Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **544,** 2737–2757 (2025) https://doi.org/10.1093/mnras/staf1836
Advance Access publication 2025 October 23

New methods of identifying AGN in the early Universe using spectroscopy and photometry in the *JWST* era

Flor Arevalo-Gonzalez [1,2,3,4★†] Titanilla Braun [1,5★] James Trussler,[1,6] Christopher J. Conselice,[1] Thomas Harvey,[1] Nathan Adams [1] Duncan Austin [1] Qiong Li [1] Ignas Juodžbalis [7,8] and Kimihiko Nakajima [9]

[1] *Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, The University of Manchester, Manchester M13 9PL, UK*
[2] *Physics Department, Tor Vergata University of Rome, Via della Ricerca Scientifica 1, I-00133 Rome, Italy*
[3] *INAF – Astronomical Observatory of Rome, Via Frascati 33, I-00040 Monte Porzio Catone, Italy*
[4] *Physics Department, Sapienza University of Rome, Piazzale Aldo Moro 5, I-00185 Rome, Italy*
[5] *Department of Physics, University of Oxford, Oxford OX1 3PU, UK*
[6] *Center for Astrophysics | Harvard & Smithsonian, 60 Garden St., Cambridge MA 02138, USA*
[7] *Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge, CB3 OHA, UK*
[8] *Cavendish Laboratory – Astrophysics Group, University of Cambridge, 19 JJ Thomson Avenue, Cambridge CB3 OHE, UK*
[9] *Institute of Liberal Arts and Science, Kanazawa University, Kanazawa, Ishikawa 920-1192, Japan*

Accepted 2025 October 21. Received 2025 October 21; in original form 2025 February 3

ABSTRACT

We explore spectroscopic and photometric methods for identifying high-redshift galaxies containing an active galactic nucleus (AGN) with *James Webb Space Telescope* (*JWST*) observations. After demonstrating the limitations of standard optical methods, which appear ineffective in the low-metallicity environment of the early universe, we evaluate alternative diagnostics using current *JWST* capabilities. Our analysis focuses on line ratios and equivalent widths (EWs) of ultraviolet (UV) lines: C IV, He II λ1640, O III] λ1665, and C III], and the faint optical line He II λ4686. We find that the most valuable diagnostic quantities for identifying AGN are the ratios (C III] + C IV)/He II λ1640 and C III]/He II λ1640, as well as the EW of He II λ1640. For more reliable AGN identification, He II λ1640 and O III] λ1665 should be detected separately. The He II λ1640/Hβ ratio effectively separates AGN from star-forming galaxies, although it depends on low dust content. We also show that to use these diagnostics effectively, future observations require longer exposure times, especially at $z > 6$. We plot three high-redshift sources with strong UV emission on these diagrams; however, further study is needed to classify them as AGN candidates due to He II + O III] blending and unreliable optical lines. Finally, we select AGN candidates in the *JWST* Advanced Deep Extragalactic Survey (JADES) Near Infrared Camera photometry using spectral energy distribution fitting with EAZY. One galaxy emerges as a strong AGN candidate, supported by both photometric selection and strong UV emission, and we present a sample of similar candidates in the JADES data using this method.

Key words: techniques: spectroscopic – galaxies: active – galaxies: high-redshift – galaxies: photometry.

1 INTRODUCTION

Thanks to the availability of new extensive observational data sets from the *James Webb Space Telescope* (*JWST*), there is now scope for a comprehensive characterization of the formation and evolution of galaxies at high redshifts. Previously, research findings were often influenced or constrained by incomplete sampling and limited data quality at specific redshifts. While *JWST* has dramatically expanded observational capabilities across a wide range of wavelengths, enabling the study of the earliest objects in our universe (J. P. Gardner et al. 2023), challenges in data quality and sampling persist at the frontier of galaxy characterization.

The correct classification and categorization of galaxies is the first step towards the study of their properties across time, bringing more clarity into what influences their formation and evolution. Trustworthy methods of identification between different types of galaxies can help us analyse more accurately their observational data and extract more precise conclusions about their properties, applying specific tailored techniques. This is even more relevant at high redshifts, given the expected different cosmic conditions at earlier stages of the universe (L. J. Kewley et al. 2013), and the difficulty in obtaining reliable and accurate data.

Active galactic nuclei (AGNs) are very compact and highly luminous sources found in the centre of galaxies that are shown to dominate the luminosity output of their host galaxy. These differ

★ E-mail: flordenisse9@gmail.com (FAG); titanilla.braun@gmail.com (TB)
† The first two authors contributed equally to this work.

from normal star-forming galaxies (referred to as SF galaxies or SFG from here on) in which the main source of the luminosity is not concentrated to such a small region and the total output of the galaxy is significantly lower considering a broad band of wavelengths.

Thus, one of the goals of galaxy formation and evolution research is to determine which galaxies contain an AGN. This is important because AGNs are powerful sources that play a crucial role in the evolution of their host galaxies, influencing processes such as triggering or quenching star formation rates (e.g. C. Dalla Vecchia et al. 2004). Additionally, the analysis of the abundance of SFG or AGN at high redshifts allows us to assess their relative impact on shaping the history of the universe; for example, by studying how they affect the reionization process during the reionization epoch (B. E. Robertson et al. 2015; K. Duncan & C. J. Conselice 2015).

It is possible to differentiate between AGN and SF galaxies based on their dominant ionizing radiation sources. In AGN, the emission is primarily powered by accretion of matter onto a central supermassive black hole (SMBH). This produces a strong, hard ionizing continuum from the accretion disc that photoionizes surrounding gas, including extended narrow-line regions (NLRs, D. Lynden-Bell 1969). In contrast, the line emission in SF galaxies is predominantly powered by young, massive stars, whose ultraviolet (UV) photons ionize the surrounding gas in ionized hydrogen (H II) regions. Although accreting stellar-mass black holes (e.g. in high-mass X-ray binaries) can also contribute to the ionizing radiation in SF galaxies, their effect is generally thought to be minor compared to the contribution from star formation (M. Lecroq et al. 2024; S. Mineo, M. Gilfanov & R. Sunyaev 2012), but this assumption is still under discussion (D. Schaerer, T. Fragos & Y. I. Izotov 2019; P. Senchyna & D. P. Stark 2019; C. Simmonds, D. Schaerer & A. Verhamme 2021; K. Garofali et al. 2024).The much harder ionizing spectra associated with AGN makes it possible to distinguish them from SF galaxies using emission-line intensity ratios.

In the conventional unified model, AGN can be further categorized into two groups based on the orientation of the accretion disc around the black hole relative to the observer (L. Lima Santos & S. B. Soltau 2024; R. Antonucci 1993). Type I AGN, also known as broad-line AGN (BL AGN), exhibit both narrow and broad emission lines since the higher velocity broad-line region (BLR), which is closer to the central SMBH, is directly detectable. In the case of type II AGN, also named narrow-line AGN (NL AGN), the BLR is obscured by a dusty medium around the accretion disc, hence only narrow emission lines are detected.

Although broad emission lines are a typical feature of AGN, they can also be produced by extreme stellar kinematics in compact galaxies, as noted in J. F. W. Baggen et al. (2024). This means that while broad lines can suggest AGN presence, diagnostic plots comparing emission-line ratios are still essential for properly distinguishing AGN from alternative explanations such as dense stellar systems, outflows, and supernova.

Line ratio plots were first proposed by J. A. Baldwin, M. M. Phillips & R. Terlevich (1981), who suggested a set of optical diagnostic diagrams, now commonly referred to as BPT diagrams, to separate emission spectra powered by different excitation mechanisms. The method was revised by S. Veilleux & D. E. Osterbrock (1987) creating the VO87 diagrams. These standard optical diagnostic diagrams are based on the line ratios [O III]/Hβ, [N II]/Hα, [S II]/Hα, and [O I]/Hα.

Using these optical diagrams, L. J. Kewley et al. (2001) defined a 'maximum starburst line' based on the upper limit of theoretical photoionization models such that galaxies above this line have a very high probability of containing an AGN. Since this is an upper limit for the SF model, it excludes possible AGN candidates below it. G. Kauffmann et al. (2003) revised this demarcation line to divide pure SFGs from the composite region between their revised line and the L. J. Kewley et al. (2001) maximum starburst line. Later the AGN region was further divided to differentiate between Seyferts and LINERs (low ionization nuclear emission-line regions) by L. J. Kewley et al. (2006). This shows the great distinguishing power of the standard BPT diagrams.

However, in the earlier stages of the Universe, astrophysical conditions differ significantly from the local Universe. The low-metallicity environment that is present in the early Universe complicates the utilization of commonly used diagnostics. According to available models and observations, there is a clear correlation between cosmic time and the metallicity of galaxies (K. Nagamine et al. 2001; R. Maiolino et al. 2008), as more recent galaxies have been subject to more consecutive processes of star formation. As metallicity decreases and the ionization parameter increases, distinguishing AGN from other sources of ionization using optical emission lines becomes more challenging (H. Übler et al. 2023).

This is supported by recent studies suggesting that the BPT and VO87 diagrams might not be effective for the identification of high-redshift AGN. A study by B. E. Backhaus et al. (2022) stated that the VO87 diagram cannot efficiently differentiate between AGN and SF galaxies at higher redshifts due to these regions overlapping. H. Übler et al. (2023) show empirically that the standard diagnostic methods cannot separate between AGN and SF galaxies in such low-metallicity systems found in the early universe. Y. Harikane et al. (2023) presented a sample of AGN with $z > 4$ which were all indistinguishable from SF galaxies on the BPT diagram due to being in the region occupied by SFGs. R. Maiolino et al. (2024) identified AGN in the *JWST* data which also would be missed in the standard BPT diagram.

While Type I AGN can still be identified by their broad lines, this makes the identification of Type II NL AGN much more difficult. Therefore, many new diagnostics have been proposed. B. E. Backhaus et al. (2022) suggested that the OHNO diagram, consisting of a plane with line ratios [O III]/Hβ and [Ne III]/[O II], can be a promising alternative. H. Übler et al. (2023) examined the potential of using the He II λ4686 emission line for the discrimination and proposed the diagrams showing the equivalent width (EW) of He II λ4686 versus the ratio He II λ4686/Hβ and He II λ4686/Hβ versus [N II] λ6584/Hα, the latter of which was first suggested by M. Shirazi & J. Brinchmann (2012). While the He II λ4686 emission line is effective in separating AGN and SF galaxies due to being independent of metallicity, it is very faint in even powerful AGN.

Another possibility is to use UV emission lines instead of optical ones to search for AGN. These might be more reliable at high redshifts as they primarily originate from high-ionization states that require the production of hard ionizing photons, and are therefore likely to trace AGN activity. This was explored by A. Feltre, S. Charlot & J. Gutkin (2016), who find that, amongst other lines, C IV $\lambda\lambda$1548,1551, O III] $\lambda\lambda$1661,1666, [C III] λ1907 + C III] λ1909, and He II λ1640 are individually good at separating photoionization by an AGN and star formation, and that possible diagrams using three of these lines can be highly effective.

In this paper, we explore some of the suggested diagnostic quantities and diagrams using photoionization models produced by K. Nakajima & R. Maiolino (2022) and where possible apply them on the *JWST* Advanced Deep Extragalactic Survey (JADES) data (A. J. Bunker et al. 2024). Our main focus is on three sources with NIRSpec (Near InfraRed Spectrograph) IDs 9422, 18846, and 10058975, which exhibit strong UV emission lines. We also comment on the

applicability and sensitivity of these diagrams and explore further diagnostic quantities, for example the possible application of UV–optical ratio diagrams. Apart from the spectroscopic identification, we also investigate a method using photometry, combining spectral energy distribution (SED) fitting and morphology, developed by I. Juodžbalis et al. (2023) to find high probability high-z AGN candidates that can then be followed up with NIRSpec spectroscopy.

The structure of this paper is as follows. Section 2 describes the observational and modelled data sets that were used. Section 3 focuses on the spectroscopic analysis of the data, discussing standard optical methods for AGN identification, along with designing alternative optical and UV diagnostic diagrams. We also detail the calculation of the EWs of emission lines in the spectra of three potential AGN sources and use these measurements to position the galaxies on the discussed diagrams. In Section 4, we carry out a photometric selection process on the JADES data. In Section 5, we examine the sensitivity of the diagnostic diagrams discussed earlier. We summarize our work and its possible extensions and draw our final conclusions in Section 6.

2 DATA

2.1 Observational data

The observational data we use can be categorized into two main groups: spectroscopic and photometric data. The spectroscopic data used throughout this paper is from the JADES released by D. J. Eisenstein et al. (2023a) using the data taken by the *JWST* NIRSpec (P. Jakobsen et al. 2022; P. Ferruit et al. 2022). In particular, we concentrated on the analysis of the 253 sources with detailed spectroscopy from A. J. Bunker et al. (2024). The latter includes a measurement of the spectroscopic redshifts (z) for 178 sources, and provides 1D spectra with a catalogue of line fluxes above a signal-to-noise ratio (S/N) of 5.

We use the 1D spectra and catalogue of the Prism/Clear data with main focus on three of the sources with NIRSpec IDs 9422, 18846, and 10058975, which exhibited strong UV emission lines: C IV λλ1549, He II λ1640 + O III] λλ1661,1666, and C III] λλ1909, in contrast to other observed objects. These lines are C IV, He II + O III] blend (or He II λ1640 and O III] λ1665 if mentioned separately), and C III], respectively, from here on. C IV and the blended He II + O III] lines are reported in all three sources with an S/N > 5, however, the C III] emission line of source 9422 is not reported in A. J. Bunker et al. (2024). To measure the emission-line flux from the spectrum of the galaxy, we use the Line Measuring package: L_IM_E. It fits the continuum and line flux to calculate the integrated flux of the line (for more details, see V. Fernández et al. 2024).

These galaxies are amongst those with the highest redshifts in the total sample, with 10058975 showing the highest redshift (z = 9.438) amongst the three (00018846 with z = 6.342 and 00009422 with z = 5.943). The spectra of the three objects, 00009422 (or 9422), 00018846 (or 18846), and 10058975, are shown in Fig. 1 with the grey dashed lines denoting important emission lines. Source 9422 was identified spectroscopically as an AGN by J. Scholtz et al. (2025), S. Tacchella et al. (2025), and Y. Li et al. (2024), but rejected as an AGN in A. J. Cameron et al. (2024). 10058975 was also identified as an AGN through spectroscopy by J. Scholtz et al. (2025). Galaxy 18846, to the best of our knowledge, has not been previously identified as an AGN.

The various identified AGN sources displayed on the optical diagrams in Sections 3.1 and 3.2.1 were collected from the literature. They are mostly confirmed as AGN based on their broad line emission (R. Maiolino et al. 2024; Y. Harikane et al. 2023; H. Übler et al. 2023; D. D. Kocevski et al. 2023; R. L. Larson et al. 2023a; V. Kokorev et al. 2023). Note that on these diagrams, we also display JADES data points from the Prism/Clear data catalogue, as well as from the R = 1000 grating catalogue.

The photometric data are from the *JWST* Near Infrared Camera (NIRCam, second data release, M. J. Rieke et al. 2023; D. J. Eisenstein et al. 2023a, b). We used the following filters from the NIRCam photometric catalogues for our objects: *F*090*W*, *F*115*W*, *F*150*W*, *F*200*W*, *F*277*W*, *F*356*W*, *F*444*W*, *F*335*M*, and *F*410*M*. All of these use the CIRC2 measurement (associated with a 0.15 arcsec radius circular aperture).

2.2 Modelled data

The theoretically modelled data sets for AGN and SF galaxies are an extended sample of the models in K. Nakajima & R. Maiolino (2022) created using CLOUDY photoionization model calculations (version 13.05; G. J. Ferland et al. 1998, 2013). These data are available in spectral form, and also in a tabulated form containing values of emission-line fluxes and some EWs.

For the SF galaxies, the SEDs are from Binary Population and Spectral Synthesis stellar evolution models (BPASS v2.2.1; J. J. Eldridge et al. 2017; E. R. Stanway & J. J. Eldridge 2018) which are based on the P. Kroupa (2001) initial mass function with the upper cut on the mass at both 100 and 300 solar masses, used in various diagnostic plots. These adopt the stellar ages 1 and 10 Myr. The SF galaxy sample used in this project contains metallicities (Z) ranging from about 0.0007 to 1 $Z_\odot$ and ionization parameters ($\log U$) from −1 to −3.

K. Nakajima & R. Maiolino (2022) use AGN models with the CLOUDY 'AGN' continuum command consisting of the 'Big Bump' as well as power-law components. These components are parametrized by the temperature of the bump (T_{bb}), ranging from 5×10^4 to 2×10^5, and the power-law index (α) which varies between −1.2 and −2. In the AGN sample used in this project, metallicities range from about 0.0007 to 2 $Z_\odot$, and ionization parameters range from −0.5 to −2.5. The full list of parameters, metallicities, and ionization values of the models we use are listed in Table 1. These models are specifically applicable to Type II AGN or the narrow-line component observed in Type I AGN.

3 SPECTROSCOPY-BASED DIAGNOSTICS

In this section, we discuss diagnostic methods relying on spectroscopy, starting with the standard BPT and VO87 diagrams that have been widely used to identify galaxies containing an AGN. After the standard diagnostic methods, we discuss the OHNO diagram and the use of the [O III] λ4363 auroral line, which have been suggested as optical alternatives to the BPT and VO87 diagrams for high-redshift systems.

We then explore the use of UV line fluxes and EWs for separating AGN and SFG, and apply some of the plots on our sources of interest. This is one of the main points of this paper – developing a new method for locating systems which contain AGN. Furthermore, after discussing the sensitivity of some of these possible diagnostics, we explore the possibility of using combined optical–UV line flux ratios considering the magnitude of the dust correction factor.

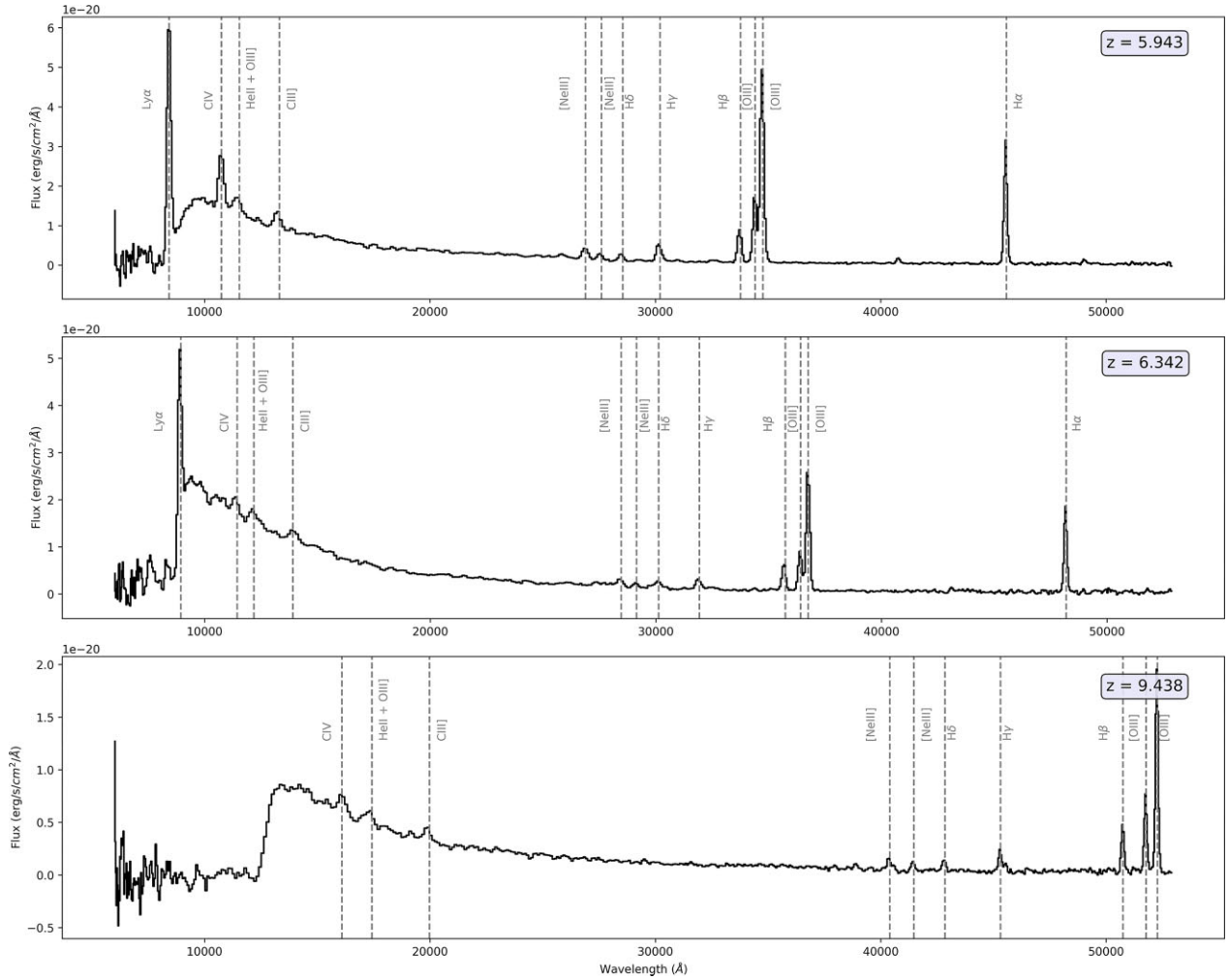

Figure 1. The NIRSpec spectra of galaxies 9422, 18846, and 10058975 are shown with spectral flux in units of erg s cm^{-2} Å^{-1} on the y-axis and wavelength in units of angstroms on the x-axis. The grey dashed lines are important emission lines found in the spectra. These spectra are from the JADES survey (A. J. Bunker et al. 2024).

Table 1. Parameters from the extended sample of the models in K. Nakajima & R. Maiolino (2022) for AGN and SFG (with α denoting the power-law index, $T_{\rm bb}$ the temperature of the bump component, Z the metallicity, and log U the ionization parameter).

AGN models		SFG models	
$\boldsymbol{\alpha}$	$-1.2, -1.6, -2.0$	Mass upper cut	100 $M_\odot$, 300 $M_\odot$
$\boldsymbol{T_{\rm bb}}$	5×10^4 K, 1×10^5 K, 2×10^5 K	Stellar age	1 Myr, 10 Myr
Z	1×10^{-5}, 0.0001, 0.00014, 0.00028, 0.0007, 0.001, 0.0014, 0.0028, 0.007, 0.014, 0.028 or in solar metallicity ($Z_\odot = 0.014$): ∼0.0007, ∼0.007, 0.01, 0.02, 0.05, ∼0.07, 0.1, 0.2, 0.5, 1, 2	Z	1×10^{-5}, 0.0001, 0.00014, 0.00028, 0.0007, 0.001, 0.0014, 0.0028, 0.007, 0.014 or in solar metallicity ($Z_\odot = 0.014$): ∼0.0007, ∼0.007, 0.01, 0.02, 0.05, ∼0.07, 0.1, 0.2, 0.5, 1
log U	Table: $-0.5, -1, -1.5, -2, -2.5$ spectra: $-0.5, -1.5, -2.5$	log U	Table: $-1, -1.5, -2, -2.5, -3$ spectra: same as table

Notes. Note that for the AGN models, there are a few ionization values that are included in the table data, but not in the spectral data. If a method uses both types of the data, then only values included in both the tabular and the spectral data are used.

3.1 The weakness of standard optical diagrams

As previously mentioned, the BPT (J. A. Baldwin et al. 1981) and VO87 (S. Veilleux & D. E. Osterbrock 1987) diagrams have been the standard methods for identifying AGN. In this section, we research the N2–BPT, and S2–VO87 diagrams, which display the line flux ratio [O III] λ5007/Hβ against the ratios [N II] λ6584/Hα and [S II] λλ6716,6731/Hα, respectively.

Ratios for the models were calculated using the tabulated data of the models. The parameters of these models are as described previously with the exception that we do not include the lowest metallicity (0.0007 $Z_\odot$) SFG and AGN models since they fall outside of the relevant region for the data presented here. They are located far below the demarcation line in the mixed region.

The N2–BPT diagram is shown on the left side of Fig. 2. The photoionization models of SFGs and AGNs are indicated with blue

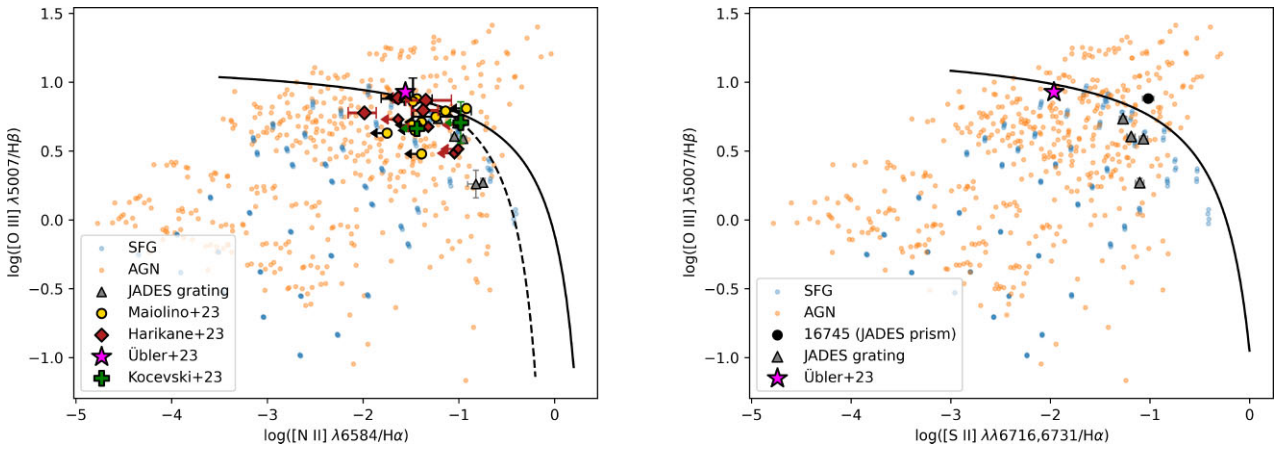

Figure 2. N2–BPT (on the left) and S2–VO87 (on the right) diagrams displaying photoionization models of AGNs (orange) and SFGs (blue) created by K. Nakajima & R. Maiolino (2022) and various AGN sources from the literature (R. Maiolino et al. 2024; Y. Harikane et al. 2023; H. Übler et al. 2023; and D. D. Kocevski et al. 2023). The solid lines are extreme starburst lines defined by L. J. Kewley et al. (2001) and the dashed classification line is from G. Kauffmann et al. (2003). Additionally, the grey triangles present JADES data points from the $R = 1000$ grating catalogue (A. J. Bunker et al. 2024). AGNs are identified above the solid and dashed lines. A source in the Prism/Clear data catalogue with NIRSpec ID 16745, which contains all the necessary lines with a minimum signal-to-noise ratio of 5, is located above the maximum starburst line on the S2–VO87 diagram. J. Scholtz et al. (2025) selected this source as an AGN using similar considerations.Narrow-line fluxes were used to create these plots for all sources, except for source 16745, where the total flux was used.

and orange points, respectively. The solid line is the extreme starburst line defined by L. J. Kewley et al. (2001). We can see that this line agrees well with the AGN and SFG models displayed on the plot. Almost all the modelled data points above the line are AGN (some SFG models with metallicities 0.5 and 1 $Z_{\odot}$ are slightly above the line), whereas both AGN and SFG points can be seen below the line. The dashed classification line was defined by G. Kauffmann et al. (2003) and it has been used to describe a composite region in the parameter space between the two lines (L. J. Kewley et al. 2006). However, a mixed region can be seen even below the G. Kauffmann et al. (2003) line. In this mixed region, the location of AGN models is consistent with what is expected for AGN hosted in metal-poor galaxies (R. Maiolino et al. 2024).

The actual AGN sources displayed on the N2–BPT diagram are from R. Maiolino et al. (2024), Y. Harikane et al. (2023), H. Übler et al. (2023), and D. D. Kocevski et al. (2023). R. Maiolino et al. (2024) present 12 AGNs at redshifts $4 < z < 7$ from the JADES survey identified by their broad Balmer lines. The 10 sources from Y. Harikane et al. (2023) are also BL AGN at $z > 4$. H. Übler et al. (2023) identify galaxy GS 3073 at $z = 5.55$ as an AGN not only from its broad lines, but also from its high He II λ4686 EW. The position of the data point was calculated using the narrow-line fluxes. Finally, D. D. Kocevski et al. (2023) report two BL AGNs, CEERS 1670 and CEERS 3210, from the Cosmic Evolution Early Release Science (CEERS) Survey at $z > 5$. We use the narrow-line flux values of the emission lines to position these sources on our plots.

We can see that most of these sources would be missed as AGN by the use of the N2–BPT diagram if they did not display broad lines. They are located in a region of the parameter space which has been thought to be populated by SFGs and which is shown to be a mixed region of SFG and low-metallicity AGN models.

Since the Hα and [N II] λ6584 lines are blended in the prism spectra, it was not possible to determine the line ratio, and hence we could not display sources from the Prism/Clear data catalogue (A. J. Bunker et al. 2024) in the N2–BPT diagram. We were able to display sources from the $R = 1000$ grating catalogue, However, these sources could not be identified as purely AGN using this diagnostic diagram, as they fall within the mixed region below the demarcation line.

On the right side of Fig. 2, we plot the S2–VO87 diagram. As before, the AGN and SFG models are represented with orange and blue points, respectively. The solid black line is the extreme starburst line from L. J. Kewley et al. (2001) and we can see that, similarly to the N2–BPT diagram, this line agrees well with the modelled data. Above the line we mostly find AGN models (some SFG models with metallicities 0.5 and 1 $Z_{\odot}$ are slightly above the line), while below it we can see both AGN and SFG models.

The galaxy from H. Übler et al. (2023) is represented the same way as on the N2–BPT diagram and is placed using the narrow-line fluxes. We can see that this AGN would be missed by the S2–VO87 diagram. The black data point is from the JADES data set and was the only system with all the necessary lines in the prism catalogue to display on the plot. It uses the total flux values of the emission lines. It is located above the demarcation line, meaning that it can be identified as an AGN. Its NIRSpec ID is 16745, and it was selected as an AGN by J. Scholtz et al. (2025) using this diagram. The grating data points are below the demarcation line, so they would not be identified as AGN by this diagram.

Fig. 2 shows why new identification methods are needed if we want to be able to effectively characterize high-redshift galaxies. The BPT and VO87 diagrams, while still very useful in some circumstances, will miss many potential high-redshift active galaxies amongst the plethora of data the *JWST* provides.

3.2 Alternative optical line flux ratio diagrams

This section explores two optical diagnostic methods which are suggested in recent previous work as alternatives of the standard methods. These are the OHNO diagram and the use of the [O III] λ4363 auroral line.

3.2.1 OHNO diagram

In this section, we examine newer methods proposed to find AGN, one of which is the so-called OHNO method. The OHNO diagram and demarcation line was suggested by B. E. Backhaus et al. (2022) and uses the line ratios [O III] λ5007/Hβ and [Ne III] λ3869/[O II] λλ3727,3729.

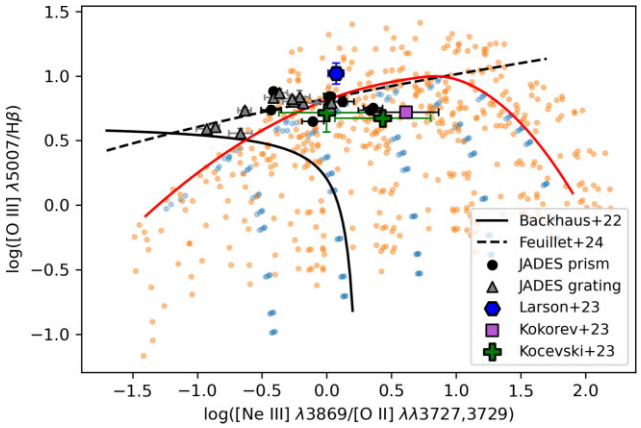

Figure 3. The OHNO diagram displaying SFG and AGN models from K. Nakajima & R. Maiolino (2022, coloured as previously), the solid and dashed black demarcation lines are based on the work of B. E. Backhaus et al. (2022) and L. M. Feuillet et al. (2024), respectively. Our newly defined solid red separation curve is defined by equations (1)–(4), and it separates a mixed region from a region only occupied by AGN models. BL AGN collected from literature (R. L. Larson et al. 2023a; V. Kokorev et al. 2023; D. D. Kocevski et al. 2023), as well as sources from the JADES Prism/Clear and $R = 1000$ grating data catalogues can be seen. Narrow-line fluxes were used to create this plot with the exceptions of the data points from R. L. Larson et al. (2023a) and the JADES PRISM where the total line flux was used.

The OHNO diagram can be seen plotted on Fig. 3. B. E. Backhaus et al. (2022) created the OHNO AGN/SF line to empirically produce similar separation between AGN and SF galaxies to another diagram called unVO87 displaying the line ratios [O III] λ5007/Hβ and [S II]/Hα + [N II]. This solid black demarcation line is meant to divide the parameter space, such that AGN sources are above it. However, it does not agree with the displayed theoretical models. Based on the models, the demarcation line does not seem to show a separation between AGN and the mixed region as the ones on the NII–BPT and SII–VO87 diagrams do.

The dashed black line is another demarcation line defined by L. M. Feuillet et al. (2024). This shows a better separation between the models, though it still fails to perfectly distinguish the regions. Our separation curve, represented by a solid red line, was determined by visually inspecting the regions. It separates a mixed region from a region exclusively occupied by AGN models. It follows the equation

$$\log\left(\frac{[\mathrm{O\,III}]}{\mathrm{H}\beta}\right) = (1-w)\,f_1 + w f_2 \tag{1}$$

when $-1.4 < \log([\mathrm{Ne\,III}]/[\mathrm{O\,II}]) < 1.9$. Logistic blending is used to smoothly transition from function f_1 to f_2. The transition is done by the sigmoid weighting function w at $\log([\mathrm{Ne\,III}]/[\mathrm{O\,II}]) = 0.8$. The functions f_1, f_2, and w are defined as follows:

$$f_1 = 1 - 0.19\left(\log\left(\frac{[\mathrm{Ne\,III}]}{[\mathrm{O\,II}]}\right) - 0.99\right)^2, \tag{2}$$

$$f_2 = 1 - 0.75\left(\log\left(\frac{[\mathrm{Ne\,III}]}{[\mathrm{O\,II}]}\right) - 0.80\right)^2, \tag{3}$$

and

$$w = \frac{1}{1+\exp\left(-10\left(\log\left(\frac{[\mathrm{Ne\,III}]}{[\mathrm{O\,II}]}\right) - 0.8\right)\right)}. \tag{4}$$

Four previously identified AGN from R. L. Larson et al. (2023a), V. Kokorev et al. (2023), and D. D. Kocevski et al. (2023) are displayed on Fig. 3. R. L. Larson et al. (2023a) identified CEERS 1019 as an AGN based on its broad lines, the presence of weak high-ionization lines, and a spatial point-source component. V. Kokorev et al. (2023) identify another BL AGN at $z = 8.5$ from the *JWST* Cycle 1 UNCOVER (Ultradeep NIRSpec and NIRCam ObserVations before the Epoch of Reionization) Treasury survey with high ionization lines and a point-source morphology. Amongst these, only the galaxy from R. L. Larson et al. (2023a) can be identified as an AGN based on this diagnostic diagram.

Furthermore, we can see nine sources from the JADES prism catalogue displayed as black data points. Amongst them is one of our galaxies of interest with NIRSpec ID 9422, however, it is located in the mixed region. Two of these data points are located above our demarcation line: 16745 and 10013620. The source 16745 has been selected as an AGN candidate by J. Scholtz et al. (2025), but galaxy 10013620 has not been identified as an AGN candidate to the best of our knowledge. This is the black data point which is above our separation curve and located on the black dashed demarcation line produced by L. M. Feuillet et al. (2024).

Lastly, we also present data points from the JADES $R = 1000$ grating catalogue. These are displayed as grey triangles and most of them are outside of the mixed region, therefore could be potential AGN candidates. The NIRSpec IDs of the sources in the region of interest are as follows: 22251, 10008071, 4270, 7938, 18090, 3892, 10009506, 10011849, 8891, and 3753.

Narrow-line fluxes were used to create this graph, except for the data points from R. L. Larson et al. (2023a) and the JADES PRISM, where the total line flux was used.

3.2.2 *The [O III] λ4363 auroral line*

In a recent study, the use of the [O III] λ4363 auroral line, which has been observed in several *JWST* sources, was presented as a possible alternative method to locate AGN by G. Mazzolari et al. (2024; see also J. Brinchmann 2023; H. Übler et al. 2023).

G. Mazzolari et al. (2024) proposed that a high [O III] λ4363/Hγ ratio is sufficient to identify the presence of an AGN, and suggested three diagnostic diagrams displaying this ratio against other ratios: [O III] λ5007/[O II] λλ3727 (O32), [Ne III] λ3869/[O II] λλ3727 (Ne3O2), and [O III] λ5007/[O III] λ4363 (O3O3). They supported these suggestions by not only empirical calibrations, but with a wide range of theoretical models. The usefulness of these diagnostics is based on the intensity of the [O III] λ4363 line relative to the Hγ line. This is influenced by the temperature and metallicity of the interstellar medium (ISM), as well as the ionization parameter (see G. Mazzolari et al. 2024, for more detail).

The [O III] λ4363 auroral line is a very promising alternative to the standard AGN diagnostic methods, however, we note that this line is quite hard to observe in the data we are considering. In our sample, we identify three sources – ID 10058975, 0021842, and 9422 – with secure detections of this line, based on the line flux catalogue. However, when comparing these detections to the criteria outlined in G. Mazzolari et al. (2024), we find that none of these galaxies meet the thresholds required to be classified as purely AGN based on their [O III] λ4363 emission. All three lie below the demarcation lines, indicating that they could plausibly be either AGN or SFG.

3.3 UV alternatives

3.3.1 *Using UV emission-line ratios*

Lines such as C IV and He II λ1640 have high potential for separating stellar population and AGN since they require extremely high energy

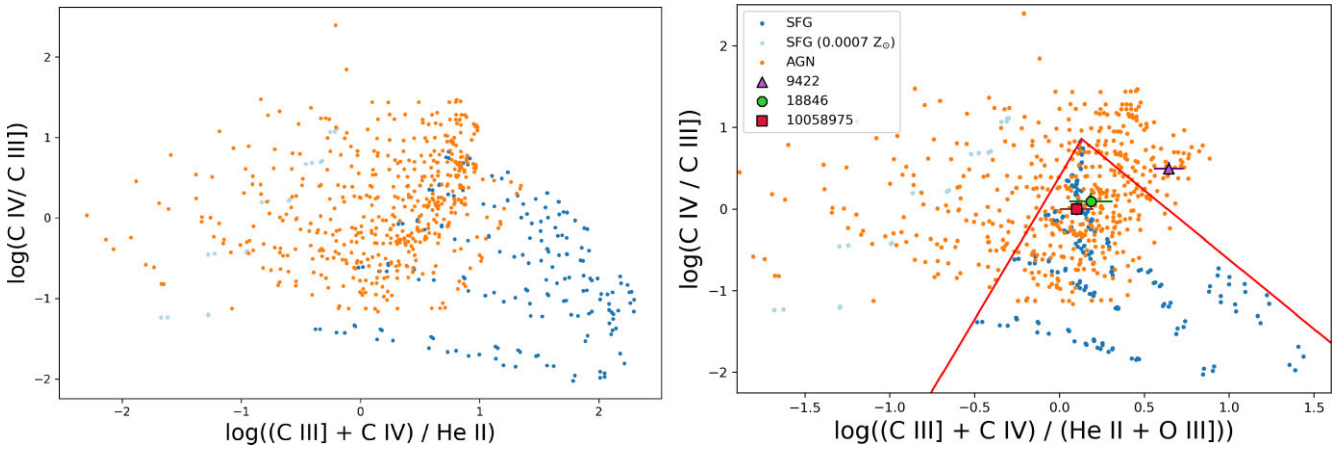

Figure 4. UV emission-line flux ratio diagrams created based on the models from K. Nakajima & R. Maiolino (2022) presented in Table 1. As in Fig. 2, the AGN models are represented with orange dots and the SF galaxy models with blue dots. The lighter blue points represent the SF galaxies model with very low metallicity (0.0007 $Z_{\odot}$). This metallicity is 3 orders of magnitude lower than metallicities of AGN models in the same region. Our galaxies of interest (with NIRSpec IDs 9422, 18846, and 10058975) are represented using three different colours (purple, green, and red). The red lines are defined by equations (5) and (6).

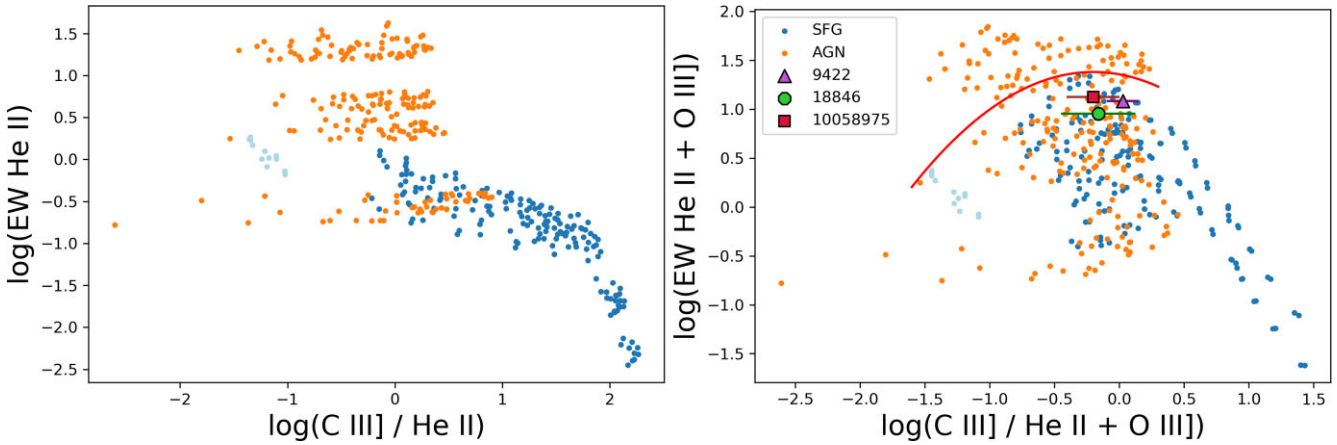

Figure 5. The EWs of the He II λ1640 and He II λ1640 + O III] lines can be seen plotted against the emission-line flux ratios C III]/He II λ1640 and C III]/(He II λ1640 + O III]) blend. The plot to the right was created by replacing the He II λ1640 line with the blended He II λ1640 + O III] line. The blending of He II λ1640 and O III] in the NIRSpec prism data reduces the diagnostic power of He II λ1640 as an AGN indicator, except for AGN with the highest He II λ1640 EW. The red curve (on the right) was computed based on equation (7).

ionizing photons. In particular, the He II λ1640 line is very sensitive to this ionization which is produced more dominantly by AGN than by stars, and thus line ratios using He II λ1640 were suggested to be useful in separating SF galaxies and AGN (A. Feltre et al. 2016; J. Gutkin, S. Charlot & G. Bruzual 2016).

A. Feltre et al. (2016) stated that the line ratios C IV/C III], C IV/He II λ1640, and C III]/He II λ1640 can help separate AGN and SF galaxies very well. In their model spectra, the C IV/He II λ1640 and C III]/He II λ1640 ratios distinguish between active and inactive galaxies even on their own, while the C IV/C III] ratio could do the separation with the help of an additional ratio.

A paper by K. Nakajima et al. (2018) also discusses possible UV diagnostic diagrams involving the lines C III], C IV, and He II λ1640. A diagnostic diagram using the line ratios C IV/C III] and C III] + C IV/He II λ1640 presented by them is called the C4C3–C34 diagram. They showed that AGN and SFG models (A. Feltre et al. 2016; J. Gutkin et al. 2016) show a remarkable difference in the C III] + C IV/He II λ1640 parameter (C34), because for AGN this value is smaller than for SFG considering a fixed value of C IV/C III] (C4C3).

Though these line ratios are very promising, in the observational data from the NIRSpec prism, He II λ1640 and O III] λ1665 are not present on their own, only as a blended line. We therefore considered whether the blended line could still provide some information and we explored these UV line ratios using the tabulated modelled data presented in Section 2.2.

As before, the plots show the AGN and SFG models in orange and blue, respectively, while the lighter blue points are the SFG model with very low metallicities (0.0007 $Z_{\odot}$). It can be seen on the left side of Fig. 4 that if the lowest metallicity SF galaxies are disregarded, then the (C III] + C IV)/He II λ1640 ratio indeed shows a very good separation between AGN and SF galaxies, even on their own. The lowest metallicity SF galaxies exhibit ratios similar to AGN, not simply due to weak metal emission lines, but also because their He II λ1640 emission can be stronger than the higher metallicity SF galaxies (as seen in Fig. 5), leading to an overlap in this diagnostic space.

Replacing He II λ1640 in the ratio presented on the left side of Fig. 4 with the blended line, an additional plot was created.

This can be seen on the right of Fig. 4, where the ratio (C III] + C IV)/(He II + O III])(CHeO) is presented on the *x*-axes. We call this new diagnostic diagram C4C3–CHeO. It is clear that on this plot the region of parameter space that is uniquely occupied by AGN is reduced. This means that the range of AGN and SFG that can be identified from this diagram is more limited. There can still be many cases, however, where we can make an SFG–AGN separations as two distinct regions remain visible. Galaxies with:

$$\log\left(\frac{\text{C IV}}{\text{C III]}}\right) > 3.5\log\left(\frac{\text{C III]} + \text{C IV}}{\text{He II} + \text{O III]}}\right) + 0.395 \quad (5)$$

when log((C III] + C IV)/(He II + O III])) < 0.13 and

$$\log\left(\frac{\text{C IV}}{\text{C III]}}\right) > -1.7\log\left(\frac{\text{C III]} + \text{C IV}}{\text{He II} + \text{O III]}}\right) + 1.08 \quad (6)$$

when log((C III] + C IV)/(He II + O III])) > 0.13 can be reliably identified as AGN, while those with values outside these thresholds fall into a mixed region that can be either SFG or AGN. The AGN in the upper right region primarily consist of those with the highest ionization parameters in the model (log$U$ = −0.5 and −1.0), with most AGN with log$U$ = −0.5 falling entirely within the region. The AGN in the left region primarily consist of those with the lowest ionization parameters in the model (log$U$ = −2.5).

This figure also displays the *JWST*/NIRSpec data sample on the diagram. As can be seen, the only three galaxies from our sample that have the needed UV lines are our galaxies of interest. The remaining galaxies did not achieve a sufficient S/N ratio, and therefore, do not appear in the emission-line flux catalogue. C IV and the blended He II + O III] lines were detected in all three sources with an S/N > 5. However, the C III] emission line was not reported for source 9422 so we use $L_I M_E$ to calculate the emission-line flux. Unfortunately, galaxies 18846 and 10058975 fell within areas of the diagnostic plot where clear identification was not possible. In contrast, galaxy 9422 was located in the AGN-only region, suggesting it could be classified as an AGN.

While this diagnostic offers a promising alternative to standard diagnostic methods, we note that the emission lines constituting these ratios are difficult to observe in the existing *JWST* data and would require additional observations, such as longer exposure time data, to be used effectively.

3.3.2 *Using UV EWs and line ratios*

K. Nakajima et al. (2018) not only discusses UV line ratios, but also suggests that the EWs of these UV lines could potentially be used as diagnostic parameters. This paper shows that the EWs of C IV and C III] plotted against the ratios C IV/He II λ1640 and C III]/He II λ1640, respectively, provide a very good separation between AGN and SF galaxies.

K. Nakajima & R. Maiolino (2022) then complemented this study by using the EW of He II λ1640. They found that two distinct regions could be seen when using the EW of He II against the ratios of either C III]/He II or O III]/He II. Unfortunately, we are unable to use these graphs with the prism data because of the blended He II + O III] line. In this section, we show how the plots change when using the blended He II + O III] line instead.

The line ratios were calculated using the tabular form of the modelled data. The EWs were calculated from a mix of the modelled spectra and the table data. The continuum was fitted by applying the fit_generic_continuum function of the specutils PYTHON package on manually selected regions around the emission lines in the spectral models.

The EWs of He II λ1640 and the blended He II + O III lines were considered and plotted against the line ratios C III]/He II λ1640, and C III]/(He II + O III), respectively. These can be seen in Fig. 5.

We can see on the left side of Fig. 5 that the EW of He II λ1640 creates three regions, one with only AGN which have the highest $EW_{He\,II}$, a region with only SFG possessing the lowest $EW_{He\,II}$ values, and a mixed region in the middle where the lowest metallicity SFG are located. Because of its high ionization potential, the recombination line He II λ1640, even on its own, is a very good diagnostic parameter to separate AGN from SF galaxies. Furthermore, the ratio C III]/He II λ1640 shows a very good separation even on its own where the SFG have the highest values and the AGN the lowest. For SFG, the He II λ1640 emission-line flux is smaller than for AGN, causing the ratio to get bigger.

These quantities could be used together as a diagnostic diagram, but as stated before, in the observational data He II λ1640 and O III] λ1665 are not present on their own, so these plots could not be applied here. A similar plot replacing the He II λ1640 line with the blended He II + O III] can be seen on the right side of Fig. 5, together with the original plot on the left. This shows significantly more mixing between AGN and SFG models than the plot that uses the separately detected lines. There is, however, still a region which would allow for the identification of AGN candidates, but this is much more limited. We define this AGN region as the part of the parameter space above the red curve defined by the following equation:

$$\log(\text{EW He II} + \text{O III]}) = 1.38 - 0.6\left(\log\left(\frac{\text{C III]}}{\text{He II} + \text{O III]}}\right) + 0.2\right)^2 \quad (7)$$

when $-1.6 < \log(\text{C III]}/\text{He II} + O\ \text{III]}) < 0.3$.

This figure also displays our galaxies of interest 9422, 18846, and 10058975. The observed EW values from the *JWST* data were converted to the rest frame equivalents using a division by $(1 + z)$. This was consistently done throughout the rest of the paper. Unfortunately, the three galaxies are in the mixed region making it impossible to clearly identify them.

3.4 EW diagrams using both optical and UV lines

As outlined previously, it is necessary to explore other methods for separating SF and AGN systems to complement diagnostics based only on flux ratios. Consequently, we also focus on the design and analysis of diagnostics based on the comparison of EWs. K. Nakajima et al. (2018) define separating regions based on the EWs of C III] and C IV and K. Nakajima & R. Maiolino (2022) on the EW of He II λ1640. In our case, we complemented the study of these emission lines with the addition of EWs of other emission lines in the optical range (such as Hβ, He II λ4686). The other diagrams we considered can be found in Appendix A, Fig. A1 and show an exhaustive list of possible diagnostic diagrams using EWs on both axes.

The plot that separates the regions the best can be seen on Fig. 6. It uses the EW of Hβ versus the EW of He II λ4686. Hα would give a similar plot, since Hβ and Hα are very similar, but Hβ is preferred due to Hα shifting out of the *JWST*/NIRSpec wavelength range at high redshifts ($z = 7$ versus 10). In the models presented by K. Nakajima & R. Maiolino (2022), the spectra consist of the NLR emission combined with the contribution from the accretion disc. This likely leads to lower EWs of Hβ in AGN models compared to those in SFG models due to the additional continuum contribution. The He II λ4686 recombination line was chosen because it appears to be a promising diagnostic due to being independent of the metallicity. It has been explored and used by M. Shirazi & J. Brinchmann (2012); J. Scholtz et al. (2025); G. Tozzi et al. (2023); H. Übler et al. (2023); and O. L. Dors et al. (2023). We define a separation line to distinguish

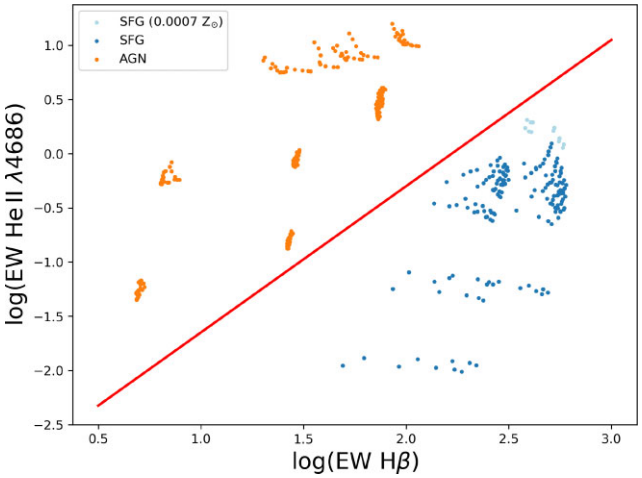

Figure 6. The modelled data on a plot displaying the EW of Hβ against the EW of He II λ4686. The red line separates the AGN and SFG regions and is defined by equation (8).

AGN from SFG as follows:

$$\log(\text{EW He II } \lambda4686) > 1.35 \log(\text{EW H}\beta) - 3. \quad (8)$$

The problem with this graph, however, is that the He II λ4686 line is very faint, so it was not detected with an S/N $>$ 5 in the observational data. An upper limit calculation provides constraints on the sources containing an AGN, however, J. Scholtz et al. (2025) stated that this approach did not give any promising results. Even though this plot cannot be used for this data, it is worth to consider what would be needed in order to better observe He II λ4686 (see Section 5).

Amongst the plots that only use UV emission-line EWs, the most promising is the one using the C IV line and the He II λ1640 line if detected separately in the observational data as we discussed in Section 3.3.2. The He II + O III] blend is reported in the data and will later be used in Section 3.5 with our galaxies of interest.

The EW of an emission line measures its strength relative to the surrounding continuum, which comes from starlight, the accretion disc, and nebular gas in both the BLR and NLR. In the BLR, additional continuum from the unobscured accretion disc reduces the broad-line EW, while in the NLR, starlight contributes to the continuum, lowering the narrow-line EW. When analysing AGN, line ratios offer a more reliable diagnostic tool compared to EWs, particularly in Type II AGN configurations. The models of K. Nakajima & R. Maiolino (2022) consist of the NLR emission combined with the accretion disc component as previously mentioned. The presence of the accretion disc component raises the continuum level, leading to a lower EW. However, observationally we aim to identify Type II AGN, which have their BLR (and accretion disc) component obscured. This discrepancy results in artificially low EWs shown by the theoretical models, as in reality the nebular lines are measured against a diminished continuum causing higher EWs. However, the line ratios remain unaffected since they depend only on the relative strengths of the emission lines.

3.5 EWs of galaxies 9422, 18846, and 10058975

In this section, we examine real spectroscopy from galaxies observed with NIRSpec. Whilst there is extensive spectroscopy of distant galaxies with *JWST*, there are few galaxies observed to date that have all of the lines in emission that can be used in the diagnostics we have mentioned. We use the spectra to determine whether these galaxies are AGN and how EWs can be used to determine this.

EW measurements of some UV and optical lines in the spectra of galaxies 9422, 18846, and 10058975 with *JWST* observations are discussed. Two optical and three UV lines are used: Hβ, [O III] λ5007, C IV, C III], and the blended He II + O III] line.

The continuum around each emission line in the galaxies was manually selected and fitted using the SPECUTILS (N. Earl et al. 2023) PYTHON package. The observed EWs were calculated by dividing the line flux of the emission line in question by the obtained continuum level.

Finally, using the rest-frame EW measurements, the three galaxies were displayed on some of the EW versus EW graphs we explored. Some of these plots can be seen in Fig. 7, and some other diagrams we considered can be found in Appendix A in Fig. A2.

We display solid red lines on the diagrams aiming to separate a mixed region from a region only occupied by AGN models. The first line (top left of Fig. 7) follows the equation

$$\log(\text{EW [O III]}\lambda5007) = 30 \log(\text{EW H}\beta) - 59. \quad (9)$$

Only SFG models are to the right of this line, and mostly AGN to its left.

The second red separation line (upper right of Fig. 7) follows the equation

$$\log(\text{EW He II + O III]}) = 0.45 \log(\text{EW C III]}) + 0.85. \quad (10)$$

Note that this red line is drawn for $-1.3 < \log(\text{EW C III]}) < 1.4$. Sources above this line are in a region fully occupied by AGN models.

The third separation line (bottom left of Fig. 7) is defined by the equation

$$\log(\text{EW C IV}) = 2.3 \log(\text{EW H}\beta) - 4.2, \quad (11)$$

such that we only see AGN models above the line.

The final separation line (bottom right of Fig. 7) is defined by the equation

$$\log(\text{EW He II + O III]}) = 2.4 \log(\text{EW H}\beta) - 4.1, \quad (12)$$

which completely separates the AGN and SFG regions of the plot.

On the plot using [O III] λ5007 and Hβ, two optical lines, all three galaxies seem to be in the SFG region. Looking at Hβ individually, values indicate that the galaxies are more similar to SFG models, especially galaxy 9422 which is the furthest from the AGN region. This is true for all other plots using Hβ. Galaxy 9422 exhibits the highest [O III] λ5007 EW of the three, closely aligning with the SFG models on both axes. In contrast, the other two galaxies lack a sufficient [O III] λ5007 EW to be classified as SFG based on that axis alone.

The plot using the He II λ1640 + O III] blend and C III] positions the galaxies in a mixed region.

On the C IV over Hβ graph, the Hβ EW shows the same as previously. The C IV EW (similarly as C III]) places the galaxies near the highest values of the SFG models, and positions galaxy 9422 even higher, indicating the possibility of it being an AGN.

The He II λ1640 + O III] blend versus Hβ plot positions galaxies 18846 and 9422 in the SFG region, while placing galaxy 10058975 on the separation line between the two regions.

Even though these diagrams were not able to conclusively identify our three sources of interest as AGN, they clearly have the potential to be very useful due to the distinct AGN and SFG regions in the new parameter space.

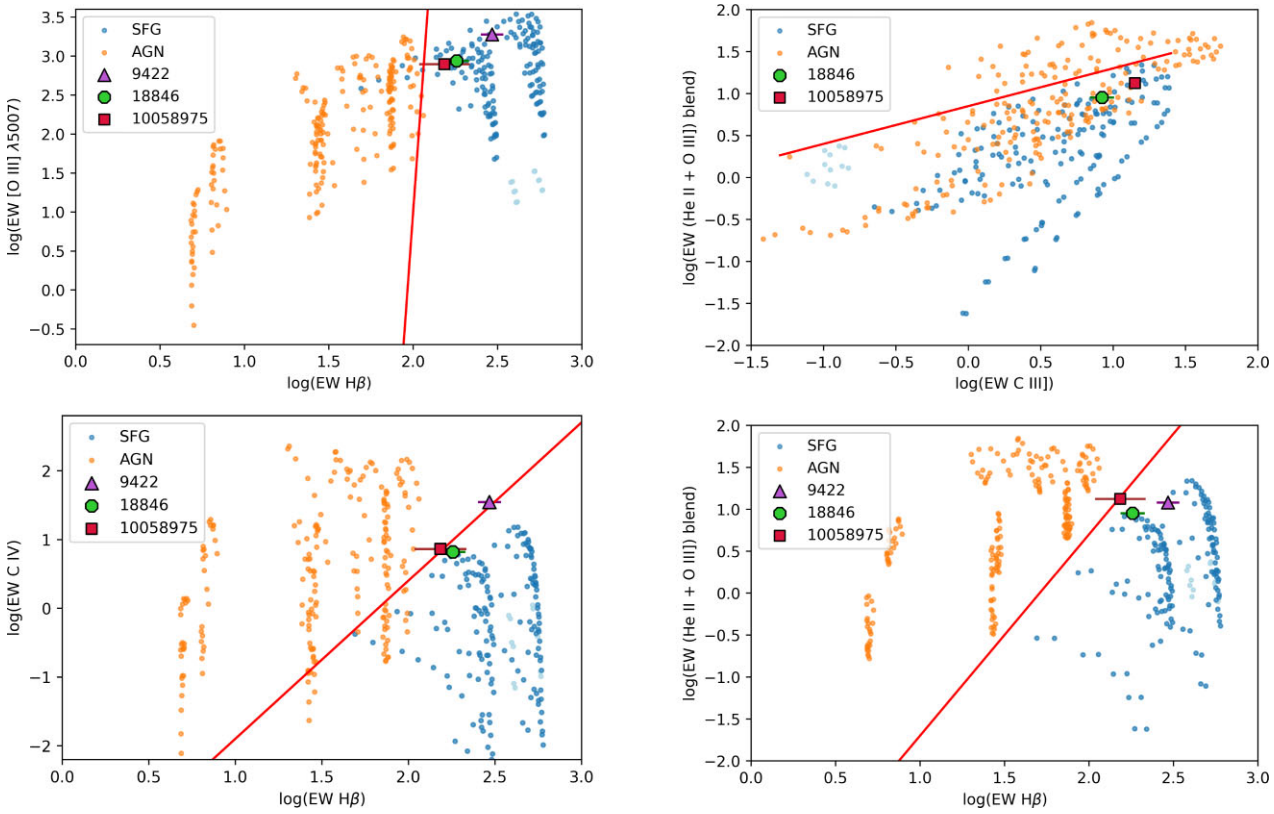

Figure 7. The galaxies with NIRSpec IDs 9422, 18846, and 10058975 are displayed on some EW versus EW plots using their rest-frame EW measurements and uncertainties. The first plot (top left) only uses optical lines, the second (top right) only UV lines, and the two bottom ones use one UV and one optical line. The red solid separation lines are defined by equations (9)–(12).

3.6 Possible use of UV–optical ratio diagrams

In this section, we discuss using UV–optical ratios: He II λ1640/Hβ versus (C III] + C IV)/Hβ. The utilization of UV and optical lines in the same ratio is not straightforward. Dust affects light differently depending on the wavelength (D. Calzetti et al. 2000). To combine them in the same graph, we would need to calculate the dust correction factor ($K_{\rm dust}$) and check if the factor is comparable to the separation between the models.

In order to estimate this dust correction factor, we used the dust attenuation curve from D. Calzetti et al. (2000) as well as the SMC (Small Magellanic Cloud) bar curve (K. D. Gordon et al. 2003) with the FM90 parametrization defined in E. L. Fitzpatrick & D. Massa (1990) for both the UV and the optical emission lines. We calculated $K_{\rm dust}$ assuming a value $E(B-V) = 0.25$ for the colour excess. For different values of the colour excess, the correction factor would scale linearly in our log–log plot. This correction factor was plotted as a black arrow for both curves considered. They can be seen in Fig. 8, such that the full extent of the arrow is log($K_{\rm dust}$) dex. In reality, dust effects on the spectrum are often less pronounced than shown here.

This diagram has very distinct AGN and SFG regions; however, it uses the He II λ1640 line on its own, while in the prism catalogue, it appears blended with O III] λ1665. Therefore, on the right side of Fig. 8, we plot a similar diagram that uses the blended line instead. Since galaxies 18846 and 10058975 had all the necessary lines in the catalogue and we estimated the C III] emission-line flux for galaxy 9422, we mark them on the plot. We display a solid red line on the diagram aiming to separate a mixed region from a region only occupied by AGN models. Galaxies with:

$$\log\left(\frac{\text{He II + O III]}}{\text{H}\beta}\right) > -0.2 \tag{13}$$

can be reliably identified as AGN while those with values below this fall into a mixed region. This diagram shows more mixing between AGN and non-AGN compared to the one presented on the left side, but all three galaxies still fall within the AGN region.

Since UV lines are more heavily affected by dust extinction compared to optical lines, the raw (uncorrected) line ratio will typically appear lower than the intrinsic (dust-free) value. This is because the UV flux will be reduced much more than the optical flux due to the higher attenuation at shorter wavelengths. Consequently, after applying dust corrections, the galaxies will shift to higher values on both axes, and will remain within the AGN region. We can thus identify 9422, 18846, and 10058975 as AGN candidates using this method.

Based on Fig. 8, the He II λ1640/Hβ versus (C III] + C IV)/Hβ diagram might be a useful diagnostic, since both axes show a good separation between the AGN and SFG models. However, it is important to note that this diagram can only be useful if the colour excess is very small. This excess would need to be accurately estimated. A possible way to include this correction into the analysis would be to use the Balmer decrement method. This method consists of measuring the observed intensities of Hα and Hβ lines and comparing them to the expected intrinsic ratio.

While the Balmer decrement is a valuable method for correcting dust extinction, it has its limitations, particularly in the context of

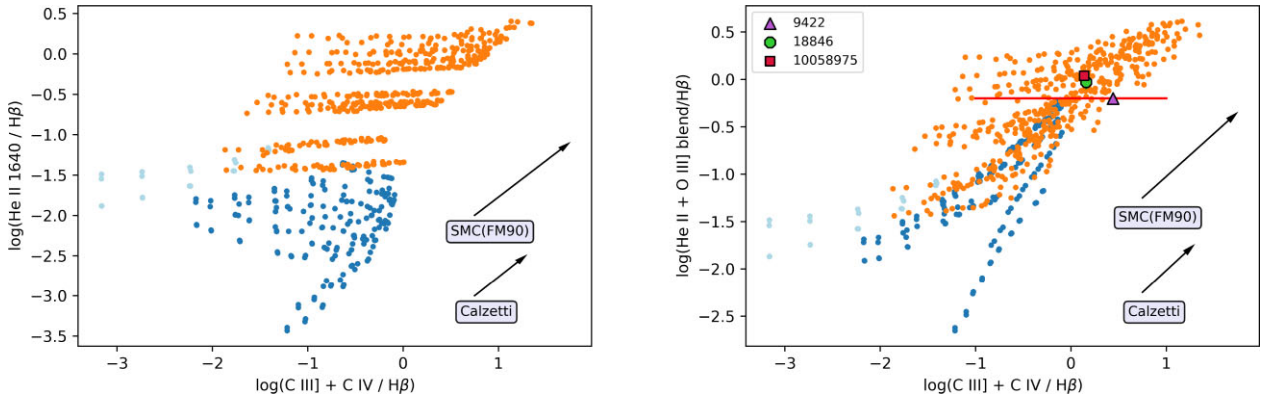

Figure 8. UV–optical ratio diagrams displaying the models of K. Nakajima & R. Maiolino (2022, coloured as previously) and arrows corresponding to the line ratio correction factor for $E(B-V) = 0.25$ assuming a Calzetti dust attenuation law (D. Calzetti et al. 2000) and an SMC bar curve with the FM90 parametrization (K. D. Gordon et al. 2003; E. L. Fitzpatrick & D. Massa 1990), respectively. The diagram on the right was created by changing line He II λ1640 on the left diagram to the blended line He II + O III] since this is what can be detected with the NIRSpec prism. Our sources of interest are marked with a purple triangle, a green circle, and a red square.

AGN. Strong ionizing radiation from the AGN's central black hole can alter the Balmer line ratios, complicating the dust extinction calculation. This makes it crucial to know the galaxy's nature before applying the method. However, since we are using this method to identify the galaxy type, it becomes more challenging to use effectively.

4 AGN DETERMINATION FROM PHOTOMETRY

In this section, we complement our analysis of the spectroscopy data with the study of the photometry data of the JADES data set, mainly focusing on sources 9422 and 18846, following similar steps to the analysis done by I. Juodžbalis et al. (2023). For the identification of AGN galaxies using photometry, we use the routine EAZY (G. B. Brammer, P. G. Dokkum & P. Coppi 2008). We apply this method to a series of SED fits to identify sources that are best fit by AGN SED templates. Furthermore, we present some values obtained with this method for a compilation of AGN sources from the literature.

4.1 EAZY selection method

I. Juodžbalis et al. (2023) developed a photometric method that can identify high probability AGN candidates in the early Universe. Their goal was to create a method which would aid the search for strong AGN candidates that could then be followed up by spectroscopy.

EAZY was run with two sets of templates. One of them, referred to as FSPS + Larson, consist of the 12 default Flexible Stellar Population Synthesis templates (C. Conroy & J. E. Gunn 2010) with the addition of six templates from R. L. Larson et al. (2023b). The other set contains nine direct collapse black hole SED templates from K. Nakajima & R. Maiolino (2022), bringing the total number of templates to 27. The accuracy of the method is highly dependent on the specific templates used with EAZY, as different templates can influence the resulting classifications and parameter estimates.

Our analysis uses SED templates based on three SFG models with strong emission lines at fixed metallicity. This approach focuses on high EW systems and does not capture the full diversity of young, metal-poor starbursts. None the less, recent work by C. Zhu et al. (2025) demonstrates that extreme [O III] EWs ($> 10^3$ Å) are primarily produced by obscured AGNs and remain effective tracers of AGN activity, even in galaxies with strong starburst-driven emission. This physical distinction supports our template strategy. By selecting high-EW systems, we preferentially target galaxies where the rest-optical emission is dominated either by young stars (as represented in our templates) or by obscured AGNs, thus naturally separating the populations of interest.

To further improve AGN selection and refine template choices in future work, complementary approaches – such as SED modelling with codes like CIGALE or AGNFITTER – could provide independent constraints on AGN activity. These tools incorporate additional physical parameters that may help break degeneracies between AGN and SFG signatures, particularly in regimes where our current templates face limitations.

The SED fitting is done twice, once only with the FSPS+Larson templates and then another time with the Nakajima & Maiolino set added on top of the FSPS + Larson set. The filters we use are: *F*090*W*, *F*115*W*, *F*150*W*, *F*200*W*, *F*277*W*, *F*356*W*, *F*444*W*, *F*335*M*, and *F*410*M*. We used the AGN fraction parameter, $f_{\rm AGN}$, defined as (I. Juodžbalis et al. 2023)

$$f_{\rm AGN} = \sum W_i, \tag{14}$$

where W_i is the weight of each AGN template. The $f_{\rm AGN}$ parameter, therefore, describes the relative weight of the AGN versus non-AGN templates in the best fit.

For the selection of candidates, reduced chi-squared values of the first and second fits, χ^2_{R1} and χ^2_{R2}, are also taken into account. The second fit is the one with the added AGN templates. The reduced chi-squared values reported by EAZY were calculated by dividing the total chi-squared by 8, corresponding to the nine photometric bands minus one. This subtraction was applied to approximate the effect of the overall flux normalization, rather than based on a formal accounting of the number of free model parameters. In template-based SED fitting, determining the true degrees of freedom is non-trivial: each galaxy's SED is modelled as a non-negative linear combination of templates, and while the template set may be large, only a few (typically 1–3) contribute significantly to the fit for a given source.

Although the number of model components can exceed the number of photometric data points, potentially raising concerns about overfitting, this situation is common in SED fitting, particularly for faint or high-redshift sources. Our approach follows standard practices in the field (e.g. G. B. Brammer et al. 2008; I. Juodžbalis et al. 2023) and allows for flexible modelling of diverse galaxy SEDs within a

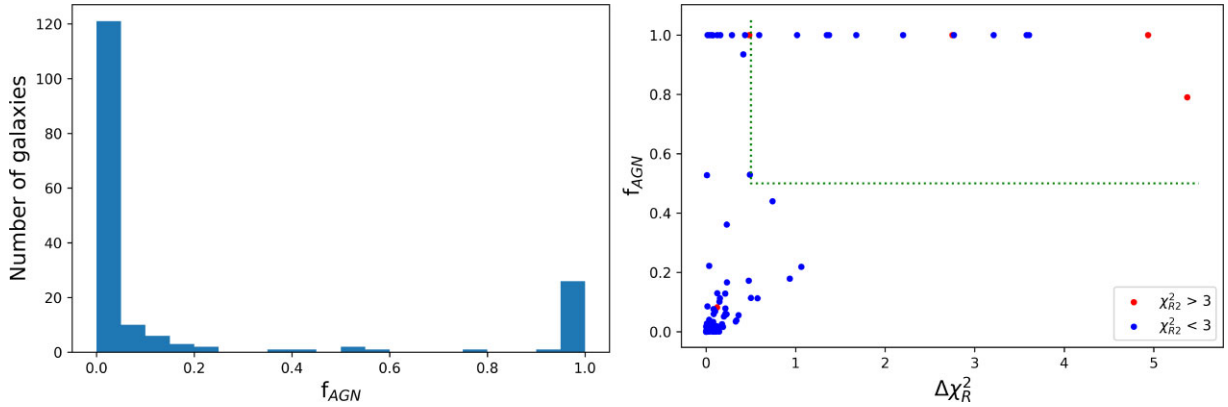

Figure 9. A histogram of $f_{\rm AGN}$ values for the JADES data set (on the left) and a plot showing the $f_{\rm AGN}$ values against the difference in χ^2_R between the two fits (on the right). Data points that satisfy the 'robust' classification by having their χ^2_{R2} value below 3, are marked with blue dots, while the the ones that do not satisfy this condition are coloured red. The green dotted lines separate the part of the parameter space where points satisfy the other two conditions: $f_{\rm AGN} > 0.5$ and $\Delta\chi^2 > 0.5$.

consistent framework. While absolute values of reduced χ^2 should be interpreted cautiously, they still provide a useful metric for evaluating relative fit quality among models applied to the same data set.

The selection criteria defined by I. Juodžbalis et al. (2023) are as follows:

(i) $\chi^2_{R2} < 3$ ('robust' classification),
(ii) $f_{\rm AGN} > 0.5$ (source is dominantly fitted by an AGN model),
(iii) $\chi^2_{R2} < \left(\chi^2_{R1} - 0.5\right)$, ensures an improved fit when the AGN models are added to the templates,

Note that this method is more sensitive to unobscured, Type I, AGN due to those being the easiest to identify with photometry and therefore, will not give a full description of the AGN population in the data. However, this methods can identify high priority sources for future spectroscopic followup.

4.2 The JADES data set

We applied the aforementioned EAZY EAZY SED fitting on the sources in the JADES NIRCam photometric catalogue, excluding the ones that do not have data in all the needed filters. This gives us 177 sources.

On the left side of Fig. 9, we plot a histogram of the $f_{\rm AGN}$ values in the JADES data. Most of the values are around 0 and with another significant fraction of them close to 1. On the right side, we show the $f_{\rm AGN}$ values plotted against the difference in the χ^2_R values between the two SED fits. Data points that satisfy the 'robust' classification by having their χ^2_{R2} value below 3, are marked with blue dots. The red data points do not satisfy this condition. The part of the parameter space where points satisfy the other two conditions: $f_{\rm AGN} > 0.5$ and $\Delta\chi^2 > 0.5$, is above and to the right of the green dotted lines.

The $f_{\rm AGN}$ values in the data are mostly concentrated close to 1 or 0, this is due to EAZY's preference towards single template models. A value close to 1, therefore, does not necessarily imply that the source contains an AGN, as the conventional FSPS + Larson templates may lack the necessary versatility in their emission-line characteristics to accurately replicate the observed data.

Out of the 177 JADES sources, 31 had an $f_{\rm AGN}$ value above 0.5. These values can be seen with the NIRSpec and NIRcam IDs of the corresponding sources in Table 2. Note that two sources were excluded because they did not have valid χ^2_R values. Sources were grouped based on the number of criteria they satisfy out of the three

Table 2. Sources with $f_{\rm AGN}$ values above 0.5 in the JADES photometric catalogue. The first 10 sources satisfy the three criteria discussed in Section 4.1, and are therefore strong AGN candidates. One of our sources of interest, 18846, is marked with a bold NIRSpec ID. The galaxies with bold NIRCam IDs have already been identified as AGN and more information can be found about them in Table 4.

NIRCam ID	NIRSpec ID	$f_{\rm AGN}$	χ^2_{R2}	$\Delta\chi^2_R$	ΔBIC
	$f_{\rm AGN} > 0.5$ and $\chi^2_{R2} < 3$ and $\Delta\chi^2_R > 0.5$				
131688	**00018846**	1.00	0.59	1.68	−6.37
201680	00007892	1.00	0.85	0.59	−15.04
110987	00008113	1.00	1.56	2.77	2.37
132780	00019342	1.00	1.59	1.37	−8.79
113461	00009343	1.00	1.74	3.61	9.07
203112	00009452	1.00	1.84	2.20	−2.18
201730	00007938	1.00	1.86	1.02	−11.65
107462	00006519	1.00	2.00	1.35	−8.98
114213	00009743	1.00	2.25	3.58	8.86
210963	00019606	1.00	2.74	3.21	5.91
	$f_{\rm AGN} > 0.5$ and $\chi^2_{R2} < 3$, but $\Delta\chi^2_R < 0.5$				
98452	00003322	0.53	0.85	0.49	−15.86
197791	00004009	1.00	0.06	0.43	−16.30
113523	10001946	1.00	0.15	0.16	−18.50
111668	10009320	1.00	1.09	0.12	−18.78
198974	00005040	1.00	0.79	0.07	−19.18
139168	–	1.00	0.56	0.07	−19.23
109645	00007507	1.00	0.71	0.06	−19.26
128470	00017251	1.00	0.99	0.06	−19.30
96216	–	1.00	0.98	0.04	−19.46
131067	10014177	1.00	0.64	0.02	−19.64
103332	10000618	0.53	0.71	0.01	−19.85
105635	00005759	0.94	1.31	0.42	−16.45
107324	–	1.00	1.56	0.16	−18.51
130158	10014220	1.00	1.89	0.29	−17.46
128771	–	1.00	1.94	0.02	−19.64
	$f_{\rm AGN} > 0.5$, but $\chi^2_{R2} > 3$				
103349	10000626	1.00	3.16	0.48	−15.91
209707	00018090	1.00	9.55	4.93	19.70
114790	00010073	1.00	9.99	2.75	2.23
110905	00008083	0.79	10.62	5.37	23.21

Notes. Note that the $f_{\rm AGN}$ values are written to a precision of three significant digits. The ΔBIC column shows the difference in Bayesian Information Criterion between the SF and AGN models ($\rm BIC_{SF} - BIC_{AGN}$).

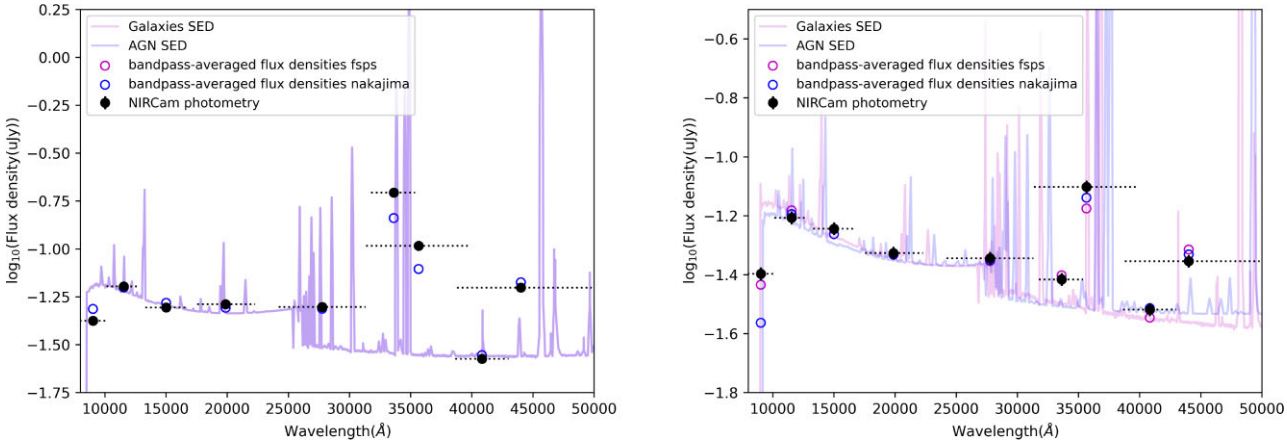

Figure 10. SED fits for our two sources of interest, 9422 (on the left) and 18846 (on the right). The y-axis shows the flux density in μJy, while the x-axis is the observed wavelength in Å. The FSPS + Larson and the added Nakajima & Maiolino fits are represented by the faint magenta and blue curves. The black data points are from the NIRCam filter measurements and the black horizontal dotted lines showcase the wavelength range of the filters. The magenta and blue circles correspond to the bandpass-averaged flux densities predicted by the galaxy and the AGN fits, respectively.

mentioned in Section 4.1. The first 10 galaxies satisfied all three criteria, and therefore, can be considered strong AGN candidates. Amongst these, we can see one of our sources of interest which has a NIRSpec ID of 18846 as the candidate with the lowest reduced chi squared. The second group satisfy 2 out of the 3 criteria, because the improvement of the fit is not sufficient, and the sources in the third group cannot be classified as 'robust' due to their χ^2_{R2} values being above 3.

4.2.1 Bayesian Information Criterion analysis

Following I. Juodžbalis et al. (2025), we evaluate the necessity of AGN templates using the Bayesian Information Criterion (BIC), which balances goodness of fit (χ^2) against model complexity (number of parameters). A positive ΔBIC ($\mathrm{BIC_{SF}} - \mathrm{BIC_{AGN}}$) > 0 indicates statistical preference for the AGN model. While this provides an additional check on template selection, we emphasize that ΔBIC is complementary to – not superseding – our primary criteria ($f_{\mathrm{AGN}} > 0.5$, $\chi^2_{R2} < 3$, and $\Delta\chi^2_R > 0.5$). The joint application of these metrics ensures robust AGN identification while accommodating cases where different indicators may disagree (e.g. objects with clear spectral AGN features but marginal ΔBIC values).

As shown in Table 2, only 4 of our 10 primary AGN candidates exhibit ΔBIC>0, formally favouring the AGN template under the criterion of I. Juodžbalis et al. (2025). However, the remaining candidates with negative ΔBIC values remain strong AGN candidates in independent diagnostics – for instance, in Fig. 11, both galaxy 6519 and 7938 are in the AGN region, while their ΔBIC are -8.98 and -11.65-, respectively. This discrepancy highlights that, while ΔBIC provides a useful statistical measure of model preference, it is not universally decisive. Thus, we maintain these candidates in our sample, as their multiwavelength AGN indicators collectively outweigh isolated ΔBIC results.

4.3 NIRSpec sources 9422 and 18846

The f_{AGN} values found for the sources with NIRSpec spectra, including those with IDs 9422 and 18846 (NIRCam IDs 113585 and 131688) are 0.0 and 0.9999461. This shows that 18846 is an excellent AGN candidate. We note again that the 0.0 value for galaxy 9422 does not mean that it does not have an AGN, only that this method was not able to identify it as a strong candidate.

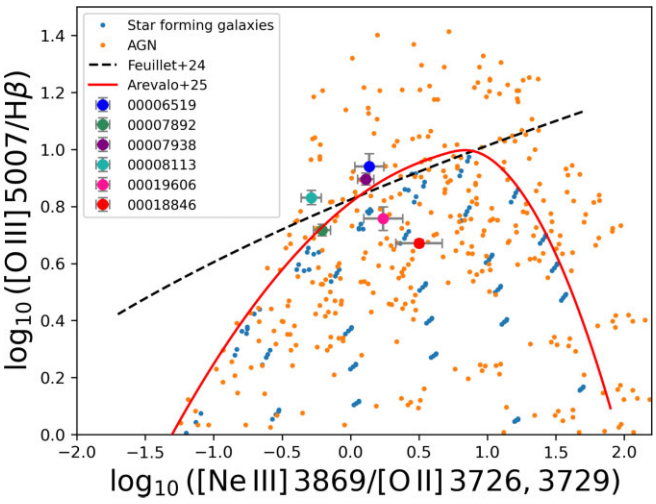

Figure 11. The OHNO diagram displaying SFG and AGN models from K. Nakajima & R. Maiolino (2022, coloured as previously), the dashed black demarcation line is based on the work of L. M. Feuillet et al. (2024). Our newly defined solid red separation curve is defined by equations (2) and (3), and it separates a mixed region from a region only occupied by AGN models. Our photometric AGN candidates are the plotted data points.

For the specific case of galaxy 9422, A. J. Cameron et al. (2024) analysis its spectrum. It shows a steep turnover in the UV continuum, which is indicative of two-photon emissions from neutral hydrogen and it is not typically seen in AGN or X-ray binaries. On the other hand, J. Scholtz et al. (2025) classify it as an AGN using the two Helium lines, S. Tacchella et al. (2025) classify it as a galaxy with an obscured type-2 AGN, and Y. Li et al. (2024) state that it is combination of young metal-poor stars and a low-luminosity AGN.

For 18846, the first SED fit with the FSPS + Larson templates gave a χ^2_R of 2.25, while the second fit with the added Nakajima & Maiolino templates gave a better value of 0.59. The estimated photometric redshifts were 6.34, and 6.5 for the two fits, respectively. These compare well to the actual spectroscopic redshift of 6.342. Therefore, 18846 satisfies all the criteria in Section 4.1 for being a strong AGN

Table 3. Average flux densities for our two sources in units of μJy across the filters used in the SED fitting with EAZY. Errors in the flux values are displayed, however, note that the fitting was run with a minimum flux error of 5 per cent to account for equipment calibration uncertainties.

Filter	9422 [μJy]	18846 [μJy]
*F*090*W*	0.0421 ± 0.0021	0.0401 ± 0.0020
*F*115*W*	0.0636 ± 0.0032	0.0620 ± 0.0031
*F*150*W*	0.0494 ± 0.0025	0.0570 ± 0.0028
*F*200*W*	0.0514 ± 0.0026	0.0471 ± 0.0024
*F*277*W*	0.0497 ± 0.0025	0.0453 ± 0.0023
F356W	0.1037 ± 0.0052	0.0790 ± 0.0040
*F*444*W*	0.0627 ± 0.0031	0.0442 ± 0.0022
*F*335*M*	0.1965 ± 0.0098	0.0384 ± 0.0019
*F*410*M*	0.0266 ± 0.0013	0.0303 ± 0.0015

candidate. This result gives motivation to further investigate galaxy 18846. Especially considering its spectral resemblance to sources 9422 and 10058975 which have already been selected as AGN by multiple studies (J. Scholtz et al. 2025; S. Tacchella et al. 2025; Y. Li et al. 2024).

The flux values in the eight filters used for the SED fitting for our two sources can be seen in Table 3. Note, however, that the fitting was run with a minimum flux error of 5 per cent, accounting for uncertainties in the calibration.

The SED fits for the two galaxies of interest can be seen on Fig. 10. Note that the units and the scale of the *y*-axes are different than in Fig. 1. The black data points are the NIRCam photometry values for the mentioned filters and the black horizontal dotted lines showcase the wavelength range of these filters. The circles represent the bandpass-averaged flux densities predicted by the SED fits. The blue line and circles represent the result of the second fit done with the added Nakajima & Maiolino templates, while the magenta line and circles show the results of the first fit.

We can see that for 9422, the two fits are overlapping, meaning that adding the AGN templates did not change the best fit, hence the low $f_{\rm AGN}$ value. For galaxy 18846, the addition of the Nakajima & Maiolino set improved the fit, although the improvement is marginal, primarily due to the filters affected by emission lines like *F*356*W* and *F*444*W*. As a result, it has a high $f_{\rm AGN}$ value.

4.4 Spectroscopic validation of photometric AGN candidates

The spectroscopic analysis proved challenging for most candidates. Of the 10 galaxies meeting all photometric criteria, only one (ID 18846) exhibited UV/optical lines of sufficient quality for direct placement on standard diagnostic diagrams. The remaining candidates presented several complications: Four galaxies (6519, 7892, 7938, and 8113) showed completely blended [O III] λ5007 and [O III] λ4959 lines. Following established theoretical work (P. J. Storey & C. J. Zeippen 2000), we adopt a flux ratio of 2.98 for these transitions when attempting to de blend the features.

Three additional candidates (9452, 9743, and 19342) lacked detectable [O II] or [Ne III] emission entirely, preventing meaningful upper limit estimates. Galaxy 19606 had all required lines detected except [O II], which showed an S/N of 3.78 when analysed using LIME and can be included in the OHNO diagram. The spectrum of 9343 proved unusable due to corruption in the Prism/clear data.

Classification using the OHNO diagram in Fig. 11 yielded important insights: three of the six analysable candidates fall in the AGN-dominated region above both demarcation lines, while one occupies the intermediate zone between our selection boundary and that of L. M. Feuillet et al. (2024). The remaining two candidates resided in the ambiguous mixed region below both lines, preventing a clear classification.

4.5 Previously identified AGN

Furthermore, we compiled a list of AGN that have been identified using various methods (such as the BPT and the VO87 diagrams or the presence of broad lines) and obtained $f_{\rm AGN}$ values for them. These are presented in Table 4. The sources with high $f_{\rm AGN}$ values are written in bold.

We can see that only 4 of the 32 confirmed AGN sources have high $f_{\rm AGN}$ values. This shows once more that the $f_{\rm AGN}$ does not correspond to a physical quantity, and galaxies with low values can still be identified as AGN. As is the case of galaxy 9422 which, though has a very low $f_{\rm AGN}$ value, is claimed to be an obscured AGN by S. Tacchella et al. (2025) and Y. Li et al. (2024), as already mentioned. This method might be able to identify some sources which have a high probability of containing an AGN, but it also disregards other possible candidates.

In conclusion, although this photometry analysis cannot be used as a completely independent diagnostic tool for finding AGN in galaxies, it can still be a good resource to accompany spectroscopic analysis. By looking first at the spectroscopy of the galaxies that pass our criteria, we can prioritize the galaxies that have a higher probability of being an AGN. None the less, we should not claim that a galaxy is not an AGN just by the results of the photometric analysis, but would need further confirmation from spectroscopy to back up the finding of AGN in these systems.

5 SENSITIVITY OF DIAGRAMS

In this section, we aim to further discuss the use of the diagrams proposed in the previous sections (Sections 3.3–3.5) by exploring their sensitivity. We do this by evaluating the needed signal-to-noise ratio (SNR) that would make the parameter space a useful diagnostic, then we calculate the needed source brightness required for detections with such precision. Finally, we make assumptions related to the exposure time dependence of this required source brightness.

5.1 Required signal-to-noise ratio values

We found that an SNR of 5 on the line fluxes and the continuum levels is sufficient to motivate the use of all previously discussed diagrams, and that for many, 3σ line detections could still give useful insight. Data points with these uncertainties provide useful information about the ionization sources of the galaxies in question. This is because the error bars of the diagnostic quantities are sufficiently small with respect to the separation between SFG and AGN in the diagrams.

The actual UV SNRs for our three target sources are listed in Table 5. As shown, each galaxy has an SNR above 5 for several lines, making them well suited to provide meaningful information.

5.2 Required brightness versus exposure time

For a correct assessment of our diagnostics in the spectroscopic analysis, it is important to have an evaluation of the number of hours needed for a correct detection of the emission lines. We calculated the minimum continuum brightness ($C_{\rm min}$) a source must have for its emission line with a rest frame EW of $\rm EW_{rest}$ to be 5σ detected by using the 10σ line sensitivity (in units of erg s cm^{-2}) for the PRISM

Table 4. List of identified AGN with the obtained f_{AGN} values (to three significant digits), identification methods, and references. Sources that satisfy the $f_{AGN} > 0.5$ criterion are written in bold. The first three of these also appear in Table 2.

IDs		f_{AGN}	Selection method	References
NIRCam ID	NIRSpec ID			
113585	00009422	0.00	He II1640 and He II4686	J. Scholtz et al. (2025)
103366	00004902	0.11	S2–VO87∗	J. Scholtz et al. (2025)
108715	00007099	0.02	S2–VO87∗	J. Scholtz et al. (2025)
110238	00007762	0.18	High ion	J. Scholtz et al. (2025)
110905	**00008083**	**0.79**	High ion and He II 4686 & broad lines	J. Scholtz et al. (2025); R. Maiolino et al. (2024)
111661	00008456	0.01	He II 4686	J. Scholtz et al. (2025)
112500	00008880	0.00	N2–BPT	J. Scholtz et al. (2025)
203112	**00009452**	**1.00**	N2–BPT	J. Scholtz et al. (2025)
114790	**00010073**	**1.00**	He II4686	J. Scholtz et al. (2025)
208642	00 016 745	0.11	S2–VO87∗	J. Scholtz et al. (2025)
208919	00 017 072	0.05	He II1640	J. Scholtz et al. (2025)
209348	00 017 670	0.00	He II4686	J. Scholtz et al. (2025)
137667	00 021 842	0.00	high ion	J. Scholtz et al. (2025)
103349	**10000626**	**1.00**	High ion and He II 4686	J. Scholtz et al. (2025)
198071	10 035 295	0.00	He II 1640	J. Scholtz et al. (2025)
200002	10 036 017	0.01	N2–BPT and S2–VO87	J. Scholtz et al. (2025)
197348	10 013 704	0.00	Broad lines	R. Maiolino et al. (2024)
GLASS 160133		**1.00**	Broad lines	Y. Harikane et al. (2023)
GLASS 150029		0.00	Broad lines	Y. Harikane et al. (2023)
CEERS 00746 or CEERS 3210		0.00	Broad lines	Y. Harikane et al. (2023); D. D. Kocevski et al. (2023)
CEERS 00672		0.00	Broad lines	Y. Harikane et al. (2023)
CEERS 02782 or CEERS 1670		0.01	Broad lines	Y. Harikane et al. (2023); D. D. Kocevski et al. (2023)
CEERS 00397		0.01	Broad lines	Y. Harikane et al. (2023)
CEERS 01236		0.00	Broad lines	Y. Harikane et al. (2023)
GOODS-S−13971		0.00	Broad lines	J. Matthee et al. (2024)
MSA 10686		0.00	Broad lines	J. E. Greene et al. (2024)
MSA 13821		0.00	Broad lines	J. E. Greene et al. (2024)
MSA 23608		0.00	Broad lines	J. E. Greene et al. (2024)
MSA 32265		0.00	Broad lines	J. E. Greene et al. (2024)
MSA 33437		0.00	Broad lines	J. E. Greene et al. (2024)
MSA 35488		0.00	Broad lines	J. E. Greene et al. (2024)
MSA 38108		0.00	Broad lines	J. E. Greene et al. (2024)

Table 5. S/Ns of the UV line fluxes and their continuum level in our three sources of interest (NIRSpec IDs 9422, 18846, and 10058975). The line flux SNRs were calculated using the *JWST*/Prism line flux catalogue, while the continuum level SNRs were calculated by dividing our continuum level estimate with the standard deviation of the continuum sample points from our continuum fit. The C III] line of galaxy 9422 was not reported in the *JWST*/Prism line flux catalogue.

	9422		18846		10058975	
Emission line	Line flux SNR	Continuum SNR	Line flux SNR	Continuum SNR	Line flux SNR	Continuum SNR
He II + O III]	14.9	21.2	6.0	20.3	7.2	23.6
C IV	12.8	24.3	6.7	23.2	6.0	24.0
C III]	–	29.4	5.3	40.9	17.1	19.8

instrument provided by D. J. Eisenstein et al. (2023a). For the case of high-resolution grating, in the range of the UV lines, grating provides a superior line flux detection limit compared to the PRISM. This is due to the grating's higher spectral resolution offering better contrast against the low background at these wavelengths. None the less, we made a conscious decision to focus our analysis on the prism sensitivity. This choice was driven by the scope of our analysis, which is based on PRISM-derived data.

We used the 5σ limits relying on the argument on the needed SNRs in Section 5.1. These values can be found in Table 6 where $F_{5\sigma}(z)$ is the 5σ detection limit obtained from D. J. Eisenstein et al. (2023a) assuming a redshift $z = 6$. In order to estimate the source brightness needed for a 5σ detection with a different exposure time observation, we assumed that the line sensitivity scales as $1/\sqrt{t}$. This is a reasonable assumption since for background-limited observations,

Table 6. 5σ NIRSpec prism detection limits for various emission lines assuming a source at redshift $z = 6$ and an exposure time of 28 h. These values were calculated using the prism sensitivity line in D. J. Eisenstein et al. (2023a), and they represent the minimum line fluxes needed for a 5σ detection.

Emission line	$F_{5\sigma}$ (10^{-18} erg s cm^{-2})
He II λ1640	1.35
C IV	1.55
C III]	0.85
O III] λ1665	1.25
He II λ4686	0.15
Hβ	0.15
[O III] λ5007	0.14

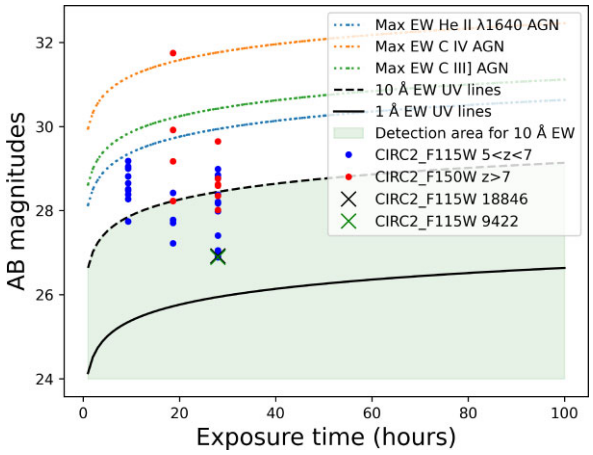

Figure 12. Diagrams displaying the exposure time dependence of the required brightness (in AB mag) a source at $z = 6$ needs to have in order to achieve 5σ line flux detections for the PRISM in the case of rest-frame EWs of 1, 10 Å, or the maximum EW value of the AGN models. The 1 Å UV average curve is displayed as a black line. The 10 Å UV average curve is displayed as a black dashed line, and the dotted coloured lines correspond to the maximum EWs in the AGN models for some of the lines in question. These maximum EW values are around 228, 55, and 42 Å for C IV, C III], and He II λ1640, respectively. The green-shaded region represents the 5σ line flux detection area for UV lines with an EW of 10 Å. Blue and red data points show the CIRC2 NIRCam measurements for the JADES sample, where the *F*115*W* filter was used for sources at $5 < z < 7$, and *F*150*W* for sources at $z > 7$. Two of our galaxies of interest, 9422 and 18846 (with NIRCam IDs 113585 and 131688, respectively), are located by overlapping green and black crosses, and are amongst the brightest sources).

like these faint high-redshift galaxies, the SNR scales as $\sqrt{t}$ and is proportional to the flux.

The required brightness vs. exposure time diagram for the 5σ detection can be seen on Fig. 12. The 1 Å UV average curve is displayed as a black line. The 10 Å UV average curve is displayed as a black dashed line. The coloured dotted lines correspond to the maximum EW values of the AGN models for some of the emission lines. These maximum EW values are around 228, 55, and 42 Å for C IV, C III], and He II λ1640, respectively. The green-shaded region represents the 5σ line flux detection area for UV lines with an EW of 10 Å. We can see that with longer exposure times, the sources need to be less bright to achieve the same precision.

We have considered exposure times from 1 h (quite commonly adopted with NIRSpec) to 100 h (the maximum that anyone will probably ever get). Increasing the time exposure in this time range can reduce the required brightness by two magnitudes. The more intense reduction in brightness would be reflected in the first 10 h of exposition. Additionally, to resolve an EW of 1 Å requires orders of brightness that are 2.5 AB magnitudes lower than for EW of 10 Å.

On Fig. 12, we also display JADES sources using the NIRCam CIRC2 measurement associated with a 0.15 arcsec radius circular aperture. In order for the UV emission lines to fall in the appropriate region of the observed spectrum, the *F*115*W* filter was used for sources at $5 < z < 7$, and *F*150*W* for sources at $z > 7$. These are represented with blue and red data points, respectively. Two of our sources of interest, 9422 and 18 846 (NIRCam IDs 113585 and 131688, respectively), can be seen marked with overlapping green and black crosses. 9422 has a value of 26.89 AB magnitude and 18846 has a value of 26.92 AB magnitude. Unfortunately, photometry was not available for galaxy 10058975. We can see that galaxies 9422 and 18846 are amongst the brightest sources in the data.

Note that many of the observational data points lie just below the maximum curve for the He II λ1640 line, meaning they would only be bright enough to exhibit this line to a sufficient SNR if they were extremely powerful AGN. For better observations of these extremely faint lines, such as He II λ1640 or He II λ4686, longer exposure times are required on brighter sources. With the current data, primarily targeting faint galaxies, we cannot effectively distinguish between AGN and SFG using UV-diagnostics in most cases. The need for the observation of these faint lines is supported by the analysis of J. Scholtz et al. (2025) who identified galaxies 9422 and 10058975 as AGN using He II λ1640 and He II λ4686 diagnostics.

We saw in Fig. 12 that JADES sources are well above the 1 Å line. This shows that it would be unreasonable to expect detections of 1 Å rest-frame EWs to a sufficient precision for these diagrams. This reduces the parameter space of the proposed UV EW diagrams significantly. On Figs 5–7, we cannot expect to see observed data points with EW values less than 1 Å, meaning no observations below a value 0 on the EW axes of these diagrams.

In some cases, this means that the expected detection region of the diagram completely excludes the SFG models. This happens for diagrams using the EW of He II λ1640 or He II λ4686, further showing why the detection of these two lines indicate the presence of an AGN.

6 CONCLUSIONS

In this paper, we explore how to identify AGNs within galaxies with data from the *JWST*. We show that existing optical diagrams such as the BPT diagram and the VO87 diagram pose limitations when analysing AGN at high redshifts. These limitations include significant regions where AGN and SFG overlap, and missing lines outside the *JWST*'s detection range, rendering analysis impossible. Therefore, new methods and ideas are needed to make progress within this very important topic.

We thus test and use a variety of new potential diagnostics, that can be used to distinguish AGNs and SFGs in the early universe. We use photoionization models created by K. Nakajima & R. Maiolino (2022) and the *JWST*/NIRSpec prism data (A. J. Bunker et al. 2024). Amongst the diagrams are some that only use UV emission-line ratios, others use EWs and line ratios, and some solely use EWs of both optical and UV lines. We examine where the galaxies with JADES NIRSpec IDs 9422, 18846, and 10058975 fall on our new EW plots, to determine their nature. We then evaluate the sensitivity and applicability of these possible diagnostics with the current observational strategy and also extend our analysis following a photometric selection method for which we use the *JWST* NIRCam photometry catalogue (M. J. Rieke et al. 2023).

Our conclusions about the explored spectroscopic diagnostics are the following:

(i) The standard diagnostic diagrams (BPT and VO87) are inefficient for the identification of AGN in the Early Universe, therefore new diagnostic methods need to be considered.

(ii) The maximum starburst line on the N2–BPT and S2–VO87 diagrams agree well with the presented SFG and AGN photoionization models, however the line thought to mark the lower limit of the mixed region does not.

(iii) The AGN/SFG demarcation line defined by B. E. Backhaus et al. (2022) on the OHNO diagram is inconsistent with the presented photoionization models. However, with the introduction of newer

separation curves, including both our own and the one from L. M. Feuillet et al. (2024), the OHNO diagram shows great potential for effectively selecting AGN candidates. We list a few potential candidates from the JADES prism and grating catalogues.

(iv) The ratios (C III] + C IV)/He II λ1640, O III] λλ1665/He II λ1640, and C III]/He II λ1640 ratios, as well as the EWs of the He II λ1640 and He II λ4686 recombination lines, are the most efficient quantities in separating AGN and SF galaxies based on the models. Combining these quantities can create powerful diagnostics. The ratio C IV/C III can also be useful when combined with one of the previous quantities.

(v) Diagrams using the He II λ1640 + O III] lines can provide some information about the nature of the galaxy. However, for robust diagnostics, the He II λ1640 and O III] λλ1665 lines should be observed separately.

(vi) The optical EW diagram using the lines He II λ4686 versus either Hα or Hβ provides an exceptionally good separation between AGN and SF galaxies. However, the He II λ4686 line is extremely hard to detect and the Balmer lines might not be reliable diagnostic quantities at high redshifts due to the limited ability of *JWST* to probe all spectral wavelengths and resolutions.

(vii) The plots that overlay the EWs of galaxies 9422, 18846, and 10058975 on the modelled data sets yield ambiguous results. Further analysis with more UV diagrams (perhaps using He II λ1640 and O III] λ1665 as separate lines) are needed to obtain conclusive results.

(viii) In many diagrams, the only SFGs that are located in the AGN region have metallicities of 0.0007 $Z_{\odot}$. These values are extremely low and would be rarely detected, thus the diagrams presented could still be very effective in separating AGN and SF galaxies.

(ix) The future observational strategy should adapt to the need for longer exposure times in order to make the new diagnostics effective. With the current data, primarily targeting faint galaxies, we cannot effectively distinguish between AGN and SFG using UV-diagnostics in most cases.

(x) 5σ line flux observations are sufficient for the effective use of the proposed diagrams (in Sections 3.3–3.5), and in many cases, observations with a precision of 3σ can still provide useful insight.

(xi) Some UV–optical line ratios diagnostics, such as the He II λ1640/Hβ versus (C III] + C IV)/Hβ diagram, may be useful in the case of a low dust attenuation. When working with line flux ratios, it is important to be careful of dust attenuation since it affects light differently depending on the wavelength. The He II λ1640/Hβ ratio shows a very clear separation between the AGN and SFG photoionization models, while high values of the (C III] + C IV)/Hβ ratio also create a region solely occupied by AGN models. 9422, 18846, and 10058975 are identified as AGN candidates using this method.

Furthermore, our conclusions about the photometric selection method are:

(i) Galaxy 18846 is a strong AGN candidate supported by the photometric selection method relying on the work of I. Juodžbalis et al. (2023), as well as its strong UV spectra resembling some already identified AGN.

(ii) We present a sample of similarly strong AGN candidates in the JADES data supported by the same method in Table 2.

(iii) Photometric pre-selection methods, such as the one used in this project, can become useful tools for selecting a few strong candidates for spectroscopic followup from large samples of galaxies. However, it should be noted that they can be biased and should therefore be used together with other selection methods.

(iv) Some templates are more biased towards a specific type of AGN which can lead to low values of the indicator of the fraction of AGN and an incorrect identification of AGN. That is why it is important to consider in the analysis an extensive and accurate database of photoionization models that would represent the most complete description of AGN and SFG.

We focused on galaxies 9422, 18846, and 10058975 due to their strong UV lines, which are potentially indicative of AGN activity. However, the result we find are inconclusive for these particular systems, showing the difficulty in knowing if a system has an AGN or not. In the UV line ratio plots, galaxies 18846 and 10058975 fell into the mixed region, while galaxy 9422 appeared in the AGN region. On the other hand, all galaxies were placed in either the SFG or mixed region in the EW plots. Although the potential underestimation of EW values in the models suggests that line ratios diagnostics are more reliable. Additionally, we identified an intriguing UV–optical plot using He II λ1640/Hβ versus (C III] + C IV)/Hβ placing the galaxies 9422, 18846, and 10058975 in the AGN region. Any dust correction would further push them into the AGN region of the diagram.

Overall, our results provide a way forward to find AGN in distant galaxies. It is clear that even with detailed spectroscopy it will not always be easy to find these systems, although they are critical for understanding how galaxy evolution and SMBH evolution have occurred in the distant past. Our results also show how AGN can be found through different new approaches using the full wavelength span available to *JWST*. These approaches can be used in new surveys with spectroscopy.

ACKNOWLEDGEMENTS

We acknowledge support from the ERC Advanced Investigator Grant EPOCHS (788113), as well as two studentships from the STFC. JT also acknowledges support from the Simons Foundation and *JWST* program 3215. Support for program 3215 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–03127. IJ acknowledges support by the Huo Family Foundation through a P.C. Ho PhD Studentship, by the Science and Technology Facilities Council (STFC), by the ERC through Advanced Grant 695671 'QUENCH', and by the UKRI Frontier Research grant RISEandFALL. KN acknowledges support from JSPS KAKENHI grants JP20K22373 and JP24K07102. This work is based on observations made with the NASA/ESA *Hubble Space Telescope* (*HST*) and NASA/ESA/CSA *James Webb Space Telescope* (*JWST*) obtained from the `Mikulski Archive for Space Telescopes` (`MAST`) at the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–03127 for *JWST*, and NAS 5–26555 for *HST*. The authors thank all involved with the construction and operation of *JWST*, without whom this work would not be possible.

This work is based on the photoionization models created by K. Nakajima & R. Maiolino (2022) and uses observations made by the *JWST* NIRSpec (P. Jakobsen et al. 2022; P. Ferruit et al. 2022; D. J. Eisenstein et al. 2023a) as well as the NIRCam (M. J. Rieke et al. 2023; D. J. Eisenstein et al. 2023a). We thank Professor Roberto Maiolino for his valuable feedback, which contributed to the improvement of this work.

Furthermore, the following PYTHON libraries were used: ASTROPY (Astropy Collaboration 2022); SPECUTILS (N. Earl et al. (2023),

MATPLOTLIB (J. D. Hunter 2007), SCIKIT-LEARN (F. Pedregosa et al. 2011), NUMPY (C. R. Harris et al. 2020), PANDAS (The pandas development team 2020), and LIME (V. Fernández et al. 2024).

DATA AVAILABILITY

The photoionization models in this work are an extended sample from K. Nakajima & R. Maiolino (2022). The observational data are based on measurements taken with the *JWST* NIRSpec (P. Jakobsen et al. 2022; P. Ferruit et al. 2022; D. J. Eisenstein et al. 2023a) and the NIRCam (M. J. Rieke et al. 2023; D. J. Eisenstein et al. 2023a). The data used in this study are publicly available at https://archive.stsci.edu/hlsp/jades. For any remaining data associated with this study, please contact the first authors.

REFERENCES

Antonucci R., 1993, ARA&A, 31, 473
Astropy Collaboration, 2022, ApJ, 935, 167
Backhaus B. E. et al., 2022, ApJ, 926, 161
Baggen J. F. W. et al., 2024, ApJL, 977, L13
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
Brinchmann J., 2023, MNRAS, 525, 2087
Bunker A. J. et al., 2024, A&A, 690, A288
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Cameron A. J., Katz H., Witten C., Saxena A., Laporte N., Bunker A. J., 2024, MNRAS, 534, 523
Conroy C., Gunn J. E., 2010, Astrophysics Source Code Library, record ascl:1010.043
Dalla Vecchia C., Bower R. G., Theuns T., Balogh M. L., Mazzotta P., Frenk C. S., 2004, MNRAS, 355, 995
de Lima Santos L., Soltau S. B., 2024, The Unified Era: An understanding journey from observations to the Unified Model of Active Galactic Nuclei , preprint (arXiv:2403.17739)
Dors O. L., Cardaci M. V., Hägele G. F., Ilha G. S., Oliveira C. B., Riffel R. A., Riffel R., Krabbe A. C., 2023, MNRAS, 527, 8193
Duncan K., Conselice C. J., 2015, MNRAS, 451, 2030
Earl N. et al., 2023, astropy/specutils: v1.12.0 (v1.12.0). Zenodo. Available at: https://doi.org/10.5281/zenodo.10016569
Eisenstein D. J. et al., 2023a, Overview of the JWST Advanced Deep Extragalactic Survey (JADES) , preprint (arXiv:2306.02465)
Eisenstein D. J. et al., 2023b, preprint (arXiv:2310.12340)
Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, Publ. Astron. Soc. Aust., 34, e058
Feltre A., Charlot S., Gutkin J., 2016, MNRAS, 456, 3354
Ferland G. J. et al., 2013, Rev. Mex. Astron. Astrofís., 49, 137
Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, PASP, 110, 761
Fernández V., Amorín R., Firpo V., Morisset C., 2024, A&A, 688, A69
Ferruit P. et al., 2022, A&A, 661, A81
Feuillet L. M., Meléndez M., Kraemer S., Schmitt H. R., Fischer T. C., Reeves J. N., 2024, ApJ, 962, 104
Fitzpatrick E. L., Massa D., 1990, ApJS, 72, 163
Gardner J. P. et al., 2023, PASP, 135, 068001
Garofali K. et al., 2024, ApJ, 960, 13
Gordon K. D., Clayton G. C., Misselt K. A., Landolt A. U., Wolff M. J., 2003, ApJ, 594, 279
Greene J. E. et al., 2024, ApJ, 964, 39
Gutkin J., Charlot S., Bruzual G., 2016, MNRAS, 462, 1757
Harikane Y. et al., 2023, ApJ, 959, 39
Harris C. R. et al., 2020, Nature, 585, 357
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jakobsen P. et al., 2022, A&A, 661, A80
Juodžbalis I. et al., 2023, MNRAS, 525, 1353
Juodžbalis I. et al., 2025, preprint (arXiv:2504.03551)
Kauffmann G. et al., 2003, MNRAS, 346, 1055
Kewley L. J., Dopita M. A., Leitherer C., Davé R., Yuan T., Allen M., Groves B., Sutherland R., 2013, ApJ, 774, 100
Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, ApJ, 556, 121
Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, MNRAS, 372, 961
Kocevski D. D. et al., 2023, ApJL, 954, L4
Kokorev V. et al., 2023, ApJ, 957, L7
Kroupa P., 2001, MNRAS, 322, 231
Larson R. L. et al., 2023a, ApJ, 953, L29
Larson R. L. et al., 2023b, ApJ, 958, 141
Lecroq M. et al., 2024, MNRAS, 527, 9480
Li Y., Leja J., Johnson B. D., Tacchella S., Naidu R. P., 2024, ApJL, 969, L5
Lynden-Bell D., 1969, Nature, 223, 690
Maiolino R. et al., 2008, A&A, 488, 463
Maiolino R. et al., 2024, A&A, 691, A145
Matthee J. et al., 2024, ApJ, 963, 129
Mazzolari G. et al., 2024, A&A, 691, A345
Mineo S., Gilfanov M., Sunyaev R., 2012, MNRAS, 419, 2095
Nagamine K., Fukugita M., Cen R., Ostriker J. P., 2001, ApJ, 558, 497
Nakajima K. et al., 2018, A&A, 612, A94
Nakajima K., Maiolino R., 2022, MNRAS, 513, 5134
Pedregosa F. et al., 2011, J. Mach. Learning Res., 12, 2825
Rieke M. J. et al., 2023, PASP, 135, 028001
Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, ApJ, 802, L19
Schaerer D., Fragos T., Izotov Y. I., 2019, A&A, 622, L10
Scholtz J. et al., 2025, A&A, 697, A175
Senchyna P., Stark D. P., 2019, MNRAS, 484, 1270
Shirazi M., Brinchmann J., 2012, MNRAS, 421, 1043
Simmonds C., Schaerer D., Verhamme A., 2021, A&A, 656, A127
Stanway E. R., Eldridge J. J., 2018, MNRAS, 479, 75
Storey P. J., Zeippen C. J., 2000, MNRAS, 312, 813
Tacchella S. et al., 2025, MNRAS, 540, 851
The pandas development team, 2020, pandas-dev/pandas: Pandas (v2.2.2). Zenodo. Available at: https://doi.org/10.5281/zenodo.3509134
Tozzi G., Maiolino R., Cresci G., Piotrowska J. M., Belfiore F., Curti M., Mannucci F., Marconi A., 2023, MNRAS, 521, 1264
Übler H. et al., 2023, A&A, 677, A145
Veilleux S., Osterbrock D. E., 1987, ApJS, 63, 295
Zhu C. et al., 2025, ApJ, 982, 27

APPENDIX: ADDITIONAL DIAGRAMS

The following diagrams (Figs A1, A2, A3) provide additional support for the analysis discussed in the main text by presenting an extended set of diagnostic plots. These include further diagrams of EW and line ratios for both UV and optical lines, as well as data for the three galaxies of interest.

We do not include these plots in the main paper as we did not find that these plots and relationships between these quantities are superior to the plots we consider in the main paper. We include these here for completeness and in case these plots may be useful in the future for those who are interested in particular lines and combinations of lines for distinguishing AGN from SFGs.

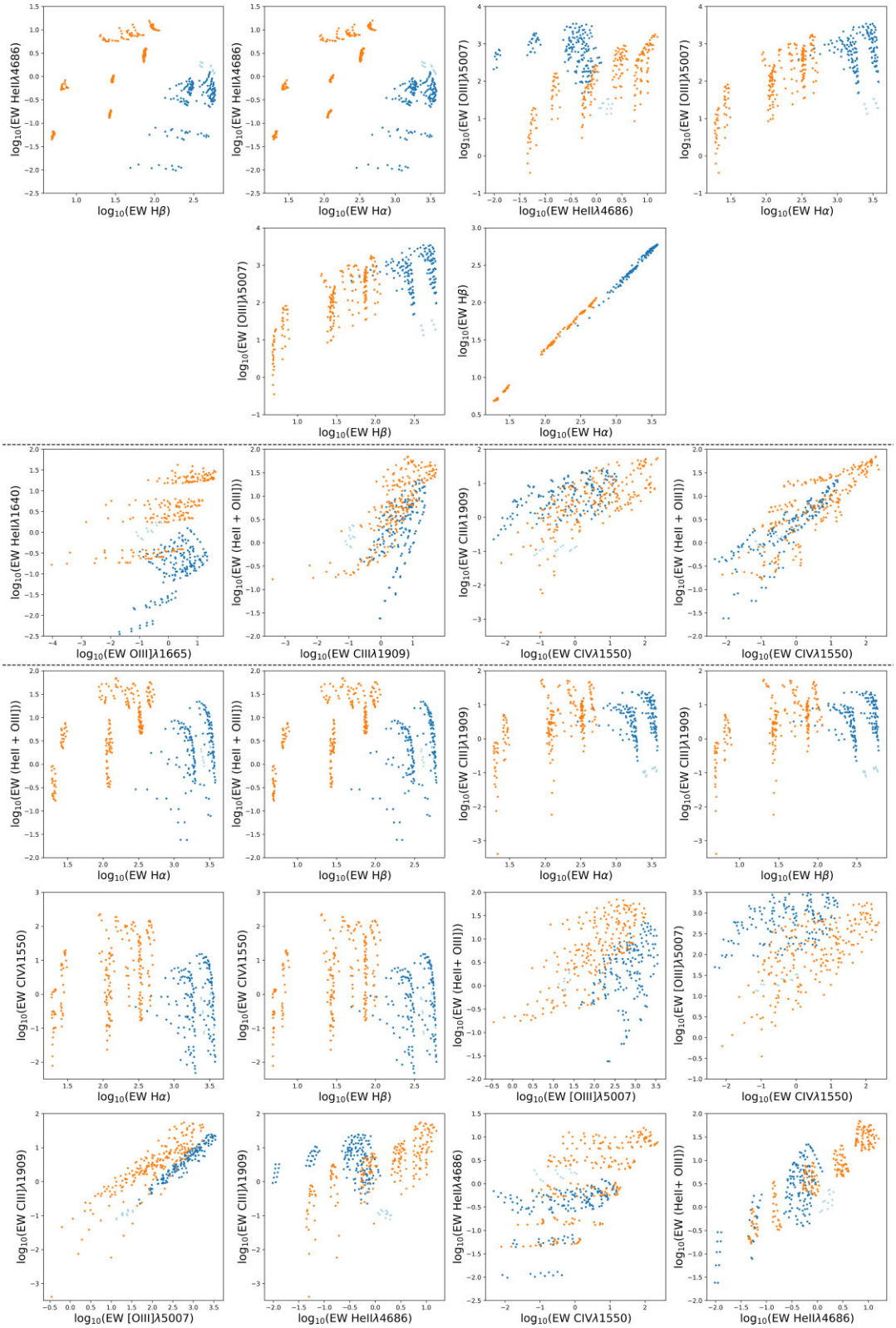

Figure A1. Plots displaying EWs calculated from the modelled data on both axes are separated into three groups: ones only using optical lines (top), ones only using UV lines (middle), and the ones using one of each (bottom). The colour coding is the same as in previous figures. Note that the first plot in the first row (upper left) also appears in Fig. 6 and it is included here for display purposes.

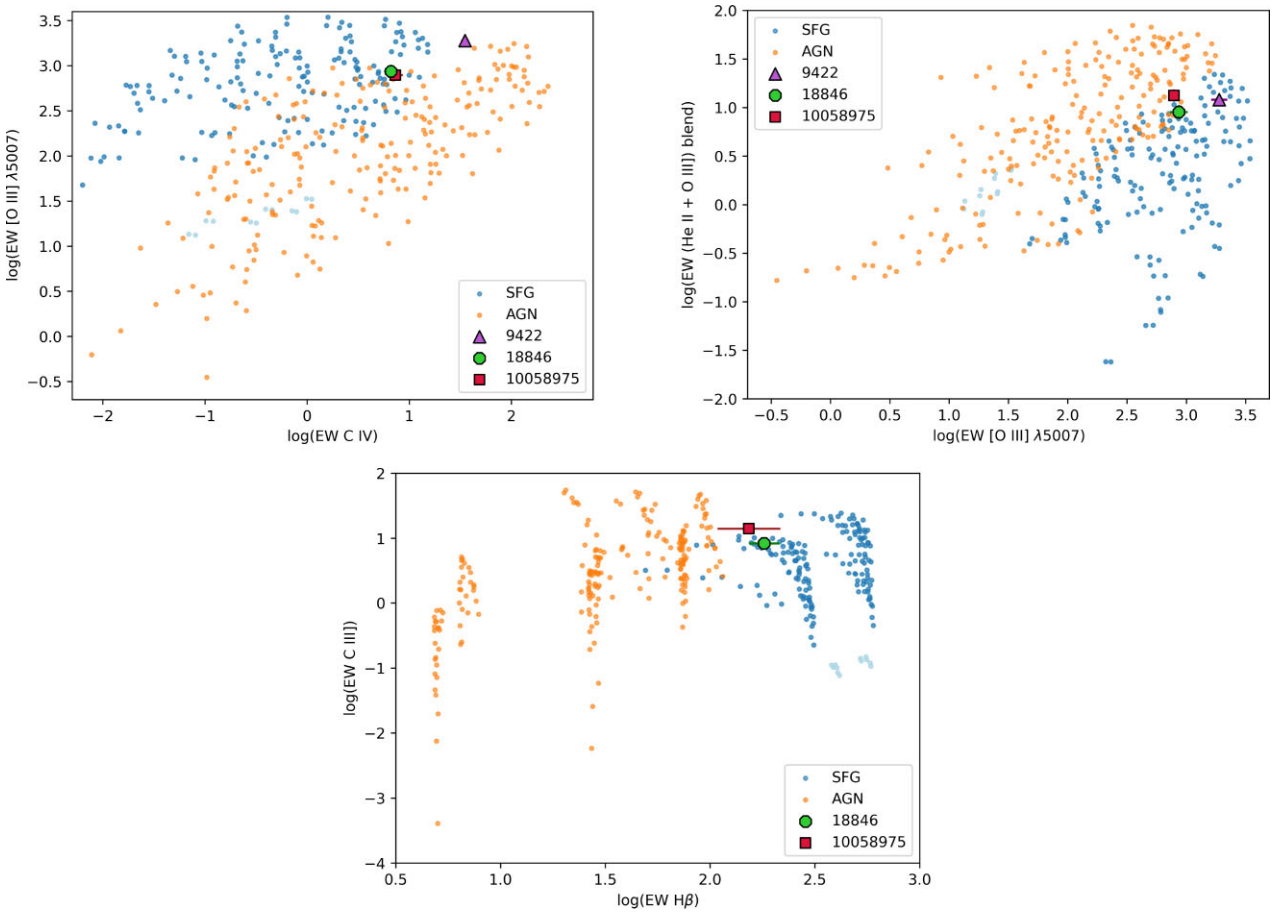

Figure A2. The galaxies with NIRSpec IDs 9422, 18846, and 10058975 are displayed on some of the EW versus EW plots from Fig. A1 using their rest-frame EW measurements and uncertainties. The plots use one UV and one optical line. This figure is an extension of Fig. 7.

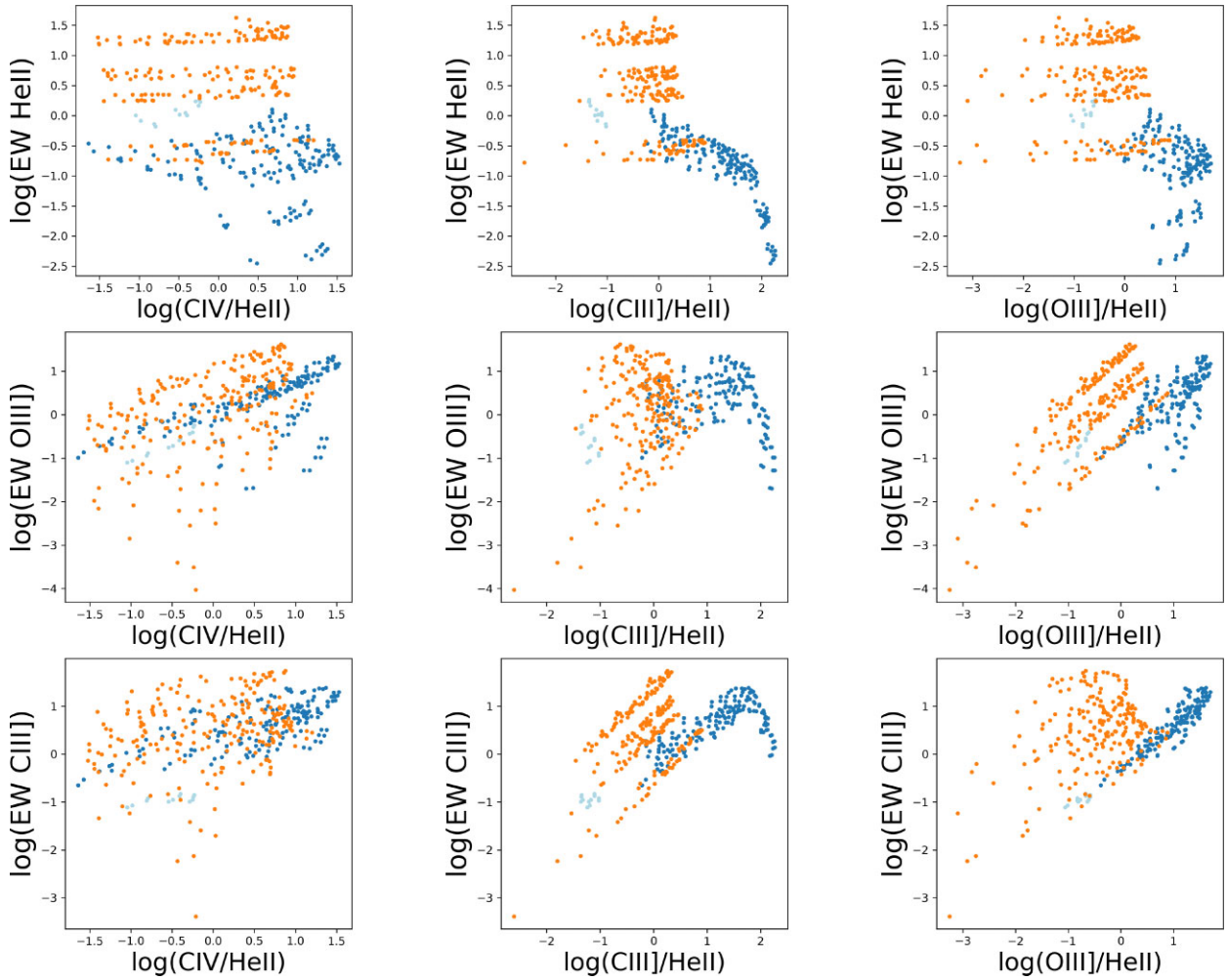

Figure A3. The (base 10) logarithmic EWs of the UV emission lines He II λ1640, O III] λ1665, and C III] plotted against the line ratios C IV/He II λ1640, C III]/He II λ1640, and O III] λ1665/He II λ1640. The colour coding is the same as in previous figures. In terms of EWs, He II λ1640 shows the most effective separation between AGN and SFG. Note that the middle plot in the upper row also appears in Fig. 5 and it is included here for display purposes.

This paper has been typeset from a TEX/LATEX file prepared by the author.