS, 402, 1467
Reddy, N. A., Kriek, M., Shapley, A. E., et al. 2015, ApJ, 806, 259
Reddy, N. A., Pettini, M., Steidel, C. C., et al. 2012, ApJ, 754, 25
Reddy, N. A., Steidel, C. C., Pettini, M., et al. 2008, ApJS, 175, 48
Rix, S. A., Pettini, M., Leitherer, C., et al. 2004, ApJ, 615, 98
Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, ApJL, 802, L19
Rockosi, C., Stover, R., Kibrick, R., et al. 2010, Proc. SPIE, 7735, 0
Rudie, G. C., Steidel, C. C., Shapley, A. E., & Pettini, M. 2013, ApJ, 769, 146
Sana, H., de Mink, S. E., de Koter, A., et al. 2012, Sci, 337, 444
Sanders, R. L., Shapley, A. E., Kriek, M., et al. 2015, ApJ, 799, 138
Sanders, R. L., Shapley, A. E., Kriek, M., et al. 2016, ApJ, 816, 23
Shapley, A. E., Reddy, N. A., Kriek, M., et al. 2015, ApJ, 801, 88
Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588, 65
Shirazi, M., & Brinchmann, J. 2012, MNRAS, 421, 1043
Shirazi, M., Brinchmann, J., & Rahmati, A. 2014, ApJ, 787, 120
Shivaei, I., Kriek, M., Reddy, N. A., et al. 2016, ApJL, 820, L23
Smith, N. 2014, ARA&A, 52, 487
Stanway, E. R., Eldridge, J. J., & Becker, G. D. 2016, MNRAS, 456, 485
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
Yoon, S.-C., & Langer, N. 2005, A&A, 443, 643