The Astrophysical Journal, 882:116 (9pp), 2019 September 10 https://doi.org/10.3847/1538-4357/ab33fd

Prospects for Extending the Mass–Metallicity Relation to Low Mass at High Redshift: A Case Study at $z \sim 1$

Alex J. Cameron[1,2], Tucker Jones[3], Tiantian Yuan[2,4], Michele Trenti[1,2], Stephanie Bernard[1], Alaina Henry[5], Austin Hoag[6], and Benedetta Vulcani[7]
[1] School of Physics, The University of Melbourne, Parkville, VIC 3010, Australia; alexc@student.unimelb.edu.au
[2] ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia
[3] Department of Physics, University of California, Davis, 1 Shields Avenue, Davis, CA 95616, USA
[4] Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia
[5] Space Telescope Science Institute, 3700 San Martin Drive Baltimore, MD 21218, USA
[6] Department of Physics and Astronomy, University of California, Los Angeles, 430 Portola Plaza, Los Angeles, CA 90095, USA
[7] INAF—Osservatorio astronomico di Padova, Vicolo Osservatorio 5, I-35122 Padova, Italy
Received 2019 February 7; revised 2019 June 23; accepted 2019 July 18; published 2019 September 9

Abstract

We report *J*-band MOSFIRE spectroscopy of a low-mass ($\log(M_*/M_\odot) = 8.62^{+0.10}_{-0.06}$) star-forming galaxy at $z = 0.997$ showing the detection of [N II] and [S II] alongside a strong Hα line. We derive a gas-phase metallicity of $\log(\mathrm{O/H}) = 7.99^{+0.13}_{-0.23}$, placing this object in a region of M_*–Z space that is sparsely populated at this redshift. Furthermore, many existing metallicity measurements in this M_*–z regime are derived from only [N II]/Hα (N2), a diagnostic widely used in high-redshift metallicity studies despite the known strong degeneracy with the ionization parameter and resulting large systematic uncertainty. We demonstrate that even in a regime where [N II] and [S II] are at the detection limit and the measurement uncertainty associated with the [N II]/[S II] ratio is high (S/N $\approx$ 3), the more sophisticated Dopita et al. diagnostic provides an improved constraint compared to N2 by reducing the systematic uncertainty due to the ionization parameter. This approach does not, however, dispel uncertainty associated with stochastic or systematic variations in the nitrogen-to-oxygen abundance ratio. While this approach improves upon N2, future progress in extending metallicity studies into this low-mass regime will require larger samples to allow for stochastic variations, as well as careful consideration of the global trends among dwarf galaxies in all physical parameters, not just metallicity.

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

1. Introduction

Measurements of the gas-phase oxygen abundance (metallicity hereafter) in galaxies provide powerful insights into the galaxy-scale star formation and gas-flow processes that have shaped the development of the galaxy population across cosmic time. Despite extensive studies into the tight correlation between metallicity and stellar mass in galaxies, the so-called mass–metallicity relation (MZR), much debate still exists as to its origin (Lequeux et al. 1979; Skillman et al. 1989; Tremonti et al. 2004; Berg et al. 2012; Andrews & Martini 2013; Yabe et al. 2014; Sanders et al. 2015; Maiolino & Mannucci 2019). The MZR is observed to have evolved with redshift, with lower average metallicities observed at earlier cosmic times for a given stellar mass (Savaglio et al. 2005; Erb et al. 2006a; Maiolino et al. 2008). However, constraints on the evolution of shape and scatter of the MZR are less clear, due in particular to the difficulties associated with making individual measurements of low-mass galaxies at high redshift. While average metallicity evolution is an important input into galaxy evolution models, extending constraints on chemical evolution to lower-mass objects promises key insights into the evolution of the galaxy population.

Existing studies suggest that the shape of the MZR is not constant across cosmic time (Zahid et al. 2013, 2014). These studies support a downsizing scenario in which low-mass galaxies enrich onto the local MZR at later times. However, despite predictions from theory suggesting that measuring the shape and scatter of the MZR below $\log(M_*/M_\odot) < 9.0$ provide the best prospects for disentangling the driving forces behind this evolution (e.g., Davé et al. 2012), high-redshift studies rarely extend into this mass regime. Thus, the MZR is poorly constrained for low-mass galaxies at high-redshift ($z \gtrsim 1$) as these observations are difficult to carry out, leading to small sample sizes of often loosely constrained measurements (Zahid et al. 2011; Wuyts et al. 2012, 2014, 2016; Belli et al. 2013; Henry et al. 2013a, 2013b; Yuan et al. 2013; Amorín et al. 2014; Maseda et al. 2014). Larger samples of low-mass measurements are required to disentangle the impact of different processes on the evolution of galaxies across cosmic time.

An additional challenge in studies of the MZR below $\log(M_*/M_\odot) < 9.0$ lies with uncertainties in how the metallicities are derived. Diagnostics based on electron temperature (T_e) are widely considered the most reliable measures of metallicity (e.g., review by Maiolino & Mannucci 2019). However, the range of application of this "direct" method is limited to objects in which the weak [O III] λ4363 emission line can be observed, making it unfeasible beyond moderate redshifts (Jones et al. 2015a; Ly et al. 2016a, 2016b; Calabrò et al. 2017). Motivated by these difficulties, numerous diagnostics have been developed based on ratios of the most easily detected strong rest-frame optical emission lines, calibrated from stellar population synthesis and photoionization models (e.g., Kewley & Dopita 2002) or T_e measurements taken in either the local universe (e.g., Pettini & Pagel 2004) or at moderate redshifts (Jones et al. 2015a). Alternatively, a number of generalized approaches exist that compare a range of strong-line fluxes to photoionization models to simultaneously

fit for metallicity alongside other key physical parameters (e.g., Pérez-Montero 2014; Blanc et al. 2015; Vale Asari et al. 2016). As metallicity studies are extended to higher redshifts and lower masses, these strong-line methods become an essential tool for understanding the galaxy population.

Strong-line diagnostics greatly extend the range of stellar masses and redshifts over which metallicities can be derived; however, questions remain about their reliability (Kewley & Ellison 2008; Steidel et al. 2014). In particular, when the number of observed emission lines is small, strong-line diagnostics often fail to disentangle the degeneracy between metallicity and the effects of other physical parameters such as ionization parameter, electron density, hardness of the ionizing sources, and relative abundance ratios (e.g., Morales-Luis et al. 2014; Maiolino & Mannucci 2019). These issues are omnipresent in studies targeting low-mass galaxies at high redshift where observational challenges frequently limit the range of available emission lines, meaning metallicities are often derived simply from [N II]/Hα (N2; Pettini & Pagel 2004).

The N2 ratio has its advantages in that the two lines are close in wavelength, such that they can be obtained in a single exposure and the ratio is independent of reddening. Thus, it has been useful in expanding measurements of the metallicity to faint galaxies (e.g., Erb et al. 2006a; Yabe et al. 2014). However, metallicities derived from this line ratio contain large systematic uncertainties due to strong degeneracy with the ionization parameter. Additionally, this ratio is primarily sensitive to nitrogen abundance, whereas oxygen abundance is derived with some assumed N/O ratio, introducing further uncertainty (e.g., Pérez-Montero & Contini 2009; Pérez-Montero 2014).

In general, measuring larger suites of emission lines will likely be critical to provide robust metallicity measurements, thereby improving constraints on the chemical evolution of galaxies. In line with this, Dopita et al. (2016; D16 hereafter) have proposed that the set of Hα, [N II] λ6584 and [S II] $\lambda\lambda$6717, 6731 rest-frame optical lines will prove convenient in this pursuit. The relatively narrow wavelength range covered by Hα, [N II], and [S II] lends itself kindly to high-redshift studies, as derived line ratios are almost independent of reddening and can typically be observed in one spectroscopic exposure. Additionally, provided the [N II]/[S II] line ratio can be adequately constrained, systematic variation in the derived metallicity caused by degeneracy with ionization parameter and interstellar medium (ISM) pressure is reduced, significantly improving uncertainty as compared to methods utilizing N2 in isolation.

In this contribution, we report *J*-band spectroscopy taken with *Keck*/MOSFIRE of a low-mass ($\log(M_*/M_\odot) \sim 8.6$) star-forming galaxy at $z \sim 1$ covering the [N II] and [S II] emission lines. With a moderate integration time, we achieve an improved constraint on metallicity using the D16 diagnostic compared to a diagnostic based on N2 alone (once systematic uncertainty is considered). In addition, the combination of ground and space-based spectroscopy covers a suite of line-flux measurements that is unique for a $z \sim 1$ dwarf galaxy, allowing us to test the consistency of a handful of metallicity diagnostics. Based on this finding, we suggest that targeted surveys utilizing existing cutting-edge instruments could leverage this diagnostic to place powerful constraints on the processes that govern the evolution of galaxies across cosmic time.

The paper is structured as follows. In Section 2 we provide details on the collection of the near-infrared spectral data. Section 3 describes our analysis of the data. Section 4 presents a brief discussion of the results before we sum up in Section 5. Throughout this letter we adopt the Planck Collaboration et al. (2016) cosmology: $\Omega_\Lambda = 0.692$, $\Omega_M = 0.308$, $\sigma_8 = 0.815$, and $H_0 = 67.8\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$. All magnitudes are quoted in the *AB* magnitude system (Oke & Gunn 1983). Unless otherwise stated, [N II] and [S II] refer to [N II] λ6584 and [S II] $\lambda\lambda$6717, 6731, respectively.

2. Data

MACS0744_667.0 is a star-forming galaxy at $z = 0.997$ with $m_{\mathrm{AB}} = 23.35 \pm 0.02$ in *J*-band (*Hubble Space Telescope* (*HST*)/WFC3 F125W) and $R_{\mathrm{eff}} = 2.49$ kpc, magnified 1.4× by cluster MACS0744 (lens modeling from Hoag et al. 2019). The source was selected from the *HST* Grism Lens-Amplified Survey from Space (GLASS; Treu et al. 2015) as an intermediate-redshift target for spectroscopic follow-up with MOSFIRE at Keck (program #Z045M, PI Trenti). Observations were carried out on 2016 March 20th under good seeing conditions (∼0.″4–0.″7 in *J* band), low atmospheric attenuation ($\Delta m < 0.1$), and minimal airmass (1.05–1.15) for a total of 8457 s, divided into individual exposures of 120 s each. An ABBA dither pattern with 3″ nodding along the slit was employed and observations started at 19:35HST after acquisition of a standard star during twilight. The MOSFIRE mask included two stars ($m_J = 16.1$ and $m_J = 16.3$) inside the GLASS *HST*/WFC3 field of view, that were used to verify source alignment.

The MOSFIRE data were reduced using the publicly available data reduction pipeline (DRP[8]). The DRP performs wavelength calibration, rectification, background subtraction, and skyline subtraction for each 2D slit in the multi-object slitmask. The resulting outputs of the DRP are individual 2D signal and noise spectra in electrons per second for each slit on the mask.

3. Analysis

3.1. 1D Spectrum Calibration and Construction

The sensitivity curve for our observations has been derived from a 1D *J*-band MOSFIRE spectrum of an A5-type reference star (observed with MOSFIRE on 2016 June 13th) obtained by integrating over the FWHM spatial extent of the 2D spectrum. We remove the intrinsic stellar spectrum shape by dividing the measured stellar spectrum by an A5-type reference spectrum from the CALSPEC Calibration Database.[9] Thus, the shape of this sensitivity curve gives the response versus wavelength of the setup, due to both the instrument itself and the atmosphere. A faint continuum detection for our object allows us to use its known *J*-band magnitude (from *HST* imaging) to derive a normalizing factor to calibrate the flux in physical units. Applying this sensitivity curve and normalization factor to the 1D signal extracted by integrating the 2D MOSFIRE spectrum along the 1.″12 spatial extent of the [S II] signal, we recover our 1D *J*-band spectrum of MACS0744_667.0, shown in Figure 1.

[8] https://www2.keck.hawaii.edu/inst/mosfire/drp.html
[9] http://www.stsci.edu/hst/observatory/crds/calspec.html

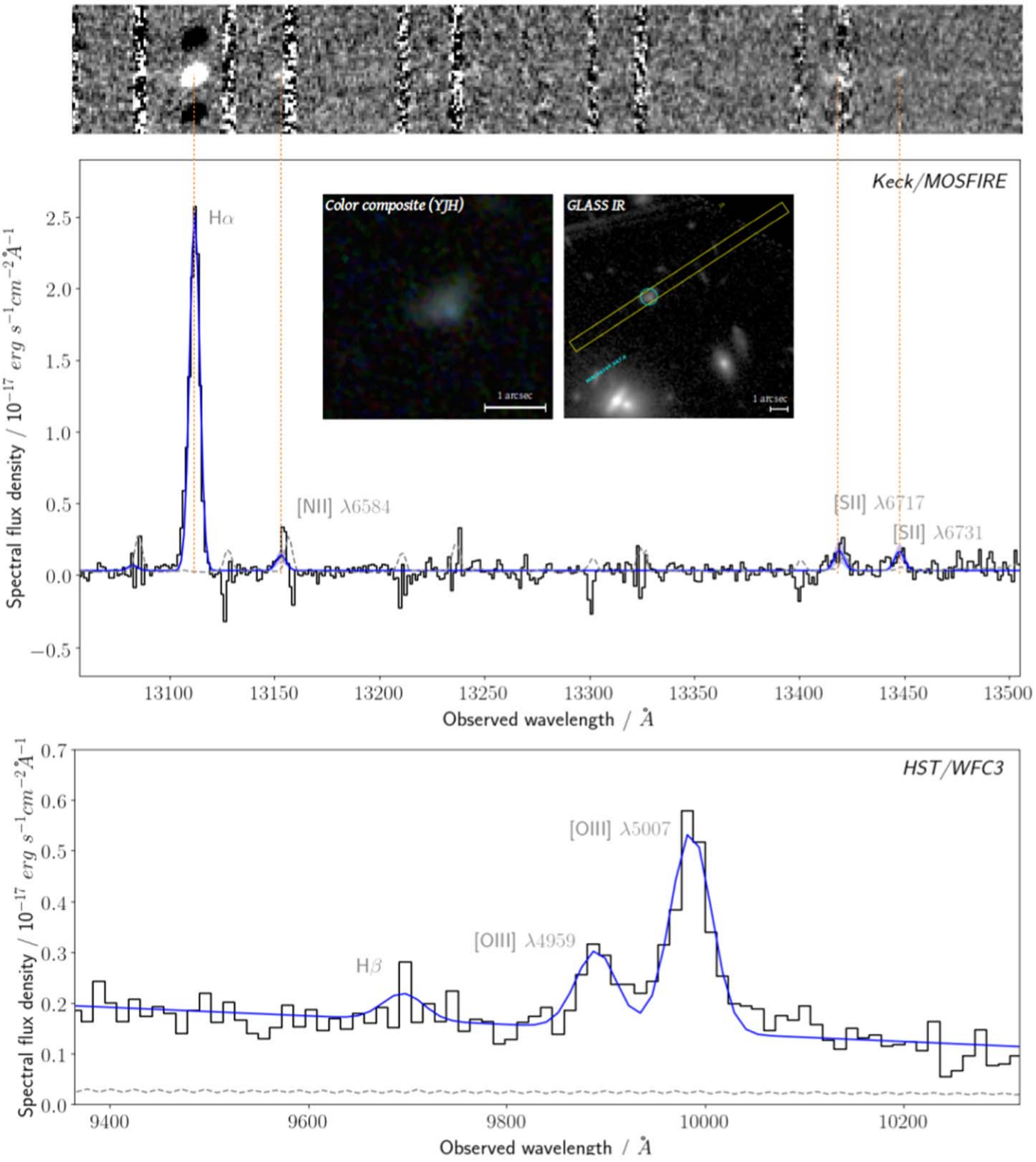

Figure 1. Top: zoom of on an area of interest in the 2D *J*-band spectrum from MOSFIRE. Panel 1: zoom of the red end of the MOSFIRE *J*-band spectrum of MACS0744_667.0. The black step-plot shows an integrated 1D MOSFIRE signal. The gray dashed line shows the 1σ uncertainty associated with each wavelength bin. The blue line shows our best-fit spectrum with the 2σ uncertainty depicted by the lighter blue shaded region. The inset is a color composite from CLASH *HST Y*-, *J*-, and *H*-band imaging as well as the GLASS alignment reference image showing the MOSFIRE mask slit placement. Panel 2: zoom of lines of interest from integrated 1D *HST*/WFC3 *G*102 grism spectrum of MACS0744_667.0 from the GLASS survey. The color-coding is the same as that for panel 1.

3.2. Stellar Mass

We derive the stellar mass of MACS0744_667.0 from its spectral energy distribution, following the methods described in our previous analyses of GLASS targets (e.g., Jones et al. 2015b; Hirtenstein et al. 2019). We use 16-band *HST* photometry spanning observed wavelengths 0.2–1.6 micron

Table 1
Fluxes of Prominent Spectral Lines and Derived Properties of MACS0744_667.0

Spectral Line[a]	Flux[b]
Hβ	2.72 ± 0.27
[O III] λ4959	7.93 ± 0.10
[O III] λ5007	20.51 ± 0.80
Hα	14.94 ± 0.11
[N II] λ6584	0.63 ± 0.137
[S II] λ6717	0.84 ± 0.228
[S II] λ6731	0.79 ± 0.161
Derived Properties	
$z_{\rm MOSFIRE}$	$0.997 \pm (3 \times 10^{-6})$
$z_{\rm GLASS}$	0.994 ± 0002
$\log(M_*/M_\odot)$	$8.62^{+0.10}_{-0.06}$
$\sigma_{\rm H\alpha}/({\rm km\ s^{-1}})$	24.1 ± 0.5
$12 + \log({\rm O/H})$	7.99 ± 0.13[c]
$n_e/({\rm cm^{-3}})$	$\lesssim$1542[d]

Notes.
[a] Hβ and [O III] lines obtained with *HST*/WFC3 G102 grism spectroscopy from the GLASS survey. Remaining lines are from *J*-band MOSFIRE spectroscopy.
[b] Fluxes are in units of 10^{-17} erg s^{-1} cm^{-2}.
[c] Determined with the D16 diagnostic. The quoted uncertainty does not include systematic effects. Refer to Sections 3.5 and 4 for more information.
[d] Refer to Section 3.6 for details.

from the Cluster Lensing and Supernova Survey with *Hubble* (CLASH; Postman et al. 2012). The contributions of strong emission lines [O III], Hβ, Hα+[N II], and [S II] given in Table 1 are subtracted from the broadband continuum fluxes. Emission-line-corrected photometry is then fit with the stellar population synthesis code FAST (Kriek et al. 2009). We adopt Bruzual & Charlot (2003) spectral templates with a Chabrier IMF, solar metallicity, Calzetti et al. (2000) dust attenuation curve, and an exponentially declining star-formation history. Our analysis is relatively insensitive to adopted stellar metallicity. Assuming a sub-solar $Z = 0.2Z_\odot$, comparable to the derived gas-phase value, changes the best-fit stellar mass by only 0.01 dex, which is negligible compared to the uncertainty. Because formal statistical photometric uncertainties typically underestimate the total error (e.g., Ilbert et al. 2006), we scale the uncertainties in flux by a multiplicative factor such that the best-fitting template has a reduced $\chi^2_\nu = 1$. This increases the typical photometric uncertainty from $\sim$0.03 to 0.08 mag. The resulting best-fit stellar mass is $\log(M_*/M_\odot) = 8.62^{+0.10}_{-0.06}$ after correcting for lensing magnification (MACS0744 cluster lens modeling obtained from Hoag et al. 2019). Note that the subtraction of strong emission lines from the broadband fluxes reduced the best-fit stellar mass by 0.07 dex.

3.3. Line Fitting

Our line-fitting procedure was run on a wavelength subset of the full *J*-band spectrum bounded by the bright sky lines at 13055 and 13505 Å. Given that the continuum in this region is only very tenuously detected, we assume it to be flat at a level taken as the median flux value of the fitting region. After removing this continuum, we obtain line fluxes by fitting a five-peaked Gaussian simultaneously to Hα, [N II] λ6548, [N II] λ6584, and [S II] $\lambda\lambda$6717, 6731 with a χ^2 minimization procedure. To minimize free parameters during fitting, we link all peak centroids and fit only for redshift and assume all peaks have equal line width. Additionally, the peak height of [N II] λ6548 is assumed to be one-third that of [N II] λ6584. Thus, we fit for a total of six parameters.

Given the faint [N II] and [S II] emission lines and their proximity to sky lines, we carried out a robust determination and characterization of the errors affecting these flux measurements. In particular, the red shoulders of the [N II] λ6548, [N II] λ6584, and [S II] λ6717 lines were subject to high rms error according to the MOSFIRE DRP due to possible skyline contamination.

We estimate our uncertainties with a so-called bootstrapping method by perturbing the one-dimensional calibrated spectrum at each wavelength by drawing from a normal distribution with a standard deviation equal to the rms error assigned to that wavelength by the MOSFIRE DRP. We perform our line-fitting procedure on 1000 realizations of these perturbed "synthetic" spectra. The mean and standard deviation on the resulting distribution of line fluxes can be adopted as the measured line flux and its 1σ uncertainty.

We find that when applied to the full fitting range, this bootstrapping method tends to overestimate the flux by around 2σ when compared to the standard chi-squared fit. However, if we fit again for our set of emission lines, discarding the potentially contaminated values on the red shoulders of the [N II] λ6548, [N II] λ6584, and [S II] λ6717 lines (λ discarded if $\lambda_{\rm centroid} < \lambda < \lambda_{\rm centroid} + 3 \times \sigma_{\rm H\alpha}$; less than 20 values discarded in total), both the bootstrapping and standard chi-squared agree within 1σ of the lower value from the original chi-squared fit using all of the values (as opposed to the overestimated bootstrapping mean value). Thus, we adopt the values and the uncertainty from the bootstrapping method as applied to this amended wavelength set without the discarded points. These line fluxes are given in Table 1.

Upon measuring line fluxes, we verify the flux calibration by comparing with measurements from GLASS *HST* data, noting that the Hα and [N II] emission lines are blended in the *HST* spectra. The fluxes are consistent to within $\sim$1.5σ and the line ratio [S II]/(Hα+[N II]) agrees to within <0.1 dex.

After subtracting instrument dispersion, estimated at 28.7 km s^{-1} for MOSFIRE *J*-band, we obtain a rest-frame Hα velocity dispersion of 26.7 ± 0.5 km s^{-1}. This includes a natural line width of $\sigma_0 = 3.2$ km s^{-1} for Hα, and thermal-broadening estimated conservatively as $\sigma_{\rm th} = 11 \pm 2$ km s^{-1} (corresponding to nebular temperature $T = (1\text{–}2) \times 10^4$ K; García-Díaz et al. 2008). Subtracting these effects in quadrature, the intrinsic velocity dispersion of the galaxy is 24.1 ± 0.9 km s^{-1}. Given the MOSFIRE slit setting (0.″7) and the effective radius of the galaxy (0.″3, derived from *HST* imaging), we do not expect significant systematic error from the good seeing conditions ($\sim$0.″4–0.″7) during the run as the source image was filling the slit. In addition, the Hα emission line width is resolved at high significance compared to sky lines, and the low derived dispersion supports a small dynamical mass. From the half-light radius measured with SourceExtractor (Bertin & Arnouts 1996), and following Erb et al. (2006b), we find a dynamical mass of $\log(M_{\rm dyn}/M_\odot) \approx 9.0$. These dynamical results support the SED-based stellar mass derived in Section 3.2, with $\log(M_*/M_{\rm dyn}) \approx -0.4$.

3.4. GLASS Line Fluxes

In addition to our MOSFIRE *J*-band observations, we obtained line fluxes for Hβ and [O III] from the GLASS slitless spectroscopic observations. The 1D grism spectra from GLASS are included in the high-level science products publicly released by the GLASS team, available from STScI/MAST.[10]

The line-fitting procedure follows a process similar to that outlined in Section 3.3. We fit over the observed wavelength range 9361 Å $\leqslant\lambda\leqslant$ 10313 Å, modeling the continuum as a best-fit linear function over this range. We then fit a three-peaked Gaussian profile to the G102 GLASS spectrum (which resolves the [O III] emission), minimizing free parameters by fitting for redshift, line width (assumed equal for all lines), and the areas of each peak.

The low wavelength resolution of the *HST*/WFC3 G102 grism creates difficulties when fitting a continuum, thus uncertainties in the grism line fluxes are likely dominated by uncertainties in the continuum. The uncertainties quoted in Table 1 are obtained by propagation of the 1σ values obtained for each fit parameter from the covariance matrix output by the line-fitting function.

In the context of the BPT diagram we find a very high log([O III]λ5007/Hβ) ratio, perhaps caused by continuum-fitting uncertainties. Although the measured position of MACS0744_667 on the BPT diagram is broadly consistent with high ionization $z\sim 2-3$ galaxies observed by Strom et al. (2018).

The 1D grism spectra from the GLASS data products were flux-calibrated independently of this analysis and direct comparison with MOSFIRE line fluxes derived in Section 3.3, e.g., the high apparent Balmer decrement measured (Hα/Hβ = 5.6), may not be reliable.

The redshift fit obtained from the GLASS data ($z_{\rm GLASS}$ = 0.994) is slightly offset from that of the MOSFIRE data ($z_{\rm MOSFIRE}$ = 0.997). Given the superior wavelength resolution of MOSFIRE, we take $z_{\rm MOSFIRE}$ to be the source redshift.

3.5. Metallicity

The suite of measured line fluxes available to us is quite unique for a galaxy at $z\sim 1$ with $\log(M_*/M_\odot)\leqslant 9.0$, affording us a range of available metallicity diagnostics. We derive metallicities from diagnostics employing the following line ratios: N2 = log([N II] λ6584/Hα), O3N2 = log (([O III] λ5007/Hβ)/([N II]λ6584/Hα)), and N2S2Hα = log([N II] λ6584/[S II] $\lambda\lambda$6717, 6731) + 0.265 × N2. N2 and O3N2 are translated into metallicities using calibrations from Pettini & Pagel (2004) based on a sample of H II regions with direct (T_e) metallicity measurements, while metallicity is inferred from N2S2 using the Dopita et al. (2016) diagnostic, based on theoretical models. These calibrations are as follows:

$$12+\log({\rm O/H})=8.90+0.57\times {\rm N2}, \quad (1)$$

$$12+\log({\rm O/H})=8.73-0.32\times {\rm O3N2}, \quad (2)$$

$$12+\log({\rm O/H})=8.77+{\rm N2S2H}\alpha. \quad (3)$$

Applying these diagnostics to our measured line ratios yields values of $Z_{\rm N2}=8.11\pm0.05$, $Z_{\rm O3N2}=8.01\pm0.03$ and $Z_{\rm D16}=7.99\pm0.13$, where $Z=12+\log({\rm O/H})$, as given in Table 2. Uncertainties quoted here are strictly measurement uncertainties; systematic uncertainties are discussed in Section 4.

[10] https://archive.stsci.edu/prepds/glass/

Table 2
Metallicity as Derived by Different Available Diagnostics

Line ratios[a]	Calibration reference	12 + log(O/H)
N2	Pettini & Pagel (2004)	8.11 ± 0.05
O3N2	Pettini & Pagel (2004)	8.01 ± 0.03
N2S2Hα	Dopita et al. (2016)	7.99 ± 0.13

Notes. Quoted uncertainties do not include systematic effects.
[a] Definitions of listed ratio names given in Section 3.5.

3.6. Electron Density

The ratio between the [S II] λ6717 and [S II]λ6731 in the [S II] doublet is the most widely used measure of electron density in H II regions. Our detection of this doublet allows us to put constraints on the electron density in this target. We calculate a ratio of [S II] λ6717/[S II]λ6731 = 1.06 ± 0.36. According to the calibration provided by Proxauf et al. (2014), this places a 1σ upper limit on the density of $n_e\leqslant 1542$ cm^{-3}.

4. Discussion

Strong-line methods are currently the only feasible route to metallicity studies with large samples at high redshifts, particularly for low-mass galaxies. In addition to measurement uncertainty, which is present at some level in any observation, strong-line measurements in particular suffer from systematic uncertainties caused by degeneracy between metallicity and other physical parameters (ionization parameter, N/O abundance ratio, etc.) on the line ratios being employed by the diagnostic. These uncertainties can arise from both stochastic variations of these physical parameters, as well as any systematic variations that may be present in the high-redshift universe. Understanding and minimizing the uncertainties associated with these methods is therefore a critical and open issue.

As a result of the observational challenges associated with assembling large suites of emission lines for low-mass galaxies, the [N II]/Hα diagnostic (N2; Pettini & Pagel 2004) is widely used in high-redshift studies as the required lines are relatively strong and are close in wavelength (e.g., Erb et al. 2006a; Wuyts et al. 2012; Yabe et al. 2014). By contrast, many widely used strong-line diagnostics, such as O3N2 (Pettini & Pagel 2004) or R_{23} (Zaritsky et al. 1994), can deliver lower systematic uncertainty by better accounting for degeneracy with other physical parameters, but are considerably more challenging to obtain at high redshifts due to the wavelength coverage required. However, the reliance of N2 on a single line ratio makes it vulnerable to systematic uncertainties; in particular those caused by degeneracy with ionization parameter and N/O abundance ratio.

4.1. Ionization Parameter Dependence

Strom et al. (2018) showed that ionization parameter is the physical parameter to which nebular spectra respond most sensitively. Accordingly, the dependence of the N2 ratio on ionization parameter introduces systematic uncertainty into metallicities derived from N2. The detection of the [S II] doublet in this object allows us to adopt the D16 approach of

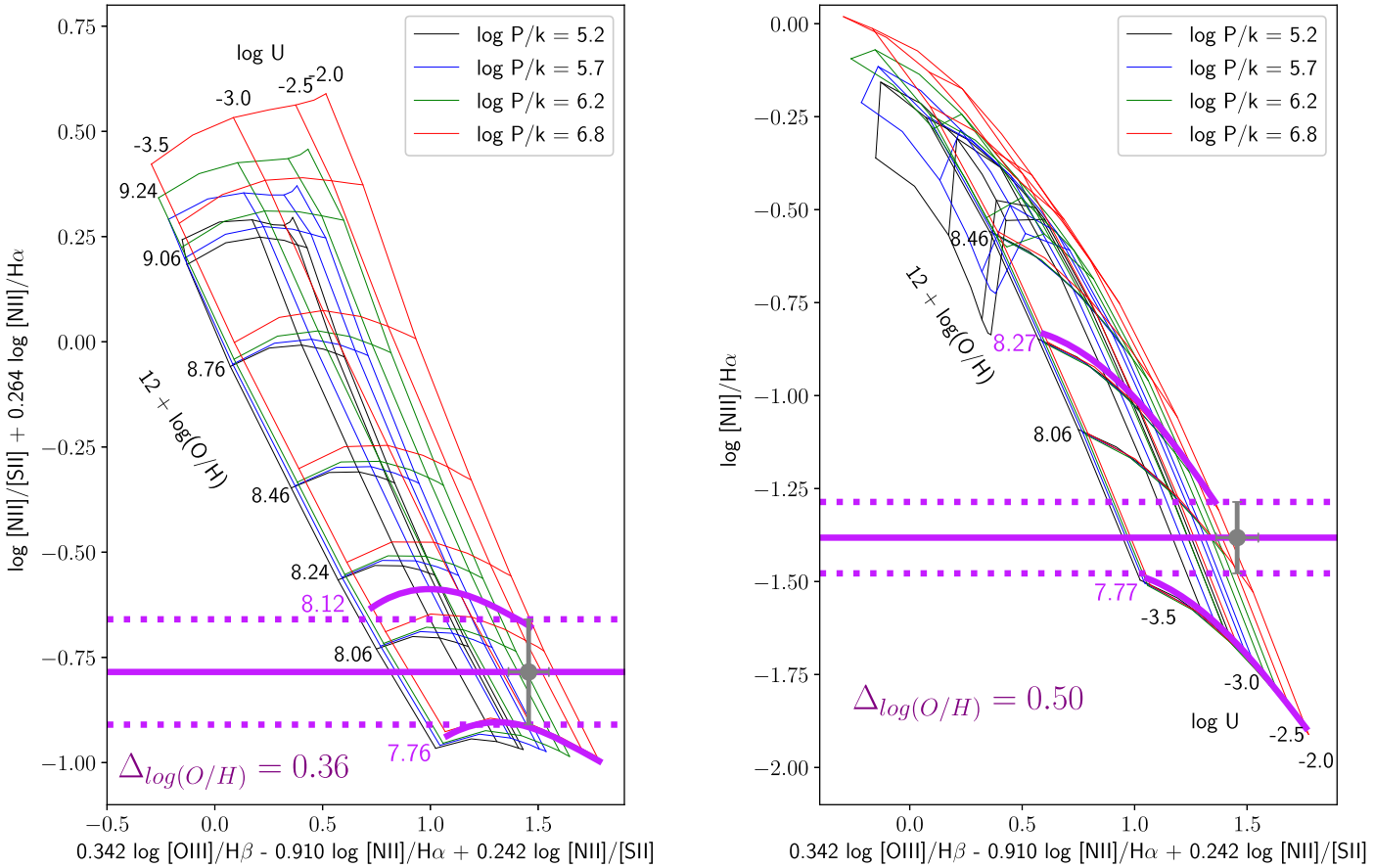

Figure 2. Two projections of theoretical grids (obtained from M. A. Dopita 2019, private communication) showing the improved systematic uncertainty afforded by the D16 diagnostic over the classic N2 diagnostic. The grid lines depict metallicity (roughly horizontal) and ionization parameter (steep diagonal) with different grid colors corresponding to different ISM pressures. The vertical axes in each panel are the set of line ratios required by each respective diagnostic (D16 on the left). The horizontal axis is a combination of line ratios that corresponds neatly with ionization parameter with the D16 y-axis at fixed metallicity and pressure; this is convenient for visual clarity. In both cases the gray point represents measurements reported here for MACS0744_667.0. The horizontal purple lines emphasize the 1σ measurement uncertainty obtained for each quantity given by the vertical axes. These uncertainty intervals clearly indicate the improvement of the D16 diagnostic over the simpler N2 diagnostic. Although the D16 line ratio value itself is not as well constrained (due to the lower S/N of the [N II]/[S II] ratio), the reduced systematic uncertainty allows for tighter constraints to be placed on the actual metallicity of this object (1σ upper and lower bounds depicted by purple contours). A first-order approximation with these grids suggests that knowledge of the N2 ratio alone is unable to constrain metallicity tighter than $8.27 \leqslant 12 + \log(\mathrm{O/H}) \leqslant 7.77$, while the D16 diagnostic constrains the metallicity to within $8.12 \leqslant 12 + \log(\mathrm{O/H}) \leqslant 7.76$ at the 1σ level.

incorporating [N II]/[S II] in addition to N2, which drastically reduces the degeneracy between metallicity, ionization parameter, and ISM pressure.

Grids obtained from M. A. Dopita (2019, private communication) plotted in Figure 2 clearly highlight the shortcomings of the classic N2 diagnostic. The grids show how the expected line ratios vary with ionization parameter and ISM pressure at fixed metallicity. We plot MACS0744_667.0 as reported in this letter onto these grids to illustrate the tighter constraint afforded using the D16 diagnostic. Although the line ratio itself has a larger uncertainty, the vastly reduced ionization dependence leads to an overall better metallicity measurement (constrained to a range of 0.36 dex for D16 see 0.50 for N2; 1σ upper and lower bounds indicated by the purple curves in Figure 2). Furthermore, unlike systematic uncertainty, measurement uncertainty reduces with higher signal-to-noise, meaning the prospects for placing tight constraints on the metallicities of low-mass galaxies at high redshift are greatly improved for D16 compared to when the N2 line ratio is used in isolation. This approach is made feasible in this low-mass and high-redshift regime by virtue of the spectral proximity of the [S II] doublet to the Hα and [N II]λ6584 lines, meaning they can be obtained without having to essentially double time-on-target requirements by observing with additional filters.

Figure 3 shows metallicity, including estimated systematic uncertainty due to ionization parameter (dotted error bars), plotted against the stellar mass (see Section 3.2) for this object and a few available in the literature in a similar redshift range ($z \sim 1$). The black triangles in Figure 3 show the $z \sim 0$ MZR derived from the SAMI survey (Sánchez et al. 2019). Given that chemical evolution of low-mass galaxies is expected to be more significant at later times (e.g., Henry et al. 2013b), samples of objects with stellar masses below $\log(M_*/M_\odot) < 9.0$ beyond redshift $z \gtrsim 1$ promise to provide valuable insights into evolutionary processes driving the galaxy population. In Figure 3 the limitations caused by systematic uncertainties can be seen. Despite the low measurement uncertainty for our N2 metallicity (left panel), the large systematic uncertainty means the resultant measurement provides little insight into distinguishing between cases where significant or very little evolution occurs from $z \sim 1$ to $z \sim 0$.

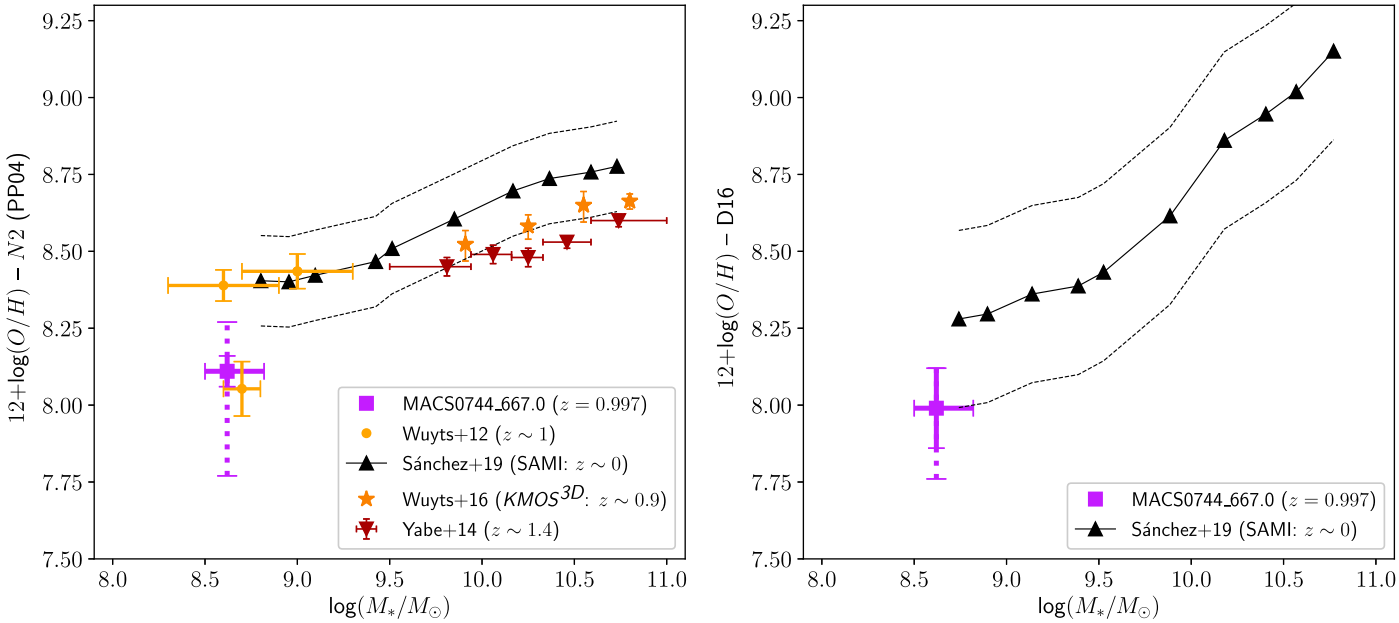

Figure 3. MACS0744_667.0, shown as the purple square, resides in a region of M_*–log(O/H) space that is sparsely populated by existing observations. The solid error bars show 1σ measurement uncertainty, while the dotted error bars depict our estimate of additional systematic uncertainty due to uncertainty in the ionization parameter (see Figure 2). Left panel: mass–metallicity relation with metallicities derived from the N2 diagnostic (Pettini & Pagel 2004). Orange stars are metallicities derived from N2 ratios of stacks of spectra at $0.6 < z < 1.1$ binned by mass from KMOS^{3D} (Wuyts et al. 2016). The orange circles are N2 metallicities from individual lensed objects at $z \sim 1$ from Wuyts et al. (2012). The red inverted triangles are N2 metallicities of $z \sim 1.4$ galaxies binned by mass from Yabe et al. (2014). The displayed error bars reflect only measurement uncertainty for these objects. The black triangles show the N2 gas-phase metallicities binned by mass measured at 1 R_e in SAMI galaxies at $z \sim 0$, with the dashed black lines indicating the average residual 1σ scatter after fitting a mass–metallicity relation (Sánchez et al. 2019). Some evolution is seen from $z \sim 1.4$ to $z \sim 0$ above log $(M_*/M_\odot) \geqslant 9.5$, however, neither large sample at high redshift is able to probe below log $(M_*/M_\odot) < 9.5$, where evolution is expected to be most significant. Despite the low measurement uncertainty associated with MACS0744_667.0, the large systematic uncertainty limits the degree to which insight can be gained from this measurement. Right panel: the mass–metallicity relation with metallicities derived from the D16 diagnostic. The black triangles show the D16 gas-phase metallicities measured at 1 R_e in SAMI galaxies at $z \sim 0$ (Sánchez et al. 2019). Here, MACS0744_667.0 has a larger measurement uncertainty; however, the additional systematic uncertainty is less problematic, suggesting larger samples using this diagnostic will be more effective in constraining redshift evolution of the mass–metallicity relation.

4.2. Nitrogen-to-oxygen Ratio Dependence

A remaining concern, however, is that the N2 ratio is primarily sensitive to the nitrogen abundance. Thus, oxygen abundances can only be inferred using some (implicitly or explicitly) assumed N/O ratio. Systematic uncertainty on measurements of oxygen abundances conducted in this way can therefore be introduced in two main ways. First, even assuming an appropriate N/H-to-O/H conversion can be applied, stochastic variations of log(N/O) at fixed log(O/H) add to the overall uncertainty of the final metallicity measurement. In a sample of objects with log(N/O) and 12+log(O/H) direct measurements (Berg et al. 2012; Pilyugin et al. 2012), we found that among the objects with $7.8 < 12 + \log(\mathrm{O/H}) < 8.2$, the standard deviation in log(N/O) was $\sigma_{\log(\mathrm{N/O})} = 0.13$ dex, comparable to line ratio measurement uncertainties. Although strong-line measurements at $z \sim 2.3$ by Strom et al. (2018) suggest this scatter could be as large as 0.8 dex in high-redshift galaxies. This stochastic variation limits the accuracy of individual metallicity measurements; larger samples are required to account for this effect. Second, some authors suggest that the N/O abundance ratio has undergone evolution with redshift (e.g., Masters et al. 2014). Although, Steidel et al. (2016) suggest that $z \sim 2$ galaxies on average lie within the same trend as local galaxies. This will critically affect metallicities derived using nitrogen lines. However, it is a difficult issue to address, as it requires large samples of high-quality spectra at high redshift. If, indeed, high-redshift galaxies do exhibit different N/O ratios to local galaxies the assumed N/O parameterization may be inappropriate. Modest samples of "direct" T_e measurements at $z > 1$ may help to further understand this.

4.3. Future Prospects

While better constraints on the mass–metallicity relation below log $(M_*/M_\odot) \leqslant 9.0$ at $z \gtrsim 1$ promise unique insights into the evolution of the galaxy population, progress has been limited by the associated observational challenges. Strong-line metallicity measurements are the only feasible approach to making progress in this area; however, the systematic uncertainties associated with the simple N2 diagnostic clearly limit its effectiveness in distinguishing between different evolutionary processes.

An additional source of uncertainty not discussed here is the contribution to the measured [SII]sHα ratio from diffuse ionized gas (DIG). The consequence of this is that global metallicities derived from N2S2Hα will be sensitive to variations in the fraction of DIG (f_{DIG}) in the galaxy population. As highlighted in Shapley et al. (2019), if high-redshift galaxies follow the same relation between f_{DIG} and Σ_{SFR} as local galaxies, N2S2Hα would vary systematically with redshift and thus not be appropriate for comparing high- and low-redshift samples. Thus, further observations are required to determine the degree to which variations in f_{DIG} would affect the systematic uncertainties induced in a sample at fixed redshift beyond $z \gtrsim 1$.

Approaches that better constrain degeneracy by including more emission-line ratios are certainly preferable yet challenging

in this regime, where long integration times are required to accurately measure even the strongest metal emission lines. Approaches that use photoionization models to simultaneously fit for all of these physical parameters, including metallicity (e.g., Pérez-Montero 2014; Blanc et al. 2015), appear to lend themselves naturally to this context; however, we did not include these in this discussion, as the low measured S/N for MACS0744_667.0 implies that both the specific details of the input model adopted and the parameter priors could have a substantial impact on the inference. While future facilities will certainly aid progress in this area, applying the D16 diagnostic to deep observations with existing instruments can improve the ionization dependence of existing N2-based constraints without requiring the factor of 2–3 increase in time-on-target associated with many other strong-line methods.

5. Conclusion

Extending constraints on the high redshift ($z \gtrsim 1$) MZR to masses below $\log(M_*/M_\odot) < 9.0$ promises powerful insights into the evolutionary processes that govern the galaxy population. Currently, strong-line methods are the only viable approach for expanding metallicity measurements to high-redshift dwarf galaxies.

However, particularly when the number of available emission lines is small, metallicity measurements made with strong-line methods may suffer from degeneracy with other physical parameters such as ionization parameter, chemical abundance ratios, and ISM pressure. Derived metallicities can be affected by either stochastic or systematic variations in these properties among the high-redshift galaxy population.

In this contribution we have presented MOSFIRE *J*-band spectroscopy of MACS0744_667.0, a low-mass ($\log(M_*/M_\odot) = 8.62^{+0.10}_{-0.06}$) star-forming galaxy at redshift $z = 0.997$ magnified $1.4\times$ by CLASH cluster MACS0744, in which we observe detection of [N II] λ6584 and [S II] $\lambda\lambda$6717, 6731 alongside strong Hα detection. Additionally, we derive Hβ, [O III] λ4959 and [O III] λ5007 line fluxes from *HST*/WFC3 G102 grism spectroscopy from the GLASS data release. Access to this set of emission lines is quite unique for a galaxy of this mass at this redshift.

We derive metallicity from N2 ($12 + \log(\mathrm{O/H}) = 8.11 \pm 0.05$ with statistical uncertainty; $12 + \log(\mathrm{O/H}) = 8.11^{+0.16}_{-0.34}$ including additional systematic uncertainty), as well as N2S2Hα ($12 + \log(\mathrm{O/H}) = 7.99 \pm 0.13$; $12 + \log(\mathrm{O/H}) = 7.99^{+0.13}_{-0.23}$). While the inclusion of the [N II]/[S II] ratio in N2S2Hα increases the measurement uncertainty, we find that even in this case where $\mathrm{S/N_{N2S2}}$ is small, the improved mitigation of the dependency on ionization parameter reduces the overall uncertainty on the metallicity measurement. We estimate that uncertainty due to the N/O abundance ratio is likely of comparable order to the measurement uncertainty of N2S2Hα. Large samples of high-quality spectra of high-redshift dwarf galaxies are needed to assess if there is a systematic variation of this abundance ratio at high redshift and at what level the stochastic variation impacts dwarf galaxy metallicities.

Further progress in extending the high-redshift MZR to dwarf galaxies requires careful consideration of the global trends among dwarf galaxies in all physical parameters, not just metallicity. Deeper surveys targeting low-mass objects at $z \gtrsim 1$ employing existing multiplexed NIR instruments (e.g., *Keck*/MOSFIRE or *VLT*/KMOS) will improve understanding of the stochastic variations in these properties among the high-redshift dwarf population, providing unique insights into the evolutionary processes that govern the galaxy population.

We are grateful to the late M. A. Dopita for providing us with the theoretical grids presented in Dopita et al. (2016) and would like to acknowledge his prolific and extensive contributions to the field spanning many decades. This research was supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. A.J.C. acknowledges support from an Australian Government Research Training Program (RTP) Scholarship.

Facility: Keck (MOSFIRE).

ORCID iDs

Alex J. Cameron https://orcid.org/0000-0002-0450-7306
Tucker Jones https://orcid.org/0000-0001-5860-3419
Tiantian Yuan https://orcid.org/0000-0002-9211-3277
Michele Trenti https://orcid.org/0000-0001-9391-305X
Stephanie Bernard https://orcid.org/0000-0003-0956-0728
Alaina Henry https://orcid.org/0000-0002-6586-4446
Austin Hoag https://orcid.org/0000-0001-8989-2567
Benedetta Vulcani https://orcid.org/0000-0003-0980-1499

References

Amorín, R., Sommariva, V., Castellano, M., et al. 2014, A&A, 568, L8
Andrews, B. H., & Martini, P. 2013, ApJ, 765, 140
Belli, S., Jones, T., Ellis, R. S., & Richard, J. 2013, ApJ, 772, 141
Berg, D. A., Skillman, E. D., Marble, A. R., et al. 2012, ApJ, 754, 98
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Blanc, G. A., Kewley, L., Vogt, F. P. A., & Dopita, M. A. 2015, ApJ, 798, 99
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Calabrò, A., Amorín, R., Fontana, A., et al. 2017, A&A, 601, A95
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Davé, R., Finlator, K., & Oppenheimer, B. D. 2012, MNRAS, 421, 98
Dopita, M. A., Kewley, L. J., Sutherland, R. S., & Nicholls, D. C. 2016, Ap&SS, 361, 61
Erb, D. K., Shapley, A. E., Pettini, M., et al. 2006a, ApJ, 644, 813
Erb, D. K., Steidel, C. C., Shapley, A. E., et al. 2006b, ApJ, 646, 107
García-Díaz, M. T., Henney, W. J., López, J. A., & Doi, T. 2008, RMxAA, 44, 181
Henry, A., Martin, C. L., Finlator, K., & Dressler, A. 2013a, ApJ, 769, 148
Henry, A., Scarlata, C., Domínguez, A., et al. 2013b, ApJL, 776, L27
Hirtenstein, J., Jones, T., Wang, X., et al. 2019, ApJ, 880, 54
Hoag, A., Bradač, M., Huang, K., et al. 2019, ApJ, 878, 12
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Jones, T., Martin, C., & Cooper, M. C. 2015a, ApJ, 813, 126
Jones, T., Wang, X., Schmidt, K. B., et al. 2015b, AJ, 149, 107
Kewley, L. J., & Dopita, M. A. 2002, ApJS, 142, 35
Kewley, L. J., & Ellison, S. L. 2008, ApJ, 681, 1183
Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221
Lequeux, J., Peimbert, M., Rayo, J. F., Serrano, A., & Torres-Peimbert, S. 1979, A&A, 80, 155
Ly, C., Malhotra, S., Malkan, M. A., et al. 2016a, ApJS, 226, 5
Ly, C., Malkan, M. A., Rigby, J. R., & Nagao, T. 2016b, ApJ, 828, 67
Maiolino, R., & Mannucci, F. 2019, A&A, 27, 3
Maiolino, R., Nagao, T., Grazian, A., et al. 2008, A&A, 488, 463
Maseda, M. V., van der Wel, A., Rix, H.-W., et al. 2014, ApJ, 791, 17
Masters, D., McCarthy, P., Siana, B., et al. 2014, ApJ, 785, 153
Morales-Luis, A. B., Pérez-Montero, E., Sánchez Almeida, J., & Muñoz-Tuñón, C. 2014, ApJ, 797, 81
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Pérez-Montero, E. 2014, MNRAS, 441, 2663
Pérez-Montero, E., & Contini, T. 2009, MNRAS, 398, 949
Pettini, M., & Pagel, B. E. J. 2004, MNRAS, 348, L59
Pilyugin, L. S., Grebel, E. K., & Mattsson, L. 2012, MNRAS, 424, 2316
Planck Collaboration, Adam, R., Ade, P. A. R., et al. 2016, A&A, 594, A1
Postman, M., Coe, D., Benítez, N., et al. 2012, ApJS, 199, 25

Proxauf, B., Öttl, S., & Kimeswenger, S. 2014, A&A, 561, A10
Sánchez, S. F., Barrera-Ballesteros, J. K., López-Cobá, C., et al. 2019, MNRAS, 484, 3042
Sanders, R. L., Shapley, A. E., Kriek, M., et al. 2015, ApJ, 799, 138
Savaglio, S., Glazebrook, K., Le Borgne, D., et al. 2005, ApJ, 635, 260
Shapley, A. E., Sanders, R. L., Shao, P., et al. 2019, ApJL, 881, L35
Skillman, E. D., Kennicutt, R. C., & Hodge, P. W. 1989, ApJ, 347, 875
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
Strom, A. L., Steidel, C. C., Rudie, G. C., Trainor, R. F., & Pettini, M. 2018, ApJ, 868, 117
Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898
Treu, T., Schmidt, K. B., Brammer, G. B., et al. 2015, ApJ, 812, 114
Vale Asari, N., Stasińska, G., Morisset, C., & Cid Fernandes, R. 2016, MNRAS, 460, 1739
Wuyts, E., Kurk, J., Förster Schreiber, N. M., et al. 2014, ApJL, 789, L40
Wuyts, E., Rigby, J. R., Sharon, K., & Gladders, M. D. 2012, ApJ, 755, 73
Wuyts, E., Wisnioski, E., Fossati, M., et al. 2016, ApJ, 827, 74
Yabe, K., Ohta, K., Iwamuro, F., et al. 2014, MNRAS, 437, 3647
Yuan, T.-T., Kewley, L. J., & Richard, J. 2013, ApJ, 763, 9
Zahid, H. J., Geller, M. J., Kewley, L. J., et al. 2013, ApJL, 771, L19
Zahid, H. J., Kashino, D., Silverman, J. D., et al. 2014, ApJ, 792, 75
Zahid, H. J., Kewley, L. J., & Bresolin, F. 2011, ApJ, 730, 137
Zaritsky, D., Kennicutt, R. C., Jr., & Huchra, J. P. 1994, ApJ, 420, 87