A preview of JWST metallicity studies at cosmic noon: the first detection of auroral [O II] emission at high redshift

Sanders, R.L.; Shapley, A.E.; Clarke, L.; Topping, M.W.; Reddy, N.A.; Kriek, M.T.; ... ; Tang, M.

The Astrophysical Journal, 943:75 (14pp), 2023 February 1 https://doi.org/10.3847/1538-4357/aca9cc

A Preview of JWST Metallicity Studies at Cosmic Noon: The First Detection of Auroral [O II] Emission at High Redshift*

Ryan L. Sanders[1,7], Alice E. Shapley[2], Leonardo Clarke[2], Michael W. Topping[3], Naveen A. Reddy[4], Mariska Kriek[5], Tucker Jones[1], Daniel P. Stark[3], and Mengtao Tang[6]
[1] Department of Physics and Astronomy, University of California, Davis, One Shields Ave, Davis, CA 95616, USA; rlsand@ucdavis.edu
[2] Department of Physics & Astronomy, University of California, Los Angeles, 430 Portola Plaza, Los Angeles, CA 90095, USA
[3] Steward Observatory, University of Arizona, 933 N Cherry Avenue, Tucson, AZ 85721, USA
[4] Department of Physics & Astronomy, University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA
[5] Leiden Observatory, Leiden University, P.O. Box 9513, NL-2300 AA Leiden, The Netherlands
[6] Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
Received 2022 July 26; revised 2022 November 22; accepted 2022 December 6; published 2023 January 27

Abstract

We present ultradeep Keck/MOSFIRE rest-optical spectra of two star-forming galaxies at $z = 2.18$ in the COSMOS field with bright emission lines, representing more than 20 hr of total integration. The fidelity of these spectra enabled the detection of more than 20 unique emission lines for each galaxy, including the first detection of the auroral [O II]$\lambda\lambda$7322, 7332 lines at high redshift. We use these measurements to calculate the electron temperature in the low-ionization O^+ zone of the ionized interstellar medium and derive abundance ratios of O/H, N/H, and N/O using the direct method. The N/O and α/Fe abundance patterns of these galaxies are consistent with rapid formation timescales and ongoing strong starbursts, in accord with their high specific star formation rates. These results demonstrate the feasibility of using auroral [O II] measurements for accurate metallicity studies at high redshift in a higher-metallicity and lower-excitation regime previously unexplored with the direct method in distant galaxies. These results also highlight the difficulty in obtaining the measurements required for direct-method metallicities from the ground. We emphasize the advantages that the JWST/NIRSpec instrument will bring to high-redshift metallicity studies, where the combination of increased sensitivity and uninterrupted wavelength coverage will yield more than an order of magnitude increase in efficiency for multiplexed auroral-line surveys relative to current ground-based facilities. Consequently, the advent of JWST promises to be the beginning of a new era of precision chemical abundance studies of the early universe at a level of detail rivaling that of local galaxy studies.

Unified Astronomy Thesaurus concepts: High-redshift galaxies (734); Metallicity (1031); Chemical abundances (224); Abundance ratios (11); Galaxy spectroscopy (2171)

1. Introduction

The gas-phase oxygen abundance (O/H) in a galaxy provides an essential indication of its evolutionary state, specifically the aggregate effects of chemical enrichment from past star formation, and dilution from inward and outward gas flows. The strong correlations between gas-phase oxygen abundance and other galaxy properties such as stellar mass, star formation rate (SFR), and gas content has been comprehensively demonstrated in the local universe (e.g., Tremonti et al. 2004; Mannucci et al. 2010; Zahid et al. 2014). The form of these relationships has been shown to provide important constraints on the parameters of star formation feedback in galaxies, such as the dependence of mass outflow rate on galaxy mass and the timescales over which SFRs and metallicities vary relative to their equilibrium values at a given mass (e.g., Peeples & Shankar 2011; Andrews & Martini 2013; Davé et al. 2017; Torrey et al. 2018, 2019).

The challenge associated with gas-phase oxygen abundance measurements in galaxies lies with translating spectroscopic measurements of multiple nebular emission-line strengths into estimates of metallicity. One of the most robust methods for estimating oxygen abundances is the so-called "direct method," in which the electron temperature (T_e) is derived from the ratio between weak, upper-level auroral line(s) and stronger, intermediate-level line(s) (e.g., [O III]λ4364/[O III]$\lambda\lambda$4960, 5008), and the electron density (n_e) is inferred from the doublet ratio of features such as [S II]$\lambda\lambda$6718, 6733. These ionized-gas properties are then used to translate the strengths of strong oxygen lines relative to hydrogen Balmer lines into the abundance of oxygen ions relative to hydrogen nuclei (e.g., Peimbert et al. 2017). The direct method has been applied to hundreds of individual H II regions in the Milky Way and nearby galaxies (e.g., Bresolin et al. 2009; Berg et al. 2020), as well as the integrated spectra of nearby star-forming galaxies (Izotov et al. 2006). The direct method has also been used to estimate average oxygen abundances in stacked spectra of large sample of galaxies drawn from the Sloan Digital Sky Survey (SDSS; Andrews & Martini 2013; Curti et al. 2017). With such measurements of direct metallicities across the entire population of local star-forming galaxies, it is possible to calibrate O/H as a function of the ratio of strong emission lines

* The data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation.
[7] NHFP Hubble Fellow.

(e.g., [N II]λ6585/Hα, R23 = ([O III]$\lambda\lambda$4960, 5008+[O II]$\lambda\lambda$3727, 3730)/Hβ).

Since at high redshift (i.e., $z > 1$), these strong-line ratios are much easier to measure than those involving the faint auroral lines, a calibration between strong-line ratio and oxygen abundance is essential for metallicity studies of the early universe. While direct-method metallicity calibrations from the local universe are commonly applied to interpret nebular emission-line ratios in high-redshift galaxies, they may in fact yield biased results for the inferred oxygen abundances. Such biases may occur due to the evolving physical conditions in the ionized interstellar medium (ISM) of high-redshift galaxies (e.g., Steidel et al. 2014; Shapley et al. 2015, 2019; Sanders et al. 2020a). It would therefore be ideal to construct direct-method calibrations of strong emission-line ratios for star-forming galaxies based on direct-method oxygen abundances that are also measured at high redshift. As reviewed by Sanders et al. (2020a), a small sample of ~20 direct metallicity measurements exists at $z = 1.7$–3.6. However, all of these measurements are based on auroral [O III] features, either [O III] λ4364 in the rest-optical or O III]$\lambda\lambda$1661, 1666 in the rest-ultraviolet.

Given that rest-optical and rest-ultraviolet auroral [O III] emission is easiest to detect in relatively metal-poor, high-excitation galaxies, the $z > 1$ sample for which such measurements have been performed is not representative of the full star-forming galaxy population over a wide range in stellar mass, SFR, and metallicity. Furthermore, while it is possible to trace the relationship between strong-line ratios and metal abundance over a wide range of metallicity in the local universe (e.g., Curti et al. 2017; Bian et al. 2018), the small dynamic range of galaxy properties probed with auroral [O III] lines at high redshift means that similarly representative correlations for distant galaxies do not yet exist.

In order to extend strong-line metallicity calibrations at $z \sim 2$–3 toward higher (i.e., solar) metallicity, we now require the detection of low-ionization auroral emission lines. Such emission can be detected in the low-excitation portions of cooler, more metal-rich star-forming regions. The brightest of these features is [O II]$\lambda\lambda$7322, 7332. Both Andrews & Martini (2013) and Curti et al. (2017) use the [O II]$\lambda\lambda$7322, 7332 feature to constrain the electron temperature in $z \sim 0$ SDSS galaxies extending up to solar metallicity. However no such measurements of [O II]$\lambda\lambda$7322, 7332 exist in the literature above $z = 1$.

In this paper, we present the *first* [O II]$\lambda\lambda$7322, 7332 auroral-line measurements outside the low-redshift universe. We have detected these features for two star-forming galaxies at $z = 2.18$, located in the COSMOS field. These galaxies were first observed with the MOSFIRE spectrograph (McLean et al. 2012) on the Keck I telescope as part of the MOSFIRE Deep Evolution Field (MOSDEF) survey (Kriek et al. 2015). These targets were selected for deep follow-up Keck/MOSFIRE spectroscopy on the basis of their strong rest-frame optical emission lines that suggested they would have relatively strong [O II]$\lambda\lambda$7322, 7332, which we indeed confirm in this paper. Measurements such as these will be essential as we attempt to construct direct-method abundance calibrations for high-redshift rest-frame optical strong-line measurements collected with JWST.

In Section 2, we describe our new MOSFIRE observations and emission-line measurements. In Section 3, we present the physical conditions and chemical abundances of the two galaxies studied here. In Section 4, we discuss the implications of our results for chemical abundance calibrations at high redshift; the prospects for conducting direct-metallicity studies with the new capabilities of JWST; the constraints obtained for the chemical abundance patterns among oxygen, nitrogen, and iron; the (lack of) significant active galactic nucleus (AGN) activity in our target galaxies; and, finally, the detection of broad rest-frame optical emission, signaling the presence of galaxy-scale outflows. In Section 5, we summarize our key results and conclusions. Throughout, we adopt cosmological parameters of $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_m = 0.30$, and $\Omega_\Lambda = 0.7$, and a Chabrier (2003) initial mass function (IMF). Rest-frame wavelengths of emission lines are given in the vacuum. We adopt solar abundance values of $12+\log(\mathrm{O/H})_\odot = 8.69$, $12+\log(\mathrm{N/H})_\odot = 7.83$, $\log(\mathrm{N/O})_\odot = -0.86$, and $\log(\mathrm{Fe/H})_\odot = 7.46$ as number density fractions, and the bulk metallicity by mass fraction $Z_\odot = 0.0139$ (Asplund et al. 2021).

2. Observations and Measurements

2.1. Targets, Observations, and Data Reduction

Detecting faint temperature-sensitive auroral emission lines at high redshifts requires deep near-infrared spectroscopy of targets displaying bright line emission. Accordingly, we obtained ultradeep observations of targets in the COSMOS CANDELS (Grogin et al. 2011; Koekemoer et al. 2011) field using the MOSFIRE instrument (McLean et al. 2012) on the 10 m Keck I telescope. The primary targets were selected from the MOSDEF survey (Kriek et al. 2015), which obtained rest-optical spectroscopy of ~1500 galaxies at $z = 1.4$–3.8 using MOSFIRE, with a typical depth of 2 hr per filter. Fluxes and ratios of strong emission lines measured in the MOSDEF spectra were used to predict the flux of auroral [O III]λ4364 and [O II]$\lambda\lambda$7322, 7332. A slitmask was then designed that maximized the number of targets with predicted auroral-line fluxes that could be detected in feasible integration times with MOSFIRE. The final mask in COSMOS included five galaxies at $z = 2.0$–3.5 targeted for [O III]λ4364 that will be presented in a future work (L. Clarke et al. 2022, in preparation) and two galaxies at $z = 2.18$ targeted for [O II]$\lambda\lambda$7322, 7332 that are the subject of the current analysis. The relatively high excitation level (typically associated with larger T_e) and strong reddening-corrected [O II]$\lambda\lambda$3727, 3730 flux displayed in the MOSDEF spectra of the latter targets suggested bright auroral [O II] $\lambda\lambda$7322, 7332 lines that fall in the *K*-band filter at this redshift. These two galaxies are identified by their ID numbers 19985 and 20062 in the v4.1 photometric catalogs of the 3D-HST survey (Skelton et al. 2014; Momcheva et al. 2016). Their coordinates are given in Table 1. These targets are among the brightest emission-line galaxies in the MOSDEF survey data set based on their observed Hα and [O III]λ5008 fluxes.

The MOSFIRE mask was observed for a total of 15.4 hr on five nights spanning 2019 January 13 to 2021 March 4, with integrations of 1.9 hr in the *J* band, 8.0 hr in the *H* band, and 5.5 hr in the *K* band. The median seeing, measured from the spatial profile of a star on the mask, was $0\farcs79 \pm 0\farcs11$, 0.51 ± 0.11, and $0\farcs44 \pm 0\farcs16$ in *J*, *H*, and *K*, respectively, where the quoted error is the standard deviation of the seeing measured in the individual exposures. We adopted the standard observing strategy used in the MOSDEF survey, specifically an

Table 1
Target Properties

ID	19985	20062
R.A. (J2000)	10:00:14.484	+02:22:57.98
decl. (J2000)	10:00:16.436	+02:23:00.79
z	2.18796	2.18541
log $(M_*/M_\odot)$	10.12 ± 0.04	10.10 ± 0.07
log (t_{age}/yr)[a]	7.80 ± 0.13	8.00 ± 0.20
log (τ/yr)[a]	8.2 ± 1.0	8.0 ± 1.0
$A_{V,\mathrm{stars}}$	1.4 ± 0.07	1.2 ± 0.08
SFR $(M_\odot\ \mathrm{yr}^{-1})$	206 ± 8	259 ± 8
sSFR (Gyr^{-1})	15.6 ± 2.1	20.6 ± 3.9
R_{eff} (kpc)[b]	1.34	1.46
Σ_{SFR}[c] $(M_\odot\ \mathrm{yr}^{-1}\ \mathrm{kpc}^{-2})$	18.3 ± 0.7	19.3 ± 0.6

Notes.
[a] Stellar population age and timescale in a delayed-τ star formation history of the form SFR $\propto t_{age} e^{-t_{age}/\tau}$.
[b] Half-light elliptical semimajor axis from van der Wel et al. (2014).
[c] $\Sigma_{\mathrm{SFR}} = \mathrm{SFR}/2\pi R_{\mathrm{eff}}^2$.

ABA′B′ dither pattern with an inner/outer nod of 1.″2/1.″5, 2 minute exposures in the J and H bands, and 3 minute exposures in the K band. Slit widths were 0.″7, yielding spectral resolutions of $R \sim 3300$, 3650, and 3600 in J, H, and K, respectively. The data were reduced using a custom IDL pipeline designed for the MOSDEF survey and described in Kriek et al. (2015), the product of which is fully calibrated two-dimensional (2D) science spectra for each slit on the mask. The program `bmep`[8] (Freeman et al. 2019) was used to obtain one-dimensional (1D) science and error spectra using an optimal extraction (Horne 1986).

In the MOSDEF reduction pipeline, observations of a standard star are used for response and telluric absorption correction, while the absolute flux calibration is achieved via a star placed on one of the MOSFIRE slits on the multiobject science mask. In each filter, the measured continuum spectrum for the slit star is scaled to match the cataloged photometry of the star, and the scaling factor is then applied to all targets on a mask. The slit star on our target mask unfortunately dithered directly on top of a galaxy of comparable brightness separated by 2.″7. As a result, the stellar spectrum was oversubtracted and the flux calibration of the mask was biased. However, the targets of this analysis were also observed on a MOSDEF survey mask that has an isolated slit star without this issue, such that the MOSDEF flux calibration is reliable. Both sets of observations were corrected for slit losses as described in Reddy et al. (2015) and Kriek et al. (2015). To achieve a robust flux calibration of the new observations, for each target we measured the fluxes of the brightest emission line in each filter ([O II]$\lambda\lambda$3727, 3730 in J, [O III]λ5008 in H, and Hα in K) in both sets of spectra, noting that each line is detected at $>25\sigma$. In each filter, we then scaled the new spectrum by a multiplicative factor such that the line flux matched that measured from the MOSDEF spectrum.

To maximize the signal-to-noise ratio (S/N), we combined the MOSDEF spectra with our new observations. The final 1D science spectra were constructed by taking the inverse-variance weighted mean of the flux density at each wavelength pixel. We note that the MOSDEF K-band spectra only extend to 2.315 μm (≈7270 Å rest frame), such that the reddest part of the K-band spectrum covering [O II]$\lambda\lambda$7322, 7332 includes only the new spectroscopy. All other measured spectral features are covered by both data sets. The final spectra thus have effective total integration times of 3.8 hr, 10.0 hr, and 7.5 hr in J, H, and K, respectively, reaching typical 3σ limiting line fluxes of 7×10^{-18}, 3×10^{-18}, and 4×10^{-18} ergs^{-1} cm^{-2} in spectral regions free of strong sky lines. These spectra, in which a number of strong and weak emission lines are visible, are presented in Figure 1.

2.2. Spectral Energy Distribution Fitting

Galaxy properties were inferred from fitting stellar population models to photometry measured in 43 filters spanning rest-frame 1200 Å to 2.5 μm, as cataloged by the 3D-HST survey team (Skelton et al. 2014; Momcheva et al. 2016). We used the spectral energy distribution (SED)-fitting code FAST (Kriek et al. 2009) in combination with the flexible stellar population synthesis (FSPS) models of Conroy & Gunn (2010). A delayed-τ star formation history, a Calzetti et al. (2000) dust attenuation curve, and a Chabrier (2003) IMF were assumed. The delayed-τ parameterization (SFR $\propto t_{age} e^{-t_{age}/\tau}$) was chosen because of its ability to capture a range of recent formation histories based on the ratio between the stellar population age (t_{age}) and e-folding timescale (τ), from rising ($t_{age} \leqslant \tau$) to falling ($t_{age} > \tau$). The FSPS models are only available in a grid of discrete metallicity values. We adopted the stellar metallicity values ($Z_* = 0.2$ and $0.5\ Z_\odot$ for 19985 and 20062, respectively) closest to the direct-method gas-phase metallicities derived below ($Z_{\mathrm{gas}} = 0.16$ and $0.35\ Z_\odot$). Before fitting, the observed photometric measurements were corrected for the contribution from nebular line and continuum emission as described in Appendix, such that the FSPS models are fit to photometry capturing only the stellar continuum emission.

This process yielded estimates of stellar mass, stellar continuum reddening, stellar population age, and a best-fit model of the stellar continuum. In the H and K bands, where the continuum is detected significantly (S/N per pixel ≈5), the best-fit SED model agrees with the observed MOSFIRE spectrum to better than 15%, which validates our absolute and band-to-band flux calibration. The best-fit stellar population parameters are given in Table 1.

2.3. Emission-line Measurements

Emission-line fluxes were measured by fitting Gaussian line profiles to the 1D science spectra. The continuum under the lines was taken to be the best-fit stellar continuum model from SED fitting. Similar results are obtained if we instead fit a constant or linear local continuum under each line. Using the best-fit SED model has the advantage of self-consistently accounting for absorption under hydrogen Balmer lines. All emission features were fit with single Gaussian profiles except for [O II]$\lambda\lambda$3727, 3730, [S II]$\lambda\lambda$6718, 6733, and [O II]$\lambda\lambda$7322, 7332[9] that were simultaneously fit with a double Gaussian, and Hα and [N II]$\lambda\lambda$6550, 6585 that were simultaneously fit with a triple Gaussian. During fitting, [O III]λ5008/[O III]λ4960 was fixed to the intrinsic ratio of 2.98 (Storey et al. 2014), while

[8] https://github.com/billfreeman44/bmep

[9] [O II]$\lambda\lambda$7322, 7332 is actually a quadruplet with vacuum centroids 7321.0, 7322.0, 7331.7, and 7332.8 Å. The separation of lines in the blue and red pair is ∼3× smaller than what MOSFIRE can resolve, such that the multiplet is effectively a doublet in the observed spectra.

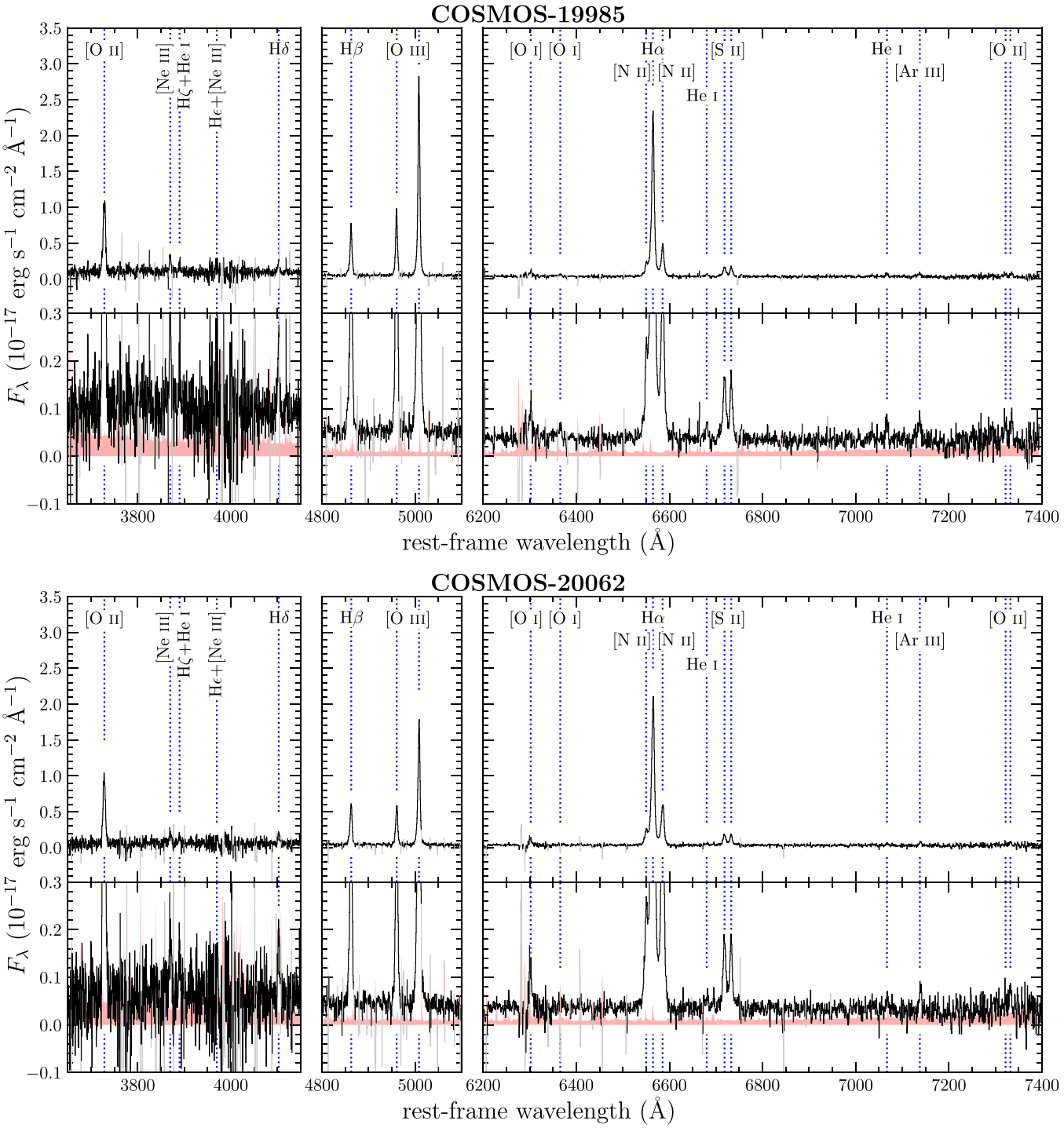

Figure 1. Near-infrared spectra of COSMOS 19985 (top) and 20062 (bottom), showing spectral features as a function of rest-frame wavelength observed in the *J* (left), *H* (middle), and *K* (right) bands. Regions affected by strong sky lines are grayed out. The shaded light red region shows the 1σ error spectrum. For each target, the upper panels present the full dynamic range of the spectrum, while the lower panels show the same spectrum zoomed in to highlight weak emission lines. Detected emission lines are labeled according to the ionic species. Blue dashed lines indicate the vacuum wavelengths of each transition.

[N II]λ6585/[N II]λ6550 was likewise fixed to 2.94 (Tayal 2011). The systemic redshifts were taken from the average redshift of Hα and [O III]λ5008, and are consistent with the previously published MOSDEF redshifts. The centroids and velocity widths of weak lines were set by the redshift and velocity width measured for Hα and [O III]λ5008. Uncertainties on line fluxes were estimated by perturbing the spectrum according to the error spectrum and remeasuring the line fluxes 500 times, where the 1σ uncertainties were taken to be half of the 16th-to-84th percentile width of the resulting flux distributions. Emission-line fluxes and uncertainties are presented in Table 2. More than 20 unique emission lines are detected at $\geqslant 3\sigma$ significance for each target.

A single line flux is reported for blended features. While the [O II]$\lambda\lambda$3727, 3730 doublet centroids are resolved at the MOSFIRE spectral resolution ($R > 3000$), the relatively broad line widths (FWHM ~ 300 km s^{-1}) prevent cleanly separating the doublet components. As such, we report only the total [O II]

Table 2
Observed Line Fluxes and Reddening-corrected Line Ratios

$F_{obs}(\lambda)$ (10^{-17} erg s^{-1} cm^{-2})		
Line	19985	20062
[O II] λ3727, 3730[a]	21.17 ± 0.49	19.58 ± 0.99
[Ne III] λ3870	3.48 ± 0.22	2.38 ± 0.20
Hζ+He I λ3890[a]	2.17 ± 0.27	1.64 ± 0.21
Hζ λ3890[b]	1.16 ± 0.32	0.81 ± 0.27
Hϵ+[Ne III] λ3969[a]	2.27 ± 0.34	1.67 ± 0.33
Hϵ λ3971[c]	1.22 ± 0.35	0.95 ± 0.32
[S II] λ4078	<0.47	<0.67
Hδ λ4103	2.92 ± 0.25	2.39 ± 0.17
He II λ4686	<0.27	<0.52
Hβ λ4863	11.80 ± 0.19	10.10 ± 0.12
[O III] λ4960	15.20 ± 0.04	9.78 ± 0.03
[O III] λ5008	45.20 ± 0.11	29.10 ± 0.09
[O I] λ6302	1.54 ± 0.15	1.68 ± 0.23
[S III] λ6314	<0.71	<0.77
[O I] λ6366	0.49 ± 0.10	<0.50
[N II] λ6550	3.96 ± 0.02	5.34 ± 0.04
Hα λ6565	49.20 ± 0.15	47.80 ± 0.14
[N II] λ6585	11.70 ± 0.07	15.70 ± 0.10
He I λ6680	0.62 ± 0.06	0.77 ± 0.11
[S II] λ6718	3.40 ± 0.12	3.92 ± 0.15
[S II] λ6733	2.77 ± 0.09	3.32 ± 0.13
He I λ7067	0.88 ± 0.12	0.55 ± 0.17
[Ar III] λ7138	0.91 ± 0.17	0.92 ± 0.18
[O II] λ7322	0.77 ± 0.18	<0.70
[O II] λ7332	0.80 ± 0.20	0.98 ± 0.22
Reddening-corrected Line Ratios		
$E(B-V)_{gas}$	0.39 ± 0.02	0.51 ± 0.01
log([O III]λ5008/Hβ)	0.56 ± 0.01	0.43 ± 0.01
log([O II]λ3728/Hβ)	0.44 ± 0.02	0.52 ± 0.02
log(R23)[d]	0.88 ± 0.01	0.84 ± 0.01
log([O III]λ5008/[O II]λ3728)	0.12 ± 0.01	-0.09 ± 0.02
log([Ne III]λ3870/[O II]λ3728)	-0.80 ± 0.03	-0.94 ± 0.04
log([N II]λ6585/Hα)	-0.62 ± 0.01	-0.49 ± 0.01
log(O3N2)[e]	1.19 ± 0.01	0.92 ± 0.01
log([S II]$\lambda\lambda$ 6718, 6733/Hα)	-0.91 ± 0.01	-0.83 ± 0.01
log([O I]λ6302/Hα)	-1.49 ± 0.04	-1.43 ± 0.06
$\frac{[O\,II]\lambda7322,7332}{[O\,II]\lambda3727,3730}$	0.029 ± 0.005	0.021 ± 0.005

Notes. For nondetected lines, 3σ upper limits are given.
[a] Blended lines.
[b] Derived by removing the blended He I λ3890 flux estimated using He I λ6680 and He I λ7067.
[c] Derived by removing the blended [Ne III]λ3969 flux estimated using [Ne III] λ3870.
[d] R23 = ([O III]$\lambda\lambda$ 4960, 5008+[O II]$\lambda\lambda$ 3727, 3730)/Hβ.
[e] O3N2 = ([O III]λ5008/Hβ)/([N II]λ6585/Hα).

$\lambda\lambda$3727, 3730 flux and refer to this sum as [O II]λ3728. Hϵ is blended with [Ne III]λ3969. We derived the Hϵ flux by subtracting the blended [Ne III]λ3969 flux, where the latter was inferred from [Ne III]λ3870, leveraging the fixed ratio of the two lines of [Ne III]λ3870/λ3969 = 3.32 calculated with `pyneb` (Luridiana et al. 2013, 2015). We estimated the He I λ3890 flux from the detected He I $\lambda\lambda$6680, 7067 lines by combining the reddening-corrected fluxes with the emissivity ratios of these transitions relative to He I λ3890. Emissivities were calculated using `pyneb`, assuming the electron temperature and density derived below. The resulting He I λ3890 estimate was used to infer the deblended Hζ flux.

2.4. Reddening Correction, Line Ratios, and Star Formation Rate

Nebular reddening, $E(B-V)_{gas}$, was derived assuming a Cardelli et al. (1989) Milky Way extinction curve with $R_V = 3.1$, using an inverse-variance weighted combination of measured Balmer line ratios (Hα/Hβ, Hδ/Hβ, and Hϵ/Hβ). The intrinsic ratios were calculated with `pyneb`, assuming the electron temperature and density derived below. Line fluxes were corrected for reddening using $E(B-V)_{gas}$ and the Cardelli et al. (1989) extinction curve. The reddening-corrected line ratios are presented in Table 2. SFR was calculated using the dust-corrected Hα luminosity assuming the conversion factor of Hao et al. (2011) adjusted to a Chabrier (2003) IMF.

2.5. Atomic Data

When calculating temperatures, densities, and chemical abundances below with `pyneb`, we adopted the atomic data recommended by Berg et al. (2015) as follows. For [O II] and [O III], we used the collision strengths from Kisielius et al. (2009) and Storey et al. (2014), respectively. The collision strengths of Tayal & Zatsarinny (2010) were used for [S II], while those of Tayal (2011) were used for [N II]. The radiative transition probabilities were taken from Froese Fischer & Tachiev (2004) for all ions. When using other atomic data sets available in `pyneb`, we found that the derived temperatures, densities, and abundances changed by $\leqslant$0.1 dex relative to our fiducial set, smaller than the derived uncertainties on these properties. Systematic uncertainties associated with atomic data are thus not a major contribution to the error budget in this study.

3. Physical Conditions and Chemical Abundances

3.1. Auroral [O II]$\lambda\lambda$7322, 7332 Lines

Auroral [O II] emission lines are detected in the spectra of both targets, the first time these temperature-sensitive lines have been detected beyond the low-redshift universe. Figure 2 shows the auroral [O II] lines in the 1D and 2D science spectra. Both components are detected in the spectrum of 19985, with significances of 4.3σ and 4.0σ for [O II]λ7322 and [O II]λ7332, respectively, and a combined significance of 5.8σ for the doublet. For 20062, [O II]λ7322 is formally undetected (1.7σ), while [O II]λ7332 is detected at the 4.4σ level.

The electron temperature in the O^+ zone, $T_e(O^+)$, can be calculated from the auroral-to-strong-line ratio [O II]$\lambda\lambda$7322, 7332/[O II]$\lambda\lambda$3727, 3730. Since only [O II]λ7332 is detected in the spectrum of 20062, we need an estimate of the total [O II] $\lambda\lambda$7322, 7332 flux to perform this calculation. The two auroral [O II] lines originate from the same upper level and therefore have a fixed flux ratio of [O II]λ7322/[O II]λ7332 = 1.17 determined by the ratio of radiative transition probabilities. While the [O II]λ7322 flux can be estimated by multiplying the detected [O II]λ7332 flux of 20062 by 1.17, this method yields a [O II]λ7322 flux that is 50% larger than the 3σ upper limit reported in Table 2, suggesting that the [O II]λ7332 flux of 20062 may be overestimated. To infer the total [O II]$\lambda\lambda$7322, 7332 flux including information from the [O II]λ7322 upper limit, we instead fit a model in which the two doublet components are fixed to the theoretical ratio and perform a χ^2 minimization to the flux and uncertainty of each line inferred from Gaussian fitting. Since the doublet ratio is fixed, the only

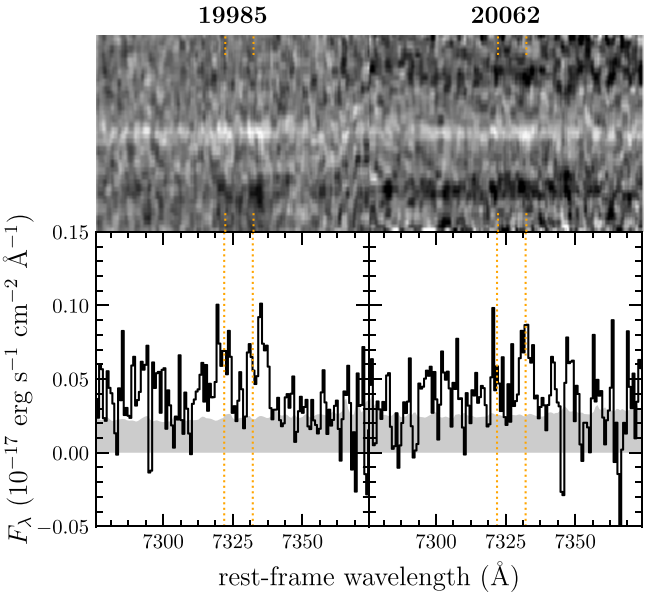

Figure 2. Spectral region around the auroral [O II]$\lambda\lambda$7322, 7332 emission lines for 19985 (left) and 20062 (right). The bottom panels show the 1D science spectra, with the error spectrum displayed in gray. The vacuum centroid wavelengths of the transitions are marked in orange. Both lines are detected at $>3\sigma$ for 19985, while only [O II]λ7332 is formally detected for 20062. The 2D spectrum (top) confirms these detections.

free parameter in this model is the total auroral [O II] flux. 20062 has observed Gaussian-fit fluxes of $4.1(\pm2.4)\times10^{-18}$ and $9.8\ (\pm2.2)\times10^{-18}$ erg s^{-1} cm^{-2} for [O II]λ7322 and [O II]λ7332, respectively. Fitting with the fixed flux ratio model yields a best-fit total [O II]$\lambda\lambda$7322, 7332 flux of $1.39\ (\pm\ 0.31)\times10^{-17}$ erg s^{-1} cm^{-2}. The best-fit flux of the individual doublet components from the fixed flux ratio model are within 1.5σ of the Gaussian-fit line fluxes.

For consistency, we use this same process to infer the total [O II]$\lambda\lambda$7322, 7332 flux of 19985, noting that 19985 has a directly measured doublet ratio of [O II]λ7322/[O II] $\lambda7332 = 0.97\pm0.36$, fully consistent with the theoretical ratio of 1.17. This target has observed Gaussian-fit fluxes of 7.7 $(\pm1.8)\times10^{-18}$ and 8.0 $(\pm2.0)\times10^{-18}$ erg s^{-1} cm^{-2} for [O II] λ7322 and [O II]λ7332, respectively, and a summed doublet flux of 1.57 $(\pm0.27)\times10^{-17}$ erg s^{-1} cm^{-2}. The best-fit total flux from the fixed flux ratio model is 1.58 $(\pm0.28)\times10^{-17}$ erg s^{-1} cm^{-2}. These best-fit total flux values were used in the calculation of the [O II]$\lambda\lambda$7322, 7332/[O II]$\lambda\lambda$3727, 3730 ratios given in Table 2.

3.2. Temperatures and Densities

We use the python package `pyneb` to derive electron temperatures (T_e) and densities (n_e), presented in Table 3. Electron density is calculated from the [S II] doublet using the ratio [S II]λ6718/[S II]λ6733.[10] We find values of $n_e\approx200$ cm^{-3} for both targets, consistent with the typical density of 200–300 cm^{-3} found for galaxies at $z\sim2-3$ (e.g., Steidel et al. 2014; Shimakawa et al. 2015; Sanders et al. 2016; Kaasinen et al. 2017; Strom et al. 2017) and elevated relative to $z\sim0$ main-sequence galaxies that typically have $n_e<$ 100 cm^{-3} (Kashino & Inoue 2019).

[10] The [O II]$\lambda\lambda$3727, 3730 doublet was not robusty deblended due to the relatively broad line widths in these targets, and thus does not provide reliable n_e constraints.

Table 3
Electron Temperatures and Densities, and Ionic and Total Elemental Abundances Calculated Using the Direct Method

ID	19985	20062
$T_e(O^+)$ (K)	13110 ± 1800	10530 ± 1420
$T_e(O^{2+})$[a] (K)	14440 ± 2680	10750 ± 2330
$n_e(S^+)$ (K)	190 ± 70	230 ± 70
$12+\log(O^+/H^+)$	7.55 ± 0.20	8.00 ± 0.24
$12+\log(O^{2+}/H^+)$	7.63 ± 0.22	7.87 ± 0.31
$12+\log(O/H)$	7.89 ± 0.20	8.24 ± 0.27
$12+\log(N^+/H^+)$	6.86 ± 0.15	7.22 ± 0.17
$12+\log(N/H)$	7.21 ± 0.15	7.46 ± 0.20
$12+\log(N/O)$	-0.69 ± 0.10	-0.78 ± 0.14
$12+\log(S^+/H^+)$	5.68 ± 0.14	5.98 ± 0.16

Note.
[a] Derived from $T_e(O^+)$ assuming the relation of Campbell et al. (1986).

The temperature in the low-ionization O^+ zone, $T_e(O^+)$, was calculated using the [O II]$\lambda\lambda$7322, 7332/[O II]$\lambda\lambda$3727, 3730 ratios given in Table 2. We find $T_e(O^+) = 13{,}110\pm1800$ K and $10{,}530\pm1420$ K for 19985 and 20062, respectively. That 19985 has a higher $T_e(O^+)$ than 20062 is in accord with the excitation-sensitive line ratios of these galaxies. 19985 has higher [O III]/Hβ, [O III]/[O II], and [Ne III]/[O II] than 20062, indicating a higher level of excitation and ionization that are usually associated with lower metallicity and higher temperature. These values of $T_e(O^+)$, near 10^4 K, are typical of moderately subsolar-metallicity ($\sim0.2-0.5\ Z_\odot$) H II regions in the local universe (Berg et al. 2015, 2020; Croxall et al. 2015, 2016; Rogers et al. 2021).

For other species in the low-ionization zone, namely N^+ and S^+, we also use the O^+ temperature:

$$T_e(N^+) = T_e(S^+) = T_e(O^+). \quad (1)$$

Following Rogers et al. (2021), we add the intrinsic scatter observed in the relations between these temperatures in local H II region samples in quadrature with the uncertainty propagated from $T_e(O^+)$, adopting an intrinsic scatter of 1000 K (Berg et al. 2020).

We cannot directly calculate the temperature in the high-ionization O^{2+} zone, $T_e(O^{2+})$, because [O III]λ4364 at $z=2.18$ falls at a wavelength of low atmospheric transmission and beyond the reach of ground-based observatories. Instead, we adopt the $T_e(O^+)-T_e(O^{2+})$ relation of Campbell et al. (1986):

$$T_e(O^+) = 0.7\times T_e(O^{2+}) + 3000\ \text{K}. \quad (2)$$

It has recently been shown that there is considerable scatter in the $T_e(O^+)-T_e(O^{2+})$ relation, with an intrinsic scatter of $\approx$1300 K in $T_e(O^{2+})$ at fixed $T_e(O^+)$ (Berg et al. 2020; Rogers et al. 2021). Following Rogers et al. (2021), we add an uncertainty of 1300 K in quadrature with the error propagated from $T_e(O^+)$ when calculating $T_e(O^{2+})$ using Equation (2). Partially due to the large intrinsic scatter, there is also considerable uncertainty about the shape of the $T_e(O^+)-T_e(O^{2+})$ relation. If we instead adopt a 1:1 relation, the derived O/H changes by only $\approx$0.1 dex.

3.3. Ionic and Total Abundances

Ionic abundances are calculated using `pyneb`, assuming the temperatures appropriate to each ion as described above. We assume that the density is constant throughout the nebula, adopting $n_e(S^+)$ for all ions. The derived ionic and total abundances are presented in Table 3.

For oxygen, we calculate O^+/H^+ using dust-corrected [O II] $\lambda3728/H\beta$ and $T_e(O^+)$, and O^{2+}/H^+ using [O III]$\lambda5008/H\beta$ and $T_e(O^{2+})$ estimated with Equation (2). We take the total oxygen abundance to be the sum of these two phases:

$$\frac{O}{H} = \frac{O^+}{H^+} + \frac{O^{2+}}{H^+}. \quad (3)$$

With an ionization energy of 54.9 eV, O^{3+} is found to be only a small ($\lesssim$5%) contribution even in extremely high-ionization sources (Berg et al. 2018, 2021) and can safely be ignored. We find the direct-method oxygen abundances of 19985 and 20062 to be $12+\log(O/H) = 7.89 \pm 0.20$ ($0.16^{+0.09}_{-0.06}$ $Z_\odot$) and $12+\log(O/H) = 8.24 \pm 0.27$ ($0.35^{+0.31}_{-0.16}$ $Z_\odot$), respectively.

Since we have directly constrained $T_e(O^+)$, we can calculate direct-method abundances for the low-ionization species N^+ and S^+. N^+/H^+ is derived using [N II]$\lambda6585/H\alpha$ and $T_e(O^+)$. To calculate the total N/H, an ionization correction is required to account for N^{2+} for which no associated lines are observed. We adopt the commonly used correction factor that leverages the similar ionization potential energies of O and N (Peimbert 1967):

$$\mathrm{ICF(N)} = \frac{N}{N^+} = \frac{O}{O^+}. \quad (4)$$

Under the same assumption, we calculate the abundance ratio N/O as

$$\frac{N}{O} = \frac{N^+}{O^+}. \quad (5)$$

Modern photoionization models have confirmed the validity of this ionization correction factor (Amayo et al. 2021). We find that 19985 and 20062 have $\log(N/O) = -0.69 \pm 0.10$ and -0.78 ± 0.14, respectively, close to the solar value of $\log(N/O)_\odot = -0.86$ (Asplund et al. 2021). S^+/H^+ is calculated using $T_e(O^+)$ and [S II]$\lambda\lambda$ 6718, 6733/Hα. Since 19985 and 20062 have comparable amounts of O in O^+ (13.6 eV) and O^{2+} (35.1 eV), there is likely a significant amount of S in S^{2+} (23.3 eV) and S^{3+} (34.8 eV) in addition to S^+. Reliable ionization correction factor prescriptions for S are available when S^+ and S^{2+} are observed (e.g., Thuan et al. 1995; Dors et al. 2016). However, S^{2+} requires observation of one of the [S III]$\lambda\lambda$ 9071, 9533 lines which fall at $\approx$3 μm at $z = 2.18$ and thus are only accessible from space. Accordingly, we cannot derive a total S/H abundance with the current MOSFIRE data set.

4. Discussion

4.1. Implications for Metallicity Calibrations at High Redshift

Determining the form of strong-line metallicity calibrations at high redshift is one of the most pressing matters for galaxy-evolution studies in the next decade. Such relations are required to take full advantage of data from large spectroscopic galaxy surveys, including existing spectra for thousands of galaxies at $z \sim 1-4$ (e.g., Steidel et al. 2014; Kriek et al. 2015; Momcheva et al. 2016; Kashino et al. 2019), and upcoming data from JWST at $z > 4$ and reaching into the Epoch of Reionization. In Figure 3, we show a range of strong-line ratio metallicity indicators plotted against direct-method metallicity for COSMOS 19985 and 20062 (red stars). For comparison, we include the compilation of Sanders et al. (2020a) of 18 galaxies at $z \sim 2.2$ with direct-method metallicities based on O^{2+} measurements (from either [O III]λ4364 or O III]$\lambda\lambda$ 1661, 1666), as well as a $z = 2.59$ dwarf galaxy with detected [O III]λ4364 from Gburek et al. (2019; blue points). We also show the calibrations of Curti et al. (2020) derived from stacked $z \sim 0$ SDSS spectra and those of Bian et al. (2018) derived from stacked spectra of $z \sim 2$ analogs selected from the SDSS. The running median and scatter of the local H II region sample presented in Sanders et al. (2020a) is also displayed.

The [O III]-based sample drawn from the literature ubiquitously displays a high degree of excitation and ionization ($\log([O\,III]/H\,\beta) \gtrsim 0.7$, $\log([O\,III]/[O\,II]) \gtrsim 0.5$, $\log(R23) \gtrsim 0.9$) and low metallicity ($12+\log(O/H) \lesssim 8.1$). In contrast, 19985 and 20062 have lower levels of excitation and ionization, and 20062 has a higher metallicity. This comparison demonstrates the potential of auroral [O II] to provide direct-method abundances in regions of parameter space where auroral [O III] is weaker and thus harder to detect. In local H II regions, it has long been known that [O III]λ4364 becomes difficult to detect at high metallicities due to the combined effect of low O^{2+}/O and cool T_e (e.g., Berg et al. 2015; Croxall et al. 2015, 2016). In contrast, [O II]$\lambda\lambda$7322, 7332 remains relatively strong over a wide range in metallicity, including at low metallicities despite the fact that it traces a low-ionization species. For example, in the sample of Izotov et al. (2006) selected from SDSS based on [O III]λ4364 detection, which spans $12+\log(O/H) \sim 7.7$–8.5, nearly every galaxy also has a detection of [O II]$\lambda\lambda$7322, 7332. Constructing a representative sample of high-redshift galaxies with direct-method abundances spanning a wide range of metallicities will thus require a sample based on a mixture of high- and low-ionization auroral emission lines. It is clear that relying on the commonly used [O III]λ4364 alone will likely result in a sample that is biased toward high excitation and may fail to span a wide dynamic range in O/H. Indeed, in order to trace the actual shape of the calibration between strong-line ratios and metallicities, a significantly wider range in $12+\log(O/H)$ must be probed than is present in the current [O III]-based sample.

Considering the $z \sim 2$ [O III]- and [O II]-auroral samples as a whole, we begin to resolve the qualitative shape of the relations between strong-line ratios and direct-method O/H for the first time at high redshift. We find that [O III]/Hβ and R23 are double valued with a turnover region around $12+\log(O/H) \sim 8.0$. [O II]/Hβ increases with increasing metallicity. [O III]/[O II] and [Ne III]/[O II] both decrease with increasing metallicity, though these ratios display a large scatter at fixed O/H. [N II]/Hα increases and O3N2 decreases with increasing O/H. All of these trends are in qualitative agreement with the shape of empirical and theoretical metallicity calibrations constructed for use in the low-redshift universe (e.g., McGaugh 1991; Kewley & Dopita 2002; Pettini & Pagel 2004; Maiolino et al. 2008; Pilyugin & Grebel 2016; Curti et al. 2017, 2020; Kewley et al. 2019). However, as noted in Sanders et al. (2020a), evolution in normalization at fixed O/H appears to be present on average for some line ratios, as evidenced by

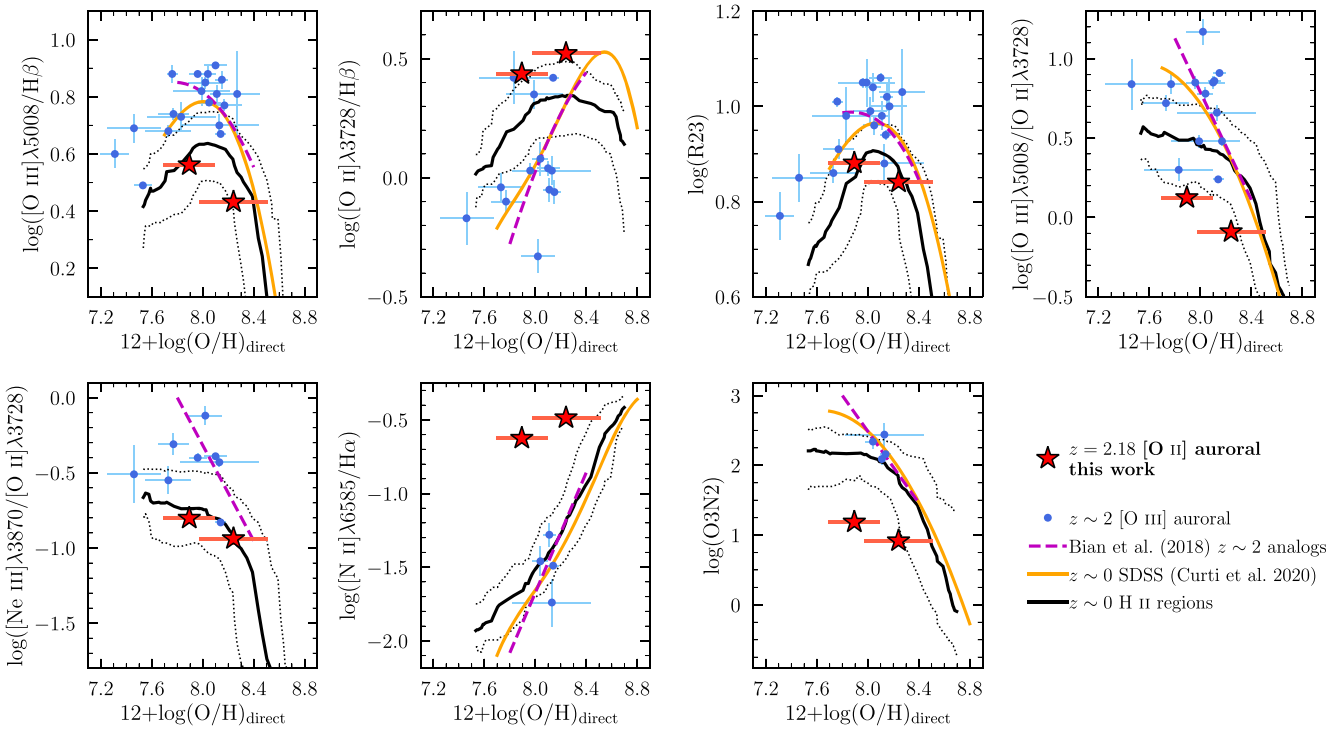

Figure 3. Strong-line ratios vs. direct-method oxygen abundance. The two $z = 2.18$ targets of this analysis, with metallicities based on [O II] auroral lines, are displayed as red stars. Blue circles denote $z > 1$ sources drawn from the literature with direct-method metallicities based on [O III] auroral lines (Gburek et al. 2019; Sanders et al. 2020a). The black solid and dotted lines show the median and 1σ scatter of the distribution of $z = 0$ H II regions from the sample presented in Sanders et al. (2017; see also Pilyugin & Grebel 2016). The solid orange line shows the calibration set derived by Curti et al. (2020) based on stacked spectra of $z \sim 0$ star-forming galaxies from the SDSS. The dashed purple line presents the high-redshift analog calibrations of Bian et al. (2018), constructed by stacking spectra of a sample of local galaxies selected to lie in the same region of the [O III]/Hβ vs. [N II]/Hα Baldwin, Phillips & Terlevich diagram as $z \sim 2$ star-forming galaxies.

the high values of R23 and [O III]/Hβ that typical local samples fail to reach.

Despite the significant observational investment that has enabled this combined high-redshift auroral-line sample, both the sample size and precision of individual metallicity determinations are clearly too low to draw any quantitative conclusions about the form of high-redshift calibrations and robustly constrain their evolution with respect to local relations. Even when pushing to the limits of what is currently feasible with 8–10 m ground-based telescopes, auroral lines can only be detected for the brightest line emitters at $z \sim 1$–3 and in many cases require rare strong gravitational lensing of targets, in particular redshift intervals. Continued progress in this area clearly requires new facilities with improved capabilities.

4.2. Prospects for Direct-method Abundances with JWST

With the recent advent of JWST science operations, galaxy-evolution studies will be revolutionized by the unrivaled infrared spectroscopy capabilities the telescope offers. Indeed, the promise of JWST/Near Infrared Spectrograph (NIRSpec) to transform high-redshift direct-method metallicity studies is already being demonstrated with the recent detection of [O III] λ4363 in three lensed $z > 7$ galaxies (Arellano-Córdova et al. 2022; Brinchmann 2022; Carnall et al. 2022; Curti et al. 2022; Katz et al. 2022; Schaerer et al. 2022; Trump et al. 2022). One of the key improvements offered by JWST for near-infrared spectroscopic observations is the removal of the challenges associated with observing through the Earth's atmosphere. These include the bright and strongly wavelength-dependent background noise, and the significant wavelength gaps in atmospheric transmission. Given the faintness of auroral emission lines, an increase in sensitivity is the key requirement to improve upon the current high-redshift direct-method metallicity sample. JWST/NIRSpec will yield a factor of several gain in sensitivity relative to Keck/MOSFIRE for emission-line studies, with the largest improvements at redder wavelengths.

As an example, for the targets of this analysis, measuring the [O II]$\lambda\lambda$7322, 7332 lines at $\lambda_{\rm obs} \sim 2.3\,\mu$m with fluxes of $\sim 8 \times 10^{-18}$ erg s^{-1} cm^{-2} at 3–4σ required 5.5 hr of integration with Keck/MOSFIRE. The JWST exposure-time calculator suggests that NIRSpec microshutter array spectroscopy would enable detection of these lines at 5σ in an on-source integration time of only $\sim$10 minutes with the G235M/F170LP setting. Likewise, unlensed galaxies in the current [O III] auroral sample have [O III]λ4364 fluxes of $\sim 5 \times 10^{-18}$ erg s^{-1} cm^{-2} (Sanders et al. 2020a). At $z = 2.2$ ($\lambda_{\rm obs} \sim 1.4\,\mu$m), this line would be detected by JWST/NIRSpec at 5σ in $\sim$30 minutes on source in the G140M/F100LP setting. This exercise demonstrates how quickly JWST/NIRSpec's capabilities could reproduce the existing $z \sim 2$ auroral-line sample, representing many nights of 8–10 m telescope time, with a higher average S/N. We note, however, that both the targets of the current analysis and those in the auroral [O III] literature sample have emission-line fluxes considerably brighter than what is typical for $z \sim 2$ main-sequence galaxies. Consequently, deeper integrations with JWST/NIRSpec of up to several hours will enable the detection of auroral emission lines in more representative $z \sim 2$ galaxies, moving beyond the highly biased sample that is available from ground-based observations. Indeed, current 8–10 m class ground-based facilities are simply not capable of establishing direct-method

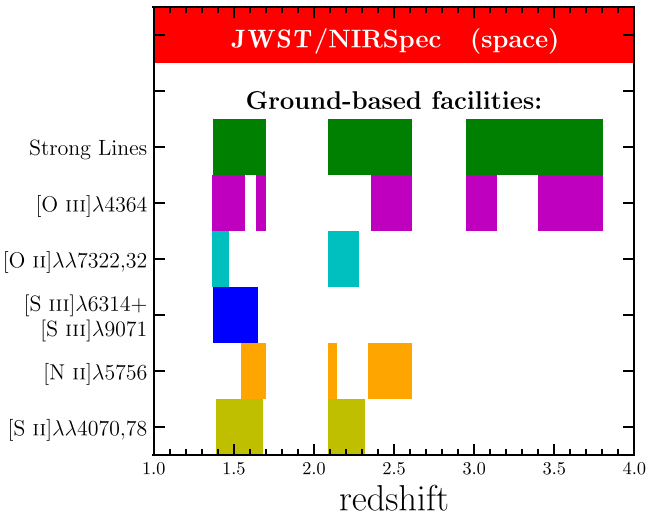

Figure 4. The redshift ranges across which the strong rest-optical emission lines ([O II]λ3728, Hβ, [O III]λ5008, Hα, [N II]λ6585, [S II]$\lambda\lambda$ 6718, 6733) and various auroral emission lines can be observed simultaneously. The green shaded regions show where the full suite of strong lines can be accessed from the ground in near-infrared windows of atmospheric transmission. The shaded regions in each row below that present the range at which auroral lines of various ionic species can be observed alongside the strong lines. Thanks to its continuous 0.7–5.0 μm wavelength coverage, JWST/NIRSpec can access all of these lines across the full $z = 1$–4 redshift range.

metallicity calibrations for a large and representative sample of star-forming galaxies at $z \sim 2$.

In addition to an increase in sensitivity, the move to space with JWST provides a critical improvement for auroral-line studies by removing the restriction of observing only at wavelengths inside the near-infrared windows of high atmospheric transmission. Deriving direct-method metallicities requires simultaneous coverage of at least one auroral line (e.g., [O III]λ4364, [O II]$\lambda\lambda$7322, 7332) and a suite of strong rest-optical emission lines (i.e., Hα, Hβ, [O III]λ5008, [O II]$\lambda\lambda$ 3727, 3730, [S II]$\lambda\lambda$6718, 6733). Fitting this combination of lines inside ground-based near-infrared bands drastically restricts the range of allowed redshifts, resulting in a vastly reduced number of potential targets on top of the requirement of extraordinarily bright line emission (indeed, 19985 and 20062 are among the brightest $z \sim 2$ line emitters in the CANDELS fields). Figure 4 displays the redshift ranges in which the rest-optical strong lines and auroral lines of [O III], [O II], [S III], [N II], or [S II] are accessible in the ground-based near-infrared bands. It is clear that the redshift range of possible targets is severely limited from the ground for this science case. There is a particularly drastic limitation for all auroral lines except [O III]λ4364, partially explaining why only auroral lines of [O III] had been detected at $z > 1$ prior to this work. Furthermore, the [O III] and [O II] auroral emission lines can only be simultaneously accessed from the ground in the highly restrictive redshift range of $z = 1.37$–1.47.

In contrast, JWST/NIRSpec's continuous high-sensitivity spectral coverage across 0.7–5 μm enables simultaneous measurements of all rest-optical auroral lines along with the strong lines. This capability has two major advantages for direct-method metallicity studies. First, the number of potential auroral-line targets within the range of JWST/NIRSpec's increased sensitivity is greatly increased, making efficient multiplexing possible. Such a gain in multiplexing provides significant improvement over ground-based observations with, for example, Keck/MOSFIRE, for which pointings with even a handful of bright enough targets within the required redshift range are exceedingly rare. Second, the continuous wavelength coverage afforded by JWST provides the ability to directly constrain the electron temperature in both the low- and high-ionization zones simultaneously by detecting auroral lines of low-ionization (e.g., [O II], [N II], [S II]) and high-ionization (e.g., [O III], [S III]) energy species in individual targets. This step is required to bring high-redshift chemical abundance studies on par with the level of precision regularly reached by local studies, and is only possible within extremely limited redshift ranges from the ground. In this work, as in all past analyses of high-redshift auroral detections, T_e is only measured in one zone while the T_e of the other zone is inferred through empirical or theoretical relations between low- and high-ionization temperatures (e.g., Campbell et al. 1986; Izotov et al. 2006). This approach introduces systematic uncertainties associated with both the unknown form of T_e(high)–T_e(low) relations at high redshift and the scatter in these relations, the latter of which is significant at $z \sim 0$ ($\sim$1000 K; Croxall et al. 2016; Berg et al. 2020; Rogers et al. 2021). Inferring direct-method O/H from a single-zone T_e measurement has been shown to introduce large biases when the predominant ionic form is in the unobserved zone (Yates et al. 2020). While such systematic effects are not a dominant source of uncertainty in this analysis since the [O II]$\lambda\lambda$7322, 7332 lines are detected at low significance (3–5σ), they will become important as JWST enables high-S/N auroral-line measurements at $z > 1$.

To demonstrate the feasibility of detecting multiple auroral lines for bright targets with JWST/NIRSpec, we combined $T_e(O^+)$ and $T_e(O^{2+})$ reported in Table 3 with the strong-line fluxes and $E(B-V)_{\rm gas}$ to estimate the observed (reddened) fluxes of auroral lines that are inaccessible from the ground for COSMOS 19985 due to its redshift. For these calculations, we assume a ratio of [S III]$\lambda\lambda$ 9071, 9533/[S II]$\lambda\lambda$6718, 6733 = 1.0 (Sanders et al. 2020b) since the strong [S III] lines are also unreachable from the ground at this redshift. In units of 10^{-18} erg s^{-1} cm^{-2}, we predict that $F_{\rm obs}$([O III]λ4364) = 6.3, $F_{\rm obs}$([N II]λ5755) = 2.7, $F_{\rm obs}$([S III]λ6312) = 2.1, and $F_{\rm obs}$([S II]λ4076) = 0.6. Thus, the [O III], [N II], [S III], and [O II] auroral lines would be detectable at $\gtrsim 5\sigma$ with JWST/NIRSpec Micro-Shutter Assembly integrations of $\sim$30 minutes in G140M/F100LP and $\sim$1 hr in G235M/F170LP for targets like 19985, noting that this source has atypically bright emission lines and that integrations of several hours would be required to obtain similar results for galaxies closer to the star-forming main sequence. This exercise also demonstrates that the [O III] and [O II] auroral lines are typically the brightest T_e diagnostics at low and moderate metallicities most relevant for high-redshift studies ($\lesssim 0.5\ Z_\odot$) and present a technically feasible avenue to simultaneous low- and high-ionization constraints. Indeed, for 19985, we would have detected [O III]λ4364 at $\sim 5\sigma$ based on the depth of our H-band MOSFIRE observations, but this line fell in the atmospheric gap between the H and J bands.

In summary, the increased sensitivity and wavelength coverage of JWST/NIRSpec will yield more than an order of magnitude increase in the efficiency of detecting auroral emission lines at $z \sim 1$–4 relative to the ability of ground-based 8–10 m class telescopes. This performance promises to usher in an era of precision chemical evolution studies of high-redshift galaxies early in the mission lifetime of JWST.

4.3. Minimizing Systematic Uncertainties on Auroral [O II]-based Metallicities

Basing direct-method metallicities on [O II]$\lambda\lambda$7322, 7332 presents some challenges that may systematically impact the outcome but can be addressed with upcoming observations. First, auroral [O II]$\lambda\lambda$7322, 7332 and strong [O II]$\lambda\lambda$3727, 3730 are widely separated in wavelength such that the strong-to-auroral ratio and derived $T_e(O^+)$ are sensitive to the reddening correction. For example, assuming the Cardelli et al. (1989) extinction curve,[11] we find that changing $E(B-V)_{gas}$ by 0.1 mag results in a change to [O II]$\lambda\lambda$7322, 7332/[O II]$\lambda\lambda$3727, 3730 of 30%, which would strongly affect the derived $T_e(O^+)$ and O^+/H^+. We found that the $E(B-V)_{gas}$ derived from higher-order Balmer lines is consistent with the value based on Hα/Hβ, suggesting a robust reddening correction for [O II]$\lambda\lambda$3727, 3730. Nevertheless, significant uncertainty about the high-redshift nebular extinction curve and its variation among individual galaxies remains. Estimates of dust reddening and determinations of the nebular dust law will soon be significantly improved via JWST/NIRSpec's long-wavelength coverage that provides access to relatively unreddeneed Paschen series lines (Reddy et al. 2020).

Temperature determinations based on [O II]$\lambda\lambda$7322, 7332 emission can also be biased by dielectric recombination into the upper level that produces these transitions, leading to an overestimate of $T_e(O^+)$ (Rubin 1986; Liu et al. 2000). However, this effect is strongest at low temperatures ($\sim$5000–8000 K) and high metallicities ($\gtrsim Z_\odot$), as well as high densities ($n_e > 1000$ cm^{-3}). This effect is thus not expected to be important at high redshift where galaxies are typically relatively metal-poor and have $n_e \sim 200$–300 cm^{-3} (e.g., Shimakawa et al. 2015; Sanders et al. 2016). Using the O II recombination coefficients from Storey et al. (2017), we find that the emissivity of the [O II]$\lambda\lambda$7322, 7332 lines resulting from recombination are $\gtrsim$100$\times$ lower than the emissivity due to collisional excitation at the $T_e(O^+)$ and n_e derived for these targets. Accordingly, recombination effects have a negligible impact on our results.

4.4. N/O and α/Fe Abundance Patterns

Chemical abundance patterns of non-α elements (e.g., N, C, Fe) relative to O can provide strong constraints on the formation histories of high-redshift systems (e.g., Berg et al. 2016, 2019; Steidel et al. 2016; Topping et al. 2020a, 2020b; Cullen et al. 2021; Strom et al. 2022). We have determined N/H and N/O using the direct method for the first time at high redshift. Figure 5 shows N/O versus O/H for 19985 and 20062, along with a sample of local H II regions from the CHAOS survey (Berg et al. 2020). Both galaxies appear to be offset from the mean $z = 0$ N/O–O/H relation toward higher N/O at fixed O/H. 20062 lies at the upper envelope of the distribution of local H II regions, though its offset from the mean relation is only $\sim 1\sigma$ significant. 19985, on the other hand, presents evidence for elevated N/O at fixed O/H, being 3σ inconsistent in O/H with the mean $z = 0$ relation at fixed N/O or 5σ inconsistent in N/O at fixed O/H. The anomalously high N/O of both galaxies explains why they are outliers in [N II]/Hα and O3N2 versus O/H (Figure 3), and highlights the potential pitfall of N-based metallicity indicators that are highly sensitive to N/O variations.

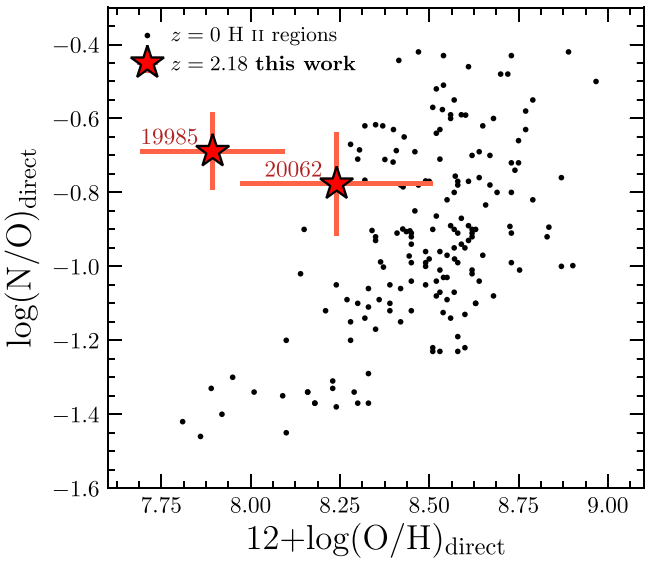

Figure 5. N/O vs. O/H for the $z = 2.18$ galaxies analyzed in this work (red stars) and local H II regions from the CHAOS survey (black points; Berg et al. 2020). All abundances are determined using the direct method. 19985 displays a significant offset from the $z = 0$ relation.

It is expected that rapidly forming galaxies will have low N/O abundance ratios because nitrogen enrichment, predominantly originating from intermediate-mass stars, occurs on longer timescales than that of α elements, which predominantly come from core-collapse supernovae. Furthermore, low-metallicity systems ($12+\log(O/H) \lesssim 8.2$) typically lie on the "primary nucleosynthesis" plateau at $\log(N/O) \approx -1.4$ (e.g., Berg et al. 2019), where the increased presence of oxygen in higher-metallicity systems catalyzes CNO reactions resulting in an increased output of nitrogen relative to oxygen with increasing metallicity. Given the extremely high specific SFR (sSFR ~ 20 Gyr^{-1}) and relatively low O/H ($12+\log(O/H) \sim 7.9$–8.2) of 19985 and 20062, it is surprising that these galaxies present near-solar N/O.

An overabundance of Wolf–Rayet stars could account for an excess of nitrogen produced by young massive stars (Masters et al. 2014), though this scenario requires significant changes to either the upper end of the initial mass function or to stellar evolutionary processes. Another possibility is that these galaxies recently experienced a strong gas-accretion event (possibly through a gas-rich merger of a lower-metallicity companion). If a galaxy quickly accretes gas with a significantly lower metallicity than its preceding ISM metallicity, ISM metals will be diluted such that O/H decreases while heavy element abundance ratios including N/O will remain relatively unchanged (Koppen & Hensler 2005). This interpretation is consistent with the presence of a strong starburst (fueled by the accreted gas) in 19985 and 20062, which lie an order of magnitude above the mean $z \sim 2$ star-forming main sequence (Speagle et al. 2014), and with their large gas fractions as inferred from Σ_{SFR} and the Kennicutt–Schmidt relation (Kennicutt & De Los Reyes 2021) as well as a gas mass measurement for 19985 based on CO ($M_{gas,CO}/M_* \sim 10$; Sanders et al. 2022). The best-fit SED models also suggest rising star formation histories (i.e., $t_{age} \leqslant \tau$; see Table 1).

Topping et al. (2020b) recently reported constraints on the Fe/H of young stars in 19985 and 20062 based on modeling

[11] Reddy et al. (2020) found that the nebular attenuation curve in a sample of $z \sim 2$ star-forming galaxies from the MOSDEF survey is similar to the Cardelli et al. (1989) Milky Way extinction law on average.

the rest-ultraviolet continuum, finding that both systems host low-metallicity massive stars with $Z_* = 0.001 = 0.07\,Z_\odot$ (i.e., $12+\log(\mathrm{Fe/H}) = 6.31$). Under the assumption that the chemical composition of young stars is identical to the instantaneous gas-phase ISM composition, comparing Fe/H from Topping et al. (2020b) to direct-method O/H from this work provides a measure of the α/Fe ratio in these systems. We find $\alpha/\mathrm{Fe} = 2.3^{+1.4}_{-0.9} \times \alpha/\mathrm{Fe}_\odot$ and $5.0^{+4.4}_{-2.3} \times \alpha/\mathrm{Fe}_\odot$ for 19985 and 20062, respectively, consistent with α enhancement relative to the solar abundance pattern for both galaxies.[12] This result agrees with other recent studies that find typical star-forming galaxies at $z \sim 2$–3 are α-enhanced by factors of ∼2–5 relative to the solar abundance pattern (Steidel et al. 2016; Strom et al. 2018, 2022; Cullen et al. 2021; Sanders et al. 2020a; Topping et al. 2020a, 2020b; Reddy et al. 2022). Such α enhancement occurs due to their rapid formation timescales that favor significant α enrichment from prompt core-collapse supernovae but have not yet been significantly enriched in Fe-peak elements by Type Ia supernovae that occur on longer timescales. This α enhancement (or iron deficit) has important implications for the hardness of the ionizing spectrum in H II regions, a crucial consideration for interpreting emission-line ratios, since Fe-peak line blanketing is a major factor governing the ionizing photon output of massive stars.

These results demonstrate the power of chemical abundance patterns to shed light on galaxy formation histories. Novel measurements with JWST will provide the necessary improvement on the precision of direct abundancess to transform such analyses from a qualitative to quantitative regime at high redshift. The wide and continuous spectral coverage of JWST/NIRSpec, alongside the ability to simultaneously constrain T_e in both the low- and high-ionization zones, will enable the determination of multielement chemical abundance patterns at $z > 1$ approaching the level of detail present in local studies (e.g., Izotov et al. 2006).

4.5. Is the Line Emission Powered by Active Galactic Nuclei?

It is of interest to consider whether the line emission in these galaxies is powered by the accreting black hole of an AGN instead of photoionization by massive stars in H II regions, as has been implicitly assumed in this analysis. Both targets are offset from the mean $z \sim 0$ sequence of star-forming galaxies in the [O III]/Hβ versus [N II]/Hα "Baldwin, Phillips & Terlevich" (BPT) diagram (Baldwin et al. 1981), falling in the "composite" region between local star-forming galaxies and AGN (Kewley et al. 2001, 2006; Kauffmann et al. 2003) but squarely within the distribution of typical $z \sim 2$ star-forming galaxies (e.g., Steidel et al. 2014; Shapley et al. 2015; Strom et al. 2017; Runco et al. 2022). However, these targets are both within the pure star-forming region of the [O III]/Hβ versus [S II]/Hα and [O I]/Hα diagnostic diagrams, where local AGNs are observed to have $\log([\mathrm{S\,II}]/\mathrm{H}\,\alpha) > -0.5$ and $\log([\mathrm{O\,I}]/\mathrm{H}\,\alpha) > -1.4$ at similar [O III]/Hβ as 19985 and 20062 (Veilleux & Osterbrock 1987; Kewley et al. 2006). The [O III]/Hβ versus [O II]$\lambda\lambda$7322, 7332/Hα diagram has also been proposed as a diagnostic to distinguish between star formation and AGN-ionized sources (Osterbrock et al. 1992). We measure dust-corrected $\log([\mathrm{O\,II}]\lambda\lambda 7322, 7332/\mathrm{H}\alpha) = -1.56 \pm 0.08$ and -1.61 ± 0.09 for 19985 and 20052, respectively. These values again place our targets within the distribution of star-forming galaxies and H II regions. Low-redshift AGN and composite objects have $\log(\mathrm{He\,II}\,\lambda4686/\mathrm{H}\,\beta) > -1.4$ (e.g., Shirazi & Brinchmann 2012). While He II λ4686 is not detected in either spectrum, we place stringent 3σ upper limits of $\log(\mathrm{He\,II}\,\lambda4686/\mathrm{H}\,\beta) < -1.60$ and < -1.25 for 19985 and 20062, respectively.

Furthermore, 19985 and 20062 are not detected in X-rays in the Chandra observations of the COSMOS field, nor do they display Spitzer infrared colors indicative of AGN-heated dust (Coil et al. 2015; Azadi et al. 2017). Both targets additionally have deep Keck/LRIS spectra covering the rest-frame ultraviolet over 1000−2300 Å (Topping et al. 2020a, 2020b), in which no high-ionization emission lines typical of AGNs (e.g., N V, He II, C IV) are detected. In summary, a wide range of diagnostics including X-ray, infrared, and rest-frame optical and ultraviolet data strongly suggest that AGN ionization is not a significant contributor to the total line emission in these galaxies and thus does not affect our results.

4.6. The Presence of Outflows and Broad Emission

Strongly concentrated star formation has been found to be associated with strong and efficient gas outflows (e.g., Steidel et al. 2010; Newman et al. 2012; Heckman et al. 2015; Davies et al. 2019; Schreiber et al. 2019; Weldon et al. 2022). The extremely high SFR surface densities ($\Sigma_{\mathrm{SFR}} = 18.3$ and $19.3\,M_\odot\,\mathrm{yr}^{-1}\,\mathrm{kpc}^{-2}$) of 19985 and 20062 suggest that the starbursts they are currently experiencing may be driving powerful outflows. Indeed, both galaxies are observed to have blueshifted and asymmetric rest-ultraviolet absorption lines with extended blue wings (Weldon et al. 2022). Furthermore, Leung et al. (2017) fit rest-optical emission lines in the MOSDEF spectra of 19985 and 20062 with double-Gaussian profiles including a broad and narrow component,[13] detecting broad components in both galaxies with FWHM $\sim 500\,\mathrm{km\,s}^{-1}$ and offset from the systemic redshift by $\sim -50\,\mathrm{km\,s}^{-1}$. In the strong rest-optical lines ([O II], Hβ, [O III], Hα, and [N II]), we observe wide asymmetric emission with a more prominent blue wing that the single-Gaussian profiles fail to fit, though this "missed" emission is $\lesssim$5% of the single-Gaussian line flux.

It is thus clear that strong star formation is driving significant outflows in both galaxies, detected in both neutral and ionized phases. We will present a more detailed analysis of the outflow properties of these targets in future work. What concerns the current analysis is whether the presence of such outflows biases the inference of physical properties and chemical abundances from integrated emission-line measurements. The answer to this question ultimately comes down to the fraction of the total line fluxes comprised by the broad component, and the magnitude of the difference between the physical conditions in the outflowing gas and H II regions. If the broad component contributes a small fraction of the total flux, it is unlikely to strongly bias properties derived from the integrated lines even if the physical conditions differ significantly. However, if the outflowing component contributes a fraction of the total flux comparable to the narrow ISM component, inferences from integrated lines may be biased. A more detailed analysis

[12] If the T_e-based metallicities are converted to the O recombination line abundance scale, as has been suggested by some recent studies (e.g., Blanc et al. 2015; Steidel et al. 2016; Sanders et al. 2020a), the O/H and inferred α/Fe values would increase by a factor of 1.7.

[13] Leung et al. (2017) analyzed these targets as AGNs based on a BPT diagram selection. With additional data, we now find strong evidence against the presence of AGNs in these galaxies (Section 4.5).

beyond the scope of this work is required to robustly address these questions. As JWST enables the measurement of temperature and density diagnostics at higher redshifts when galaxies had higher SFRs and smaller radii (and thus are expected to have stronger outflows), it will become increasingly important to understand whether outflowing gas biases abundance determinations from integrated line measurements.

5. Summary and Conclusions

In this paper, we presented ultradeep rest-optical spectroscopy of two star-forming galaxies at $z = 2.18$ in the COSMOS field with bright rest-optical emission lines, representing more than 20 hr of on-source integration with Keck/MOSFIRE. The high fidelity of the resulting spectra enabled the first detections of the auroral [O II]$\lambda\lambda$7322, 7332 emission-line doublet outside of the low-redshift universe. In turn, we used these data to obtain constraints on the electron temperature in the low-ionization zone of the ionized ISM and calculate the gas-phase O/H, N/H, and N/O abundance ratios via the direct method. Our main results are summarized as follows:

1. We measured low-ionization temperatures of $T_e(O^+) = 13{,}110 \pm 1800$ K and $10{,}530 \pm 1420$ K and gas-phase oxygen abundances of $12+\log(O/H) = 7.89 \pm 0.20$ and 8.24 ± 0.27 for COSMOS 19985 and 20062, respectively. These detections of [O II]$\lambda\lambda$7322, 7332 at $z \sim 2$ demonstrate the feasibility of using [O II] auroral lines for direct-method abundance studies of distant galaxies in the early universe. The NIRSpec instrument on JWST would take only $\sim$10 minutes of exposure at $R \sim 1000$ to detect [O II]$\lambda\lambda$7322, 7332 as bright as observed in these targets. As such, auroral [O II] can be detected in the brightest high-redshift line emitters with only shallow integrations with JWST/NIRSpec.
2. We placed the new [O II]-based direct-method O/H measurements alongside a sample of $\sim$20 galaxies at $z > 1$ with direct-method metallicities based on auroral [O III] on diagrams of strong-line ratio versus $12+\log(O/H)_{direct}$. We found that the [O II]-based targets have lower levels of ionization and excitation and lie at higher metallicity than the bulk of the [O III]-based sample. This result suggests that obtaining a representative direct-method abundance sample, essential for producing accurate strong-line calibrations for use at high redshifts, will ultimately require a combination of low- and high-ionization auroral-line measurements to populate a sufficient dynamic range in metallicity and excitation.
3. JWST/NIRSpec promises to simultaneously detect low- and high-ionization auroral lines of relatively bright individual targets in moderately deep exposures ($\geqslant$1 hr), with [O III] and [O II] being the most observationally accessible combination. The gain in efficiency of JWST/NIRSpec over ground-based facilities for auroral-line surveys is twofold, benefiting from both an increase in sensitivity and from uninterrupted wavelength coverage over 1–5 μm (where ground-based facilities can only observe in segmented near-infrared atmospheric transmission windows). These combined effects significantly increase the on-sky density of potential targets for which the required auroral and strong lines can be simultaneously observed, significantly improving multiplexing efficiency.
4. We investigated the abundance patterns of non-α elements, including nitrogen and iron. Both galaxies display near-solar N/O despite having <40% solar O/H. This unexpected composition may indicate a recent accretion of a large mass of relatively unenriched gas. This scenario is consistent with the apparent starburst nature of these targets. We found evidence of supersolar α/Fe in both objects, suggestive of rapid formation timescales for these systems, consistent with other studies at $z \sim 2$–3. The ability of JWST/NIRSpec to measure auroral emission lines of multiple species promises constraints on gas-phase chemical abundance patterns approaching the level of detail present in studies of local H II regions and star-forming galaxies.
5. The depth of the spectra analyzed here enabled a number of diagnostic tests to determine whether accreting supermassive black holes contribute significantly to the line emission in these sources. We found no evidence of AGN activity based on the position of our targets in the [O III]/Hβ versus [N II]/Hα, [S II]/Hα, and [O I]/Hα BPT diagrams. This conclusion was further confirmed by the measured [O II]$\lambda\lambda$7322, 7332/Hα ratio and stringent upper limits on the He II λ4686/Hβ ratio, both lower than the value displayed by active galaxies.
6. We found evidence for powerful gas outflows based on broad blueshifted wings around strong emission lines, revealed in the deep rest-optical spectra. The presence of ionized emission from such strong outflows may bias inferences on the physical conditions and abundances of the ISM based on integrated line emission. This question requires more analysis and will become increasingly important at higher redshifts, where star formation becomes more concentrated and is more likely to drive massive galaxy-scale outflows.

This analysis has demonstrated the feasibility of detecting auroral [O II]$\lambda\lambda$7322, 7332 to constrain the chemical abundances of high-redshift star-forming galaxies across a range of metallicities. However, a significant observational investment was required to obtain these results on auroral [O II] emission for two galaxies and also assemble the small, biased, and low-precision sample of high-redshift auroral [O III] targets, representing many nights on 8–10 m ground-based facilities. The significant cost of these efforts clearly demonstrates that current ground-based telescopes cannot provide the quality of near-infrared spectroscopy required to robustly understand the absolute metallicities of high-redshift sources. The advent of JWST, and moderate-resolution spectroscopy with its NIRSpec instrument in particular, represents an unprecedented leap forward in our ability to efficiently detect temperature-sensitive auroral emission lines at high redshift, ushering in an era of precision chemical abundance studies in the early universe.

Support for this work was provided through the NASA Hubble Fellowship grant #HST-HF2-51469.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555. We also acknowledge support from NSF AAG grant Nos. AST-1312780, 1312547, 1312764, 1313171, 2009313, and 2009085, grant No. AR-13907 from the Space Telescope

Science Institute, and grant No. NNX16AF54G from the NASA ADAP program. We wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

Facility: Keck:I (MOSFIRE).

Software: `pyneb` (Luridiana et al. 2015).

Appendix
Nebular Emission Correction to Measured Photometry

We applied a correction to the measured photometric flux densities to remove contributions from nebular line and continuum emission. This correction is necessary because the SED-fitting code FAST employed in Section 2.2 only fits models of stellar emission, such that any significant nonstellar contributions to the broadband photometry will bias the inferred stellar population parameters. The emission-line correction was achieved using initial line flux estimates from the MOSFIRE spectroscopy following Sanders et al. (2021). Briefly, each emission line detected at $S/N \geqslant 3$ was added to a noise-free emission-line-only model spectrum in which each line is represented as a single Gaussian based on the measured line centroid, width, and flux. This spectrum was then passed through the transmission curve of each photometric filter to infer the flux density contributed by line emission, and this value was subtracted from the measured photometric flux density to remove the emission-line contribution. The uncertainty on the line flux is propagated into the uncertainty of the corrected photometry.

Nebular continuum models were computed using Cloudy photoionization models (Ferland et al. 2017). The stellar ionizing spectrum input was derived from BPASS v2.2.1 binary models (Stanway & Eldridge 2018) assuming a constant star formation history that has reached equilibrium in the ionizing spectral shape (achieved at ages $\gtrsim$10 Myr). The stellar metallicity and ionization parameter were taken to be the values derived for 19985 and 20062 from spectral fitting of the rest-ultraviolet continuum by Topping et al. (2020b): $Z_* = 0.001$ and $\log(U) = -2.7$. The stellar population parameters used in the Cloudy models need not match the assumptions used in the SED fitting because the former requires the iron abundance of massive stars formed over the last $\sim$10 Myr that dominate the nebular emission production, while the latter requires the bulk metallicity of the total stellar population. The nebular metallicity, which controls the strength of the nebular continuum breaks (Byler et al. 2017), was set to the direct-method values derived from the rest-optical emission lines ($Z_{\rm gas} = 0.16$ and $0.35\,Z_\odot$). The model nebular continuum flux density is normalized by the total Hβ flux of the model, then multiplied by the dust-corrected Hβ flux measured from the MOSFIRE spectra to infer the intrinsic nebular continuum spectrum. This intrinsic component is then reddened assuming the Cardelli et al. (1989) extinction curve based on the observed Hα/Hβ ratio. For each photometric filter, the reddened nebular continuum component is then passed through the response curve to estimate the flux density contribution to each photometric measurement and subtracted off. The uncertainty in the nebular continuum flux density due to the uncertainty on $Z_{\rm gas}$ (0.2–0.3 dex) is $\approx$10%, which is propagated into the final uncertainty of the corrected photometry. Due to the large rest-optical emission-line equivalent widths in our targets, the correction for emission lines is important, accounting for approximately 20%, 40%, and 60% of the photometric flux density in the near-IR *J*-, *H*-, and *K*-band filters. The contribution from nebular continuum was $\approx$2–4% bluewards of rest-frame 2600 Å and $\approx$5–10% redwards of that wavelength. The emission-line correction was thus significantly more important for deriving accurate stellar population properties from SED fitting. Derived stellar masses are $\sim$0.5 dex higher without the nebular emission correction to the photometry.

ORCID iDs

Ryan L. Sanders https://orcid.org/0000-0003-4792-9119
Alice E. Shapley https://orcid.org/0000-0003-3509-4855
Leonardo Clarke https://orcid.org/0000-0003-1249-6392
Naveen A. Reddy https://orcid.org/0000-0001-9687-4973
Mariska Kriek https://orcid.org/0000-0002-7613-9872
Tucker Jones https://orcid.org/0000-0001-5860-3419
Mengtao Tang https://orcid.org/0000-0001-5940-338X

References

Amayo, A., Delgado-Inglada, G., & Stasinska, G. 2021, MNRAS, 505, 2361
Andrews, B. H., & Martini, P. 2013, ApJ, 765, 140
Arellano-Córdova, K. Z., Berg, D. A., Chisholm, J., et al. 2022, ApJL, 940, L23
Asplund, M., Amarsi, A. M., & Grevesse, N. 2021, A&A, 653, A141
Azadi, M., Coil, A. L., Aird, J., et al. 2017, ApJ, 835, 27
Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
Berg, D. A., Chisholm, J., Erb, D. K., et al. 2021, ApJ, 922, 170
Berg, D. A., Erb, D. K., Auger, M. W., Pettini, M., & Brammer, G. B. 2018, ApJ, 859, 164
Berg, D. A., Erb, D. K., Henry, R. B. C., Skillman, E. D., & McQuinn, K. B. W. 2019, ApJ, 874, 93
Berg, D. A., Pogge, R. W., Skillman, E. D., et al. 2020, ApJ, 893, 96
Berg, D. A., Skillman, E. D., Croxall, K. V., et al. 2015, ApJ, 806, 16
Berg, D. A., Skillman, E. D., Henry, R. B. C., Erb, D. K., & Carigi, L. 2016, ApJ, 827, 126
Bian, F., Kewley, L. J., & Dopita, M. A. 2018, ApJ, 859, 175
Blanc, G. A., Kewley, L., Vogt, F. P. A., & Dopita, M. A. 2015, ApJ, 798, 99
Bresolin, F., Gieren, W., Kudritzki, R. P., et al. 2009, ApJ, 700, 309
Brinchmann, J. 2022, arXiv:2208.07467
Byler, N., Dalcanton, J. J., Conroy, C., & Johnson, B. D. 2017, ApJ, 840, 44
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Campbell, A., Terlevich, R., & Melnick, J. 1986, MNRAS, 223, 811
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Carnall, A. C., Begley, R., McLeod, D. J., et al. 2023, MNRAS, 518, L45
Chabrier, G. 2003, PASP, 115, 763
Coil, A. L., Aird, J., Reddy, N., et al. 2015, ApJ, 801, 35
Conroy, C., & Gunn, J. E. 2010, FSPS: Flexible Stellar Population Synthesis, Astrophysics Source Code Library, ascl:1010.043
Croxall, K. V., Pogge, R. W., Berg, D. A., Skillman, E. D., & Moustakas, J. 2015, ApJ, 808, 42
Croxall, K. V., Pogge, R. W., Berg, D. A., Skillman, E. D., & Moustakas, J. 2016, ApJ, 830, 4
Cullen, F., Shapley, A. E., McLure, R. J., et al. 2021, MNRAS, 505, 903
Curti, M., Cresci, G., Mannucci, F., et al. 2017, MNRAS, 465, 1384
Curti, M., D'Eugenio, F., Carniani, S., et al. 2023, MNRAS, 518, 425
Curti, M., Mannucci, F., Cresci, G., & Maiolino, R. 2020, MNRAS, 491, 944
Davé, R., Rafieferantsoa, M. H., Thompson, R. J., & Hopkins, P. F. 2017, MNRAS, 467, 115
Davies, R. L., Förster Schreiber, N. M., Übler, H., et al. 2019, ApJ, 873, 122
Dors, O. L., Pérez-Montero, E., Hägele, G. F., Cardaci, M. V., & Krabbe, A. C. 2016, MNRAS, 456, 4407
Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, RMxAA, 53, 385
Förster Schreiber, N. M., Übler, H., Davies, R. L., et al. 2019, ApJ, 875, 21
Freeman, W. R., Siana, B., Kriek, M., et al. 2019, ApJ, 873, 102
Froese Fischer, C., & Tachiev, G. 2004, ADNDT, 87, 1
Gburek, T., Siana, B., Alavi, A., et al. 2019, ApJ, 887, 168
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Hao, C.-N., Kennicutt, R. C., Johnson, B. D., et al. 2011, ApJ, 741, 124

Heckman, T. M., Alexandroff, R. M., Borthakur, S., Overzier, R., & Leitherer, C. 2015, ApJ, 809, 147
Horne, K. 1986, PASP, 98, 609
Izotov, Y. I., Stasińska, G., Meynet, G., Guseva, N. G., & Thuan, T. X. 2006, A&A, 448, 955
Kaasinen, M., Bian, F., Groves, B., Kewley, L. J., & Gupta, A. 2017, MNRAS, 465, 3220
Kashino, D., & Inoue, A. K. 2019, MNRAS, 486, 1053
Kashino, D., Silverman, J. D., Sanders, D., et al. 2019, ApJS, 241, 10
Katz, H., Saxena, A., Cameron, A. J., et al. 2023, MNRAS, 518, 592
Kauffmann, G., Heckman, T. M., White, D. S. M., et al. 2003, MNRAS, 341, 33
Kennicutt, R. C., & De Los Reyes, M. A. C. 2021, ApJ, 908, 61
Kewley, L. J., & Dopita, M. A. 2002, ApJS, 142, 35
Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, ApJ, 556, 121
Kewley, L. J., Groves, B., Kauffmann, G., & Heckman, T. 2006, MNRAS, 372, 961
Kewley, L. J., Nicholls, D. C., & Sutherland, R. S. 2019, ARA&A, 57, 511
Kisielius, R., Storey, P. J., Ferland, G. J., & Keenan, F. P. 2009, MNRAS, 397, 903
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Koppen, J., & Hensler, G. 2005, A&A, 434, 531
Kriek, M., Shapley, A. E., Reddy, N. A., et al. 2015, ApJS, 218, 15
Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221
Leung, G. C. K., Coil, A. L., Azadi, M., et al. 2017, ApJ, 849, 48
Liu, X. W., Storey, P. J., Barlow, M. J., et al. 2000, MNRAS, 312, 585
Luridiana, V., Morisset, C., & Shaw, R. A. 2013, PyNeb: Analysis of emission lines, Astrophysics Source Code Library, ascl:1304.021
Luridiana, V., Morisset, C., & Shaw, R. A. 2015, A&A, 573, A42
Maiolino, R., Nagao, T., Grazian, A., et al. 2008, A&A, 488, 463
Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., & Gnerucci, A. 2010, MNRAS, 408, 2115
Masters, D., McCarthy, P., Siana, B., et al. 2014, ApJ, 785, 153
McGaugh, S. S. 1991, ApJ, 380, 140
McLean, I. S., Steidel, C. C., Epps, H. W., et al. 2012, Proc. SPIE, 8446
Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, ApJS, 225, 27
Newman, S. F., Genzel, R., Förster-Schreiber, N. M., et al. 2012, ApJ, 761, 43
Osterbrock, D. E., Tran, H. D., & Veilleux, S. 1992, ApJ, 389, 196
Peeples, M. S., & Shankar, F. 2011, MNRAS, 417, 2962
Peimbert, M. 1967, ApJ, 150, 825
Peimbert, M., Peimbert, A., & Delgado-Inglada, G. 2017, PASP, 129, 082001
Pettini, M., & Pagel, B. E. J. 2004, MNRAS, 348, L59
Pilyugin, L. S., & Grebel, E. K. 2016, MNRAS, 457, 3678
Reddy, N. A., Kriek, M., Shapley, A. E., et al. 2015, ApJ, 806, 259
Reddy, N. A., Shapley, A. E., Kriek, M., et al. 2020, ApJ, 902, 123
Reddy, N. A., Topping, M. W., Shapley, A. E., et al. 2022, ApJ, 926, 31
Rogers, N. S. J., Skillman, E. D., Pogge, R. W., et al. 2021, ApJ, 915, 21
Rubin, R. H. 1986, ApJ, 309, 334
Runco, J. N., Shapley, A. E., Sanders, R. L., et al. 2022, MNRAS, 517, 4337
Sanders, R. L., Jones, T., Shapley, A. E., et al. 2020b, ApJL, 888, L11
Sanders, R. L., Shapley, A. E., Jones, T., et al. 2021, ApJ, 914, 19
Sanders, R. L., Shapley, A. E., & Jones, T. 2023, ApJ, 942, 24
Sanders, R. L., Shapley, A. E., Kriek, M., et al. 2016, ApJ, 816, 23
Sanders, R. L., Shapley, A. E., Reddy, N. A., et al. 2020a, MNRAS, 491, 1427
Sanders, R. L., Shapley, A. E., Zhang, K., & Yan, R. 2017, ApJ, 850, 136
Schaerer, D., Marques-Chaves, R., Barrufet, L., et al. 2022, A&A, 665, L4
Shapley, A. E., Reddy, N. A., Kriek, M., et al. 2015, ApJ, 801, 88
Shapley, A. E., Sanders, R. L., Shao, P., et al. 2019, ApJL, 881, L35
Shimakawa, R., Kodama, T., Steidel, C. C., et al. 2015, MNRAS, 451, 1284
Shirazi, M., & Brinchmann, J. 2012, MNRAS, 421, 1043
Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, ApJS, 214, 24
Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15
Stanway, E. R., & Eldridge, J. J. 2018, MNRAS, 479, 75
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., Strom, A. L., Pettini, M., et al. 2016, ApJ, 826, 159
Storey, P. J., Sochi, T., & Badnell, N. R. 2014, MNRAS, 441, 3028
Storey, P. J., Sochi, T., & Bastin, R. 2017, MNRAS, 470, 379
Strom, A. L., Rudie, G. C., Steidel, C. C., & Trainor, R. F. 2022, ApJ, 925, 116
Strom, A. L., Steidel, C. C., Rudie, G. C., et al. 2017, ApJ, 836, 164
Strom, A. L., Steidel, C. C., Rudie, G. C., Trainor, R. F., & Pettini, M. 2018, ApJ, 868, 117
Tayal, S. S. 2011, ApJS, 195, 12
Tayal, S. S., & Zatsarinny, O. 2010, ApJS, 188, 32
Thuan, T. X., Izotov, Y. I., & Lipovetsky, V. A. 1995, ApJ, 445, 108
Topping, M. W., Shapley, A. E., Reddy, N. A., et al. 2020a, MNRAS, 495, 4430
Topping, M. W., Shapley, A. E., Reddy, N. A., et al. 2020b, MNRAS, 499, 1652
Torrey, P., Vogelsberger, M., Hernquist, L., et al. 2018, MNRAS, 477, L16
Torrey, P., Vogelsberger, M., Marinacci, F., et al. 2019, MNRAS, 484, 5587
Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898
Trump, J. R., Arrabal Haro, P., Simons, R. C., et al. 2022, arXiv:2207.12388
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
Veilleux, S., & Osterbrock, D. E. 1987, ApJS, 63, 295
Weldon, A., Reddy, N. A., Topping, M. W., et al. 2022, MNRAS, 515, 841
Yates, R. M., Schady, P., Chen, T. W., Schweyer, T., & Wiseman, P. 2020, A&A, 634, A107
Zahid, H. J., Dima, G. I., Kudritzki, R.-P., et al. 2014, ApJ, 791, 130