JADES
Probing interstellar medium conditions at z 5.5-9.5 with ultra-deep JWST/NIRSpec spectroscopy
Cameron, Alex J.; Saxena, Aayush; Bunker, Andrew J.; D'eugenio, Francesco; Carniani, Stefano; Maiolino, Roberto; Curtis-lake, Emma; Ferruit, Pierre; Jakobsen, Peter; Arribas, Santiago; Bonaventura, Nina; Charlot, Stephane; Chevallard, Jacopo; Curti, Mirko; Looser, Tobias J.; Maseda, Michael V.; Rawle, Tim; Rodríguez Del Pino, Bruno; Smit, Renske; Übler, Hannah; Willott, Chris; Witstok, Joris; Egami, Eiichi; Eisenstein, Daniel J.; Johnson, Benjamin D.; Hainline, Kevin; Rieke, Marcia; Robertson, Brant E.; Stark, Daniel P.; Tacchella, Sandro; Williams, Christina C.; Willmer, Christopher N. A.; Bhatawdekar, Rachana; Bowler, Rebecca; Boyett, Kristan; Circosta, Chiara; Helton, Jakob M.; Jones, Gareth C.; Kumari, Nimisha; Ji, Zhiyuan; Nelson, Erica; Parlanti, Eleonora; Sandles, Lester; Scholtz, Jan; Sun, Fengwu
Published in: Astronomy & Astrophysics
DOI: 10.1051/0004-6361/202346107
Publication date: 2023
Document version Publisher's PDF, also known as Version of record
Document license: CC BY
JADES: Probing interstellar medium conditions at $z \sim 5.5$–9.5 with ultra-deep JWST/NIRSpec spectroscopy


ABSTRACT

We present emission-line ratios from a sample of 27 Lyman-break galaxies from $z \sim 5.5$–9.5 with $M_{500} < -20.4$, measured from ultra-deep JWST/NIRSpec multi-object spectroscopy from the JWST Advanced Deep Extragalactic Survey (JADES). We used a combination of 28 deep PRISM/CLEAR and 7 deep G395M/F290LP observations to measure, or place strong constraints on, ratios of widely studied rest-frame optical emission lines including Hα, Hβ, [O ii], [Ne iii], [S iii], [O iii], [Si ii], [S ii], and [N ii] (e.g. Peimbert et al. 2017). Ratios of emission line fluxes are sensitive to abundances of metals and their ionisation states, but also the physical conditions of the emitting nebulae, such as temperature and density, which are in turn linked to the nature of the ionising sources powering the emission (e.g. Osterbrock & Ferland 2006). Emission-line ratios therefore have great diagnostic power for understanding properties of the interstellar medium (ISM) and stellar populations in galaxies (see Kewley et al. 2019 for a review).

However, observed emission lines in galaxies can arise from other phenomena, including active galactic nuclei (AGN; Kewley et al. 2006; Feltre et al. 2016) or shock-heated gas (Dopita & Sutherland 1996; Allen et al. 2008; Sutherland & Dopita 2017). Meanwhile, stellar populations other than the young O and B stars typically associated with HII regions, such as high-mass X-ray binaries (HMXBs; Sanchyana et al. 2020) or post-AGB stars (Byler et al. 2019), can power emission in photoionised regions of gas. Each of these heating sources yields characteristically different emission spectra. An observed global emission spectrum of a star-forming galaxy is typically a superposition of emission from many HII regions with different conditions as well as possible contributions from these additional sources, making the interpretation of emission spectra challenging.

Despite this complexity, galaxies have been observed to follow remarkably tight trends in emission line-ratio space. The so-called BPT and VO87 diagrams, which relate the ratio of [O iii]/Hβ to ratios of either [N ii]/Hα or [O iii]/[O ii], indicating significant diversity in the ISM conditions within the sample. We find no convincing detections of [N ii]/Hα in our sample, either in individual galaxies, or a stack of G395M/F290LP spectra. The emission-line ratios observed in our sample are generally consistent with galaxies with extremely high ionisation parameters (log $U \sim -1.5$), and a range of metallicities spanning from $-0.1 \times Z_\odot$ to higher than $-0.3 \times Z_\odot$, suggesting we are probing low-metallicity systems undergoing periods of rapid star formation, driving strong radiation fields. These results highlight the value of deep observations in constraining the properties of individual galaxies, and hence probing diversity within galaxy population.

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

1. Introduction

Emission lines can be some of the most prominent features of galaxy spectra. Regions of photoionised gas around young stars (HII regions) routinely exhibit bright emission from hydrogen and helium recombination, as well as collisionally excited emission from numerous metal ions, especially Oii, Neiii, Sii, and Nii (e.g. Peimbert et al. 2017). Ratios of emission line fluxes are sensitive to abundances of metals and their ionisation states, but also the physical conditions of the emitting nebulae, such as temperature and density, which are in turn linked to the nature of the ionising sources powering the emission (e.g. Osterbrock & Ferland 2006). Emission-line ratios therefore have great diagnostic power for understanding properties of the interstellar medium (ISM) and stellar populations in galaxies (see Kewley et al. 2019 for a review).

However, observed emission lines in galaxies can arise from other phenomena, including active galactic nuclei (AGN; Kewley et al. 2006; Feltre et al. 2016) or shock-heated gas (Dopita & Sutherland 1996; Allen et al. 2008; Sutherland & Dopita 2017). Meanwhile, stellar populations other than the young O and B stars typically associated with HII regions, such as high-mass X-ray binaries (HMXBs; Sanchyana et al. 2020) or post-AGB stars (Byler et al. 2019), can power emission in photoionised regions of gas. Each of these heating sources yields characteristically different emission spectra. An observed global emission spectrum of a star-forming galaxy is typically a superposition of emission from many HII regions with different conditions as well as possible contributions from these additional sources, making the interpretation of emission spectra challenging.

Despite this complexity, galaxies have been observed to follow remarkably tight trends in emission line-ratio space. The so-called BPT and VO87 diagrams, which relate the ratio of [O iii]/Hβ to ratios of either [N ii]/Hα or [O iii]/[O ii], indicating significant diversity in the ISM conditions within the sample. We find no convincing detections of [N ii]/Hα in our sample, either in individual galaxies, or a stack of G395M/F290LP spectra. The emission-line ratios observed in our sample are generally consistent with galaxies with extremely high ionisation parameters (log $U \sim -1.5$), and a range of metallicities spanning from $-0.1 \times Z_\odot$ to higher than $-0.3 \times Z_\odot$, suggesting we are probing low-metallicity systems undergoing periods of rapid star formation, driving strong radiation fields. These results highlight the value of deep observations in constraining the properties of individual galaxies, and hence probing diversity within galaxy population.

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

Received 8 February 2023 / Accepted 17 July 2023
have been observed to form a tight sequence in the [N ii]–BPT diagram, with a general anti-correlation between the two ratios (e.g. Kauffmann et al. 2003; Kewley et al. 2006). This tight correlation is the result of the fact that the two parameters that most strongly dictate the emission spectra of galaxies, metallicity, and ionisation parameter (U), are themselves generally correlated in galaxies (Dopita & Evans 1986; Carton et al. 2017). The position within this sequence is thus dominated by these two parameters, although variations in other properties including the N/O abundance ratio, electron density, and radiation hardness also contribute (e.g. Kewley et al. 2013a; Bian et al. 2016; Curti et al. 2022). These correlations are not unique to the BPT and VO87 diagrams, and indeed other combinations of emission lines, such as ([O II] + [O III])/Hβ versus [O III]/[O II] (‘R23–O32’), have been observed to exhibit similar trends (e.g. Veilleux & Osterbrock 1987; Kewley et al. 2006).

Existing studies of earlier cosmic epochs show that, compared to the loci inhabited by local galaxies in line-ratio space, z ∼ 2 galaxies are generally offset to higher excitation and higher ionisation (e.g. Strom et al. 2017; Runco et al. 2021). This is perhaps best exhibited by the widely studied offset in the [N ii]–BPT diagram (Kewley et al. 2013b; Steidel et al. 2014, 2016; Runco et al. 2022), where z ∼ 2 galaxies are observed with typically higher [N ii]/Hα at a fixed [O III]/Hβ. This is often explained as a combination of higher ionisation parameters, higher densities, and higher ionising spectra at high redshift (Brinchmann et al. 2008; Kewley et al. 2013a; Steidel et al. 2016; Sanders et al. 2016; Hirschmann et al. 2017). Many authors have suggested that super solar α/Fe abundance ratios at these earlier times play an important role by driving harder stellar ionisation fields (e.g. Strom et al. 2017; Topping et al. 2020a,a). However, the N/O abundance ratio is clearly another important parameter to consider, with some studies suggesting this offset can be explained as the result of higher N/O at fixed [O III]/Hβ (Masters et al. 2016).

Despite the wealth of rest-frame optical studies highlighting the different ISM conditions in galaxies at z ∼ 2, emission-line studies of galaxies at z ≥ 6 were, until recently, very limited. Broad-band excesses had already suggested an abundance of galaxies at this epoch with very high equivalent width emission from [O III] + Hβ (e.g. Endsley et al. 2021, 2023; Fujimoto et al. 2023). However, constraints on their ISM conditions were limited to studies of far-infrared lines, such as [O III]88 μm and [C II]158 μm from the Atacama Large Millimeter Array (ALMA), which suggested high ionisation parameters and enhanced α/Fe abundance ratios (e.g. Hashimoto et al. 2019; Carniani et al. 2020; Harikane et al. 2020; Fujimoto et al. 2022; Wiitsok et al. 2022).

The recent commissioning of JWST is transforming our ability to probe ISM physics in the early Universe. Already, emission-line measurements from medium-depth observations from the initial months of JWST operations have observed z ≥ 6 galaxies to exhibit generally low metallicity, high ionisation, high excitation, and high temperature (Curti et al. 2023; Katz et al. 2023; Sun et al. 2023; Rhoads et al. 2023; Tacchella et al. 2023; Trump et al. 2023; Langeroodi et al. 2022; Williams et al. 2023; Heintz et al. 2023; Matthee et al. 2023; Fujimoto et al. 2023; Mascia et al. 2023; Sanders et al. 2023; Nakajima et al. 2023; Tang et al. 2023). Furthermore, the observed emission spectra of some of these targets have been found to be consistent with the presence of AGN activity (Brinchmann 2023), while Katz et al. (2023) suggested that the presence of high-mass X-ray binaries could help to explain the observed emission-line ratios.

Although much insight has already been gained from these early observations, they have largely been based on relatively shallow spectroscopy. Here we leverage data taken from the ‘Deep’ spectroscopic tier of the JWST Advanced Deep Extragalactic Survey (JADES), the deepest spectroscopic observations yet taken with NIRSpec, to provide a more detailed look at the emission-line ratios of galaxies at z ∼ 5.5–9.5. These multi-object spectroscopy (MOS) observations reached exposure times of up to 28 h in the PRISM/CLEAR (R ∼ 100) and up to 7 h in each of the three medium resolution gratings (R ∼ 1000), providing unprecedented new insights into the ISM properties of galaxies within the first Gyr of the Universe’s history. The structure of this paper is as follows. In Sect. 2 we describe the observations, data reduction and measurement of emission line fluxes. In Sect. 3 we present our emission-line ratios on various diagnostic diagrams and provide comparisons with literature samples. Section 4 then presents a discussion of the implications of these observations. We summarise in Sect. 5. Throughout this paper we adopt the Planck Collaboration VI (2020) cosmology: ΩΛ = 0.6847, Ωm = 0.3153, H0 = 67.36 km s⁻¹ Mpc⁻¹. All magnitudes are quoted in the AB magnitude system (Oke & Gunn 1983).

2. Data and analysis

2.1. Observations

The data presented in this paper were obtained via multi-object spectroscopic observations from JWST/NIRSpec using the micro-shutter assembly (MSA; Jakobsen et al. 2022; Ferruit et al. 2022). Observations were carried out as part of JADES in three visits between 21 and 25 Oct. 2022 (Programme ID: 1210; PI: N. Lutzgendorf) in the Great Observatories Origins Deep Survey South (GOODS-S) legacy field in a region overlapping with the Hubble Ultra Deep Field (HUDF; Koekemoer et al. 2013). Each visit consisted of 33,613 s integration in the PRISM/CLEAR low-resolution setting and 8403 s integration in each of G140M/F070LP, G235M/F170LP, G395M/F290LP, and G395H/F290LP filter and grating settings. Across three visits, this totals 28 h of integration in the PRISM, which provides continuous spectral coverage from 0.6 to 5.3 μm at R ∼ 30–300, and ∼7 h in each of the medium resolution gratings, which combine to provide R ∼ 700–1300 across the full spectral range of NIRSpec, plus 7 h in the high-resolution grating which provides R ∼ 2700 from ∼2.8 to 5.1 μm. In this paper we present only measurements from the PRISM/CLEAR and G395M/F290LP observations. Results obtained from other configurations will be presented in forthcoming work.

Observations within each visit were performed as a 3-shutter nod. The central pointing of each visit was dithered by <1 arcsec such that common targets were observed in different shutters and different detector real-estate. Thus, each visit had a unique MSA configuration, although target allocation (performed with the eMPT) was optimised for maximising target commonality between all three dither positions. A total of 253 unique targets were observed in the PRISM configuration with the three dithers featuring 145, 155, and 149 targets respectively. 67 targets were triply exposed in the PRISM and reached full

A115, page 2 of 19

http://github.com/esdc-esac-esa-int/eMPT_v1
depth, while 62 targets featured in only two dithers and a further 124 received only one-third total depth. The hierarchical allocation of shutters means that the double- and triple-exposed targets are biased towards the highest priority (and hence highest redshift) targets. Indeed, of the 27 galaxies presented in this work, 18 received triple coverage, six received double, and three received only single coverage. Details of the target selection scheme are presented in Bunker et al. (2023).

The low dispersion associated with the PRISM mode means that all targets are observed with non-overlapping spectra. However, in the higher resolution modes, individual spectra are dispersed over a much larger extent of detector real-estate. To minimise the possibility of contaminating emission, in our grating observations we isolate our highest priority targets by closing the shutters of low-priority targets on the same row (i.e. targets that could cause overlapping spectra). Thus, for our grating spectra we observe only 198 unique targets (119, 121, and 111 in each dither). However, the high-priority \( z \sim 5.5-9.5 \) galaxies presented in this work are almost unaffected by this measure with only two exceptions. One target (JADES-GS+53.11351–27.77284) that was triply exposed in PRISM was only singly exposed in the gratings, while one target (JADES-GS+53.17582–27.77446) was not observed in the medium resolution setting at all. The total integration times obtained on each target are summarised in Table A.1.

### 2.2. Data reduction

These observations were processed by adopting algorithms developed by the ESA NIRSpec Science Operations Team (SOT) and the NIRSpec GTO Team. Details of the data-processing workflow will be presented in a forthcoming NIRSpec GTO collaboration paper. As this work uses the observations carried out with only two spectral configurations of the program, we describe the main steps for the PRISM/CLEAR and G395M/F290LP filter and grating settings. Upon retrieving the level-1a data from the Mikulski Archive for Space Telescopes (MAST), we estimated the count rate per pixel by using the unsaturated groups in the ramp and removing jumps due to cosmic rays identified by estimating the slope of the individual ramps. During this first stage, we also performed the master bias and dark subtraction, corrected snowball artefacts, and flagged saturated pixels. We then performed the pixel-by-pixel background subtraction by combining the three nod exposures of each point. We note that for some targets we exclude one of the 3-shutter nod in the background subtraction stage due to the serendipitous presence of a source contaminating the open background shutter. We then created two-dimensional (2D) cutouts of each 3-shutter slit and performed the flat-field, spectrograph optics, and disperser corrections. We then ran the absolute calibration and corrected the 2D spectra for path-losses. Path-loss corrections are dependent on the morphology of the target and its relative position within the shutter. In this paper, we apply the path-loss corrections assuming a point-like source at the location of the target centroid. The high-redshift targets presented in this paper are generally smaller, or comparable in size, to the angular resolution of the telescope at the observed wavelength of emission lines used in this study (\( \lambda \geq 3 \mu m \)). We estimate that for the sample presented here, even the ratio with the longest wavelength baseline (\( [O\text{III}] \lambda 5007/[O\text{II}] \lambda 3726, 3729 \)) would be affected only at the 1% (0.01 dex) level by adopting an extended morphology based on fitting to broad-band imaging. Since it is not a priori known how closely the spatial distribution of emission lines trace the starlight observed in broad-band imaging, we decide to adopt the point-like assumption, but acknowledge that this effect introduces some systematic uncertainty to our measurements.

We rectified and interpolated the 2D continuum map onto a regular grid for the G395M/F290LP observation, and an irregular grid for the PRISM/CLEAR to avoid an oversampling of the line spread function at short wavelengths. Finally, the 1D spectra were extracted from the 2D map adopting a box-car aperture as large as the shutter size, centred on the relative position of the target in the shutter. For each target, we combined all 1D spectra and removed any bad pixels with sigma-clipping. Cutouts of five example representative 1D spectra are shown in Fig. 1.

As a verification of our flux calibration, we compared the NIRSpec spectrophotometry to broad-band photometry from NIRCam and found the differences to be on average less than 10% even before considering the effects of slit-losses due to the MSA. Details of this comparison are presented in Appendix C.

### 2.3. Emission line ratio sample

Although 253 unique targets were observed in these observations, this study focuses only on a rest-frame ultraviolet (UV) selected sample of confirmed \( z > 5.5 \) galaxies with robustly detected emission lines. The criteria used to select targets for this programme are described in detail in Bunker et al. (2023). We describe here only the aspects most relevant to this work.

The highest priorities in target assignment were reserved for galaxies at the highest redshifts, with a preference for the UV-bright galaxies. Candidate \( z > 5.5 \) galaxies were primarily drawn from a compilation of existing literature catalogues based on Hubble Space Telescope (HST) broad-band imaging (e.g. Bunker et al. 2004; Bouwens et al. 2015, 2021; Finkelstein et al. 2015; Harikane et al. 2016), with a number of the highest redshift galaxies presented here having been identified with HST (e.g. Bouwens et al. 2011, 2016; Ellis et al. 2013; McLure et al. 2013; Oesch et al. 2013). In fact, only five galaxies presented in this work were newly identified from JWST/NIRCam imaging (see Bunker et al. 2023 for details). In total, 47 galaxies were assigned MSA shutters as \( z > 5.7 \) candidates out of 299 possible such candidates within the MSA footprint. Of these, 28 passed redshift-dependent rest-UV cuts to place them in our highest priority classes \( PC = 1 \) and \( PC = 2 \) (\( PC = 2 \) and three contained only five targets between them, none of which featured here; see Bunker et al. 2023 for class descriptions). For this analysis, we consider only galaxies with robust spectroscopically confirmed redshifts in which \( H\beta \) is detected with \( S/N > 5 \). Of the 28 rest-UV-selected targets in \( PC = 1 \) and \( PC = 2 \), 18 met these criteria.

We also include six (one) galaxies selected as \( z > 5.7 \) candidates that had failed the rest-UV selection cut, but were included in \( PC = 6.1 \) (\( PC = 6.2 \)) as a result of meeting the magnitude criterion of \( m_{F160W} < 29 \) \( m_{F160W} > 29 \) and turned out to match our \( H\beta \) \( S/N \) criterion for this paper. Finally, three galaxies selected in \( PC = 7.5 \) (\( 4.5 < z < 5.7 \), \( m_{F160W} < 29 \)) had \( z_{\text{spec}} > 5.5 \) and \( S/N_{H\beta} > 5 \), and we included these in our sample as well. One galaxy exhibited a broad-component under \( H\alpha \) in \( R \sim 1000 \) spectroscopy, suggestive of nuclear activity, and was discarded so as to avoid contamination from emission from AGN. Thus, our final sample included 27 star-forming galaxies, drawn from a largely rest-UV-selected sample.

This work focuses predominantly on the low-resolution PRISM/CLEAR spectra (‘PRISM’ hereafter), since these are significantly deeper than our medium-resolution grating spectra. The spectral resolution of the PRISM varies significantly with wavelength from \( R \sim 30 \) around 1.2 \( \mu m \) up to \( R \sim 300 \) at the long
wavelength cutoff around 5 μm (e.g., Fig. 5 in Jakobsen et al. 2022). By limiting ourselves to rest-frame optical emission lines at $z \geq 5.5$, the measurements presented in this paper are made at observed wavelengths of $\lambda_{\text{obs}} \geq 2.4 \mu$m where the resolution is $R \geq 100$, and improving steeply as a function of wavelength. Thus, in all cases from the PRISM, the [O III] $\lambda\lambda 4959, 5007$ doublet is completely resolved from Hβ. The doublet itself is always at least partially resolved, and is completely resolved in our highest redshift targets (Fig. 1). [O II] $\lambda\lambda 3726, 3729$ and [S II] $\lambda\lambda 6716, 6731$ are never resolved into two components, but the latter can easily be fully resolved from the Hα + [N II] complex over this redshift range. However, Hα and [N II] $\lambda 6583$ require resolutions of $R \sim 500$ to be fully resolved and thus are significantly blended across the full redshift range. Constraints on this ratio are instead derived from our higher resolution (but shallower) G395M/F290LP grating spectra.

Figure 2 shows the redshift histogram of galaxies in this sample. We divide galaxies into two subsamples, based on the observability of Hα. The purple ‘z ≤ 6’ sample spans over 5.507 ≤ $z \leq 6.931$ ($z_{\text{median}} = 5.943$) and includes 21 galaxies for which Hα is still within the wavelength coverage of NIRSpec. There are an additional six galaxies in our ‘z ≤ 8’ sample, spanning over 7.05 ≤ $z \leq 9.45$ ($z_{\text{median}} = 7.629$), for which we study only emission lines from [O III] $\lambda\lambda 3726, 3729$ to [O III] $\lambda 5007$. Since the G395M/F290LP spectra are dispersed across a larger area of detector real estate, they do not provide continuous spectral coverage over the nominal spectral range. For two out of 21 galaxies in the $z \sim 6$ sample, we do not have [N II]/Hα constraints because these lines fall into the detector gap of NIRSpec. We additionally have one galaxy (JADES-GS+53.17582–27.77446) which was only observed in the PRISM mode (see Sect. 2.1). For one galaxy in the $z \sim 6$ sample, the [S II] doublet falls beyond the observed spectral range.

2.4. Spectral analysis and line flux measurements

2.4.1. Calculation of $M_{1500}$

The continuum is generally well detected in the rest-frame ultraviolet (UV) across our sample. In all except our faintest target, the signal-to-noise on the continuum reaches $S/N > 3$ for individual spectral pixels longwards of the Lyman break, and in six spectra we reach in excess of $S/N \geq 10$ per pixel in the rest-frame UV. For our faintest target, JADES-GS+53.16746–27.77201, we do observe faint continuum in the rest-UV, but the signal-to-noise per pixel is only $S/N \approx 1.8$. We derive rest-frame

\[ \text{Continuum flux} = S \times N \]

where $S$ is the signal and $N$ is the noise. The flux is then converted to a luminosity using

\[ L = 4\pi d_L^2 F \]

where $d_L$ is the luminosity distance and $F$ is the flux. The resulting values are corrected for Galactic extinction.

\[ A_{1500} = 2.5 \log_{10} \left( \frac{L_{1500}}{L_{1500,\odot}} \right) \]

Finally, the far-infrared luminosity $L_{1500}$ is corrected for dust absorption using the relation

\[ L_{1500} = L_{1500,\text{cor}} \times (1 + z)^2 \]

where $L_{1500,\text{cor}}$ is the corrected luminosity.
was manually set to 0. These templates are complemented by a 5th-degree multiplicative Legendre polynomial, to take into account systematic differences between the SSPs and the data (e.g. dust, mismatch between the SSP models and high-redshift stellar populations, and residual flux calibration problems). The emission lines are modelled as pixel-integrated Gaussians, again matching the observed spectral resolution. To reduce the number of degrees of freedom, we divide the emission lines into four kinematic groups, constrained to have the same redshift and intrinsic broadening. These groups are: UV lines (blueswards of 3000 Å in the rest-frame), the Balmer series of Hydrogen, non-Hydrogen optical lines (blueswards of 9000 Å in the rest-frame), and near-infrared (NIR) lines. The stellar continuum has the same kinematics as the Balmer lines. Furthermore, we tie together doublets that have fixed ratios, and constrain variable ratio doublets to their physical ranges. In particular, this study focuses on the following lines of interest: [O II]λλ3726, 3729, [Ne III]λ3869, Hγ, Hβ, [O III]λλ4959, [O III]λ5007, [O I]λ6300, Hα, [N II]λλ6583, [S II]λλ6716, 6731. We note that, at the resolution of the PRISM/CLEAR mode, the two components of the [O II]λ3726, 3729 doublet are significantly blended and so are fit as a single Gaussian component. We visually inspect these fits to the 1D spectra to ensure they are reliable. For the faintest lines, we also individually inspect the 2D spectrum for each target and confirm that the emission line can be visually identified along the same trace to confirm the measured flux is not affected by uncorrected artefacts. Where a line is not detected, we derive an upper limit directly from the noise spectrum. The 1σ upper limit is obtained by integrating the variance spectrum over an interval spanning 1.4 × σ_{lim} either side of the expected centroid of the line (where σ_{lim} is the Gaussian line width), and multiplying this by the spectral extent of a pixel.

We note that, at the spectral resolution of our PRISM observations, [Ne III]λ3869 is partially blended with the HeIλ3889 line. From grating spectra, we expect the [Ne III] line to be brighter than HeI, typically by at least a factor of two, suggesting that performing a two-component fit can recover the [Ne III] flux (e.g. Appendix C.3 in Cameron et al. 2021). The fitting procedure outlined above simultaneously fits for both lines in this complex, and we confirm from visual inspection that these fits model the complex well in cases with a high signal-to-noise ratio. In this paper, the only ratio for which we consider the [Ne III] line is the [Ne III]λ3869/[O II]λλ3726, 3729 ratio. This ratio has a very short wavelength baseline and hence the uncertainties introduced by partial blending are somewhat offset by having almost no dependence on wavelength dependent corrections or calibrations. Nonetheless, we note that that the [Ne III]-based measurements presented in this work should be treated with caution, and that any inference requiring high precision [Ne III]-based measurements should be performed with medium- to high-resolution spectra.

2.4.2. PRISM/CLEAR emission line measurements

Emission line measurements and continuum modelling are performed simultaneously using the penalised pixel fitting algorithm, PPF/SF (Cappellari 2017, 2023). PPF/SF models the continuum as a linear superposition of simple stellar-population (SSP) spectra, using non-negative weights and matching the spectral resolution of the observed spectrum. As input, we used the high-resolution (R = 10 000) SSP library combining isochrones from MESA Isochrones and Stellar Tracks (MIST; Choi et al. 2016) and the C3K theoretical atmospheres (Conroy et al. 2019). The flux bluewards of the Lyman break UV absolute magnitudes (M_{1500}) for our sample directly from the spectra. We shift the observed spectra to the rest-frame using the spectroscopically confirmed redshift and integrate across a synthetic 200 Å wide top-hat filter centred on 1500 Å. This integrated flux is then converted into an absolute magnitude. Derived M_{1500} values are reported in Table A.1. Aside from the faintest target (JADES-GS+53.16746–27.77201 with M_{1500} = −17.01 ± 0.22), all other targets have magnitudes in the range −17.5 ≤ M_{1500} ≤ −20.4.

2.4.3. Constraints on [N II]/Hα from G395M/F290LP

As discussed in Sect. 2.3, the only line ratio for which we consider the data from our higher spectral resolution, but much shallower, G395M/F290LP spectra is [N II]λ6583/Hα. These lines are sufficiently close in wavelength that they remain at least partially blended in PRISM/CLEAR spectra across the entire redshift range considered here. However, in the R ~ 1000 G395M/F290LP spectra, this complex is always completely resolved.

We performed a visual inspection simultaneously for the PRISM and G395M/F290LP 1D and 2D spectra of each target.
to identify the possible presence of \([\text{N}\,\text{II}]\). In fact, we find no convincing evidence for the detection of \([\text{N}\,\text{II}]\) in any of our 18 galaxies, either from the resolved \(R \sim 1000\) spectra, or by way of the appearance of a ‘red wing’ in the \(H\alpha\) profile in the PRISM spectra. We derive upper-limits on the \([\text{N}\,\text{II}]/H\alpha\) ratio in the same way as described above, fitting the \(H\alpha\) line with pPXF, and integrating the variance spectrum across the expected spectral extent of \([\text{N}\,\text{II}]\), based on the redshift and line-width measured for \(H\alpha\).

### 2.4.4. Emission line ratio measurements

In this paper we focus on the following key diagnostic line ratios:

\[
\begin{align*}
N2 &= \log \left( [\text{N}\,\text{II}] / \lambda 6583/H\alpha \right), \\
S2 &= \log \left( ([\text{S}\,\text{II}] \lambda 6716, 6731 / H\alpha) \right), \\
O1 &= \log \left( ([O]\,I] \lambda 6300/H\alpha) \right), \\
R2 &= \log \left( ([\text{O}\,\text{II}] \lambda 3726, 3729/H\beta) \right), \\
R3 &= \log \left( ([O]\,I] \lambda 5007/H\beta) \right), \\
R23 &= \log \left( ([O]\,I] \lambda 4959, 5007 + [O]\,I] \lambda 3726, 3729 / H\beta) \right), \\
O32 &= \log \left( ([O]\,I] \lambda 5007/[O]\,I] \lambda 3726, 3729) \right), \\
Ne3O2 &= \log \left( ([Ne]\,III] \lambda 3869/[O]\,I] \lambda 3726, 3729) \right).
\end{align*}
\]

Several of these line ratios, such as \(N2\), \(S2\) and \(R3\), are calculated across short wavelength baselines and are thus largely insensitive to dust attenuation or wavelength-dependent calibration issues. However, to accommodate the analysis of the longer baseline ratios, we correct our emission line fluxes and upper limits for dust extinction according to the Small Magellanic Cloud (SMC) dust law from Gordon et al. (2003). We calculate the \(E(B-V)\) from spectral energy distribution (SED) fitting the continuum by masking the regions of bright emission. For the \(z \sim 6\) sample, we correct the measured emission-line ratios for the effects of dust-reddening in the same way as for the individual measurements, described in Sect. 2.4.4, using the measured \(H\alpha/H\beta\) ratio to derive \(E(B-V)\) for the \(z \sim 6\) sample and the \(H\gamma/H\beta\) ratio for the \(z \sim 8\) sample.

We repeat the stacking process and line measurements for each sample where each spectrum is instead normalised by the UV continuum flux (taken as the median flux between 1400 and 1600 Å in the rest-frame spectrum). Finally, we repeat these measurements removing the target with the highest average noise level. Although the ratios measurement from these alternative stacks do not vary so much as to change any of the key findings in this paper, the change is larger than the statistical uncertainties on the measurements. For the final dust-corrected emission-line ratios reported in Table B.2, we adopt the values measured from the Balmer-emission-normalised stacks. However, rather than adopting the statistical uncertainty, we report the systematic uncertainty, which we take as the maximum difference between the values derived from any two of the stacking regimes described above.

### 3. Diagnostic diagrams

#### 3.1. Comparison samples

In this section, we compare the emission-line ratios measured for our JADES \(z \sim 6\) and \(z \sim 8\) samples to various literature samples in several diagnostic diagrams. To form our baseline \(z \sim 0\) comparison sample, we use 0.03 < \(z < 0.1\) galaxies from SDSS MPA-JHU catalogues4 (Aihara et al. 2011) to generate 2D probability density functions (PDFs). We consider only galaxies with spectra flagged as ‘reliable’ in the public release catalogue, and select for only galaxies with prominent emission lines by making a signal-to-noise cut of \(S/N > 40\) on \(H\alpha\). For diagnostic diagrams that include the \([Ne]\,III] \lambda 3869\) line, we additionally apply a cut of \(S/N_{3869} > 5\). In the \(z \sim 8\), we select for only galaxies with prominent emission lines by making a signal-to-noise cut of \(S/N > 5\) on each emission line and plot only detections (i.e. we do not show any limits). We additionally select only galaxies with a sky line flag with <0.2 for all lines considered, as recommended by the catalogue release. For lines that are flagged as being near the edge of the spectrum, we also select only those flagged as having reliable flux. For diagrams involving \(N2\), \(S2\) or \(O1\), the MOSDEF sample spans from \(1.2 < z < 2.6\) with median \(z = 2.1\), while for diagrams involving only the bluer lines, it probes slightly higher redshifts: typically 2.0 < \(z < 3.6\) with median of \(z \sim 2.3\).

A number of line ratio measurements have already been reported from NIRSpec at \(z \sim 6\). The JWST Early Release Observations (EROS; Pontoppidan et al. 2022) provided emission line measurements for three galaxies at \(z \sim 7.5\), one of which was reported to have an extremely low metallicity of \(12 + \log(O/H) = 6.99 \pm 0.11\) (Curti et al. 2023). We use these three galaxies as comparison points in oxygen-based diagrams, adopting the recently re-measured ratio values from Nakajima et al. (2023). We compare our individual galaxy measurements with

4 https://www.sdss3.org/dr16/spectro/galaxy_mpajhu.php
5 https://mosdef.astro.berkeley.edu/for-scientists/data-releases/
individual $4.5 \leq z \leq 8.0$ galaxies targeted in the GLASS survey, which leverages gravitational lensing (Mascia et al. 2023). Additionally, we compare to composite spectra presented in Sanders et al. (2023) and Tang et al. (2023) from the Cosmic Evolution Early Release Science (CEERS) survey. The $z \sim 5.6$ and $z \sim 7.5$ redshift bins in Sanders et al. (2023) are well matched in redshift to the sample presented here. Those authors report a median stellar mass of $\log(M_*/M_{\odot}) = 8.57^{+0.04}_{-0.05}$ for their $z \sim 5.6$ composite, while they do not estimate masses in their $z \sim 7.5$ bin. Tang et al. (2023) present one composite spectrum of 16 $z \geq 7$ galaxies from CEERS with median redshift $z = 7.7$. Those authors report a median UV magnitude of $M_{UV} = -20.6$. This is considerably brighter than our median magnitude of $M_{1500} = -18.71$, indicating that the shallower CEERS spectroscopy is probing generally brighter galaxies that our JADES data.

We also consider a number of possible ‘high-redshift analogues’ identified from low-redshift samples by various authors. ‘Green pea’ galaxies are a population of galaxies at $z \sim 0$ with unusually high equivalent-width emission of [O III]λ5007 ( Cardamone et al. 2009; Yang et al. 2017a). These galaxies are characterised by masses of $M_* \sim 10^{8.5-9.5} M_{\odot}$ and high star formation rates ($\sim 10 M_{\odot}$ yr$^{-1}$), analogous to galaxies at higher redshifts. We consider also ‘blueberry’ galaxies, which are a population of $z \sim 0$ dwarfs with even more extreme properties than green peas, having masses of $M_* \sim 10^{6.5-7.5} M_{\odot}$, very high ionisation ([O III]/[O II] $\gtrsim 10$) and very low metallicities ($12 + \log(O/H) \approx 7.1-7.8$), (Yang et al. 2017b). We also consider samples of the most metal-poor galaxies known in the low-redshift Universe (Izotov et al. 2018, 2019).

The $z \sim 0$ comparison samples are corrected for dust assuming a Cardelli et al. (1989) law with $R_V = 3.1$ where the $A_V$ is derived from the $H\alpha/H\beta$ decrement assuming an intrinsic ratio of 2.86. Meanwhile, the $z \sim 2$ MOSDEF sample is corrected in an identical manner to that used for our $z \sim 6$ sample, described in Sect. 2.4.

### 3.2. Excitation properties of $z \sim 6$ galaxies: BPT and VO87 diagrams

The left panel of Fig. 3 shows our $z \sim 6$ JADES sample plotted onto the so-called N2–BPT diagram (Baldwin et al. 1981). As described in Sect. 2.4.3, we find no convincing detections of [NH]λ6583 in any of our JADES $z \sim 6$ sample (purple points). Figure 3 also shows comparison samples of $z \sim 0$ galaxies from SDSS (background 2D PDF), extremely metal-poor galaxies at $z \sim 0$ (Izotov et al. 2018, 2019), yellow pentagons), $z \sim 2$ galaxies from MOSDEF (green circles), and the $z \sim 5.6$ composite CEERS spectrum from Sanders et al. (2023).

Based on the R3 ratios alone, we can already see that these JADES $z \sim 6$ galaxies span the upper extreme of what is observed for galaxies in SDSS and MOSDEF, with the 18 galaxies shown here having a median R3 of 0.74, and all lying above R3 $> 0.53$. Considering the N2 ratio, SDSS galaxies lying along the star-forming BPT sequence with $0.72 < R3 < 0.76$, have a
Fig. 3. BPT and VO87 diagrams showing the comparison between JADES \( z \approx 6 \) galaxies and samples across various redshifts. Individual measurements are shown as solid purple points, while values derived from stacked JADES spectra are shown as the open magenta diamonds. The background grey 2D PDF shows \( z \approx 0 \) galaxies from SDSS. Extremely metal-poor \( z \approx 0 \) galaxies from Izotov et al. (2018, 2019) are shown as yellow pentagons. MOSDEF \( z \approx 2 \) galaxies are shown as green circles (Kriek et al. 2015). The \( z \approx 5.6 \) composite CEERS spectra presented by Sanders et al. (2023) are shown as the blue open circles. The solid navy lines in each panel show the theoretical maximum starburst lines from Kewley et al. (2001), while the dotted line in the left panel shows the empirical demarcation derived by Kauffmann et al. (2003). Left: classical BPT or ‘N2–BPT’ – Purple triangles provide 3-\( \sigma \) upper limits on the locations of our JADES galaxies since we are unable to detect \([\text{N} \, \text{II}] / \text{H} \alpha \) in any of the 18 galaxies for which we have \( G395M/F290LP \) spectral coverage of the \( \text{H} \alpha \) complex, although the R3 ratio is very well constrained in all cases from PRISM data. Even after stacking the grating spectra of all galaxies, we do not recover a robust detection of \([\text{N} \, \text{II}] / \text{H} \alpha \). Centre: in the \( S2–VO87 \) diagram (often referred to as ‘S2–BPT’), three out of 20 of our JADES galaxies show detections of \([\text{S} \, \text{II}] / \text{H} \alpha \) in individual spectra. The tight 3-\( \sigma \) upper limits we place on our non-detections highlight that there must be more than an order of magnitude in scatter in the \( S2 \) ratio within the sample. Right: we find one tentative detection of \([\text{O} \, \text{I}] / \text{H} \alpha \), (Fig. 3 in 01–VO87 diagram.

A115, page 8 of 19
Because the \([\text{S II}] \lambda 6716, 6731\) doublet is more widely separated in wavelength from \(H\alpha\), we can measure this ratio directly from the much deeper PRISM spectrum\(^7\). Consequently, we observe three \([\text{S II}]\) detections at greater than 3-\(\sigma\). These detections are broadly consistent with lying along the upper extension of the \(z \sim 0\) sequence and \(z \sim 2\) sequence. The deeper PRISM spectra also allow us to place much tighter upper limits on the \(S2\) ratio for the 17 out of 20 non-detections in our sample. In some cases the 3-\(\sigma\) upper limits are more than an order of magnitude lower in \(S2\) than our three detections, suggesting there is significant scatter in \(S2\) ratios within the \(z \sim 6\) population. The composite \(z \sim 5.6\) CEERS spectrum from Sanders et al. (2023) yields \(S2 \sim 1.30^{+0.17}_{-0.17}\) which is broadly in line with the three JADES galaxies with measurable \([\text{S II}]\). This suggests that stacks based on the shallower, but wider area, CEERS spectroscopy are biased towards galaxies that are more in line with the tip of the \(z \sim 0\) sequence, which are likely more evolved systems.

In the right panel of Fig. 3 we next plot \(O1 = [O I] \lambda 6300/H\alpha\) against the \(R3\) ratio. As with the \(S2\) ratio, we can again derive \([O I]/H\alpha\) constraints directly from the PRISM spectra. We find one tentative \([O I]\) detection in our sample (JADES-GS \(+53,15613 – 27.77584\)), indicating the presence of neutral gas in this galaxy. The measured \(O1\) ratio for this galaxy places it above the theoretical maximum starburst excitation limit from Kewley et al. (2001), which could suggest the presence of shock-heated gas in the ISM (Sutherland & Dopita 2017). Interestingly, this galaxy is not one of the three galaxies with a \([\text{S II}]\) detection from Fig. 3. However, it does have the second lowest \(O32\) ratio of galaxies in this sample (Table 1). Overall, the individual \(O1\) constraints presented here are not strongly constraining relative to the Kewley et al. (2001) demarcation line. The measurement from our stacked spectrum suggests the median \(O1\) in this sample could be consistent with the upper extension of the \(z \sim 0\) star-forming sequence. However, we showed above that there is significant diversity in \(S2\) ratios within our sample. If this diversity also holds for \(O1\), it is clear that larger samples probing even deeper would be required to characterise the sample properties of \(O1\) at \(z \sim 6\).

3.3. Excitation properties from \(z \sim 5.5 – 9.5\): \(R2 – R3\) diagram

In Fig. 4 we instead consider the \(R2 = [O II] \lambda 4372, 3729/H\beta\) ratio. The longer wavelength baseline of this ratio has the drawback of being much more sensitive to the reddening correction applied compared to the \(N2 – BPT, S2 – VO87, \) and \(O1 – VO87\) diagrams discussed in Sect. 3.2. However, the shorter rest-frame wavelength of these emission lines means that the \(R2 – R3\) diagram can be used to explore the ISM conditions of galaxies up to \(z \sim 9.5\) with JWST/NIRSpec. \([\text{O II}]\) is typically intrinsically much brighter than \([\text{N II}]\), \([\text{S II}]\), or \([O I]\). Moreover, like \([\text{S II}]\) and \([\text{O I}]\), ratio measurements can be taken directly from the PRISM spectrum. As a result, not only does Fig. 4 include the additional six galaxies from the \(z \sim 8\) sample, we also have a much higher detection rate, with 24 out of 27 galaxies in our combined sample having detections of \([\text{O II}]\).

We see immediately in Fig. 4 that there is more than an order of magnitude in scatter in the \(R2\) ratio, suggesting there is significant diversity in the ISM conditions of galaxies within this sample. We find a median value of \(R2 = -0.23\) with standard deviation of 0.38 among our 24 detections, significantly offset from the median value from MOSDEF of \(R2 = 0.38\). Instead, we find that the range of \(R2\) values observed for this sample are much more aligned with those observed in extreme \(z \sim 0\) dwarf starbursts such as green peas (Cardamone et al. 2009; Yang et al. 2017a; orange plusses) and blueberries (Yang et al. 2017b; blue crosses).

Both \(R2\) and \(R3\) are known to form a double-valued sequence in metallicity (e.g. Curti et al. 2020). In \(z \sim 0\) galaxies, \(R3\) does not change strongly with metallicity between \(-0.1 – 0.5 \times Z_\odot\). On the other hand, \(R2\) turns over at higher metallicities between \(-0.5 – 0.8 \times Z_\odot\), at a maximum value of \(R2 \approx 0.5\) (Curti et al. 2020). The combination of highly consistent \(R3\) ratios but wildly varying \(R2\) ratios indicates that our sample likely falls into the metallicity range of \(-0.1 – 0.5 \times Z_\odot\). We have five galaxies in our JADES sample with \(R2 > 0.0\), overlapping with the region inhabited by typical \(z \sim 2\) MOSDEF galaxies. These may represent systems which are more evolved and more metal-rich. Below its metallicity turnover, \(R2\) values in the \(z \sim 0\) sample from Curti et al. (2020) drop steeply with decreasing metallicity. The lowest \(R2\) values measured in our sample lie below \(R2 < -0.75\), below the range probed by the Curti et al. (2020) calibration. However, we note that the low \(R2\) portion of our sample exhibits a fair resemblance to blueberry galaxies from Yang et al. (2017b) which have been show to have very low metallicities \(<0.1 \times Z_\odot\). However, given that \(R2\) is strongly affected by the ionisation parameter, inferring metallicity values from this ratio is likely not robust. Nonetheless, there is clearly a large range of metallicity being probed in this sample.

We see no evidence for evolution of the \(R2\) ratio between our \(z \sim 8\) and \(z \sim 6\) samples, albeit with only six \(z \sim 8\) galaxies. We note that all of our \(R2 > 0\) galaxies are from the \(z \sim 6\) sample; however, the measurements from stacked spectra in the two samples actually yield a higher value for the \(z \sim 8\) sample in Fig. 4 (open diamonds), likely affected by uncertainty due to small sample size. This suggests that there is no rapid time evolution in this diagram at this epoch and that the scatter is instead driven by sample diversity.

3.4. Ionisation-excitation diagrams: \(R23 – O32\) and \(R23 – Ne3O2\)

Figure 5 shows the \(R23 – O32\) diagram with our JADES sample plotted against the same literature comparison samples shown in Fig. 4. \(R23\) is often taken as an indication of total excitation, as it encompasses emission lines from both singly and doubly ionised oxygen, while the \(O32\) ratio is sensitive to ionisation parameter (e.g. Kewley et al. 2019).

It is in this diagram that the offset of \(z \geq 6\) from typical \(z \sim 0\) and \(z \sim 2\) galaxies is perhaps most striking: \(O32\) ratios are clearly much higher in JADES \(z \sim 5.5 – 9.5\) galaxies compared to MOSDEF galaxies at \(z \sim 2\). Across our combined sample we find a median \(O32\) value of 0.98 with a standard deviation of 0.36, with 18 of our galaxies have robust detections showing \(O32 > 0.75\). This is much higher than MOSDEF galaxies plotted in Fig. 5 which have a median \(O32\) value 0.18 and are already tracing the upper end of the \(z \sim 0\) relation. This indicates that galaxies across our sample exhibit very high ionisation parameters, much larger than what are seen in typical galaxies at \(z \sim 0\). However, we can see in Fig. 5 that our JADES sample largely overlaps with the region spanned by blueberries (Yang et al. 2017b) and green peas (Yang et al. 2017a), which

\(^7\) We note that, although \(H\alpha\) would in theory be blended with \([\text{N II}]\) at this resolution, as highlighted above, we see no evidence for the presence of \([\text{N II}]\) in either the PRISM or grating spectra of any galaxies in our sample. From our grating upper limits we can constrain that blending between \(H\alpha\) and \([\text{N II}]\) would contribute no more than 6\% (0.02 dex) to a bias the observed \(S2\) ratio.
are examples of $z \sim 0$ galaxies with extreme star formation activity. Comparable O32 ratios have also been observed in $z \sim 2$–4 Lyman-α emitters (Nakajima et al. 2016; Tang et al. 2021) and $z \sim 1.3$–2.4 extreme [OIII] emitters (Tang et al. 2019). It is therefore likely that the high O32 ratios observed in our sample are indicative of intense star formation activity.

We find that our $z \sim 6$ sample is significantly offset from the $z \sim 5.6$ stack presented in Sanders et al. (2023), with our composite O32 value found to be more than a factor of four larger. Indeed where individual detections from CEERS are reported in Sanders et al. (2023) they typically have O32 < 0.75. Mirroring what was observed in our R2 sequence from Fig. 4, we find six galaxies with much lower O32, more in line with the composite spectrum from Sanders et al. (2023). Given the resemblance between the upper end of our O32 sample and dwarf starbursts such as blueberries, it is likely that this portion of our sample is probing higher star formation rate and lower-metallicity galaxies (analogous to blueberries), while the handful of lower O32 galaxies presented here in alignment with the Sanders et al. (2023) $z \sim 5.6$ composite could be probing more evolved systems.

On the other hand, the composite CEERS values at $z \sim 7.5$ and $z \sim 7.7$, from Sanders et al. (2023) and Tang et al. (2023) respectively, are in closer agreement with our composite spectra, albeit still offset towards slightly higher R23. This could suggest that the fraction of more evolved, metal-enriched systems is increasing rapidly from $z \sim 8$ to $z \sim 6$, despite the bulk of our UV-selected $z \sim 6$ sample showing no significant evolution from our $z \sim 8$ sample. Addressing this question in more detail requires larger samples of deep spectroscopy where properties of individual galaxies can be isolated.

By leveraging gravitational lensing, the measurements of 4.5 $\leq z \leq$ 8 galaxies from Mascia et al. (2023) can more readily probe fainter and lower mass systems. These are similarly offset from the Sanders et al. (2023) $z \sim 5.6$ composite and largely overlap with our JADES sample, spanning a large range in R23. Two galaxies are reported by Mascia et al. (2023) as having R23 $\leq 0.6$ despite having high O32 comparable with our JADES sample. This indicates that gravitational lensing may be important for studying low metallicity galaxies (below the R23 turnover) and not simply ultra-deep observations like those presented here.

We note that, as with Fig. 4, this diagram relies on ratios with long baselines in wavelength and that dust reddening can have a significant effect. The dust-correction (described in Sect. 2.4.4) is most significant on the O32 ratio here, with larger corrections serving to reduce O32 ratios. From the corrections we apply, all but two galaxies have corrections that shift O32 by less than 0.15. The offset between the Sanders et al. (2023) $z \sim 5.6$ composite and our $z \sim 6$ stack cannot be explained by the dust-corrected...
correction; ‘correcting’ our value down to \( \text{O32} \approx 0.5 \) would require \( E(B - V) \gtrsim 1 \), significantly larger than our measured value of \( E(B - V) = 0.06 \).

Nonetheless, in Fig. 6 we show our JADES \( z \sim 6 \) and \( z \sim 8 \) samples with Ne3O2 replacing O32 to provide the ionisation axis, which provides the bonus that Ne3O2 is essentially unaffected by dust extinction and the associated correction uncertainties\(^8\). Despite the lower detection rate, and generally much larger measurement uncertainties associated with switching from [O III] \( \lambda 5007 \) to the intrinsically much fainter [Ne III] \( \lambda 3869 \), we see the same story reflected here with the JADES sample exhibiting large scatter (median Ne3O2 = \(-0.11\) with standard deviation 0.29) and generally offset from SDSS and MOSDEF galaxies, instead being more aligned with blueberry dwarf starbursts. This again supports the conclusion that our JADES sample is probing a diverse sample of galaxies, many of which exhibit very high ionisation parameters analogous to those observed in local compact starbursts. Tang et al. (2023) presented Ne3O2 measurements from their \( z \sim 7.7 \) composite CEERS spectrum. These are marginally lower than our Ne3O2 composite values. As in Fig. 5, the larger offset is again the 0.1 dex difference in R23. This offset could reflect an increase in metallicity from our sample to the Tang et al. (2023) galaxies which are reported as having typically brighter UV magnitudes (median \( M_{\text{UV}} = -20.6 \) from Tang et al. 2023).

Finally, we note that [Ne III] \( \lambda 3869 \) emission may become important for confirming spectroscopic redshifts at \( z \gtrsim 10 \) where [O III] \( \lambda 4959, 5007 \) is no longer observable within the wavelength coverage of NIRSpec. We find the [Ne III] \( \lambda 3869 \) emission to be brighter than the [O II] doublet in at least 8 out of 27 galaxies in our sample. Given that these cases should represent the more extreme examples (highest ionisation parameters and lowest metallicities), [Ne III] \( \lambda 3869 \) may be one of the most accessible emission lines in searches for the highest redshift galaxies. However, existing deep spectroscopic observations of such galaxies have already demonstrated that observing any emission lines at those redshifts can be challenging (Curtis-Lake et al. 2023; Robertson et al. 2023).

4. Discussion

4.1. ISM conditions in the epoch of reionisation

Section 3 showed that our \( z \sim 5.5-9.5 \) sample from JADES exhibits significant diversity in ISM conditions. The measured ratios highlight a clear difference between this high-redshift sample and typical galaxies at \( z \sim 0-3 \), although parallels can be drawn instead with extreme starburst populations of low-redshift galaxies, such as blueberries and green peas. In this section

---

\(^8\) We note that at the spectral resolution of our PRISM observations, [Ne III] \( \lambda 3869 \) is partially blended with the He I \( \lambda 3889 \) line. See discussion in Sect. 2.4.
Ne$^3$O$_2 = \log \frac{[\text{Ne}^3]}{[\text{O}]} \lambda \lambda 3869, 3726, 3729$

... between 30, 40 \times Z_{\odot}

... x_{\text{eff}} \geq 0.07, 0.15, 0.22, 0.30, 0.45, 0.60, 0.75, 1.0, 1.5, 2.2, 3.0, assuming Z_\odot = 0.0134; Asplund et al. 2009).

Metal abundances ratios are assumed to follow solar abundance patterns (Asplund et al. 2009), and the depletion of metals onto dust is accounted for.

Figure 7 highlights that photoionisation models driven by SEDs of young stars can reproduce the parameter space populated by our JADES galaxies with only fairly standard assumptions. This indicates we do not need to invoke alternative heating sources (such as AGN or shocks) to explain our findings. We see from this simple comparison that the large scatter observed in our emission line ratio sample likely reflects a diversity in metallicity and ionisation parameter among these galaxies. Our most extreme galaxies (i.e. highest O32) are only reproduced by the models with very low-metallicity (0.07–0.15 \times Z_{\odot}) and very high ionisation parameters (-2.0 < \log U < -1.0). These ionisation parameters are significantly higher than those typically observed in $z \sim 0$ HII regions (Dopita et al. 2000) or $z \sim 2$–3 galaxies (Sanders et al. 2016). We note that we have not explored the effect of harder radiation fields, which would increase O32 at a fixed ionisation parameter (e.g. Sanders et al. 2016). Indeed, many studies at $z \sim 2$–3 have argued that harder radiation fields, driven by super-solar $\alpha$/Fe abundance ratios, are a dominant effect in driving the $z \sim 2$ N2–BPT offset (Steidel et al. 2016; Strom et al. 2017; Topping et al. 2020b,a; Cullen et al. 2021). Given the long timescale of iron enrichment via Type Ia supernovae, one would of course expect that these $z \gtrsim 5.5$ galaxies would be $\alpha$-enhanced, and models with non-solar abundance ratios would be more appropriate for parameter estimation.

We can see from these comparisons that our sample likely spans quite a range in metallicity. At lower O32 values, we see a turnover in the models where metallicity and ionisation parameter become highly degenerate. However, model metallicities of at least 0.45 \times Z_{\odot} are required to approach our highest R2 and R23 values. The R2, R3, and R23 ratios are all known to exhibit double-valued relations (e.g. Curti et al. 2020) in metallicity. We tentatively observe that our highest O32 galaxies tend to have the lowest R23, which reverses the trend seen in SDSS galaxies (grey 2D PDF), suggesting we are sampling the turnover of this metallicity range.
relation. Some of our ‘mildest’ galaxies (i.e. highest R2, lowest O32) have R2 ≈ 0.4 and R23 ≈ 1.0 which are broadly consistent with the turnover observed for these ratios in z ≈ 0 galaxies from Curti et al. (2020). This suggests these galaxies are the most chemically evolved systems in our sample and may have metallicities as high as 0.3–0.5 × Z⊙. Indeed, these are also the points that are in best agreement with the Sanders et al. (2023) z ∼ 5.6 composite spectrum. This suggests the shallower CEERS spectroscopy is probing on average more chemically evolved systems than those targeted in JADES.

Overall, it is clear from this simple model comparison that a range of ISM conditions is required to explain the diversity of emission-line ratios observed in this sample. As noted in Sect. 3, our sample largely spans the same regions of emission-line-ratio-space inhabited by low-redshift extreme starbursts such as blueberries and green peas. Thus, the emerging picture is that the galaxies in our sample are likely low-metallicity (−0.1–0.3 × Z⊙) systems undergoing intense periods of star formation driving strong radiation fields with high ionisation parameters (−2.0 ≤ log U ≤ −1.0) in the ISM.

4.2. High O32: implications for escape fraction and reionisation

The high O32 ratios presented in Fig. 5 could have implications for interpreting the ionising photon escape fraction of these z ∼ 5.5–9.5 galaxies. When and how the Universe was reionised remains one of the most important open questions in high-redshift astrophysics. The efficiency with which ionising photons are produced within galaxies and the fraction of these which escape (fesc) are two important quantities required to understand the contribution of galaxies to this process. Unfortunately, fesc
can never be directly measured within the epoch of reionisation itself because any light escaping the galaxy bluewards of the Lyman-limit will interact with intervening neutral hydrogen. Thus indirect probes of fesc are highly sought after.

High O32 has been proposed as being an indicator of higher fesc (e.g. Nakajima & Ouchi 2014; Faisst 2016). The reasoning is that the high O32 ratio selects for highly ionised systems which have a higher likelihood of having density-bounded channels through which ionising photons can escape. However, some studies have shown that O32 does not necessarily correlate well with fesc; instead finding that results are very dependent on the observed line of sight, and that high O32 can be produced in systems without Lyman continuum (LyC) escape (Paalvast et al. 2018; Bassett et al. 2019; Katz et al. 2020). Nonetheless, samples of Lyman continuum (LyC) leaking galaxies at low-redshift generally show that the fraction of galaxies with high fesc increases towards higher O32, even if the correlation is not tight (e.g. Izotov et al. 2016; Flury et al. 2022). Indeed, the limit itself because any light escaping the galaxy bluewards of the Lyman-limit will interact with intervening neutral hydrogen.

The fainter [O iii] λ5007, 3729 and [Ne III] λ3869 are much more readily detected throughout the sample. Measured R2, O32 and Ne302 ratios particularly highlight the diversity in this sample; we find median values and standard deviations of $R2-z = -0.23 \pm 0.38$, $O32 = 0.98 \pm 0.36$, and $Ne302 = -0.03 \pm 0.27$. We find that these ratios are generally consistent with very high ionisation parameters up to log $U \sim -1.0$. The sample appears to span a significant range in metallicity of $-0.1 - 0.3 \times Z_\odot$, depending on metallicity calibrator assumptions. These ratios are found to be generally consistent with extreme low-redshift populations, particularly green pea and blueberry dwarf starburst samples. This is consistent with the suggestion that these $z \sim 5.5-9.5$ galaxies are low-mass, low-metallicity galaxies undergoing periods of rapid star formation, driving strong radiation fields.

We first compare our sample to composite spectra at $z \sim 5.6$ and $z \sim 7.5$ from CEERS (Sanders et al. 2023; Tang et al. 2023). We find significant offset between our $z \sim 6$ sample median and the $z \sim 5.6$ CEERS composite. We instead observe good agreement between this $z \sim 5.6$ CEERS composite measurement and the low O32 end of our sample. Assuming these galaxies represent the more evolved, metal-enriched extent of our sample, it is clear that the shallower CEERS spectra are preferentially picking up on the more evolved portion of the $z \sim 6$ population. We find no evidence for significant evolution between our $z \sim 6$ and $z \sim 8$ samples (albeit with a very small sample). Both our $z \sim 6$ and $z \sim 8$ sample medians are found to be in somewhat closer agreement with the CEERS $z \sim 7.5$ composite spectra, which showed large offset from the CEERS $z \sim 5.6$ composite spectrum. This could indicate that the fraction of moderately metal-enriched systems increases rapidly from $z \sim 8$ to $z \sim 5.5$, even if our sample median does not show evidence of significant redshift evolution.

Within its first six months of observations, JWST/NIRSpec has already shed considerable light on the ISM conditions of $z \gtrsim 5.5$ galaxies. The diversity of ISM conditions presented in this work highlights the need for observations that span the full range of galaxy properties. This will require a combination of deep observations reaching down to even fainter brightness limits than those presented here, as well as wider area surveys probing rarer, more evolved systems at early epochs. Characterising the ISM conditions in larger samples of galaxies across this broader parameter space will help chart galaxy assembly across the first ~1 Gyr of cosmic time.

5. Conclusions

We have presented rest-frame optical emission line ratio measurements in 27 individual galaxies at $5.5 < z < 9.5$ spanning over $\sim 7.0 < M_{1500} < -20.4$, observed as part of the ‘Deep’ tier of the JADES program, which represent the deepest integrations yet taken with JWST/NIRSpec. We find that galaxies in this sample occupy regions of line-ratio space that are offset from those inhabited by ‘typical’ galaxies at $z \sim 0$ or $z \sim 2$, although generally aligned with more extreme low-redshift populations such as ‘blueberry’ and ‘green pea’ dwarf starbursts (Yang et al. 2017a,b).

The high signal-to-noise achieved in individual spectra from our 28 h deep PRISM/CLEAR integrations reveal significant intrinsic scatter in the observed line ratios, with our JADES $z \sim 5.5-9.5$ galaxies spanning a larger range than that observed in high-redshift typically have high O32 ratios (Vanzella et al. 2016, 2020, 2022; Fletcher et al. 2019; Saxena et al. 2022). Extreme [O III] emitters at $z \sim 2-3$ have also been shown to exhibit high O32 ratios ([O III]/[O II] $\sim 10$), while also preferentially selecting for high equivalent-width Lyman-α emission (Tang et al. 2019, 2021). This evidence therefore seems to suggest that high O32 is a necessary but not sufficient condition for high fesc.

We have shown in Sect. 3.4 that our JADES $z \sim 5.5-9.5$ galaxies have characteristically high O32 ratios, with 21 out of 27 of the sample having [O III]/[O II] $> 5$ and a median value of 9.4. Owing to the large scatter in O32–fesc relations, we do not attempt to derive constraints on the fesc of these galaxies. However, the high fraction of our sample that meets the criterion of high O32 suggests that the fraction of potential LyC leakers might be increasing to high redshift.

Acknowledgements. A.J.C., A.S., A.J.B., J.C.C., and G.C.J acknowledge funding from the “FirstGalaxies” Advanced Grant from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant agreement No. 789056). F.D.E., R.M., M.C., T.J.L., I.W., L.S., and J.S. acknowledge support by the Science and Technology Facilities Council (STFC), ERC Advanced Grant 695671 “QUENCH”. R.M. also acknowledges funding from a research professorship from the Royal Society. J.W. also acknowledges support from the Fondation MERAC. S.C. acknowledges support by European Union’s H2020 Starting Grant No. 101040227 – WINGS. E.C.L. acknowledges support of an STFC Webb Fellowship (ST/W001438/1). The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant no.140. S.A. and B.R.D.P. acknowledge support from the research project PID2021-12771NB-00 of the Spanish Ministry of Science and Innovation/State Agency of Research (MICIN/AEI). R.S. acknowledges support from a STFC Ernest Rutherford Fellowship (ST/S004831/1). H.U. gratefully acknowledges support by the Isaac Newton Trust and by the Kavli Foundation through a Newton-Kavli Junior Fellowship. D.J.E. is supported as a Simons Investigator and by JWST/NIRSpec. We find that galaxies in this work highlights the need for observations that span the full range of galaxy properties. This will require a combination of deep observations reaching down to even fainter brightness limits than those presented here, as well as wider area surveys probing rarer, more evolved systems at early epochs. Characterising the ISM conditions in larger samples of galaxies across this broader parameter space will help chart galaxy assembly across the first ~1 Gyr of cosmic time.
**Appendix A: Table of selected targets**

Table A.1 gives an overview of the targets selected in our JADES z ~ 5.5 – 9.5 sample.

**Table A.1.** Summary of coordinates, magnitudes, and total integration times on each of the targets presented in this sample.

<table>
<thead>
<tr>
<th>ID</th>
<th>NIRSpec_ID</th>
<th>R.A.</th>
<th>Decl.</th>
<th>zPRISM</th>
<th>t\text{int. (hr)} PRISM</th>
<th>t\text{int. (hr)} G395M</th>
<th>M\text{1500}</th>
</tr>
</thead>
<tbody>
<tr>
<td>JADES-GS+53.11243-27.77461</td>
<td>10058975</td>
<td>53.112434</td>
<td>-27.774609</td>
<td>9.438</td>
<td>28.0 7.0*</td>
<td>-20.36 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.16446-27.80218</td>
<td>8013</td>
<td>53.164464</td>
<td>-27.802180</td>
<td>8.479</td>
<td>28.0 7.0*</td>
<td>-17.55 ± 0.84</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.15682-27.76716</td>
<td>21842</td>
<td>53.156825</td>
<td>-27.767159</td>
<td>7.981</td>
<td>28.0 7.0*</td>
<td>-18.43 ± 0.03</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.16746-27.77201</td>
<td>10013682</td>
<td>53.167464</td>
<td>-27.772006</td>
<td>7.277</td>
<td>18.7 4.7*</td>
<td>-17.01 ± 0.22</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.11833-27.76901</td>
<td>10013905</td>
<td>53.118327</td>
<td>-27.769010</td>
<td>7.206</td>
<td>28.0 7.0*</td>
<td>-18.88 ± 0.05</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.13423-27.76891</td>
<td>20961</td>
<td>53.134229</td>
<td>-27.768913</td>
<td>7.051</td>
<td>18.7 4.7*</td>
<td>-18.82 ± 0.06</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.11730-27.76408</td>
<td>10013609</td>
<td>53.117303</td>
<td>-27.764082</td>
<td>6.931</td>
<td>9.3 2.3</td>
<td>-18.71 ± 0.06</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.15579-27.81520</td>
<td>4297</td>
<td>53.155788</td>
<td>-27.815202</td>
<td>6.718</td>
<td>9.3 2.3</td>
<td>-18.58 ± 0.07</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.15138-27.81917</td>
<td>3334</td>
<td>53.151381</td>
<td>-27.819165</td>
<td>6.709</td>
<td>28.0 7.0</td>
<td>-17.80 ± 0.07</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.16904-27.77884</td>
<td>16625</td>
<td>53.169042</td>
<td>-27.778838</td>
<td>6.631</td>
<td>28.0 7.0</td>
<td>-18.75 ± 0.03</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.13492-27.77271</td>
<td>18846</td>
<td>53.134917</td>
<td>-27.772709</td>
<td>6.342</td>
<td>28.0 7.0</td>
<td>-20.14 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.17582-27.77446</td>
<td>18179</td>
<td>53.175816</td>
<td>-27.774464</td>
<td>6.353</td>
<td>18.7 0.0</td>
<td>-18.71 ± 0.06</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.16660-27.77240</td>
<td>18976</td>
<td>53.166602</td>
<td>-27.772403</td>
<td>6.329</td>
<td>28.0 7.0</td>
<td>-18.61 ± 0.04</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.15613-27.77584</td>
<td>17566</td>
<td>53.156128</td>
<td>-27.775841</td>
<td>6.105</td>
<td>18.7 4.7</td>
<td>-19.05 ± 0.04</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.16062-27.77161</td>
<td>19342</td>
<td>53.160620</td>
<td>-27.771613</td>
<td>5.981</td>
<td>28.0 7.0</td>
<td>-18.67 ± 0.04</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.11911-27.76080</td>
<td>10013618</td>
<td>53.119112</td>
<td>-27.760802</td>
<td>5.948</td>
<td>28.0 7.0</td>
<td>-19.42 ± 0.02</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.12176-27.79763</td>
<td>9422</td>
<td>53.121755</td>
<td>-27.797634</td>
<td>5.943</td>
<td>28.0 7.0*</td>
<td>-18.91 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.11041-27.80892</td>
<td>6002</td>
<td>53.110411</td>
<td>-27.808923</td>
<td>5.941</td>
<td>28.0 7.0</td>
<td>-18.60 ± 0.03</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.12259-27.76057</td>
<td>10013620</td>
<td>53.122590</td>
<td>-27.760569</td>
<td>5.920</td>
<td>28.0 7.0</td>
<td>-19.64 ± 0.02</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.17655-27.77111</td>
<td>19066</td>
<td>53.176550</td>
<td>-27.771108</td>
<td>5.893</td>
<td>9.3 2.3</td>
<td>-18.75 ± 0.06</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.11351-27.77264</td>
<td>10056849</td>
<td>53.113511</td>
<td>-27.772836</td>
<td>5.822</td>
<td>28.0 2.3</td>
<td>-18.16 ± 0.05</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.16730-27.80287</td>
<td>10005113</td>
<td>53.167303</td>
<td>-27.802870</td>
<td>5.820</td>
<td>28.0 7.0*</td>
<td>-17.85 ± 0.08</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.15407-27.76607</td>
<td>22251</td>
<td>53.154070</td>
<td>-27.766072</td>
<td>5.804</td>
<td>18.7 4.7</td>
<td>-18.93 ± 0.04</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.11537-27.81477</td>
<td>4404</td>
<td>53.115372</td>
<td>-27.814771</td>
<td>5.775</td>
<td>28.0 7.0</td>
<td>-19.31 ± 0.03</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.13002-27.77839</td>
<td>16745</td>
<td>53.130023</td>
<td>-27.778393</td>
<td>5.574</td>
<td>28.0 7.0</td>
<td>-19.60 ± 0.01</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.12972-27.80818</td>
<td>6246</td>
<td>53.129722</td>
<td>-27.808177</td>
<td>5.570</td>
<td>28.0 7.0</td>
<td>-18.01 ± 0.05</td>
<td></td>
</tr>
<tr>
<td>JADES-GS+53.11572-27.74966</td>
<td>10016374</td>
<td>53.115717</td>
<td>-27.74955</td>
<td>5.507</td>
<td>18.7 4.7</td>
<td>-18.66 ± 0.02</td>
<td></td>
</tr>
</tbody>
</table>

The six targets above the horizontal rule comprise the JADES 'z ~ 8' sample and do not have coverage of any of H\text{β}, [N\text{II}] or [S\text{II}] in either PRISM or G395M/F290LP data. The remainder are referred to as the 'z ~ 6' sample. 'NIRSpec_ID' refers to the integer ID used in the JADES data release (Bunker et al. 2023). * Target does not have coverage of the H\text{β+[NII]} complex in G395M/F290LP observation.
Appendix B: Emission line ratio measurements from stacked spectra

In Section 2.5 we described our procedure for measuring emission-line ratios from stacked spectra. Here we briefly summarise the values obtained from this analysis. Table B.1 outlines the number of galaxies included in each stack and median properties of these samples. Table B.2 provides the dust-corrected emission-line ratios obtained from the resultant stacked spectra. These values are shown as open diamonds in Figures 3–6.

Table B.1. Properties of samples underpinning stacked spectra.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Disperser</th>
<th>N\textsubscript{galaxies}</th>
<th>(z)†</th>
<th>(M_{1500}^{*})</th>
</tr>
</thead>
<tbody>
<tr>
<td>JADES (z \sim 6)</td>
<td>PRISM/CLEAR</td>
<td>21</td>
<td>5.943</td>
<td>-18.71</td>
</tr>
<tr>
<td>JADES (z \sim 6)</td>
<td>G395M/F290LP</td>
<td>18</td>
<td>5.945</td>
<td>-18.73</td>
</tr>
<tr>
<td>JADES (z \sim 8)</td>
<td>PRISM/CLEAR</td>
<td>6</td>
<td>7.629</td>
<td>-18.82</td>
</tr>
</tbody>
</table>

† Median value of all galaxies included in the stack.

Table B.2. Emission line ratio measurements from our stacked JADES spectra.

<table>
<thead>
<tr>
<th>Ratio</th>
<th>JADES (z \sim 6)</th>
<th>JADES (z \sim 8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N2)</td>
<td>(&lt; -1.44)</td>
<td>(&lt; )</td>
</tr>
<tr>
<td>(S2)</td>
<td>(&lt; -1.11)</td>
<td>(&lt; )</td>
</tr>
<tr>
<td>(O1)</td>
<td>(-2.20 \pm 0.37)</td>
<td>(&lt; )</td>
</tr>
<tr>
<td>(R3)</td>
<td>(0.71 \pm 0.03)</td>
<td>(0.70 \pm 0.05)</td>
</tr>
<tr>
<td>(R2)</td>
<td>(-0.50 \pm 0.13)</td>
<td>(-0.28 \pm 0.08)</td>
</tr>
<tr>
<td>(R23)</td>
<td>(0.86 \pm 0.02)</td>
<td>(0.86 \pm 0.04)</td>
</tr>
<tr>
<td>(O32)</td>
<td>(1.21 \pm 0.16)</td>
<td>(0.97 \pm 0.13)</td>
</tr>
<tr>
<td>(Ne3O2)</td>
<td>(0.14 \pm 0.07)</td>
<td>(-0.13 \pm 0.16)</td>
</tr>
</tbody>
</table>

Ratio names are defined in Equation 1. Non-detections are quoted as 3-\(\sigma\) upper limits. Ratios in this table are derived from stacked PRISM/CLEAR spectra in all cases, with the exception of \(N2\) which is derived from stacked G395M/F290LP spectra.

Appendix C: Verification of flux calibration

As a verification of our spectral flux calibration, we compared the fluxes obtained from NIRCam photometry with those obtained from the PRISM/CLEAR spectra. Using imaging data from the JADES programme (Eisenstein et al. 2023; Rieke et al. 2023), we extract apodised fluxes for each filter by extracting the flux within the projected shutter position of each target. Assuming the MSA pointing uncertainty is negligible, this should represent the total amount of flux in each spectral region incident on the shutter area. However, the point spread functions (PSFs) of NIRCam and NIRSpec are different, and in reality the final flux recorded by NIRSpec is additionally reduced by losses due to diffraction caused by the shutter itself. Our data reduction pipeline includes a step which corrects for these slitlosses, meaning that the output spectrum should be scaled to the same flux density as the total magnitude of the source (correcting for position within the slit and flux falling outside the slit). As a further check on the relative flux calibration between NIRCam and NIRSpec, we also consider the flux from the NIRCam imaging which would fall within the NIRSpec microshutter aperture. To make a fair comparison, we first remove the slitloss correction applied by the pipeline to the NIRSpec spectrum (to consider just the uncorrected flux within the shutter) and convolve the resultant PRISM/CLEAR spectrum with the NIRCam filter transmission curves.

Figure C.2 shows the comparison between the NIRCam and NIRSpec fluxes derived in this way, with grey points representing each individual pointing of each target galaxy. Typical statistical uncertainties for individual measurements are shown by the horizontal grey marks, which include the contributions from both the NIRCam and the NIRSpec uncertainties. For a handful of galaxies, the NIRSpec continuum flux was low enough that the final measured flux did not have \(S/N > 3\), and we removed these comparison. From the large red points and errorbars, showing the median and standard deviation across all observations in a particular filter, we see that for all filters the median of the comparison points is well within one standard deviation of unity. However, on average many of these indicate that the NIRCam fluxes are up to \(~10\)% lower than the NIRSpec fluxes. In particular, we see evidence of a subtle trend with the NIRCam-to-NIRSpec flux ratio increasing slightly to long wavelengths. This arises as a result of the NIRSpec PSF degrading more significantly with increasing wavelength beyond 2.5 \(\mu m\).

Our NIRSpec data reduction pipeline does take this effect into account, in the form of the slitloss correction (which was undone for Figure C.2). These slitlosses will depend on the morphology of the galaxy and its location within the shutter, but the \(~10\)% drift seen with wavelength in the comparison in Figure C.2 is entirely consistent with the differing PSFs of NIRCam and NIRSpec as a function of wavelength. Slitlosses presented in this work were calculated assuming each galaxy is a point source at the location of the centroid derived from imaging. Table C.1 gives the half-light radii from the JADES public catalogue (Rieke et al. 2023)\(^9\) for the subset of our sample for

\(^9\) Note each target was observed in up to three independent MSA configurations (see Section 2), and that these in general have similar, but not identical, shutter projections.

\(^10\) https://archive.stsci.edu/hlsps/jades
which NIRCam imaging was available. In this study, we focused only on rest-frame optical emission lines at $z > 5.5$, meaning that the observed-frame wavelengths were all $\lambda > 2.5 \mu m$. At these wavelengths, the NIRCam PSF is typically worse than $0''1$ ($0''092$ in F277W, $0''116$ in F356W). Thus, at the wavelengths considered in this work, the assumption of a point source profile should be reasonable for all but a few of our largest targets.

The majority of the emission-line ratios studied in this work have relatively short wavelength baselines, for which the precise slitloss corrections will make a negligible difference. The $O32$ and $R23$ ratios are the notable exceptions to this. However, even before slitloss corrections are applied, Figure C.2 demonstrates that, on average, there is only a $\sim 10\%$ affect across the full wavelength range. Given that $[O\text{II}].\lambda 3726,3729$ and $[O\text{III}].\lambda 5007$ are only separated by $\sim 1.4 \mu m$ in the observed frame, the effect would be even smaller than this quoted 10%. Furthermore, the point-source slitloss corrections we do apply should be adequate for most sources. Thus, the final flux calibration uncertainty on the $O32$ ratio is likely well below 10\% which is smaller than the quoted measurements uncertainties throughout this paper.

C.1. Comparison of grating flux calibration

This paper almost exclusively makes use of the PRISM/CLEAR mode. This is because, at redshifts beyond $z > 5$, the rest-frame optical emission lines fall beyond $\lambda_{\text{obs}} > 2.5 \mu m$ where the PRISM/CLEAR spectral resolution is $R \sim 200$–$300$, meaning most of these lines are well resolved. The notable exception to this is the $[N\text{II}].\lambda 6583/H\alpha$ ratio, for which we require the $R \sim 1000$ G395M/F290LP grating. Although reported ratios are only ever measured within the same spectral observation, we briefly consider the relative flux calibration of the PRISM/CLEAR and G395M/F290LP. Figure C.2 shows the ratio of H$\alpha$ fluxes measured in the G395M/F290LP to those measured in the PRISM/CLEAR mode for galaxies in this sample which have spectral coverage in both settings. We note that the fluxes measured here are not those reported in the main text, where the continuum is modelled with stellar templates. Rather the continuum here is modelled with a simple spline fit. This is done to minimise any model-dependent effects in this comparison. The dashed magenta line shows the median value of 1.16, indicating that fluxes measured from the grating are systematically larger. This effect has been noted in Bunker et al. (2023). We note that in the PRISM/CLEAR, the reported H$\alpha$ flux would be blended with $[N\text{II}]$, whereas these lines are resolved in the G395M/F290LP measurements. However, as pointed out in Section 3, $[N\text{II}].\lambda 6583$ is not clearly detected in any of these galaxies, and this would contribute at the $<10\%$ level. Furthermore, the effect of this blending would be to increase the PRISM/CLEAR flux, but it is the G395M/F290LP flux that is determined to be larger. As outlined above, the flux calibration in the PRISM/CLEAR mode is in good agreement with broadband photometry. Thus, we consider that the PRISM/CLEAR flux calibration is likely reliable, while the grating flux may be overestimated. However, given we only make use of the G395M/F290LP data to measure the ratio of two lines very close in wavelength within the same observation, the systematically high fluxes recorded should have no effect on any of the findings reported here.

Values taken from the JADES public catalogue (Rieke et al. 2023) for the subset of our sample with NIRCam imaging was available. ‘NIRCam ID’ gives this cross-matched ID from those imaging catalogues. The horizontal rule mark the FWHM of the NIRCam PSF in the F356W filter ($0''116$); i.e. targets below this rule have half-light radii larger than the NIRCam PSF.

<table>
<thead>
<tr>
<th>ID</th>
<th>NIRCam ID</th>
<th>$R_{50}$ (&quot;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8013</td>
<td>110748</td>
<td>0.088</td>
</tr>
<tr>
<td>4297</td>
<td>101683</td>
<td>0.090</td>
</tr>
<tr>
<td>9422</td>
<td>113585</td>
<td>0.090</td>
</tr>
<tr>
<td>20961</td>
<td>135819</td>
<td>0.096</td>
</tr>
<tr>
<td>16625</td>
<td>127219</td>
<td>0.097</td>
</tr>
<tr>
<td>10005113</td>
<td>10319</td>
<td>0.098</td>
</tr>
<tr>
<td>21842</td>
<td>137667</td>
<td>0.099</td>
</tr>
<tr>
<td>18846</td>
<td>131688</td>
<td>0.101</td>
</tr>
<tr>
<td>19606</td>
<td>210963</td>
<td>0.111</td>
</tr>
<tr>
<td>18179</td>
<td>130262</td>
<td>0.114</td>
</tr>
<tr>
<td>6002</td>
<td>106292</td>
<td>0.114</td>
</tr>
<tr>
<td>4404</td>
<td>101990</td>
<td>0.115</td>
</tr>
<tr>
<td>6246</td>
<td>106885</td>
<td>0.116</td>
</tr>
<tr>
<td>19342</td>
<td>132780</td>
<td>0.121</td>
</tr>
<tr>
<td>3334</td>
<td>197088</td>
<td>0.126</td>
</tr>
<tr>
<td>22251</td>
<td>138571</td>
<td>0.132</td>
</tr>
<tr>
<td>18976</td>
<td>131971</td>
<td>0.136</td>
</tr>
<tr>
<td>10013682</td>
<td>210625</td>
<td>0.132</td>
</tr>
<tr>
<td>17566</td>
<td>209276</td>
<td>0.190</td>
</tr>
<tr>
<td>16745</td>
<td>208642</td>
<td>0.191</td>
</tr>
</tbody>
</table>

\[ \text{H}\alpha \text{ flux ratio } f_{\text{R}} \approx \frac{1000}{f_{\text{PRISM}}} \]

\[ \text{Fig. C.2. Comparison of H}\alpha \text{ flux measurements made from low-resolution PRISM/CLEAR and } R \sim 1000 \text{ G395M/F290LP measurements for galaxies in this sample. Fluxes reported in the } R \sim 1000 \text{ data are systematically higher by a median factor of 1.16.} \]

\[ \text{https://jwst-docs.stsci.edu} \]