CLASSY VIII: Exploring the Source of Ionization with UV ISM diagnostics in local High-z Analogs[*]

Matilde Mingozzi,[1] Bethan L. James,[2] Danielle A. Berg,[3] Karla Z. Arellano-Córdova,[3] Adele Plat,[4] Claudia Scarlata,[5] Alessandra Aloisi,[1] Ricardo O. Amorín,[6,7] Jarle Brinchmann,[8] Stéphane Charlot,[9] John Chisholm,[3] Anna Feltre,[10] Simon Gazagnes,[3] Matthew Hayes,[11] Timothy Heckman,[12] Svea Hernandez,[2] Lisa J. Kewley,[13] Nimisha Kumari,[2] Claus Leitherer,[1] Crystal L. Martin,[14] Michael Maseda,[15] Themiya Nanayakkara,[16] Swara Ravindranath,[17] Jane R. Rigby,[17] Peter Senchyna,[18] Evan D. Skillman,[5] Yuma Sugahara,[19,20,21] Stephen M. Wilkins,[22] Aida Wofford,[23] and Xinfeng Xu[12]

[1]Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

[2]AURA for ESA, Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

[3]Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, TX 78712, USA

[4]Steward Observatory, The University of Arizona, 933 N Cherry Ave, Tucson, AZ, 85721, USA

[5]Minnesota Institute for Astrophysics, University of Minnesota, 116 Church Street SE, Minneapolis, MN 55455, USA

[6]Instituto de Investigación Multidisciplinar en Ciencia y Tecnología, Universidad de La Serena, Raul Bitrán 1305, La Serena 2204000, Chile

[7]Departamento de Astronomía, Universidad de La Serena, Av. Juan Cisternas 1200 Norte, La Serena 1720236, Chile

[8]Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, PT4150-762 Porto, Portugal

[9]Sorbonne Université, CNRS, UMR7095, Institut d'Astrophysique de Paris, F-75014, Paris, France

[10]INAF - Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via P. Gobetti 93/3, 40129 Bologna, Italy

[11]Stockholm University, Department of Astronomy and Oskar Klein Centre for Cosmoparticle Physics, AlbaNova University Centre, SE-10691, Stockholm, Sweden

[12]Center for Astrophysical Sciences, Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA

[13]Research School of Astronomy and Astrophysics, Australian National University, Cotter Road, Weston Creek, ACT 2611, Australia; ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Canberra, ACT 2611, Australia

[14]Department of Physics, University of California, Santa Barbara, Santa Barbara, CA 93106, USA

[15]Department of Astronomy, University of Wisconsin-Madison, 475 N Charter Street, Madison, WI 53706 USA

[16]Swinburne University of Technology, Melbourne, Victoria, AU

[17]Observational Cosmology Lab, Code 665, NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA

[18]Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101, USA

[19]Institute for Cosmic Ray Research, The University of Tokyo, Kashiwa-no-ha, Kashiwa 277-8582, Japan

[20]National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

[21]Waseda Research Institute for Science and Engineering, Faculty of Science and Engineering, Waseda University, 3-4-1, Okubo, Shinjuku, Tokyo 169-8555, Japan

[22]Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK

[23]Instituto de Astronomía, Universidad Nacional Autónoma de México, Unidad Académica en Ensenada, Km 103 Carr. Tijuana-Ensenada, Ensenada 22860, México

Submitted to AASJournal ApJ

ABSTRACT

In the current JWST era, rest-frame UV spectra play a crucial role in enhancing our understanding of the interstellar medium (ISM) and stellar properties of the first galaxies in the epoch of reionization (EoR, $z > 6$). Here, we compare well-known and reliable optical diagrams sensitive to the main ionization source (i.e., star formation, SF; active galactic nuclei, AGN; shocks) to UV counterparts proposed in the literature - the so-called "UV-BPT diagrams" - using the HST COS Legacy Archive Spectroscopic SurveY (CLASSY), the largest high-quality, high-resolution and broad-wavelength range atlas of far-UV spectra for 45 local star-forming galaxies. In particular, we explore where CLASSY UV line ratios are located in the different UV diagnostic plots, taking

Corresponding author: Matilde Mingozzi

mmingozzi@stsci.edu

into account state-of-the-art photoionization and shock models and, for the first time, the measured ISM and stellar properties (e.g., gas-phase metallicity, ionization parameter, carbon abundance, stellar age). We find that the combination of C III]$\lambda\lambda$1907,9 He IIλ1640 and O III]λ1666 can be a powerful tool to separate between SF, shocks and AGN at sub-solar metallicities. We also confirm that alternative diagrams without O III]λ1666 still allow us to define a SF-locus with some caveats. Diagrams including C IV$\lambda\lambda$1548,51 should be taken with caution given the complexity of this doublet profile. Finally, we present a discussion detailing the ISM conditions required to detect UV emission lines, visible only in low gas-phase metallicity (12+log(O/H)$\lesssim$ 8.3) and high ionization parameter (log(U)$\gtrsim$ −2.5) environments. Overall, CLASSY and our UV toolkit will be crucial in interpreting the spectra of the earliest galaxies that JWST is currently revealing.

1. INTRODUCTION

Emission lines provide precious information about the conditions and ionizing sources of the interstellar medium (ISM) in galaxies across cosmic time. Several classification methods have been proposed to identify the main ionization sources: hot and young massive stars, tracing recent star formation (SF); post-asymptotic giant branch (post-AGB) stars, tracing older stellar populations; Active Galactic Nuclei (AGN), and thus the energy of the accretion disc surrounding a central supermassive black hole; low-ionization (nuclear) emission-line regions (LI(N)ERs); shocks, due to supernovae and stellar winds, outflows from starbursts/AGN activity and mergers, or from random motions of interstellar clouds; accretion on compact objects (e.g., High-Mass X-ray Binaries, HMXB); or a mixture of them (e.g., Heckman 1980; Izotov & Thuan 1999; Nagao et al. 2006; Kewley et al. 2006; Allen et al. 2008; Stasińska et al. 2015; Belfiore et al. 2016).

Historically, the main diagnostics to discriminate between different ISM ionization mechanisms are the classical optical diagnostic diagrams, first presented in Baldwin et al. (1981) (and usually referred as BPT diagrams) and then progressively updated by Keel (1983); Veilleux & Osterbrock (1987); Kauffmann et al. (2003); Kewley et al. (2006). These diagrams are constructed on intensity ratios of strong emission lines close in wavelength, namely [N II] λ6584/Hα, [S II] $\lambda\lambda$6717,31/Hα and [O I] λ6300/Hα versus [O III] λ5007/Hβ (i.e., [N II], [S II] and [O I] BPT diagrams), minimizing the effects of differential reddening by dust and flux calibration issues. They have been found useful to discriminate between H II-like sources and objects photoionised by a harder radiation field (e.g., power-law continuum by an AGN or shock excitation). However, existing BPT diagrams can be less effective in discriminating mechanisms other than SF in metal-poor galaxies (e.g., Groves et al. 2006; Reines et al. 2020; Polimera et al. 2022), since, on the one hand, the involved line ratios can be dependent on metallicity and, on the other, the hot metal-poor stellar

populations have a hard spectrum to irradiate the gas (e.g., Feltre et al. 2016; Byler et al. 2018; Xiao et al. 2018; Wofford et al. 2021). Also, these diagnostics are not optimal in discriminating shocks from other ionization sources.

Exploiting the large statistics of the SDSS DR7 (Abazajian et al. 2009), Shirazi & Brinchmann (2012) proposed a less metallicity-dependent diagnostic diagram to discriminate the ionization source: [N II]/Hα versus He II λ4686/Hβ. He II λ4686 is produced via recombination and thus indicates the existence of sources of hard ionizing radiation capable of ionizing the helium (ionization potential of 54.4 eV). Such a high radiation field can be produced by young stellar populations, including O-type stars (e.g., Sixtos et al. 2023) and some types of Wolf–Rayet (WR) stars (e.g., WC, WO, and WNE, Schaerer & Vacca 1998; see also Schmutz et al. 1992), but also other mechanisms such as AGN (Shirazi & Brinchmann 2012), X-ray binaries (Garnett et al. 1991; Mayya et al. 2023) or shocks (Thuan & Izotov 2005; Dopita & Sutherland 1996; Stasińska et al. 2015; Alarie & Morisset 2019) can play a role. Recently, Tozzi et al. 2023 investigated the Shirazi & Brinchmann (2012) diagnostic diagram with integral field spectroscopy (IFS) observations from the SDSS-IV MaNGA (Mapping Nearby Galaxies at APO; Bundy et al. 2015) survey, confirming its power in revealing AGN activity undetected from the classical BPT diagrams. Finally, to better disentangle the presence of shocks, Kewley et al. (2019) presented an overview of alternative optical diagrams, including [O III]λ5007, [O II]λ3727 and [O I]λ6300 emission lines (see also Heckman 1980; Allen et al. 2008) as well as the importance of correlations between line ratios and the velocity dispersion of the gas. Overall, the combination of (some of) these optical criteria provides a reliable powerful tool to investigate the ISM ionization sources in galaxies.

However, the well-studied optical emission lines (from [O II]$\lambda\lambda$3727 to [S III]$\lambda\lambda$9069, 9532) on which the diagnostics described in the previous paragraphs rely are not easily accessible to investigate the early phases of galaxy evolution in the epoch of reionization (EoR, $z > 6$) in the current

JWST and future extremely large ground-based telescopes (ELTs) era. For example, JWST/NIRspec does not cover anymore Hα at $z \gtrsim 6.5$ and [O III] $\lambda\lambda4959,5007$ at $z \gtrsim 10$, while JWST/MIRI is less sensitive and does not have multi-object slit capabilities. Hence, rest-frame UV emission lines have started to play a critical role in understanding EoR objects, thus making a reliable UV toolkit of ISM diagnostics essential. EoR galaxies are expected to be increasingly compact, metal-poor, with low-masses and large specific star formation rates (e.g., Wise et al. 2014; Madau & Haardt 2015; Robertson et al. 2015; Stanway et al. 2016; Stark 2016) and characterized by prominent high-ionization nebular UV emission lines (e.g., He IIλ1640, O III]$\lambda\lambda$1661,6, C III]$\lambda\lambda$1907,9, C IV$\lambda\lambda$1548,51; Stark et al. 2015; Mainali et al. 2017; 2018). The spectra of some of these objects have already started to be revealed with JWST/NIRSpec (e.g., Curti et al. 2022; Arellano-Córdova et al. 2022a; Curtis-Lake et al. 2022; Matthee et al. 2022; Trump et al. 2023; Arrabal Haro et al. 2023), with one of the furthest ($z \sim 10.603$) galaxies ever spectroscopically observed characterized by plenty of UV (from Lyα to Mg II$\lambda\lambda$2795,2802) and optical (from [O II]$\lambda\lambda$3727 to [O III]λ4363) emission lines (Bunker et al. 2023). In the foreseeable future, further JWST observations including large surveys such as GLASS (e.g., Treu et al. 2022; Roberts-Borsani et al. 2022a;b; Castellano et al. 2022), CEERS (e.g., Finkelstein et al. 2022; Arrabal Haro et al. 2023) and JADES (e.g., Curtis-Lake et al. 2022; Bunker et al. 2023) will provide more and more rest-frame UV spectra of $z > 6$ systems, exploring in depth the EoR.

Several recent works proposed UV alternatives to the optical diagnostics, either using photoionization models from SF and AGN (Feltre et al. 2016; Jaskot & Ravindranath 2016; Nakajima et al. 2018; Byler et al. 2020) and shock models (Jaskot & Ravindranath 2016), or from the comparison of either models and simulations (Hirschmann et al. 2019; 2022). For instance, Feltre et al. (2016) proposed various UV diagnostic plots, including the C III]$\lambda\lambda$1907,9/He IIλ1640 versus C IV$\lambda\lambda$1548,51/He IIλ1640 to discriminate between SF and AGN models. Jaskot & Ravindranath (2016) used the same diagnostic diagram to separate SF from shocks models, while Nakajima et al. (2018) proposed alternative models and diagrams based on C IV$\lambda\lambda$1548,51/C III]$\lambda\lambda$1907,9 and their equivalent widths (EWs). Finally, Hirschmann et al. (2019; 2022) coupled models and cosmological zoom-in simulations, exploring synthetic optical and UV emission-line diagnostic diagrams, listing the most promising in Table 1 of Hirschmann et al. (2022). However, UV diagnostics have yet to be compared against the optical ones in a systematic way, because it is not trivial to observe a full suite of optical and UV lines capable of probing the different ISM properties (e.g., density, temperature, metallicity, ionization parameter), which is fundamental for interpreting the results.

The analysis of local galaxies can provide the tools to interpret high-z galaxies, given the high signal-to-noise (S/N) of their spectra and the possibility of getting multi-wavelength coverage, which can help in the data interpretation. In Mingozzi et al. (2022) (Paper IV hereafter) we have started to create such a UV toolkit using the Cosmic Origins Spectrograph (COS) Legacy Spectroscopic SurveY (CLASSY[1]; Berg et al. 2022; James et al. 2022, Paper I and Paper II hereafter). CLASSY represents the first high-quality, high-resolution far-UV (FUV, $1150-2000$ Å) Treasury of 45 nearby ($0.002 < z < 0.182$) star-forming galaxies, with properties that make them representative at all redshifts (see Paper I Fig. 8), but generally characterized by more extreme ionization fields, lower stellar masses, and higher star formation rates (SFRs) than $z \sim 0$ objects, as typically observed in the EoR (e.g., Wise et al. 2014; Madau & Haardt 2015; Robertson et al. 2015; Stanway et al. 2016; Stark 2016). Specifically, in Paper IV we provided detailed measurements of dust attenuation ($E(B-V)$), electron density (n_e), electron temperature (T_e), gas-phase metallicity (12+log(O/H)) and ionization parameter (log(U)), using both UV and optical direct diagnostics, taking into account the different ionization zones of the ISM. Then, from the comparison of the derived properties, we provided a set of equations to estimate ISM properties only from UV emission lines. This paper presents the second part of the UV-based toolkit introduced in Paper IV and aims at exploring the well-known optical emission-line diagnostics tracing the different ionization sources and the recently proposed UV-based counterparts, to check their consistency and reliability to discriminate SF from AGN and shock ionization.

In Sec. 2 we describe the CLASSY sample, summarizing the characteristics of the UV and optical data, the main steps of the data analysis performed in Paper IV, while in Sec. 3, we give an overview of the models used in this work to interpret the optical and UV diagnostics. In Sec. 4 we present the main optical diagnostic diagrams sensitive to the ionization source to demonstrate that CLASSY galaxies are indeed dominated by SF. CLASSY represents the local reference sample of metal-poor star-forming galaxies that we can use to explore the boundaries of the star formation locus in UV diagnostic diagrams, as we show in Sec. 5, interpreting the results taking into account ISM properties and state-of-the-art photoionization and shock models. Finally, in Sec. 6 we discuss the conditions needed to observe UV emission lines and in Sec. 7 we summarize our main findings. Throughout this work we assume a flat ΛCDM cosmology (H$_0 = 70$ km/s/Mpc, $\Omega = 0.3$) and 12+log(O/H)$_\odot = 8.69$ (Asplund et al. 2009).

---

[1] https://archive.stsci.edu/hlsp/classy

2. THE CLASSY SURVEY

The 45 local galaxies from the CLASSY survey span a wide range of stellar masses ($6.2 < \log(M_\star/M_\odot) < 10.1$), SFRs ($-2 < \log(SFR/M_\odot \text{yr}^{-1}) < +1.6$), oxygen abundances ($7 < 12+\log(\text{O/H}) < 8.8$), electron densities ($10 < n_e([\text{S II}]\lambda\lambda6717,31)/\text{cm}^{-3} < 1120$), degree of ionization ($0.54 < O3O2 < 38.0$, with $O3O2 = [\text{O III}] \; \lambda\lambda5007/[\text{O II}] \; \lambda\lambda3727,9$), and reddening values ($0.02 < E(B-V) < 0.67$). In Paper I, we present our sample, explaining in detail the selection criteria along with an extensive overview of the HST/COS and archival optical spectra. The data reduction is presented in Paper II, including spectra extraction, co-addition, wavelength calibration, and vignetting. In Paper IV we explained in detail the data analysis, from the stellar continuum modeling to the emission line fitting, providing measurements for the UV redshifts (z_{UV}), the main UV and optical emission line fluxes and equivalent widths, and ISM properties. In the following subsections, we summarize the most important characteristics of the sample and the main steps of the data analysis performed in these previous works.

2.1. *The sample*

CLASSY spectra span from 1150 Å to 2100–2500 Å, combining the G130M, G160M, G185M, G225M and G140L *HST*/COS gratings, from 135 orbits of new COS data (PID: 15840, PI: Berg) and 177 orbits of archival COS data, for a total of 312 orbits. Both this paper and Paper IV focus on the analysis of all the emission lines (except for Lyα) in the range $1150 - 2000$ Å, using the so-called *High Resolution* (HR: G130M+G160M; $R \sim 10000 - 24000$) and *Moderate Resolution* (MR: G130M+G160M+G185M/G225M; $R \sim 10000 - 20000$) co-added spectra[2], presented in detail in Berg et al. (2022)[3]. We collected data for the entire CLASSY sample also in the optical wavelength regime, gathering SDSS spectra if available, integral field spectroscopy data (VLT/VIMOS, MUSE, Keck/KCWI), or long-slit spectra (MMT; LBT/MODS, see Arellano-Córdova et al. 2022b, CLASSY Paper V). A full description of the CLASSY optical dataset can be found in Paper IV[4].

In Paper IV and throughout this work, we take into account the properties of the CLASSY galaxies in terms of redshift, stellar mass $M_\star$, *SFR*, and 12+log(O/H), estimated and described in Paper I (see Sec. 4.5 and 4.7). Specifi-

cally, 12+log(O/H) measurements are based on the direct T_e method, using [S II] $\lambda6717/ \lambda6731$ and [O III] $\lambda4363/ \lambda5007$ as electron density and temperature tracers, respectively. We also take into account the ionization parameter $\log(U)$ from [O III]$\lambda5007$/[O II]$\lambda\lambda3727$ and the stellar age of our targets, determined from the analysis presented in Paper IV. For this paper, we also measured the carbon-to-oxygen abundance C/O, which we calculated following the prescriptions of Berg et al. (2016). Specifically, we calculated the ionic abundances of C^{++} and O^{++} from C III] $\lambda1907,9$/Hβ and [O III] $\lambda1666$/Hβ using PyNeb (Luridiana et al. 2015) and T_e([O III]). Then, we divided one by the other multiplying per the ionization correction factor (ICF), provided by Berg et al. (2016), taking into account $\log(U)$. We find $(\text{C/O})/(\text{C/O})_\odot$ to vary in the range 0.1–0.6, with $(\text{C/O})_\odot = 0.44$. A detailed discussion about CLASSY C/O will be given by Berg et al. (in preparation).

2.2. *Data analysis*

Here, we briefly summarize the main steps of the data analysis of UV and optical spectra performed in Paper IV (see their Sec. 3). In particular, we used a set of customized Python scripts in order to first fit and subtract the stellar component and then fit the main emission lines with multiple Gaussian components where needed.

As a preliminary step, both UV and optical CLASSY spectra were corrected for the total Galactic foreground reddening along the line of sight of their coordinates with the Cardelli et al. (1989) reddening law (see Paper I). Then, we fitted and subtracted the UV and optical stellar components, adopting the latest version of the Bruzual & Charlot (2003) models, assuming the Chabrier (2003) initial mass function (IMF), as described in detail in Paper IV Sec. 3.1. Finally, we performed the fit of the main UV and optical emission lines separately. In particular, we simultaneously fitted each spectrum taking into account a set of UV (from [N IV] $\lambda1483$, N IV] $\lambda1487$, to [C III] $\lambda1907$, C III] $\lambda1909$) and optical emission lines (from [O II] $\lambda\lambda3727$ to [S III] $\lambda9069$) with a linear baseline centered on zero and a single Gaussian, making use of the code MPFIT (Markwardt 2009), which performs a robust non-linear least squares curve fitting (see Paper IV Sec. 3.2 for more details).

We noted that the Hα line, in particular, shows a broad component in many CLASSY galaxies (as shown in the top panel of Fig. 1 for J0021+0052). Therefore, we performed 2-component Gaussian fits to the main optical emission lines (i.e., [O II] $\lambda\lambda3727,9$, Hδ, Hγ, Hβ, [O III] $\lambda\lambda4959,5007$, [O I] $\lambda\lambda6300,74$, [N II] $\lambda\lambda6548,84$, Hα, [S II] $\lambda\lambda6717,31$). In particular, we used a narrow ($\sigma < 200$ km/s) and a broad ($\sigma < 1000$ km/s) Gaussian component, tying the velocity and velocity dispersion of each component to be the same for all the emission lines. We applied a reduced chi-square $\tilde{\chi}^2$ selec-

---

[2] The instrumental broadening at the wavelengths of the emission lines taken into account in the HR and MR spectra is, respectively, $\sigma_{\text{int}} \sim 15 - 30$ km/s.

[3] Only for J1044+0353 and J1418+2102, we used the so-called *Low Resolution* (LR: G130M+G160M+G140L or G130M+G160M+G140L; $R \sim 1500 - 4000$, $\sigma_{\text{int}} \sim 80$ km/s) coadds because of the lack of higher resolution data.

[4] The instrumental broadening of the optical data is in the range $\sigma_{\text{int}} \sim 35 - 120$ km/s.
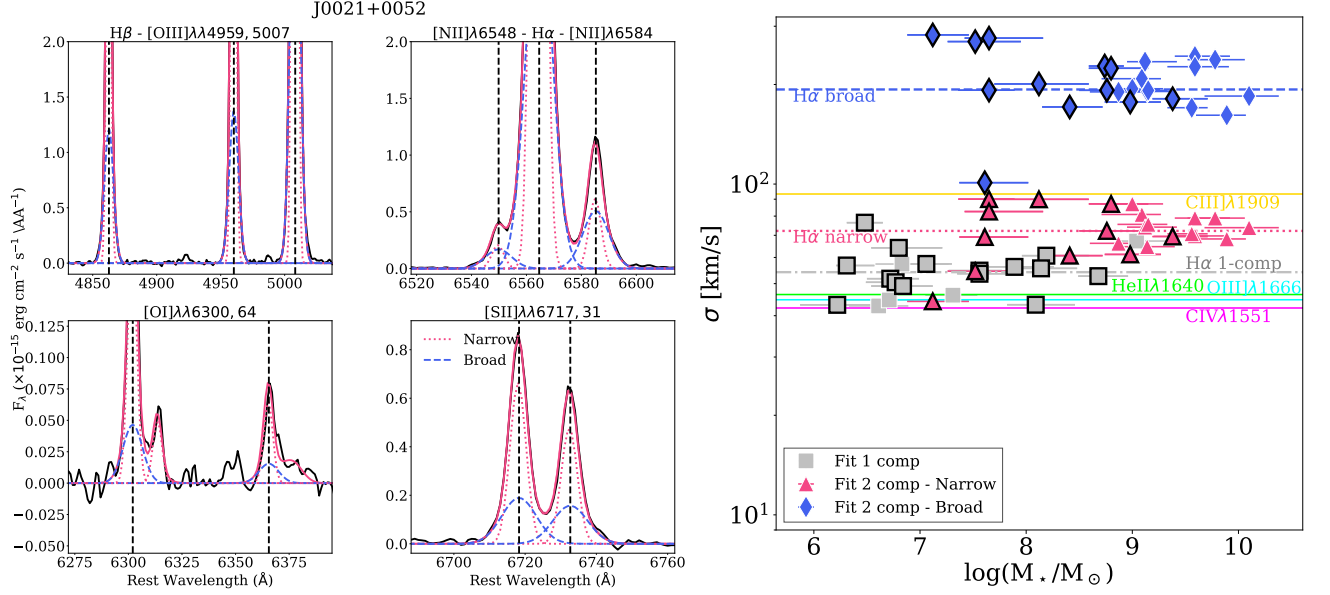
Figure 1. Left panel: optical emission lines best-fit with narrow and broad Gaussian components (dotted red and dashed blue lines, respectively; total fit shown by red solid line) for one of the CLASSY galaxies, J0021+0052. Here we show the lines needed to create the classical optical BPT diagrams (Hβ, [O III] $\lambda\lambda$4959,5007, [N II] $\lambda\lambda$6548,84, Hα, [O I] $\lambda\lambda$6300,64, [S II] $\lambda\lambda$6717,31) to discriminate among different ionization sources (see Sec. 4). The emission line next to [O I] λ6300 is [S III] λ6312, while the one next to [O I] λ6364 is Si II λ6371. Right panel: intrinsic velocity dispersion σ (y-axis in log scale) for the 1-component (grey squares) and, when needed, the 2-component best-fit (narrow and broad components shown as red triangles and blue diamonds, respectively) of optical emission lines, as a function of the stellar mass. The broad component σ (median value $\sim$ 190 km/s, dashed blue line) is generally larger than the narrow ($\sim$ 70 km/s, dotted pink line) and single ($\sim$ 55 km/s, dash-dotted grey line) components. Galaxies with at least one $S/N > 3$ UV emission line (i.e., C III]$\lambda\lambda$1907,9; 12+log(O/H) $\lesssim$ 8.3) are highlighted by a black thick edge. The UV lines intrinsic σ median values are generally consistent with the optical, as shown by the solid horizontal lines: C IV$\lambda\lambda$1548,51 (magenta), He IIλ1640 (green), O III]$\lambda\lambda$1661,6 (cyan), C III]$\lambda\lambda$1907,9 (gold).

tion in order to state if a fit with a second component better reproduces the observed spectral profiles (see Sec. 3.2.2 of Paper IV for more details). Overall, a second broad component is needed in 24 out of 45 galaxies of our sample. This component tends to be slightly blue-shifted ($\sim 50-80$ km/s) at increasing $M_\star$ and accounts for $\lesssim 30\%$ of the total emission (Paper IV).

The right panel of Fig. 1 shows the intrinsic velocity dispersion[5] σ that we calculate from the optical emission lines as a function of the galaxy stellar mass for our sample. In particular, we show either the single Gaussian σ (grey squares) or the narrow (red triangles) and broad component (blue diamonds) σ, according to the best-fit selected for each galaxy. The median values of the single, narrow, and broad components are ~ 55 km/s, ~ 70 km/s and ~ 190 km/s, respectively, as indicated by the dash-dotted grey, dotted red and dashed blue horizontal lines. The single Gaussian σ for the objects that need a 2-component fit is not shown, but its median value would be ~ 100 km/s, and thus much higher than

the value calculated for the narrow component. Overall, the broad emission component can trace gas at different velocities along the line of sight due to different mechanisms, including stellar winds, galactic outflows or turbulence, linked to different ionization sources (e.g., stellar or AGN photoionization and/or shocks; Izotov & Thuan 2007; James et al. 2009; Amorín et al. 2012; Bosch et al. 2019; Komarova et al. 2021; Hogarth et al. 2020). Hence, in Sec. 4 we use σ to explore whether correlations exist between gas kinematics and the optical diagnostics tracing the level of ionization of the gas (e.g., Kewley et al. 2019).

In Fig. 1, the CLASSY targets with at least one UV emission line (i.e., C III] $\lambda\lambda$1907,9) are highlighted by a black thick edge: they all have generally low stellar masses and 12+log(O/H) $\lesssim$ 8.3. C III] $\lambda\lambda$1907,9 is the most common UV emission line in CLASSY galaxies (observed in 28 objects) without considering Lyα (see Hu et al. 2023 submitted), as it is usually seen in low-mass SFGs (e.g., Rigby et al. 2015; Maseda et al. 2017; Du et al. 2017). Specifically, all the CLASSY galaxies showing either He II λ1640 (19) or O III] λ1666 (22) have also C III] emission (see Paper IV). Also, these three emission lines are all visible in the 9 CLASSY galaxies characterized by C IV $\lambda\lambda$1548,51 in pure emission. The colored solid horizontal lines in Fig. 1

---

[5] The emission line velocity dispersion σ refers to the Hα line for the optical and is corrected for the instrumental broadening σ_{int} of the different instruments taken into account, assuming $\sigma = \sqrt{\sigma_{obs}^2 - \sigma_{int}^2}$.

indicate their median intrinsic velocity dispersions, which we calculated, taking into account COS instrumental resolution as well as the extension of the sources derived from the COS NUV acquisition images (see Paper II Sec. 2 and Sec 3.2; James et al. 2014 Sec. 3.2). In general, He II λ1640, O III] λ1661,6 and C IV $\lambda\lambda$1548,51 share very similar values ($\sigma \sim 50$ km/s), consistent with optical emission lines at similar stellar masses. We will further comment on this in Sec. 6. We emphasize that the UV emission-line S/N is typically not high enough to detect faint broad components in the emission lines, apart from a few exceptions (Paper IV). Also, UV lines are not seen in the most massive targets in which the optical line broad components are particularly enhanced. However, the fact that σ(C III) is systematically higher than other UV lines (σ(C III) ~ 90 km/s) could indicate the presence of a hidden undetected broad component. A further discussion of the optical and UV kinematical properties of the sample can be found in App. A.

3. DIAGNOSTIC MODELS USED WITHIN THIS WORK

This work aims to understand how state-of-the-art models from the literature compare to observed line ratios of the CLASSY survey, using well-studied optical diagnostics and less well-constrained UV diagnostics. This section provides an overview of each model considered for the main sources of ionization found within galaxies: photoionization, AGN, and shocks. We provide a summary of the models taken into account in Tab. 1. In the next sections of the paper, we investigate and comment on optical and UV diagnostics to distinguish between these different ionizing mechanisms.

Concerning stellar photoionization grid models (referred to as SF models hereafter), we take into account emission from H II regions around young stars, using nebular-emission models of SFGs from Gutkin et al. (2016) (G16 hereafter) based on two limiting cases: one instantaneous burst (single stellar population; SSP models) or constant activity with time (CST models). In particular, single-burst models, in which the emission is strongest at young ages (i.e., from 3-4 Myr up to 5-7 Myr if stellar rotation is included) and disappears as the massive star population evolves, are generally used for young stellar populations and very recent starbursts, and thus more consistent with the CLASSY galaxies and EoR systems. G16 models do not take into account either stellar rotation or stellar multiplicity, which would imply a longer duration of ionizing photon production (e.g., MIST models from Byler et al. 2017 models) and, in the latter case, harder radiation field (e.g., BPASS models from Xiao et al. 2018). Other possible differences between G16 and other models in the literature concern the choice of stellar libraries, IMF or the way to scale gas-phase metallicity for elements such as nitrogen and carbon[6]. For the purpose of this study, we just take into account G16 models for SF grids, since their parameters (see Tab. 1) are well matched with the AGN and shock grids that we present below, allowing us to make a consistent comparison between optical and UV diagnostic diagrams.

As shown in Feltre et al. (2016) Fig. 1, the ionizing spectrum powered by a lower metallicity stellar population is harder with respect to a more metal-enriched one, producing a different spectral energy distribution at photon energies greater than ~ 20 eV. However, it generally cannot account for ionic species with ionization energies above ~ 50 eV. Hence, the presence of emission lines requiring such ionization energy could be a good indicator of the presence of an additional ionization mechanism than SF (e.g., AGN or shocks). In this work, we take into account the AGN models from Feltre et al. (2016) (F16 hereafter). The main parameters to take into account are the slope α of the ionizing spectrum, as well as the other quantities shown in detail in Tab. 1. Collisional excitation and photoionization from slow ($v \lesssim 150$ km/s) and fast shocks ($v \sim 150-1000$ km/s), respectively, can also produce a rich spectrum of UV and optical emission lines, since shocked regions have high electron temperature and ionization state (Dopita & Sutherland 2003). Here we take into account the shock models from the 3MdBs[7] database (Alarie & Morisset 2019; AM19). These model grids span a broad range of metallicity, matched with Gutkin et al. (2016) grid values (see Tab. 1), with respect to the well-known MAPPINGS models from Allen et al. (2008)[8], and thus are the best suited for the CLASSY sample.

---

[6] The former has secondary nucleosynthetic production at high metallicity, while part of the latter returns to the ISM through metallicity-dependent processes such as stellar winds (see e.g., Byler et al. 2018 for more details).

[7] http://3mdb.astro.unam.mx/

[8] AM19 models are made with MAPPINGS-V and agree with Allen et al. (2008) predictions at LMC metallicity (see AM19 for more details).

Table 1. Summary of the models taken into account in this work: constant (CST) and single-burst (single stellar population; SSP) Gutkin et al. (2016) grids (G16) for SF, Feltre et al. (2016) (F16) for AGN narrow-line region (NLR), Alarie & Morisset (2019) (AM19) for shocks. The parameters include the ionizing spectrum, gas-phase metallicity (Z_{ISM}), ionization parameter log(U_S), dust-to-gas ratio (ξ_d), carbon-to-oxygen abundance (C/O) and initial mass function (IMF). ξ_d accounts for the depletion of metals onto dust grains in the ionized gas. We refer to Sec. 4 of Plat et al. (2019) to see the impact of the variation of the parameters that we fix in this work, as indicated in the notes of this table.

Model	Parameters	Sampled Values	Notes
G16	Ionizing spectrum	CST SF; age = 100 Myr	Bruzual & Charlot (2003) stellar pop. synthesis models + Cloudy (Ferland et al. 2013). The values of gas-phase metallicity correspond to 12+log(O/H) = 6.64, 7.64, 8.55, 8.88 and log(Z_{ISM}/[$Z_\odot$]) −2.2–0.04. We tested all values of n_H, ξ_d and (C/O), but here we show only $n_H = 100$ cm^{-3}, $\xi_d = 0.3$ and (C/O)/[(C/O)$_\odot$] = 0.1, 0.72. Special prescriptions for C and N because scale non-linearly with O (G16 Sec. 2.3.1; see Henry et al. 2000; Berg et al. 2016).
	Z_{ISM}	0.0001,0.001,0.008,0.017; $Z_\odot$ = 0.01524	
	n_H/[cm^{-3}]	10, 100, 1000, 10000.	
	log(U_S)	−3.5, −3. , −2.5, −2. , −1.5, −1.	
	ξ_d	0.1, 0.3, 0.5	
	(C/O)/[(C/O)$_\odot$]	0.10, 0.14, 0.20, 0.27, 0.38, 0.52, 0.72, 1.00, 1.40; (C/O)$_\odot$ = 0.44	
	IMF	Chabrier (2003) (m_{min} = 0.1 M$_\odot$; m_{up} = 100.,300. M$_\odot$)	
G16	Ionizing spectrum	SSP; age = 1, 2, 3, 5, 7, 9, 10 Myr	Same as in CST G16 grids. We tested all values of age, but we mainly show in plots grids at 3 Myr and 5 Myr. Given the absence of stellar rotation and multiplicity, the ionizing radiation drops at older ages than 5 Myr.
	Z_{ISM}	Same as in CST G16	
	n_H/[cm^{-3}]	Same as in CST G16	
	log(U_S)	Same as in CST G16	
	ξ_d	Same as in CST G16	
	(C/O)/[(C/O)$_\odot$]	Same as in CST G16	
	IMF	Same as in CST G16	
F16	Ionizing spectrum	AGN NLR; α = −1.2, −1.4, −1.7, −2.0	Cloudy models with open geometry, luminosity $L_{AGN} = 10^{45}$ erg/s, inner radius from the NLR of 300 pc. Ionizing spectrum parameterized with the power-law index α of the spectral energy distribution of the incident ionizing radiation from the AGN accretion disc at UV and optical wavelengths in AGN models. All heavy elements except for nitrogen (see F16 Eq. 6) are assumed to scale linearly with oxygen abundance.
	Z_{ISM}	Same as G16	
	n_H/[cm^{-3}]	Same as G16	
	log(U_S)	Same as G16	
	ξ_d	Same as G16	
AM19	Ionizing spectrum	Fast shocks ($v_s > 100$ km/s)	Mappings-V (Sutherland & Dopita 2017) and a plane-parallel geometry. Models created using same prescriptions of the well-known Allen et al. (2008) shock models by AM19, but using the metallicity values of G16, to study variation with metallicity. We tested all values, but in this paper we show only grids with $n_H = 1$ cm^{-3}, (C/O)/(C/O)$_\odot$ = 0.26 for Z_{ISM} = 0.0001, 0.001 and (C/O)/(C/O)$_\odot$ = 1.00 for Z_{ISM} = 0.017 (trans-solar metallicity grid).
	Z_{ISM}	0.0001, 0.001, 0.008, 0.08, 0.017	
	n_H/[cm^{-3}]	1., 10., 100., 1000., 10000.	
	v_s/[km/s]	100, 125, ..., 1000	
	(C/O)/[(C/O)$_\odot$]	0.26, 1.00	
	B_0/[μG]	10^{-4}, 0.5, 1.0, 2.0, 3.23, 4.0, 5.0, 10	

The emission line spectra primarily depend on the shock velocity v_s and the magnetic field parameter B_0, which regulate the shock ionization spectrum and the effective ionization parameter. An increase in the density of the material preceding the shock ionization front (preshock density n_H) can play a role as well because it enforces the collisional de-excitation of forbidden lines. Generally, two sets of models are taken into account: shocks and shocks+precursor. The precursor emission becomes important at increasing velocities of the shock front ($v \gtrsim 170$ km/s), where the ionization front velocity exceeds v_s. Indeed, the high velocity causes the photoionization front to detach from the shock and form a "precursor", which can contribute to or even dominate the shock emission. In general, we noticed that an increase in density from $n_H \sim 1$ cm^{-3} to $n_H \sim 1000$ cm^{-3} does not lead to dramatic changes in the optical and UV line ratios discussed in the following sections unless noted in the text. Hence, in this work, we show only grids with $n_H = 1$ cm^{-3} (consistently with shock grids usually shown in other works; e.g., Allen et al. 2008). Also, at sub-solar metallicities, we show the grids with $(C/O)/(C/O)_\odot = 0.26$, which are also more consistent with the properties of the CLASSY galaxies.

4. OPTICAL DIAGNOSTICS SENSITIVE TO THE IONIZATION SOURCE

Here we present the most frequently used classification methods in the optical, showing and discussing where the CLASSY galaxies locate in different diagnostic diagrams. In Table 2, we summarize the optical diagnostics and the corresponding classifications for each galaxy, highlighting when this is different from star formation. We anticipate that we find that CLASSY galaxies are characterized by [O III]/Hβ narrow and broad components typical of star-forming galaxies. Furthermore, their broad components — if present — have enhanced [N II]/Hα (and slightly enhanced [S II]/Hα and [O I]/Hα) and [O I]/[O III] (possible shock evidence), but lower [O III]/[O II] (i.e., lower level of ionization) and do not show a trend with He IIλ4686 (that could be enhanced by shocks/AGN). We do not preclude that a few of them can have shocks (and possibly hidden AGN activity in the trans-solar galaxy J0808+3948), but, in general, the main ionizing mechanism is clearly star formation.

This conclusion is not surprising since the sample was selected excluding targets with a classification other than star-forming, but here we confirmed it by exploring different diagnostic diagrams. In particular, we compared different criteria to discriminate between ionization processes often finding that there is not a one-to-one agreement in classification, especially for low-mass and low-metallicity galaxies (e.g., Groves et al. 2006; Shirazi & Brinchmann 2012; Sartori et al. 2015; Polimera et al. 2022; Nakajima & Maiolino 2022). Overall, a combination of criteria represents the best way to

provide a robust determination of the main ionization source. This section is important since in Sec. 5 we want to test how UV diagnostic diagrams compare to the well-known and reliable optical ones presented here. In the following plots, we highlight with a black thick edge the objects with at least one UV line detection (i.e., S/N(C III) $\lambda\lambda$1907,9) > 3), since only a sub-sample of low-metallicity objects shows UV emission lines.

4.1. *The classical optical BPT diagrams*

The three panels of Fig. 2 illustrate the [N II], [S II] and [O I]-BPT diagrams for the single or narrow (red dots) and broad (blue diamonds) components of the CLASSY galaxies. The G16 models are not shown for clarity reasons but can completely cover the classical SF locus, while the F16 models predict higher [O III]/Hβ than G16, consistently with the classical AGN locus. According to these diagnostics, the majority of the CLASSY galaxies are dominated by SF, even though the classification is not always consistent among different diagnostics, with many objects at the edge of the star formation locus (indicated by the black solid maximum starburst line). In particular, at these edges, we mainly find the CLASSY objects characterized by UV emission lines (highlighted by thick black edges), that have generally higher [O III]/Hβ than the others and low gas-phase metallicity (12+log(O/H) $\lesssim$ 8.3, corresponding to $Z \lesssim 50\%$ $Z_\odot$). However, as we discuss in the following paragraphs, two main factors can affect the definition of the star formation locus and need to be taken into account: the metallicity and the hardness of the radiation field.

On the one hand, [N II]/Hα is particularly sensitive to metallicity (Kewley & Dopita 2002; Denicoló et al. 2002; Pettini & Pagel 2004; Kewley & Ellison 2008), with lower [N II]/Hα at decreasing metallicity (see e.g., the CLASSY galaxies J1323-0132 and J0808+3948 in Paper V Fig. 2). Since higher [N II]/Hα line-ratios are required to classify galaxies as AGN, the [N II]-BPT by itself is not sufficient for identifying AGN in typical low-metallicity star-forming $z \sim 0$ galaxies (e.g., Groves et al. 2006; Reines et al. 2020; Molina et al. 2021; Polimera et al. 2022). For instance, Polimera et al. (2022) showed that the [O I]-BPT[9] can identify a theoretical dwarf AGN with a spectrum characterized by $\sim 90\%$ SF contribution and classified as star-forming according to the [N II]-BPT (see also Hogarth et al. 2020). Indeed, [O I]/Hα is highly sensitive to hard radiation fields, particularly shock emission in the neutral ISM (e.g., Osterbrock & Ferland 2006; Allen et al. 2008) and is not very dependent on metallicity. Hence, an [O I]/Hα enhancement could suggest that there is an additional excitation mechanism other

---

[9] Polimera et al. (2022) also tested the [S II]-BPT diagram, which resulted in being less sensitive to SF dilution compared to the [O I]-BPT.

Table 2. CLASSY sample classification according to the optical diagnostic diagrams taken into account: [NII]-BPT, [SII]-BPT, [OI]-BPT, [OII]-diagram, [OI]-[OII]-[OIII]-diagram and HeII-diagram. For the optical diagnostics, we specify the classification of both the narrow (N) and (B) - if present - components. The 28 galaxies in bold are those which also show UV emission lines (i.e., C III]$\lambda\lambda$1907,9; 12+log(O/H) $\lesssim$ 8.3, $Z \lesssim$ 50 % Z$_\odot$).

Object		[NII]-BPT N/B	[SII]-BPT N/B	[OI]-BPT N/B	[OII] N/B	[OI]-[OII]-[OIII] N/B	HeII N/B
J0021+0052		SF/SF	SF/SF	SF/SF	*Comp/Comp*	SF/SF	SF
J0036-3333	Haro 11 knot C	SF/SF	SF/SF	SF/SF	...	SF/SF	SF
J0127-0619	Mrk 996	SF/SF	SF/NC	SF/NC	...	SF/NC	NC
J0144+0453	UM 133	SF/NC	*AGN*/NC	*AGN*/NC	Comp/NC	SF/NC	SF
J0337-0502	SBS 0335-052 E	SF/NC	SF/NC	SF/NC	...	SF/NC	SF/NC
J0405-3648		SF	SF	SF	...	SF	SF
J0808+3948		*Comp*/SF	SF/SF	SF/SF	SF/NC	*AGN-shocks*/NC	*AGN*
J0823+2806	LARS9	SF/*Comp*	SF/SF	SF/SF	*Comp/Comp*	SF/SF	SF
J0926+4427	LARS14	SF/*Comp*	*AGN*/SF	*AGN*/SF	*Comp/Comp*	SF/SF	NC
J0934+5514	Izw 18	...	...	...	*Comp*	...	
J0938+5428		SF/SF	SF/SF	SF/SF	*Comp/Comp*	SF/SF	SF
J0940+2935		SF	SF	SF	...	SF	NC
J0942+3547	CG-274, SB 110	SF	SF	SF	...	SF	SF
J0944+3442		SF	SF	SF	...	*AGN-SF*	NC
J0944-0038	CGCG007-025	SF	SF	SF	*Comp*	SF	SF
J1016+3754	1427-52996-221	SF	SF	SF	...	SF	SF
J1024+0524	SB 36	SF	SF	SF	*Comp*	SF	SF
J1025+3622		SF/SF	SF/SF	*AGN*/SF	*Comp/Comp*	SF/SF	SF
J1044+0353		SF	SF	SF	...	SF	SF
J1105+4444	1363-53053-510	SF/SF	SF/SF	SF/*AGN*	*Comp/AGN*	SF	SF
J1112+5503		*Comp*/SF	SF/SF	SF/SF	SF/SF	SF/SF	*Comp*
J1119+5130		SF	SF	SF	...	SF	SF
J1129+2034	SB 179	SF	SF	SF	...	SF	SF
J1132+1411	SB 125	SF	SF	SF	Comp	SF	SF
J1132+5722	SBSG1129+576	SF	SF	SF	...	SF	SF
J1144+4012		SF/SF	SF/SF	SF/SF	Comp/LI(N)ER-Comp	SF	NC
J1148+2546	SB 182	SF	SF	SF	Comp	SF	SF
J1150+1501	SB 126, Mrk 0750	SF	SF	SF	...	SF	SF
J1157+3220	1991-53446-584	SF	SF	SF	...	SF	SF
J1200+1343		*Comp/Comp*	SF/SF	*AGN*/NC	Comp/Comp	SF/NC	SF
J1225+6109	0955-52409-608	SF/NC	SF/NC	SF/NC	...	SF/NC	SF
J1253-0312	SHOC391	...	...	...	*Comp/Comp*	SF/SF	NC
J1314+3452	SB 153	SF	SF	SF	...	SF	SF
J1323-0132		SF	SF	SF	*Comp*	SF	SF
J1359+5726	Ly 52, Mrk 1486	SF/SF	SF/SF	SF/SF	*Comp/LI(N)ER-Comp*	SF/SF	SF
J1416+1223		*Comp*/SF	SF/SF	SF/SF	SF/SF	SF/SF	NC
J1418+2102		SF	SF	SF	...	SF	SF
J1428+1653		SF/SF	SF/SF	SF/SF	*Comp/Comp*	SF/SF	NC
J1429+0643		SF/SF	*AGN*/SF	*AGN/AGN*	*Comp/Comp*	SF/SF	SF
J1444+4237	HS1442+4250	SF	SF	SF	...	SF	NC
J1448-0110	SB 61	SF	SF	SF	*Comp*	SF	SF
J1521+0759		SF/SF	SF/SF	*AGN*/SF	*Comp/Comp*	SF/SF	NC
J1525+0757		SF/SF	SF/SF	SF/SF	SF/NC	SF/SF	NC
J1545+0858	1725-54266-068	SF/*Comp*	SF/SF	SF/*AGN*	*Comp/Comp*	SF/SF	SF
J1612+0817		*Comp/Comp*	SF/SF	SF/SF	SF/SF	SF/SF	SF
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)

NOTE—Main ionization source classification according to the [N II]-BPT (3), [S II]-BPT (4), [O I]-BPT (5), the [O II]$\lambda\lambda$3727 diagram from Lamareille et al. (2004); Lamareille (2010) (6), the [OI]-[OII]-[OIII] discussed in Sec. 4.2 (7), He IIλ diagram discussed in Sec. 4.3 (8): star formation (SF), composite (Comp), Active Galactic Nuclei (AGN), or, if one or more of the involved lines have $S/N < 3$, not classified (NC). The '...' indicates that some of the involved emission lines are out of the observed wavelength range. For further details about column (6) see App. B. None of the CLASSY galaxies is classified as a low-ionization (nuclear) region or LI(N)ER.
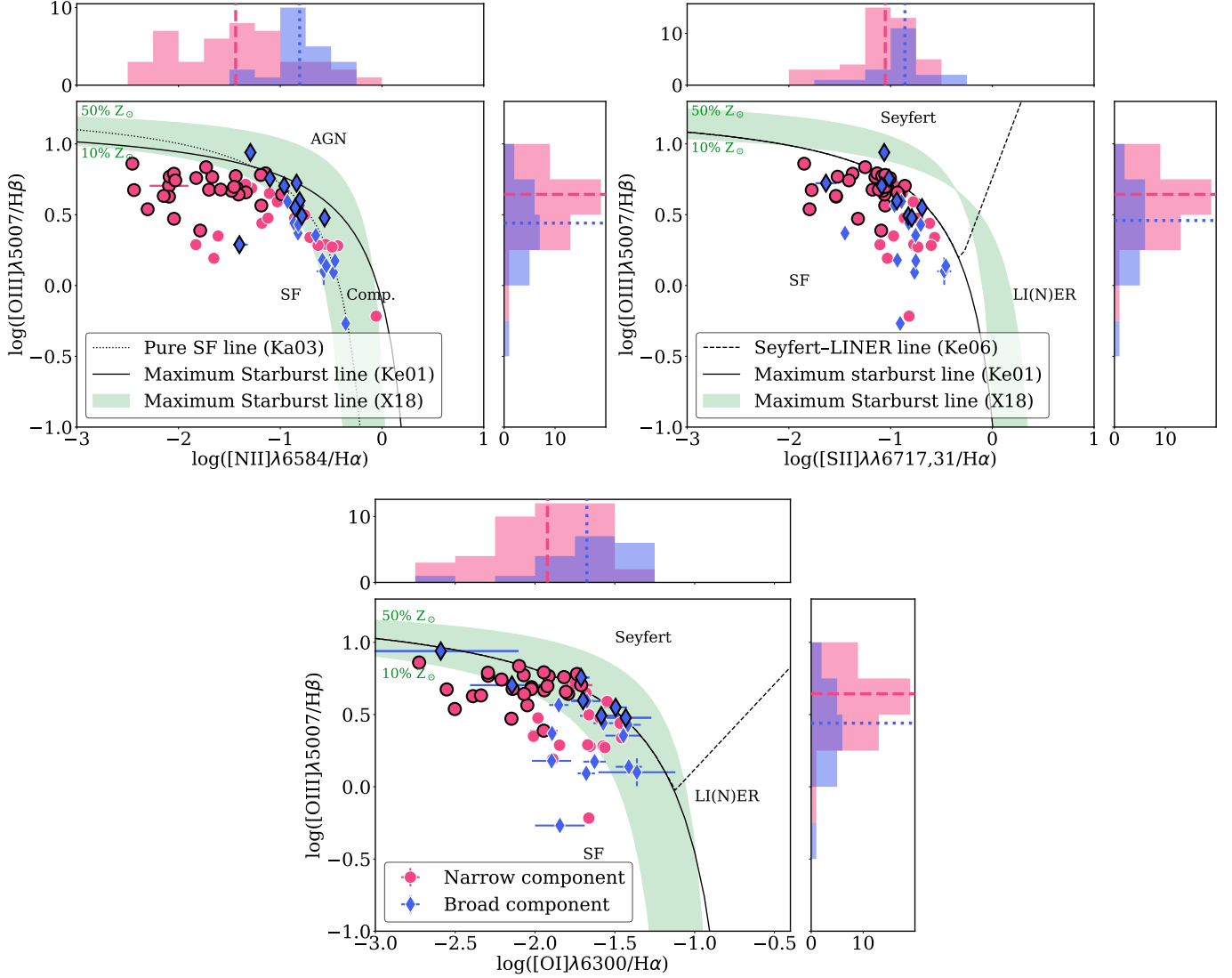
Figure 2. [N II] (top left), [S II] (top right), and [O I] (bottom) BPT diagrams for the narrow and broad components, shown as red dots and blue diamonds, respectively. All the measurements have $S/N > 3$ for all the lines involved. The black dotted curve is the empirical relation to divide pure star-forming from Seyfert–H II composite objects by Kauffmann et al. (2003), while the black solid curves are the maximum starburst line derived by Kewley et al. (2001). The black dashed lines are the Kewley et al. (2006) boundary between Seyferts and low-ionization (nuclear) emission-line regions (LI(N)ERs; see also Heckman 1980; Belfiore et al. 2016). Histograms showing the distribution of [N II]/Hα, [S II]/Hα, [O I]/Hα, and [O III]/Hβ line ratios in the narrow and broad components are shown in red and blue, respectively, on each axis, with median values shown by the dashed and dotted black lines. The broad component clearly shows higher x-axis line ratios. The green-shaded regions represent the maximal starburst predictions at metallicity between 10% $Z_{\odot}$ and 50% $Z_{\odot}$ values, using pure stellar photoionization including massive star binaries through the BPASS stellar evolution models (Xiao et al. 2018), to show to what extent the maximal starburst line can change in the metallicity range in which we observe UV emission lines. The galaxies with at least one UV emission line (i.e., C III] $\lambda\lambda1907,9$) are highlighted by a black thick edge.

than stellar photoionization, missed by the [N II]-BPT diagram. Looking at Table 2, some galaxies are classified as SF-dominated according to the [N II]-BPT and as "AGN" according to the [O I]-BPT, because they are located slightly beyond the Kewley et al. (2006) maximal starburst line. However, here we stress that the [O I]λ6300 line is fainter than [N II] and [S II], and thus the [O I]/Hα line ratio has larger error bars, as shown in Fig. 2.

On the other hand, low metallicity stellar populations produce harder radiation fields and thus higher [O III]/Hβ (and lower low-ionization line ratios) without the need to invoke other mechanisms than SF (e.g., Feltre et al. 2016; Byler et al. 2018; Xiao et al. 2018). In this context, Xiao et al. (2018) explored the variation of the *maximal starburst line* (i.e., the highest line ratios reproduced with SF models) in BPT diagrams as a function of the metallicity, using pure stellar photoionisation including massive star binaries through the BPASS (Binary Population and Spectral Synthesis; Eldridge & Stanway 2009) stellar evolution models[10]. The green shaded regions in Fig. 2 show the predictions of the Xiao et al. (2018) *maximal starburst lines* between $Z \sim 10\%$ Z$_\odot$ and $Z \sim 50\%$ Z$_\odot$ to underline how difficult the separation between different ionization mechanisms can be for galaxies that lie very close to the edges of the classical BPT diagrams in this range of metallicity (at which we observe also the UV emission lines we discuss in Sec. 5). Overall, we stress that taking into account the criteria of Xiao et al. (2018) and the gas-phase metallicity measured for each object of our sample, the classification of our galaxies is mainly SF.

Another interesting implication from Fig. 2 is that the broad component of our galaxies is usually characterized by higher [N II]/Hα, slightly higher [S II]/Hα, and [O I]/Hα line ratios, and slightly lower [O III]/Hβ line ratios. This difference (up to ~ 0.5 dex in [N II]/Hα) is shown by the red and blue histograms on each axis of Fig. 2. In particular, strong low-ionization lines could be a signature of shocks (Osterbrock & Ferland 2006). Also, the broad component has (by definition) a much larger velocity dispersion than the narrow one, as shown in Fig. 1. This σ enhancement at increasing [N II]/Hα, [S II]/Hα, and [O I]/Hα line ratios suggests that the kinematics and ionization state could be coupled, implying the same physical origin, which could be different from the one producing the narrow components. This correlation has been already found in normal star-forming galaxies and AGN (e.g., Ho et al. 2014; Mingozzi et al. 2019; Hogarth et al. 2020), and it was attributed to a shocked gas component, as stellar/AGN photoionization is not expected to cause such a trend (see e.g. Dopita & Sutherland 1995; Rich et al.

2010; McElroy et al. 2015; Kewley et al. 2019. We note that the σ enhancement is more evident in [N II]/Hα and so could also be related to a nitrogen enhancement of the more perturbed gas (e.g., James et al. 2009). We will explore this further in Arellano-Córdova et al. 2023 in prep.

Finally, in Appendix B we also comment about an alternative optical BPT diagnostic diagram that exploits only line ratios in the blue part of the optical spectrum (i.e., [O II] $\lambda\lambda$3727/Hβ versus [O III] λ5007/Hβ, Fig. 12; Lamareille et al. 2004; Lamareille 2010)[11].

To summarize, since [O III]/Hβ is in the star formation locus in the BPT diagrams (Fig 2) and is not enhanced in the broad component, we think that is unlikely that AGN are present in our sample. Here we acknowledge that recently Hatano et al. (2023) claimed that SBS 0335-052 E hosts an AGN on the basis of a recent NIR variability and the broad Hα component. We also reveal a broad Hα component ($\sigma \sim 1000$km/s) not visible in the other emission lines (this could be due to low S/N and low metallicity), but the narrow component line-ratios analyzed in this work are all classified as SF-dominated. Finally, the [O I]-excess that we observe in some galaxies of the sample, as well as the increase of [N II]/Hα, [S II]/Hα, and [O I]/Hα line ratios with velocity dispersion, indicate the possible presence of additional mechanisms such as shocks, which we further discuss in the following section.

4.2. *[OI] λ6300 as a shock indicator*

Fig. 3 shows the [O III]λ5007/[O II]$\lambda\lambda$3727 versus [O I]λ6300/[O III]λ5007 diagram ([OI]-[OII]-[OIII] diagram) for the CLASSY galaxies (narrow and broad components shown as in Fig. 2), presented as a diagnostic plot for the first time in Heckman (1980) to separate Seyferts from LINERs (see also Stasińska et al. 2015). This diagram is not very dependent on gas-phase metallicity (Stasińska et al. 2015), avoiding the drawbacks discussed for the classical [N II]-BPT diagram. Clearly, Fig. 3 represents a good tool to discriminate between SF, AGN, and shocks for our sample, as also shown by the good separation of the SF, AGN, and shock model grids, with just a slight overlap between the grids at the highest [O I]/[O III] line ratios[12]. We highlight that increasing the preshock density n_H from $n_H \sim 1$ cm^{-3} to $n_H \sim 1000$ cm^{-3} would lead to higher log([O III]/[O II]), that can increase up to $\sim 0.5 - 1.5$ for shock+precursor and pure-

---

[10] Binary-star evolution pathways produce harder radiation up to older ages than single stars mainly because of the gas accretion onto compact objects (e.g., HMXBs).

[11] This diagram is useful to check two galaxies of our sample, the famous blue compact dwarf I Zw 18 (J0934+5514) and J1253-0312, which are not shown in Fig. 2 because we do not have Hα coverage (see Paper IV Sec. 3). These two objects are classified as star-forming in Fig. 12.

[12] Kewley et al. (2019) present different shock diagnostics (see their Fig. 11), including one comparing [O III]/[O II] versus [O I]/Hα, that we tested but excluded because the AGN and SF grids of models could not be separated as well as in Fig. 3.
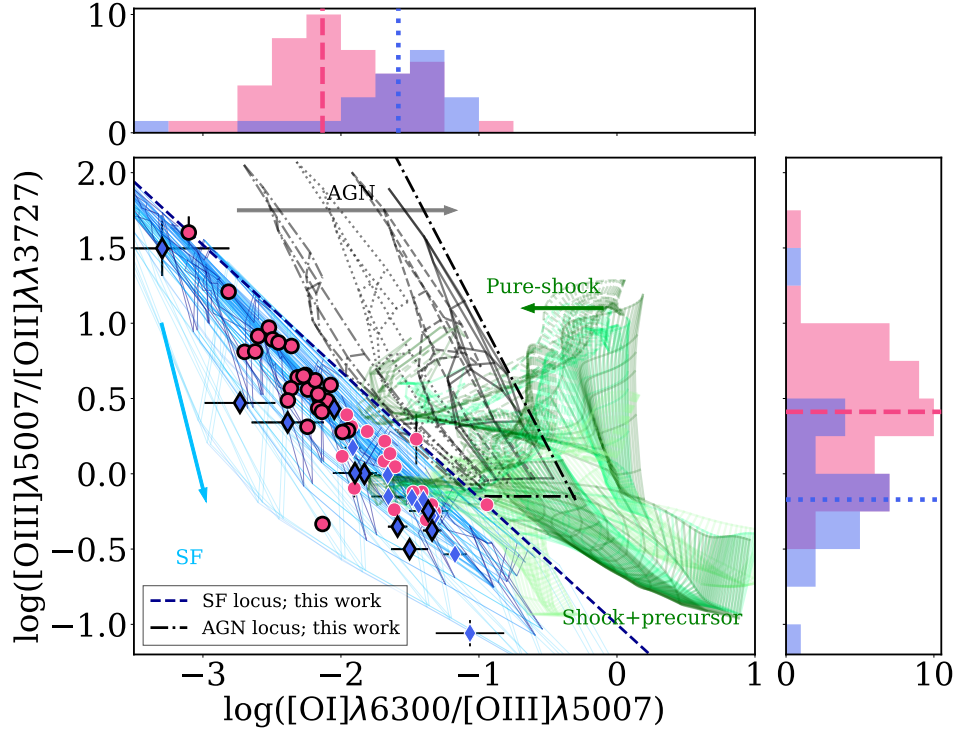
Figure 3. Location of the CLASSY galaxies in the log([O III]/[O II]) versus log([O I]/[O III]) diagnostic diagram from Kewley et al. (2019) (see also Stasińska et al. 2015). The narrow and broad components are shown as red dots and blue diamonds and the histograms on the axes show their distributions and median value as in Fig. 2. The blue grids show the G16 single-burst SF models, with ages of 3 Myr (darker blue) and 5 Myr (lighter blue). The G16 SF grids indicatively follow the blue arrow at increasing ages. The G16 constant SF models are not shown for clarity, because they completely overlap the single-burst grids. The black grids show the F16 AGN models, with the parameter α increasing from -2 (dashed-dotted grid) to -1.2 (solid grid) - and thus harder AGN ionizing radiation - in the direction of the black arrow. AM19 pure-shock (dashed) and shock+precursor (solid) model grids at varying velocities and magnetic fields are shown at different metallicities, with increasing values from lighter to darker green (following the green arrow). Shock models are more sensitive to the metallicity and have log([O III]/[O II]) $\gtrsim 0$, but lower log([O I]/[O III]) ($\lesssim 0$) than shock+precursor ones, that can reach log([O I]/[O III]) $\lesssim 1$. The dashed blue line qualitatively separates the SF from AGN/shock models, while the dash-dotted black line indicates the AGN locus. This plot represents a good tool to separate SF, AGN and shock models for our sample, which includes sub-solar to solar metallicity targets, underlining the importance of [O I] $\lambda6300$ as a good shock diagnostic at optical wavelengths. Only two CLASSY galaxies are at the edge of the star-forming locus. The galaxies with at least one UV emission line (i.e., C III]$\lambda\lambda1907,9$; 12+log(O/H) $\lesssim 8.3$) are highlighted by a black thick edge.

shock grids, respectively. Still, shock models are generally characterized by higher log([O I]/[O III]) than AGN models, reaching values up to ~ 1. The dashed blue line

$$y = -0.84x - 1.0 \qquad (1)$$

qualitatively separates the SF from AGN/shock models, while the dash-dotted black line indicates the AGN locus:

$$
\begin{aligned}
y &= -0.15 \quad and \quad -0.96 < x < -0.3 \\
y &= -1.75x - 0.70 \quad and \quad -3.5 < x < -0.3
\end{aligned}
\qquad (2)
$$

with $y = \log([\text{O III}] \ \lambda5007/[\text{O II}] \ \lambda\lambda3727)$ and $x = \log([\text{O I}] \ \lambda6300/[\text{O III}] \ \lambda5007)$.

Overall, Fig. 3 shows that the vast majority of the CLASSY galaxies can be reproduced by SF models, further confirming the classification obtained with the classical BPT diagrams (Fig. 2). In particular, the optical line broad velocity components show a significant enhancement in terms

of the [O I]/[O III] line ratio (with the broad showing a larger median value by ~ 0.5 dex; except for J0926+4427, J1148+2546 and J1253-0312) and a lower [O III]/[O II] (in all galaxies except for J1148+2546). Shocks can naturally explain high [O I]/[O III] line ratio with high [O III]/[O II], while photoionization models that produce high [O I]/[O III] show, in turn, low [O III]/[O II] (e.g., Stasińska et al. 2015; Plat et al. 2019). This suggests that also the more turbulent ionized gas has a stellar source of ionization. The narrow components of only two CLASSY galaxies are at the edge of the SF locus defined by our Eq. 1 in this diagram: J0944+3442 and J0808+3948, which could be both with AGN and shock models. J0944+3442 is classified as star-forming according to the optical BPT diagrams, while only the narrow component of J0808+3948 is classified as composite according to the [N II]-BPT diagram. J0808+3948 has also an extremely high nitrogen enhancement and high N/O (Arellano-Córdova

et al. 2023 in prep.; see also Stephenson et al. 2023). However, these galaxies are not detected in UV lines, so are not shown in the UV diagnostic diagrams shown in Sec. 5[13].

Even though the [O III]/[O II] versus [O I]/[O III] diagram can be a very good tool to understand the role of shocks and separate SF from harder mechanisms, it can also have drawbacks. [O I] $\lambda6300$ is produced in the warm transition region between the fully ionized gas and neutral gas (Osterbrock & Ferland 2006; Draine 2011). As such, it traces the external parts of H II regions, close to the ionization front, where shocks and non-equilibrium heating are important (Dopita & Sutherland 2003; Dopita et al. 2013), which makes it difficult to predict [O I] $\lambda6300$ in photoionization models (Dopita et al. 2013). Furthermore, [O I] $\lambda6300$ is generally a faint emission line (this makes it even more tricky to disentangle a broad second component). These implications affect the reliability of diagnostic plots that take [O I] $\lambda6300$ into account. Nevertheless, an [O I] $\lambda6300$ enhancement could still represent a sign of ionizing mechanisms other than SF.

4.3. *The Shirazi-Brinchmann diagram*

To further confirm the SF classification of CLASSY galaxies we also consider high-ionization lines such as He II $\lambda4686$. Indeed, as introduced in Sec. 1, Shirazi & Brinchmann (2012) proposed to use [N II]/Hα versus He II $\lambda4686$/Hβ to better constrain the ionization source. The two major advantages of this diagram (HeII diagram) are that (i) it is more sensitive to the hardness of the ionizing source than [O III]/Hβ, thus allowing a cleaner separation between star-forming and composite galaxies; and (ii) it is not particularly sensitive to the metallicity (Shirazi & Brinchmann 2012). The HeII diagram for the CLASSY sample is shown in Fig. 4 and is color-coded as a function of 12+log(O/H) (the broad component fit of He II $\lambda4686$ line is not detected because of the low S/N). Similarly to the standard BPT diagrams described in Sec. 4.1, SFGs are expected to lie below the maximum starburst (black solid) line, while AGN are located above. The G16 models are not shown for clarity reasons but can completely cover the star formation locus. Almost all the CLASSY galaxies shown in the HeII diagram are classified as star-forming. The only targets with a different classification are J0808+3948 (classified as AGN) and J1112+5503 (at the edge between the SF and composite locii). However, we emphasize that their He II lines are particularly faint (S/N$\sim 3-5$), hence this classification is not reliable by itself. Given that J0808+3948 is not located in the SF locus also in Fig. 3 and in the [N II]-BPT (Fig. 2), we conclude that it is possible that this galaxy has a fur-
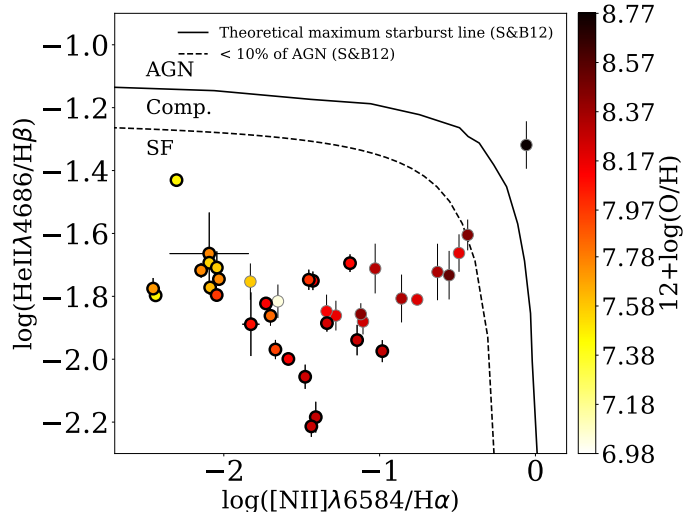
Figure 4. [N II]/Hα versus He II $\lambda4686$/Hβ diagnostic diagram for the CLASSY galaxies, shown as dots, color-coded as a function of the gas-phase metallicity. Only the measurements with $S/N > 3$ for all the lines involved are reported. The dotted and solid black lines represent the locus where $\sim 10\%$ of He II $\lambda4686$ comes from an AGN and the theoretical maximum starburst line, respectively, as defined by Shirazi & Brinchmann (2012). Only two CLASSY galaxies are outside the star-forming locus, but we note that their He II $\lambda4686$ lines are particularly faint ($S/N \lesssim 5$), hence this classification is not reliable by itself. The galaxies with at least one UV emission line (i.e., C III]$\lambda\lambda1907,9$; 12+log(O/H) $\lesssim 8.3$) are highlighted by a black thick edge.

ther ionizing mechanism than pure SF. In Sec. 4.1 and 4.2 we could not exclude the presence of shocks in our targets, but, if present, shocks should easily enhance the He II $\lambda4686$ flux (e.g., Shirazi & Brinchmann 2012; Stasińska et al. 2015). However, here we stress that we do not see a particular enhancement of He II at increasing [O I]/[O III], which we consider a good shock diagnostic (Fig. 3) or in galaxies where we identify a second broader component characterized by higher [N II]/Hα, [S II]/Hα, and/or [O I]/Hα, as shown and discussed in Fig. 2. This is a further confirmation of the SF classification of our systems.

Finally, we highlight that almost all the C III] $\lambda\lambda1907,9$-emitters are shown in Fig. 4. Those missing are: Izw 18 and J1253-0312, because without Hα measurement in our sample, but with log(He II $\lambda4686$/Hβ) $\sim -1.35, -1.86$ and 12+log(O/H) $\sim 6.98, 8.06$, respectively, and thus still lying in the SF locus; Mrk 996 (J0127-0619) with an unreliable He II $\lambda4686$ flux measurement, due to the contamination from the WR red clump, but still classified as star-forming (see James et al. 2009 for further details); J1444+4237, which does not show He II $\lambda4686$ emission (upper limit of log(He II $\lambda4686$/Hβ) $\lesssim -1.8$), even though it has bright UV emission lines. Overall, all the galaxies that we will show in

Sec. 5 are classified as SF-dominated also according to the Shirazi & Brinchmann (2012) criterion.

4.4. *The importance of optical coronal lines*

As a concluding note on optical diagnostics, we discuss the so-called coronal lines as proof of the presence of AGN activity. These are high-ionization (≥ 70 eV) forbidden transitions excited by collisions similarly to [O III] $\lambda5007$ and are usually considered undeniable evidence of AGN ionization (e.g., Korista & Ferland 1989). Their presence in low-massive metal-poor galaxies optical (and IR) spectra is a current hot topic in the hunt to reveal intermediate-mass black holes (e.g., Cann et al. 2020; 2021; Molina et al. 2021).

After a careful analysis of our spectra, we exclude the presence of these lines (i.e., [Fe VI] $\lambda5146$, [Fe XIV] $\lambda5303$, [Fe VII] $\lambda5721$, [Fe VII] $\lambda6087$, [Fe X] $\lambda6374$) in the optical spectra we analyzed[14]. In the upper panel of Fig. 1, next to [O I] $\lambda6364$ there is a "bump" that could be interpreted as a tentative [Fe X] $\lambda6374$. However, this feature is more consistent with the Si II $\lambda6371$ (actually a multiplet, Si II $\lambda\lambda6348,71$), previously identified in some CLASSY galaxies in high-resolution optical spectra (e.g., VLT/GIRAFFE SBS 0335-052 E spectrum, Izotov et al. 2006). These permitted Si II lines may have a fluorescent origin and be produced by absorption of the intense UV radiation (Grandi 1976; Izotov et al. 2001). The detection of very high-ionization emission lines could be an alternative channel to identify hidden AGN activity in dwarf galaxies, especially in the JWST era with rest-frame IR coverage (e.g., Cann et al. 2018). However, here we want to highlight that it can be easy to misidentify [Fe X] $\lambda6374$ and Si II $\lambda6371$.

5. UV DIAGNOSTIC DIAGRAMS

In the previous section, we concluded that all the CLASSY galaxies characterized by UV emission lines are dominated by star formation, using different optical criteria, highlighting all their caveats. Further confirmation of their SF-dominated nature lies in the lack of N V $\lambda\lambda1239,1243$ doublet in emission. Indeed in the UV, an unequivocal signature of AGN activity is generally represented by this very high-ionization emission doublet ($E > 77.5$ eV; Feltre et al. 2016). We do not observe the ISM N V doublet in emission in any of the CLASSY galaxies, but only in P-Cygni shape - strongly indicating the presence of massive stars (e.g., Chisholm et al. 2019). Having said that, here we want to explore if UV diagnostic diagram classification can be consistent with the optical and thus identify the best-observed UV diagnostics to discriminate SF from AGN and shock ion-

ization. Of course, without AGN within the sample, we are unable to assess the ability of UV diagnostics to correctly classify AGN, but we are able to test whether SF galaxies are incorrectly classified as AGN. We do this by comparing our sample of CLASSY galaxies with G16 constant (CST) and single-burst (SSP) SF models, AM19 shock models and F16 AGN models presented in Sec. 3. As we introduced in Sec. 1, UV diagnostics can be crucial to investigate ISM properties with JWST observations in objects above $z \sim 6$, and thus it is important to understand in detail their strengths and drawbacks.

Specifically, in the following subsections, we show the UV diagnostic plots that we found more promising among those proposed in the literature on the basis of models and simulations (Jaskot & Ravindranath 2016; Feltre et al. 2016; Nakajima et al. 2018; Hirschmann et al. 2019; 2022):

- C3He2-C4He2 (Fig. 5): C III] $\lambda\lambda1907,9$/He II $\lambda1640$ vs C IV $\lambda\lambda1549,51$/He II $\lambda1640$;

- C3He2-O3He2 (Fig. 6): C III] $\lambda\lambda1907,9$/He II $\lambda1640$ versus O III] $\lambda1666$/He II $\lambda1640$;

- C4C3He2-C4C3 (Fig. 7): C IV $\lambda\lambda1548,51$/C III] $\lambda\lambda1907,9$ versus (C IV+C III])/He II $\lambda1640$;

- EWC4 and EWC3 (Fig. 8): EW(C IV) and EW(C III]) versus C III]/He II $\lambda1640$.

We also report in App. C the diagnostics diagrams C III]/He II $\lambda1640$ versus O III] $\lambda1666$/C III] $\lambda\lambda1907,9$ (C3He2-O3C3) and C III]/He II $\lambda1640$ versus C IV $\lambda\lambda1548,51$/C III] $\lambda\lambda1907,9$ (C3He2-C4C3).

As a general note about the AM19 shock grids, in all the following plots they are often completely overlapped in terms of velocity and magnetic field and gas-phase metallicity, but we decided to show the entire range of the v and B parameter space at three representative values of metallicity ($Z = 0.05, 0.5, 1\ Z_\odot$) for completeness. These grids also cover very narrow regions in C4He2 space that correspond to the lowest velocities in the parameter space ($v < 200$ km/s, seen as spike-shaped features in grid coverage).

5.1. *C3He2-C4He2*

Fig. 5 shows the C III] $\lambda\lambda1907,9$/He II $\lambda1640$ versus the C IV $\lambda\lambda1548,51$/He II $\lambda1640$ diagnostic diagram (C3He2-C4He2) with previously described SF, shock and AGN grid models overlaid (see caption for details). This diagnostic diagram was first proposed as a good AGN and star-forming galaxies separator in F16. Jaskot & Ravindranath (2016) further supported this as a good diagnostic plot to discriminate between shocks and pure SF, where their pure photoionization models lie above the shaded gray region in Fig. 5, while most of their models with a shock

---

[14] J0944-0038 MUSE data recently presented in del Valle-Espinosa et al. (2022) reveal [Fe VI] $\lambda5146$ line in the brightest star-formation knot, only partly covered by the SDSS and LBT data that we took into account.
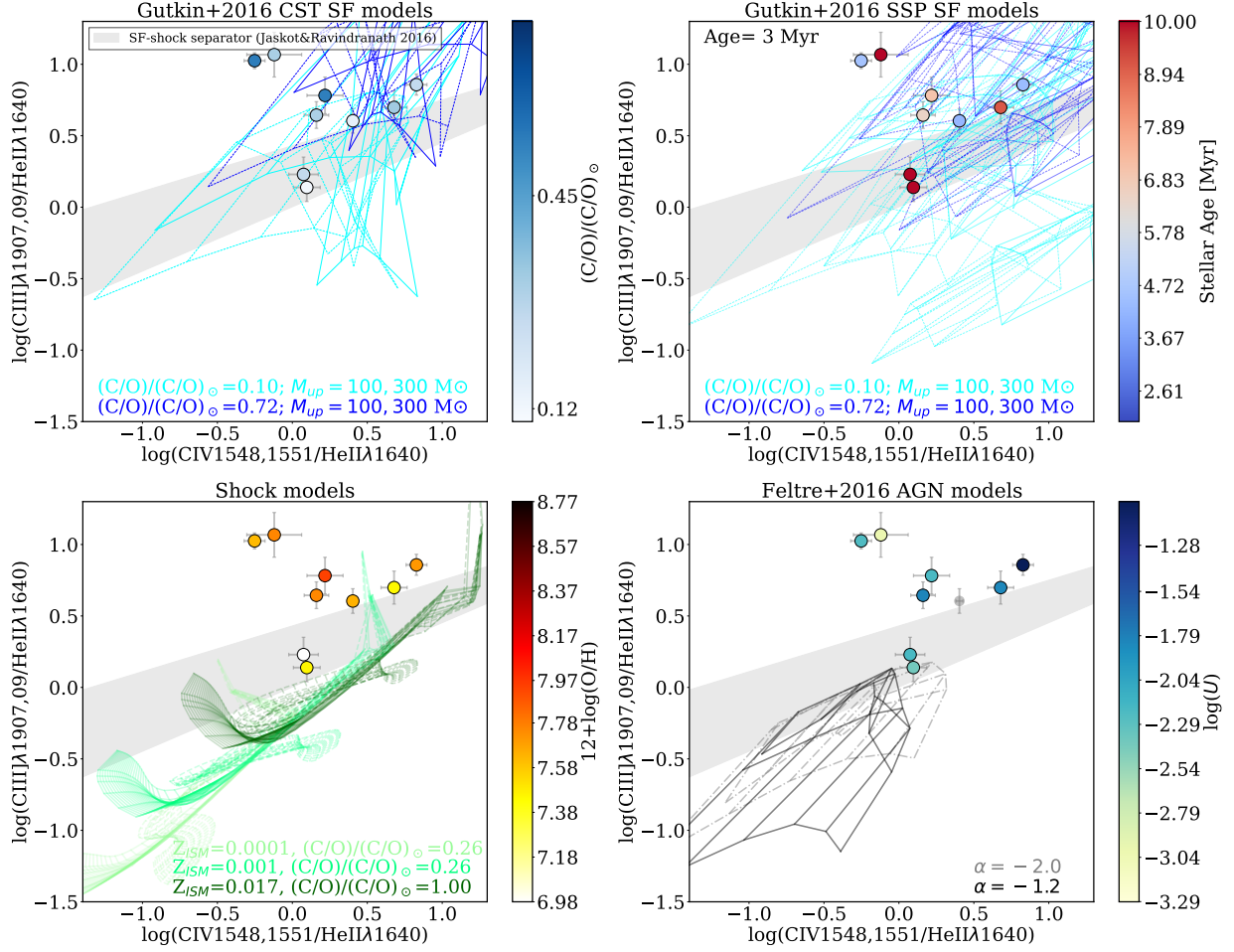
Figure 5. C3He2-C4He2 diagram: C III] $\lambda\lambda1907,9$/He II $\lambda1640$ vs C IV $\lambda\lambda1549,51$/He II $\lambda1640$ diagnostic plot (Feltre et al. 2016, Jaskot & Ravindranath 2016) to separate SF and AGN activities. The gray shaded region represents the separator proposed by Jaskot & Ravindranath (2016) between pure-SF models (located above) and models with $\geq 50\%$ shock contribution (located below). **a:** Gutkin et al. (2016) models with continuous SF are shown in light ((C/O)/(C/O)$_\odot$=0.1) and dark blue ((C/O)/(C/O)$_\odot$=0.72), with the scatter plot color-coded as a function of the (C/O)/(C/O)$_\odot$ measured for the sample. The solid and dashed grids refer to different IMF cut ($M_{up} = 100$ M$_\odot$ and $M_{up} = 300$ M$_\odot$, respectively). **b:** Gutkin et al. (2016) models with burst of SF at 3 Myr are reported as in panel a, with the scatter plot color-coded as a function of the stellar age derived from the UV stellar component fitting. **c:** Alarie & Morisset (2019) shock models with increasing metallicities from light to dark green, as shown in the legend, and (C/O)/(C/O)$_\odot$ = 0.26 are shown on a scatter plot color-coded as a function of the gas-phase metallicity 12+log(O/H) (darker green means higher metallicity, up to the trans-solar value of $Z_{ISM} = 0.017$). **d:** Feltre et al. (2016) AGN models with $\alpha = -2$ (solid - softer ionization) and $\alpha = -1.2$ (dash-dotted - harder ionization) overplotted on a scatter plot color-coded as a function of the ionization parameter log(U). The two galaxies located in the gray-shaded region are I Zw 18 and SBS 0335-052 E, which have the lowest 12+log(O/H) and C/O abundance in the sample. Overall, the C3He2-C4He2 diagram can be used to discriminate SF from AGN and shocks, with some caveats. Shocks and AGN are harder to discriminate, but still the former are expected to have higher carbon-to-helium line-ratios, especially at increasing metallicity.

contribution of $\geq 10\%$ and $\geq 50\%$ to the shown line ratios fall within and below the gray shaded region, respectively. Indeed, the He II $\lambda1640$ line can be enhanced by hard ionizing radiation with high-energy photons (> 54.4 eV), which can be produced by either AGN or shocks, causing a decrement of C III] $\lambda\lambda1907,9$/He II $\lambda1640$ and C IV $\lambda\lambda1548,51$/He II $\lambda1640$. This also implies that the C3He2-C4He2 diagram cannot clearly distinguish between AGN and shocks. Indeed, looking at the bottom panels of

Fig. 5, shock and AGN grids cover a similar C III]/He II vs C IV/He II parameter space, with the most metal-poor shock grids reaching C III]/He II and C IV/He II down to ~ -1.75 and ~ -2.5, respectively, as the $\alpha = -1.2$ AGN models (hardest AGN radiation) at the lowest values of metallicity and ionization parameter. Indeed, the strength of C III] and C IV is known to drop at decreasing metallicity (12+log(O/H) $\lesssim 7.5$; e.g., Jaskot & Ravindranath 2016; see also Paper IV) and decreasing C/O. We also notice that the AM19 shock models

shown in Fig. 5 can reach higher values than F16 AGN models in terms of C III]/He II line ratios, going beyond the gray shaded SF-shock separator region at low values of velocities ($v < 200$ km/s, spike-shaped feature in the shown grids; see also Jaskot & Ravindranath 2016 Sec. 4.4).

The scatter plots in Fig. 5, color-coded as reported in the color bar labels, show the 9 galaxies of our sample with C IV $\lambda\lambda1548,51$ in pure emission (the other galaxies show pure absorption or P-Cygni profiles). In particular, SBS 0335-052 E (J0337-0502) and I Zw 18 (J0934+5514) are located in the gray shaded region that separates the pure-SF and shock locii according to Jaskot & Ravindranath (2016). Moreover, the two objects partially overlap with the F16 AGN models. However, from the optical diagnostics, we know that these galaxies are SF-dominated (Sec. 4). Interestingly, these objects are the two lowest metallicity galaxies of the sample showing UV emission lines ($12+\log(O/H) \sim 7.46$ and ~ 6.98, respectively) with also the lowest C/O abundances ($(C/O)/(C/O)_{\odot} \sim 0.26$ and ~ 0.12, respectively). This means that the Jaskot & Ravindranath (2016) criterion can fail in classifying extremely metal-poor star-forming objects possibly characterized by very hard ionization and low C/O, which could mimic AGN or shock ionization (see also Maseda et al. 2017). In agreement with this, Gutkin et al. (2016) SF models (both CST and SSP) can fall beyond the Jaskot & Ravindranath (2016) pure-SF line when considering low C/O (cyan grids have $(C/O)/(C/O)_{\odot}=0.1$), which would reduce C III] and C IV emission. Another factor that pushes the Gutkin et al. (2016) SF models below the Jaskot & Ravindranath (2016) pure-SF locus is a higher IMF upper-cut ($M_{up} = 300$ M$_{\odot}$ instead of $M_{up} = 100$ M$_{\odot}$, dashed and solid lines, respectively), which would imply more massive stars, hence a harder ionization field (i.e., higher He II emission). Concerning Gutkin et al. (2016) single-burst SF grids, we only show the 3 Myr models for clarity reasons, but ages younger than ~ 4 Myr can reproduce the line ratios observed for the CLASSY galaxies. Since the Gutkin et al. (2016) models do not take into account either stellar rotation or binary interactions (which would amplify the galaxies ionizing fluxes for several Myr after a SF burst), the ionizing photons drop at older ages than 4 Myr, with a subsequent reduction of He II (but also C III] and C IV), inducing a shift of the grids to higher C III]/He II and C IV/He II than the range shown in Fig. 5 (see e.g., Jaskot & Ravindranath 2016 Sec. 3.1 for a detailed comparison between different model prescriptions).

Overall, to explain the locations of galaxies such as SBS 0335-052 E and I Zw 18, lying in the region separating SF from harder mechanisms, the most likely explanation is their low O/H and C/O abundances (see also Wofford et al. 2021 for SBS 0335-052 E). Hence, before using the C3He2-C4He2 diagram the user should be mindful of the metallicity and C/O abundance of the targets, in that ionization sources in low-metallicity objects are less easily distinguishable for this diagnostic.

5.2. C3He2-O3He2

In Fig. 6 we show the C III] $\lambda\lambda1907,9$/He II $\lambda1640$ versus O III] $\lambda1666$/He II $\lambda1640$ diagnostic diagram, proposed by F16 and further explored by Hirschmann et al. (2022). In contrast to the C3He2-C4He2 diagram shown in Fig. 5, C3He2-O3He2 in Fig. 6 can separate all of the three contributions of SF, AGN, and shocks quite well. This is with the exception of the G16 SF models with either low C/O or high M_{up}, which, again, cover the entirety of the parameter space (see Sec. 5.1), as we further comment below. Most importantly, AGN and shock grids are only slightly overlapped, considering sub-solar metallicities (considering $(C/O)/(C/O)_{\odot} = 0.26$), with shock models showing higher O III]/He II line ratios than AGN grids. Instead, the dark-green grid at trans-solar metallicity ($Z_{ISM} = 0.017$; $(C/O)/(C/O)_{\odot} = 1$) shifts to higher C III]/He II, overlapping more with the AGN and SF grids. In particular, He II emission is expected to increase in both AGN and shock models, as already discussed in Sec. 5.1, explaining the shift of AGN and shocks to lower C III]/He II and O III]/He II than SF grids (but $v < 200$ km/s shock models - spike-shaped in the shown grids - can reach C III]/He II comparable to SF grids as in Fig. 5). However, shocks are expected also to have an increase of C III], which generally tightly correlates with temperature-sensitive lines such as O III] $\lambda1666$ (see e.g., Jaskot & Ravindranath 2016 Fig. 12 and 9). This could explain the higher [O III]/He II for shocks than AGN models. As a reference, in Fig. 6 we added a dashed green line to separate shock models from the SF locus

$$y = 0.8x + 0.2 \qquad (3)$$

and a horizontal dash-dotted line to separate AGN models from the SF locus

$$y = 0.1 \quad and \quad -1.8 < x < -0.15 \qquad (4)$$

with $x = \log(O$ III] $\lambda1666$/He II $\lambda1640)$ and $y = \log(C$ III] $\lambda\lambda1907,9$/He II $\lambda1640)$. The AGN grids again extend beyond the shown x- and y-ranges, with $\log(O$ III]/He II) and $\log(C$ III]/He II) down to ~ -1.5 and ~ -2.5, respectively, at the highest values of metallicity and lowest of ionization parameter.

The scatter plot in Fig. 6 shows the 19 CLASSY galaxies with all the involved UV emission lines with $S/N > 3$. Interestingly, the parameter space covered by the CLASSY galaxies in this diagnostic diagram is very similar to that of the $z \sim$2–4 galaxies of Nanayakkara et al. (2019), which supports the classification of these systems as analogues to high-z SF-galaxies. As in Fig. 5, G16 SF models cover the majority of the C3He2-O3He2 plane. We notice that grids with
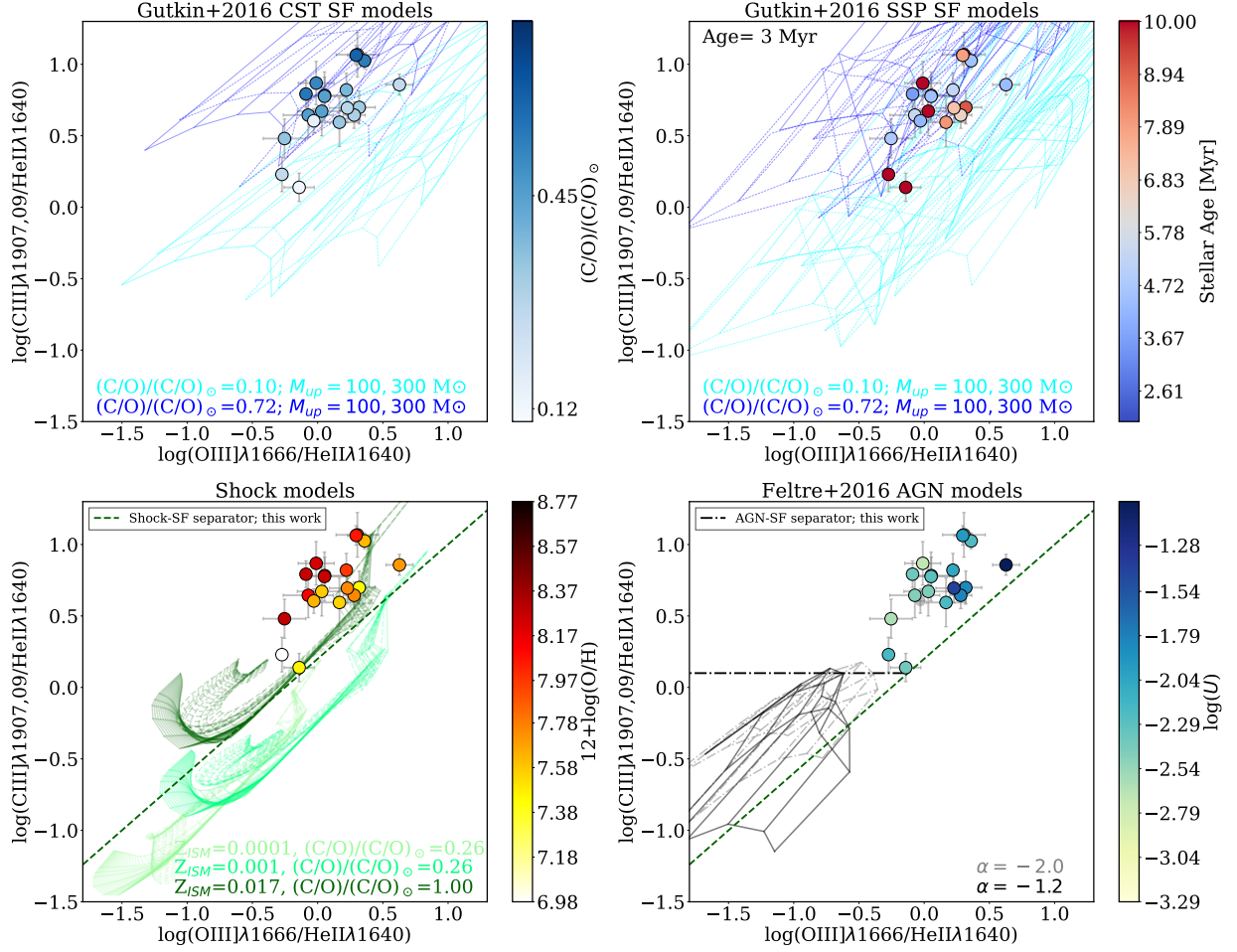
Figure 6. C3He2-O3He2 diagram: C III] λλ1907,9/He II λ1640 versus O III] λ1666/He II λ1640 diagnostic diagram, proposed by F16. The models superimposed are as explained in Fig. 5. Overall, this represents the best diagram to discriminate SF, AGN and shocks at sub-solar metallicities (i.e., excluding the dark-green shock grid), due to the minimal overlap between each set of grids and the fact that all CLASSY galaxies are correctly classified as star-forming, as shown by the dashed green and dash-dotted lines to separate shock and AGN models from the SF locus (Eq. 3 and Eq. 4), respectively.

low C/O abundances (solid and dashed cyan G16 grids in the upper left panel) are shifted to lower C III]/He II and higher O III]/He II (trend in agreement with the C/O color-coding shown in Fig. 6 upper left-hand panel). This is probably due to the fact that carbon is a key coolant in the nebula, and thus a lower C/O ratio raises the electron temperature – as commented in Jaskot & Ravindranath (2016), and demonstrated by the decrease in metallicity (which is color-coded in the bottom left-hand panel) as O3He2 increases. Hence, the increased collisional excitation rate partially compensates for the reduced C abundance, preventing an extreme drop of the C III]/He II ratio, and possibly enhancing [O III] λ1666. As a result, the low C/O G16 grids (solid and dashed cyan, according to the different IMF cut) manage to cover the lowest C III]/He II and O III]/He IISBS 0335-052 E and I Zw 18.

Also Hirschmann et al. (2019) proposed line separators to classify the main ionization mechanisms for the C3He2-

O3He2 diagnostic plot, obtained from synthetic emission line ratios computed by coupling G16 and F16 models with high-resolution cosmological zoom-in simulations of massive galaxies. However, according to their AGN-Composite and Composite-SF separators (that we do not report in Fig. 6 for clarity reasons), all our star-forming targets would be classified as composite. Hirschmann et al. (2019) also compared their simulated galaxies with observations of AGN and star-forming galaxies, finding generally good agreement with their classification, although the local and high-z low-metallicity galaxies they took into account were located mainly in the composite region as we find for the CLASSY galaxies. Hence, we conclude that alongside low-metallicity galaxies, the Hirschmann et al. (2019) composite region also hosts star-forming galaxies with more "extreme" high-z characteristics that are typical of our sample.

Overall, we conclude that the C3He2-O3He2 represents a better diagnostic diagram than the C3He2-C4He2, since it is able to clearly separate (up to solar metallicities) AGN and shocks, which are located in different locii with respect to star-forming dominated emission. Moreover, all the CLASSY galaxies, even the most metal-poor SBS 0335-052 E and I Zw 18, are covered by the SF grids and do not overlap with the AGN or shock grids.

5.3. *C4C3He2-C4C3, EWC4, and EWC3*

Fig. 7 shows the (C IV+C III])/He II vs C IV/C III] diagnostic plot proposed by Nakajima et al. (2018), with their proposed AGN locus shown in shaded gray. As in Fig. 5, SBS 0335-052 E and I Zw 18 end up in the AGN locus, overlapping slightly with the F16 AGN models. Analogously to what we commented in Sec. 5.1, their shift can be explained in terms of their low 12+log(O/H) and low C/O abundance. In this figure, it is even clearer that only the G16 SF models (both CST and SSP) with low C/O (cyan grids) can cover this sub-set of CLASSY galaxies, excluding two outliers. A higher C/O abundance grid with a high-cut in the IMF (dashed blue grids) lies close but does not cover the lowest (C IV+C III])/He II line ratios, while the higher C/O abundance grid with $M_{up} = 100$ $M_{\odot}$ (solid blue grids) does not cover the CLASSY galaxies. Overall, an increase of C/O quickly pushes the G16 grids toward values of (C IV+C III])/He II than what we typically observe in our sample. Moreover, AGN and shock grids are almost entirely overlapped (apart from the most metal-poor shock models, characterized by the lowest (C IV+C III])/He II).

Nakajima et al. (2018) proposed other two additional diagnostic plots to separate SF and AGN activities, shown in Fig. 8, with the CLASSY data color-coded as a function of the ionization parameter log(U). In particular, we notice that EW(C III]) and EW(C IV) increase at higher log(U), and also the C IV and C III] emissions are higher with respect to He II. This is clearer in the right panel since EW(C IV) provides a valid log(U) tracer as demonstrated in Paper IV. As in Fig. 7, the AGN regions according to Nakajima et al. (2018) criteria are shown in shaded gray. Nakajima et al. (2018) computed AGN photoionization models, assuming a continuum underneath the line emission dominated by the accretion disk, commenting that this is an overestimate since it does not take into account the effects of dust attenuation (i.e., presence of the torus). In general, evaluating the continuum emission to compute EWs in shock and AGN models implies many arbitrary assumptions, since the continuum can have a composite nature, coming from both the stellar component and the additional ionizing sources as well as their intrinsic characteristics. For this reason F16 and AM19 do not provide EWs for their models. On the other hand, all the displayed CLASSY galaxies are in the SF-locus defined by Nakajima

et al. (2018), with the exception of J0944-0038 in the left panel, characterized by EW(C III]) ~ 33 Å, which still lies in a region where both AGN and SF models could overlap. Also, the EW values predicted by G16 single-burst SF models with ages younger than 4 Myr (that we do not show for clarity reasons) cover the Nakajima et al. (2018) SF locus, and thus the EW(C III]) and EW(C IV) values observed for the CLASSY galaxies, while the constant SF models hardly reproduce them. As commented also for Fig. 6, CLASSY galaxies fall entirely in the region defined as composite in Hirschmann et al. (2019), confirming that star-forming objects with the characteristics of the CLASSY sample (and thus high-z analogs) could risk being misclassified according to their criteria. Overall, these diagrams could be useful tools to classify even metal-poor star-forming galaxies, but since it is not trivial to evaluate the continuum to compute EWs in shock and AGN models, they cannot be used to separate AGN from shocks.

5.4. *Summary*

All the discussed UV diagnostic diagrams in the previous subsections, with some caveats, can distinguish SF from AGN and shocks, while it is possible to disentangle AGN from shocks only by taking into account C III], He II and O III] emission lines. Here we summarize our main findings:

- The C3He2-C4He2 diagram (Fig. 5) from Jaskot & Ravindranath (2016) can be used to discriminate SF from AGN and shocks. However, the users should be mindful of the metallicity and C/O abundance of the targets, in that ionization sources in low-metallicity objects are less easily distinguishable for this diagnostic. Shocks and AGN are harder to discriminate, but still the former are expected to have higher carbon-to-helium line-ratios than the latter, especially at increasing metallicity.

- C3He2-O3He2, proposed by Feltre et al. (2016), is the only diagram able to separate all the three ionization mechanisms at sub-solar metallicities (Fig. 6 and Eq. 3, 4). In this plot, all the CLASSY galaxies, even the objects characterized by the lowest 12+log(O/H) and C/O (i.e., SBS 0335-052 E and I Zw 18), are covered by the G16 SF models and do not overlap to the AGN or shock grids, that cover similar C III]/He II but different O III]/He II line ratios (lower for AGN, higher for shocks). Other diagnostic diagrams from Feltre et al. (2016) are C3He2-C3O3 and C3He2-C4C3 (App. B), with the former able to separate SF, AGN and shocks as C3He2-O3He2 (Eq. C1, C2), while the latter can only separate SF from AGN and shocks (Eq. C3, C4).

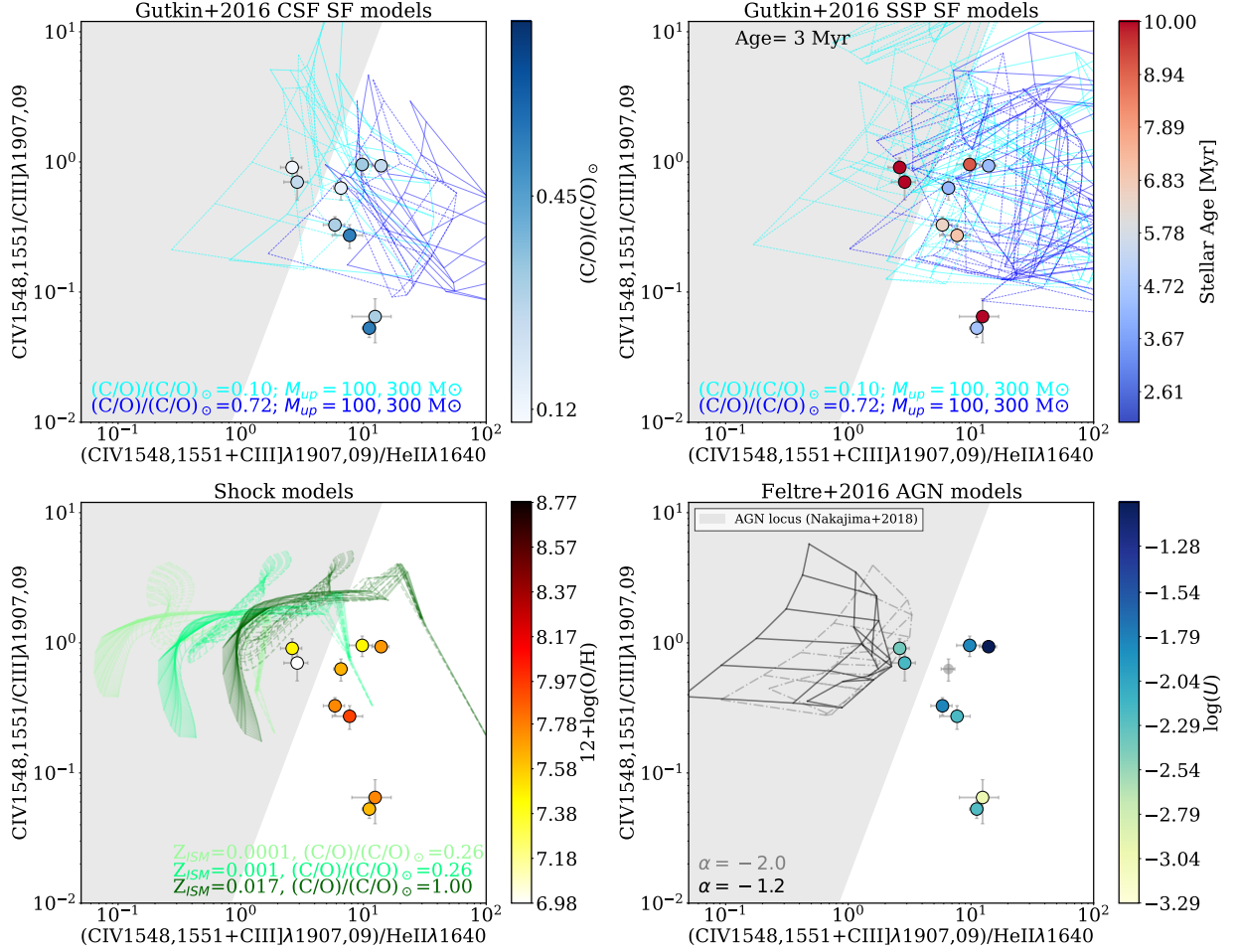- The C4C3He2-C4C3 diagram (Fig. 7) from Nakajima et al. (2018) is a good diagnostic plot to sep-

Figure 7. C4C3He2-C4C3 diagram: (C IV λλ1548,51+C III] λλ1907,9)/He II λ1640 vs C IV/C III] diagnostic plot proposed by Nakajima et al. (2018) to separate SF and AGN activities, with their proposed AGN-dominated region highlighted in shaded gray. The models superimposed are as explained in Fig. 5. Similarly to Fig. 5, this plot can be used to separate SF from AGN with some caveat. Indeed, the two galaxies which end up in the AGN locus are again I Zw 18 and SBS 0335-052 E, the two lowest 12+log(O/H) and C/O abundance galaxies of the CLASSY sample, and can be reproduced by the low C/O SF grids (in cyan).

arate SF and AGN, with the caveat of taking into account 12+log(O/H) and C/O, similarly to C3He2-C4He2. However, it is not possible to distinguish between AGN and shocks, for which O3He2-C3He2 and C3He2-C3O3 would be better suited.

- Finally, EWC4-C3He2 and EWC3-C3He2 (Fig. 8) from Nakajima et al. (2018) provide as well good diagnostic plots to separate star-forming galaxies from objects characterized by non-photoionization mechanisms via the Nakajima et al. (2018) AGN-SF separator line. This diagnostic diagram holds true even at low metallicity, low C/O abundance and high log(U), but does not allow the user to discriminate shocks from AGN models, given the difficulties in estimating the corresponding EWs.

6. DISCUSSION

In the previous sections, we demonstrated that our galaxies are SF-dominated according to optical diagnostics, and then explored the most promising UV diagnostic diagrams, highlighting strengths and weaknesses. In the following, we want to discuss the observability of UV emission lines shown in this work. We do this for two reasons: (1) to provide guidance on whether or not these lines should be expected in proposed observations of certain targets and (2) to provide context on the properties of the target if the lines *are* detected in their observations. Indeed, UV emission lines are a powerful (and at increasing redshift also the only) tool to investigate ISM conditions in galaxies.

In general, we confirmed that the main two conditions to have UV emission line detections in CLASSY (i.e., C IVλλ1548,51, He IIλ1640, O III]λλ1661,6, C III]λλ1907,9) are low metallicity and high ionization parameter, traced by the [O III]/[O II] line ratio (but also
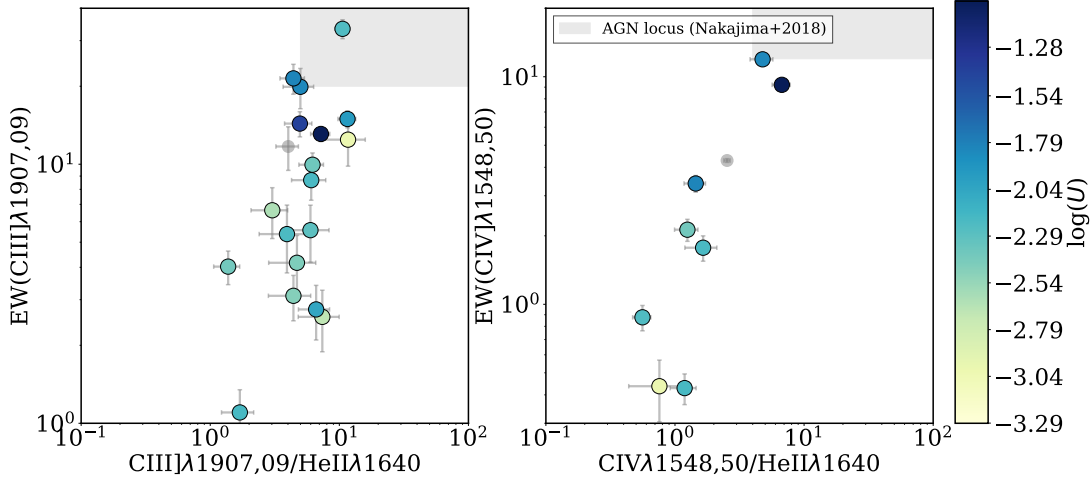
Figure 8. EW(C III]) and EW(C IV) diagnostic plots, where the gray shaded regions represent the Nakajima et al. (2018) AGN-locus, separated from the star formation locus. Interestingly, this classification scheme works well for the CLASSY galaxies, even for the most metal-poor systems.

[O III]/Hβ). Specifically, we see no C III] $\lambda\lambda$1907,9 (and thus no other UV emission lines) in CLASSY galaxies with 12+log(O/H) $\gtrsim$ 8.3 (see also Rigby et al. 2015; Jaskot & Ravindranath 2016; Ravindranath et al. 2020). This is also the reason why C III]-emitting galaxies (highlighted by thick black edges in all the figures) are systematically shifted to lower log([N II]/Hα) ($\lesssim$ −1) in Fig. 2 and Fig. 4), with the lowest metallicity targets (12+log(O/H) < 7.8), having also higher log(He IIλ4686/Hβ) line ratios ($\sim$ −1.4 to −1.8). However, the trend between the EW of UV lines and metallicity is not linear because carbon lines in our galaxies can be suppressed at 12+log(O/H) < 7.5 (see e.g., Fig. 15 in Paper IV) and at decreasing C/O (Berg et al. in preparation). On the other hand, the EW of UV lines slightly correlates also with log(U), with a significant decrease of detected UV lines at log(U) $\lesssim$ −2.5 for our sample. This is also visible from Fig. 2 and Fig. 3, in which C III]-emitters are clearly shifted to higher log([O III]/Hβ) ($\sim$ 0.5 − 1) and higher log([O III]/[O II]) ($\sim$ 0.5 − 1.5).

The physical reason why UV metal lines are visible only at low metallicity is well-explained in Jaskot & Ravindranath (2016) (see their Sec. 3.2). In particular, a decrease in metallicity implies a harder stellar ionizing radiation and a higher temperature nebula, and thus a larger amount of ionizing photons and a larger collisional excitation rate, respectively. Indeed, the UV metal lines are collisionally excited with energy levels above the ground state of E/k $\sim$ 6.5 − 8.3 eV $\sim$ 75,000 − 97,000 K, much higher than optical collisional lines that typically have E/k << 70,000 K (e.g., $\sim$ 7.5 eV $\sim$ 87,000 K for O III]$\lambda\lambda$1661,6 versus $\sim$ 2.5 eV $\sim$ 30,000 K for [O III]λ5007 and $\sim$ 5.4 eV $\sim$ 62,000 K for [O III]λ4363; Osterbrock & Ferland 2006; Draine 2011). Specifically, C III]$\lambda\lambda$1907,9 (the brightest and most common) has the low-

est E/k $\sim$ 75,000, while the less common C IV$\lambda\lambda$1548,51 and N IV]$\lambda\lambda$ have the highest E/k $\sim$ 93,000 − 97,000 K. Hence, the UV lines are extremely temperature sensitive in photoionized gas that can reach T $\sim$ 2 × 10^4 K in our systems (see also Paper IV), compensating for the lower ionic relative abundances. As discussed in previous works (e.g., Jaskot & Ravindranath 2016), also dust attenuation, density, SFR and stellar population characteristics (stellar age and metallicity) can play a role in the strength or absence of these UV lines, but we find no clear correlations comparing the derived quantities for the CLASSY galaxies. Furthermore, UV lines such as He II λ1640 and C IV $\lambda\lambda$1548,51 can have both a stellar and nebular origin, with C IV being also a resonant line affected by radiation transfer effects (e.g., Steidel et al. 2016). For CLASSY, we can confirm He II and C IV nebular origin thanks to the general consistency of their velocity dispersion values (corrected by instrumental broadening) with those of O III] $\lambda\lambda$1661,6 ($\sigma \sim$ 50 km/s; Fig. 1 bottom panel).

Concerning C IV, typical galaxies without an AGN show either C IV in absorption from the surrounding ISM or circumgalactic medium (CGM) gas or a P-Cygni profile from the stellar winds of massive stars (e.g., Leitherer et al. 2001). Nebular C IV emission is rarely observed in the literature, and comes uniquely from studies of systems with 12+log(O/H) $\lesssim$ 8 and generally high ionization parameter log(U) $\gtrsim$ −2.5 (see Fig. 16 in Paper IV), probably tracing a rapid hardening of the ionizing spectrum at low metallicities (see also Senchyna et al. 2019; Berg et al. 2021; Senchyna et al. 2023a; Schaerer et al. 2022). From a visual inspection of C IV feature, we classified the CLASSY galaxies as being either C IV emitters (9 objects), P-Cygni (10), without C IV coverage (3), and defining the remaining others as ISM absorbers (23). Specifically, we applied the classification after the subtraction of the
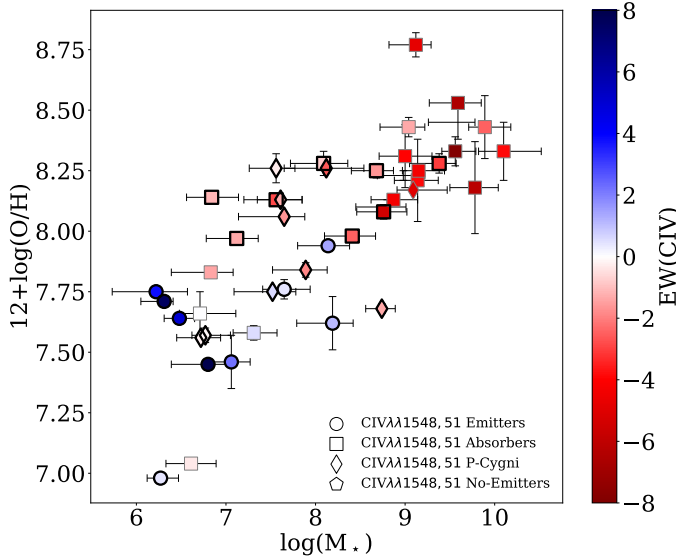
Figure 9. Mass-metallicity relation for CLASSY galaxies with C IV in emission (dots), absorption (squares) and P-Cygni (diamonds), color-coded by EW(C IV), with red and blue values indicating negative (lines in absorption) and positive (lines in emission) EWs, respectively. The galaxies with at least one UV emission line (i.e., C III] are highlighted by a black thick edge. These plots show that galaxies which are C IV emitters, or with narrow P-Cygni non-stellar profiles, tend to have lower metallicities (12+log(O/H)$\lesssim$ 8, 8.2) and lower stellar masses (log($M\star/M_\odot$) $\lesssim$ 8, 9).

stellar component, thus we are considering only the residual ISM contribution of C IV doublet. We then calculated the EW(C IV) by integrating the C IV emission in a ~ 10 Å window centered on the C IV lines, and the continuum in two featureless ~ 5 Å windows on each side. We show the results of this exercise in Fig. 9, where we color-code the CLASSY mass-metallicity relation as a function of EW(C IV). Also, galaxies with different C IV profiles are indicated using different symbols, as listed in the legend. Hence, within the CLASSY sample, C IV emitters (but also P-Cygni) are found to be more metal-poor and low-mass (12+log(O/H)$\lesssim$ 8, 8.25, log($M\star/M_\odot$) $\lesssim$ 8, 9, respectively). It should be noted here, that the P-Cygni profiles here are very narrow and non-stellar (i.e., the stellar component has been subtracted) and should not be confused with the strong stellar P-Cygni profiles seen at high-metallicity. Also, C IV absorbers within the CLASSY sample have increasing EWs at larger stellar masses and metallicities. We do not find further particular correlations with other gas and stellar properties (e.g., SFR, E(B-V), density). Of course, these inferences are being drawn from our local sample of high-z analogues and may not hold true at all redshifts.

Overall, given the complex nature of the C IV nebular emission (see also Sec. 6.4.1. of Paper IV), diagnostic diagrams involving the C IV lines (i.e., C IV/He II and C IV/C III]

line ratios, as well as EW(C IV)) should be used with caution. The complexity of this line is also evident in the velocity offset that we find for C IV lines, in that they are systematically redshifted by $\sim 25 - 100$ km/s with respect to the other UV lines (see Fig. 11 in App. A), also seen in Berg et al. (2019); Wofford et al. (2021); Senchyna et al. (2023a), which could be consistent with resonance lines in an outflowing medium. Indeed, we found that to properly fit the C IV lines, we have to employ a different value of their velocity[15] (and thus line center) with respect to He II and O III]. However, in order to fully explore and compare in detail the different line kinematics, we would need integral-field spectroscopy data in the UV to see how the spectra change as a function of the location within the galaxy.

Moving to He II instead, its stellar contribution is due to massive WR stars. Their contribution is more evident in optical spectra, where the He II $\lambda 4686$ line can be blended with several metal lines in the so-called "blue-bump" (Brinchmann et al. 2008). As discussed in Sec. 3.1.1 of Paper IV, our stellar component fitting accounts for different prescriptions for the evolution and atmospheres of massive stars, predicting He II$\lambda 1640$ stellar contribution (e.g., Senchyna et al. 2021). The best example is Mrk 996 (J0127-0619) shown in Fig. 10, which hosts a significant WR population (James et al. 2009) and a broad He II$\lambda 1640$ feature ($\sigma > 100$ km/s), that is completely stellar according to the C&B stellar component fit (blue dashed line), without signs of nebular emission. Only four CLASSY galaxies (i.e., J1129+2034, J1253-0312, J1200+1343, J1314+3452) have S/N(He II) > 3 and σ(He II) > 100 km/s (up to 200 km/s), indicating that a stellar contribution could still be present despite our subtraction of the stellar-component emission. Indeed, after a visual inspection of their optical spectra, all these four galaxies show the presence of the WR blue-bump.

Overall, we acknowledge that He II$\lambda 1640$ (both for its stellar/nebular origin and high ionization potential) can be also problematic to reproduce by photoionization models (e.g., Gutkin et al. 2016; Steidel et al. 2016; Nanayakkara et al. 2019). This led some authors (e.g., Hirschmann et al. 2019; Feltre et al. 2016) to propose UV diagnostic diagrams without He II, comparing C IV $\lambda\lambda 1549, 51$ and C III] $\lambda\lambda 1907, 9$ with other UV lines, such as N V $\lambda\lambda 1239, 1243$, N IV] $\lambda\lambda 1483, 87$ and N III] $\lambda\lambda 1747$–54, or optical lines like [O III] $\lambda 5007$ and [O I] $\lambda 6300$ (that can be a good low metallicity AGN and shock diagnostic, Sec. 4.1, 4.2). However, as shown in Paper IV, these UV lines, as well as Si III] $\lambda\lambda 1893, 92$, are very faint and observable in a handful of objects, while it is difficult to have extended coverage from the UV up to the very faint [O I] $\lambda 6300$.

---

[15] C III] were fit separately from the other UV lines, as explained in Paper IV Sec. 3.2.
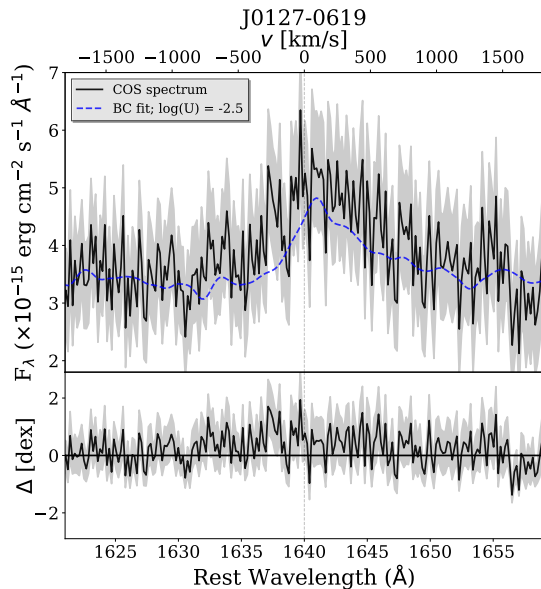
Figure 10. The He IIλ1640 feature of Mrk 996 (J0127-0619) super-imposed with the stellar component best-fit made with C&B (blue dashed line) described in detail in Sec. 3.1.1 of Paper IV. The bottom panel shows the difference between the observed spectrum and the C&B model. The very broad He II component ($\sigma > 100$ km/s) has a stellar origin, in contrast to the nebular He II emission that we generally observe in other CLASSY galaxies ($\sigma \sim 50$ km/s; see Fig. 22 of Paper IV). This figure aims to illustrate the broad stellar He II profile and guide the user in its identification.

This is the reason why we have not explored further UV diagnostic diagrams in this paper.

Up to now, the furthest spectroscopically confirmed galaxies are four young, extremely metal-poor (stellar metallicity log(Z/Z$_\odot$) $\sim -1.91 - -0.18$), low-mass ($M \sim 10^7 - 10^8$ M$_\odot$) systems at $10.3 < z < 13.2$ (Curtis-Lake et al. 2022). Their redshifts were estimated using the Lyman-break as a reference, because of the complete lack of both UV and optical emission lines, which could be due to a combination of the NIRSpec prism spectral resolution ($R \sim 100$) and their low metallicity (see also Roberts-Borsani et al. 2022a). The current furthest ($z \sim 10.603$) galaxy spectrum with clear emission lines (from Lyα to Hγ) is GN-z11, the most luminous candidate $z > 10$ Lyman break galaxy in the GOODS-North field (Bunker et al. 2023). This galaxy is compact and metal-poor (log(Z$_{neb}$/Z$_\odot$) ~ -0.92; 12+log(O/H) ~ 7.5), more massive ($\sim 10^9$ $M_\odot$) and luminous than the objects analyzed in Curtis-Lake et al. (2022) (Bunker et al. 2023; Tacchella et al. 2023). It is quite interesting that GN-z11 clearly shows both N IV]$\lambda\lambda$1483,7 and N III] λ1750 (Bunker et al. 2023), very rarely observed in the local Universe as well as in the CLASSY survey (Paper IV). In particular, N IV]$\lambda\lambda$1483,7 lines (ionization potential $E > 47.5$ eV) are rarely seen in emission in SFGs (Fosbury et al. 2003; Raiter et al. 2010;

Vanzella et al. 2010; Mingozzi et al. 2022; Harikane et al. 2023), trace very dense gas ($n_e > 10^4$ cm^{-3}) and, when present, they can also be - but not unambiguously - a signature of AGN ionization. N III] $\lambda\lambda$1750 instead is a multiplet that could have a nebular origin and be prominent in SFGs, but has been rarely observed (e.g., in the Sunburst Arc, a lensed $z \sim 2.4$ galaxy hosting a $\sim 3-4$ Myr super star cluster; Pascale et al. 2023) and is visible in only two galaxies of the CLASSY sample: Mrk 996 (i.e., J0127-0619) and J1253-0312 (Fig. 20 of Paper IV; log(N III]λ1750/He IIλ1640) $\sim$ 1.34, 0.05, respectively), both also characterized by WR features in their optical spectra, as commented above. WR stars represent the best tracer of a young ongoing starburst in galaxies, since they are uniquely characterized by very fast and strong winds (especially at high-metallicity), and thus a short lifetime ($\sim 10^5$ yr; Crowther 2007), that can lead to the production of prominent nitrogen features both at optical and UV wavelengths (e.g., Crowther & Smith 1997). However, in WR-dominated spectra, it is expected also to observe N IV] λ1719 resonance lines with a clear P-Cygni profile, which can help to discriminate between the nebular and stellar origin of the N III] and N IV] emission. Mrk 996 of the CLASSY sample shows a N IV] λ1719 P-Cygni feature, a broad He II λ1640, clear WR features in the optical spectra (James et al. 2009) as well as the N III] λ1750 multiplet, so a (partly) stellar origin of the UV N emission lines is plausible. Different is the case of GN-z11, which has opened an intense debate in the literature, with authors implying an extremely elevated nitrogen enrichment (at only 440 Myr from the Big Bang) and proposing different scenarios to explain it, including signatures of globular cluster precursors, massive star winds, runaway stellar collisions or tidal disruption events (Senchyna et al. 2023b; Cameron et al. 2023; Charbonnel et al. 2023). Recently, Maiolino et al. (2023) claimed that this object hosts AGN activity which would be the origin of the peculiar nitrogen emission lines, thus solving the issue of their presence and the puzzling N enrichment. To conclude, we need further studies on local samples where we can broadly observe these UV N features to ultimately clarify their origin and diagnostic power. Taking into account all these considerations, at the moment we consider these features not reliable diagnostics for the so-called *UV BPT diagrams*.

7. CONCLUSIONS

In this work, we investigated the optical and UV diagnostics tracing different ionization mechanisms (i.e., SF, AGN, shocks), using the CLASSY survey (Berg et al. 2022; James et al. 2022), that collects for the first time high-quality, high-resolution, and broad-wavelength-range FUV ($\sim 1200 - 2000$ Å) spectra for 45 nearby ($0.002 < z < 0.182$) star-forming galaxies, including analogs of the high-

z universe thanks to the broad ISM properties parameter space covered. This paper builds on our results shown in CLASSY Paper IV in which we measured CLASSY UV (from N IV]$\lambda\lambda$1483,7 to C III]$\lambda\lambda$1907,9) and optical (from [O II]$\lambda\lambda$3727 to [S III]λ9069) emission lines, and several ISM properties (i.e., n_e, T_e, $E(B-V)$, 12+log(O/H), log(U)).

As a first result of this paper, we confirmed the star-forming nature of our systems using different well-known optical diagnostics sensitive to the dominant ionization source, also taking into account their limitations due to the low metallicity of some of our galaxies. Then, we explored the proposed UV counterparts of these diagnostic diagrams - the so-called "UV-BPT diagrams" - from previous works based on photoionization and shock models (Gutkin et al. 2016; Jaskot & Ravindranath 2016; Hirschmann et al. 2019; 2022). To do this, we compared our measured line ratios (and their corresponding ISM properties) with a set of state-of-the-art models taken from the literature: constant and single-burst SF models from Gutkin et al. (2016); AGN models from Feltre et al. (2016); shock models from the 3MdBs[16] database (Alarie & Morisset 2019). Finally, we discussed which conditions favor the detection of UV emission lines. In the following, we summarise our findings:

1. In Sec. 4 we explored the classical BPT diagrams, as well as the use of [O I]λ6300 as a shock diagnostic and the Shirazi & Brinchmann (2012) diagram, that takes into account the He IIλ4686 emission line. In particular, we showed that CLASSY galaxies have [O III]/Hβ and He IIλ4686/Hβ typical of SFGs. Also, their broad components - if present - are still generally classified as star-forming, and, even though have enhanced [N II]/Hα, [S II]/Hα, [O I]/Hα and [O I]/[O III] and low [O III]/[O II] (possible evidence of shocks), do not correlate with He IIλ4686, that could be enhanced by shock ionization. Overall, it is not possible to completely exclude the presence of shocks (and possibly AGN activity for the only trans-solar metallicity CLASSY galaxy), but the main ionizing mechanism is clearly SF, especially for the CLASSY subsample showing UV emission lines.

2. In Sec. 5, we presented our list of reliable UV diagnostic plots that we explored for the CLASSY galaxies showing UV emission lines. Many of the discussed UV diagnostic diagrams can separate SF from AGN and shocks, with the caveat that low 12+log(O/H) and carbon abundance C/O objects can fall in the AGN/shock-locus proposed by previous authors (e.g., C3He2-C4He2, Fig. 5; C4C3He2-C4C3,

Fig. 7; C3He2-C4C3 in App.B), while they usually struggle in separating AGN from shocks. Diagnostics using carbon lines equivalent widths identify the star formation locus, even at low metallicity, low C/O abundance and high log(U) (e.g., EWC4-C3He2 and EWC3-C3He2; Fig. 8), but the modelled EWs are heavily dependent on the model assumptions of the continuum component. Also, the C IV doublet has a complex nature that still needs to be properly understood and is rarely observed to be purely nebular (see also Sec. 6), thus diagnostic diagrams to investigate the ionization source or ISM properties involving these lines should be used with caution.

3. In Sec. 5, we also concluded that the combination of C III] $\lambda\lambda$1907,9, He II λ1640 and O III] λ1666 (C3He2-O3He2, Fig. 6, see Eq. 3, 4; C3He2-C3O3 in App.B, Eq. C1, C2) represents the best diagnostic diagram able to separate all the three ionization mechanisms at sub-solar metallicities. In particular, AGN models show lower O III]/He II and O III]/C III] than SF and shocks, while SF grids are usually located at higher C III]/He II.

4. In Sec. 6, we confirmed that UV emission lines observed in CLASSY (with C III]$\lambda\lambda$1907,9, the most common FUV line, excluding Lyα) are mainly favored in ISMs with low metallicity and high ionization parameters, with none of them visible in targets with 12+log(O/H) $\gtrsim$ 8.3. In particular, C III]-emitters have systematically lower log([N II]/Hα) ($\lesssim -1$; see Fig. 2), can reach high values log(He IIλ4686/Hβ) ($\sim -1.4--1.8$; Fig. 4), and are clearly shifted to higher log([O III]/Hβ) ($\sim 0.5-1$) and higher log([O III]/[O II]) ($\sim 0.5-1.5$; Fig. 2), and thus higher log(U). We also showed that the C IV$\lambda\lambda$1548,51 and He IIλ1640 emission lines measured in CLASSY are mainly nebular, given the consistency of their velocity dispersion with O III]λ1666 ($\sigma \sim 50$ km/s; Fig. 1). Finally, we underlined the almost total absence of nitrogen UV lines (i.e., N IV]$\lambda\lambda$1483,7, N III]λ1750) in the CLASSY survey. Hence, we need further studies of samples where we can observe these features before attesting their reliability as diagnostics to investigate the main source of ionization.

Overall, this paper together with Paper IV provides the tool-kit to investigate ISM properties using UV emission lines from the CLASSY survey. This can be particularly important to explore the $z > 6$ rest-frame UV spectra in the JWST era, in which an unprecedented number of EoR galaxies have already been revealed (e.g., Curtis-Lake et al. 2022; Curti et al. 2022; Arellano-Córdova et al. 2022a; Bunker et al. 2023; Cameron et al. 2023; Roberts-Borsani et al. 2022a).

REFERENCES

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543, doi: 10.1088/0067-0049/182/2/543

Alarie, A., & Morisset, C. 2019, RMxAA, 55, 377, doi: 10.22201/ia.01851101p.2019.55.02.21

Allen, M. G., Groves, B. A., Dopita, M. A., Sutherland, R. S., & Kewley, L. J. 2008, ApJS, 178, 20

Amorín, R., Vílchez, J. M., Hägele, G. F., et al. 2012, ApJL, 754, L22, doi: 10.1088/2041-8205/754/2/L22

Arellano-Córdova, K. Z., Berg, D. A., Chisholm, J., et al. 2022a, ApJL, 940, L23, doi: 10.3847/2041-8213/ac9ab2

Arellano-Córdova, K. Z., Mingozzi, M., Berg, D. A., et al. 2022b, ApJ, 935, 74, doi: 10.3847/1538-4357/ac7854

Arrabal Haro, P., Dickinson, M., Finkelstein, S. L., et al. 2023, arXiv e-prints, arXiv:2304.05378, doi: 10.48550/arXiv.2304.05378

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481

Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5

Belfiore, F., Maiolino, R., Maraston, C., et al. 2016, MNRAS, 461, 3111, doi: 10.1093/mnras/stw1234

Berg, D. A., Chisholm, J., Erb, D. K., et al. 2019, ApJL, 878, L3, doi: 10.3847/2041-8213/ab21dc

—. 2021, arXiv e-prints, arXiv:2105.12765. https://arxiv.org/abs/2105.12765

Berg, D. A., Skillman, E. D., Henry, R. B. C., Erb, D. K., & Carigi, L. 2016, ApJ, 827, 126

Berg, D. A., James, B. L., King, T., et al. 2022, ApJS, 261, 31, doi: 10.3847/1538-4365/ac6c03

Bosch, G., Hägele, G. F., Amorín, R., et al. 2019, MNRAS, 489, 1787, doi: 10.1093/mnras/stz2230

Brinchmann, J., Kunth, D., & Durret, F. 2008, A&A, 485, 657. https://arxiv.org/abs/0805.1073

Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000, doi: 10.1046/j.1365-8711.2003.06897.x

Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, ApJ, 798, 7, doi: 10.1088/0004-637X/798/1/7

Bunker, A. J., Saxena, A., Cameron, A. J., et al. 2023, arXiv e-prints, arXiv:2302.07256, doi: 10.48550/arXiv.2302.07256

Byler, N., Dalcanton, J. J., Conroy, C., & Johnson, B. D. 2017, ApJ, 840, 44

Byler, N., Dalcanton, J. J., Conroy, C., et al. 2018, ApJ, 863, 14, doi: 10.3847/1538-4357/aacd50

Byler, N., Kewley, L. J., Rigby, J. R., et al. 2020, ApJ, 893, 1, doi: 10.3847/1538-4357/ab7ea9

Cameron, A. J., Saxena, A., Bunker, A. J., et al. 2023, arXiv e-prints, arXiv:2302.04298, doi: 10.48550/arXiv.2302.04298

Cann, J. M., Satyapal, S., Abel, N. P., et al. 2018, ApJ, 861, 142, doi: 10.3847/1538-4357/aac64a

Cann, J. M., Satyapal, S., Bohn, T., et al. 2020, ApJ, 895, 147, doi: 10.3847/1538-4357/ab8b64

Cann, J. M., Satyapal, S., Rothberg, B., et al. 2021, ApJL, 912, L2, doi: 10.3847/2041-8213/abf56d

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245

Castellano, M., Fontana, A., Treu, T., et al. 2022, arXiv e-prints, arXiv:2212.06666, doi: 10.48550/arXiv.2212.06666

Chabrier, G. 2003, PASP, 115, 763, doi: 10.1086/376392

Charbonnel, C., Schaerer, D., Prantzos, N., et al. 2023, arXiv e-prints, arXiv:2303.07955, doi: 10.48550/arXiv.2303.07955

Chisholm, J., Rigby, J. R., Bayliss, M., et al. 2019, ApJ, 882, 182, doi: 10.3847/1538-4357/ab3104

Concas, A., Maiolino, R., Curti, M., et al. 2022, MNRAS, 513, 2535, doi: 10.1093/mnras/stac1026

Crowther, P. A. 2007, 45, 177

Crowther, P. A., & Smith, L. J. 1997, A&A, 320, 500

Curti, M., D'Eugenio, F., Carniani, S., et al. 2022, arXiv e-prints, arXiv:2207.12375. https://arxiv.org/abs/2207.12375

Curtis-Lake, E., Carniani, S., Cameron, A., et al. 2022, arXiv e-prints, arXiv:2212.04568. https://arxiv.org/abs/2212.04568

del Valle-Espinosa, M. G., Sanchez-Janssen, R., Amorin, R., et al. 2022, arXiv e-prints, arXiv:2212.10585, doi: 10.48550/arXiv.2212.10585

Denicoló, G., Terlevich, R., & Terlevich, E. 2002, MNRAS, 330, 69, doi: 10.1046/j.1365-8711.2002.05041.x

Dopita, M. A., & Sutherland, R. S. 1995, ApJ, 455, 468, doi: 10.1086/176596

—. 1996, ApJS, 102, 161, doi: 10.1086/192255

—. 2003, Astrophysics of the diffuse universe

Dopita, M. A., Sutherland, R. S., Nicholls, D. C., Kewley, L. J., & Vogt, F. P. A. 2013, ApJS, 208, 10

Draine, B. T. 2011, Physics of the Interstellar and Intergalactic Medium

Du, X., Shapley, A. E., Martin, C. L., & Coil, A. L. 2017, ApJ, 838, 63, doi: 10.3847/1538-4357/aa64cf

Eldridge, J. J., & Stanway, E. R. 2009, MNRAS, 400, 1019

Feltre, A., Charlot, S., & Gutkin, J. 2016, MNRAS, 456, 3354

Ferland, G. J., Porter, R. L., van Hoof, P. A. M., & others. 2013, RMxAA, 49, 137

Finkelstein, S. L., Bagley, M. B., Arrabal Haro, P., et al. 2022, ApJL, 940, L55, doi: 10.3847/2041-8213/ac966e

Fosbury, R. A. E., Villar-Martín, M., & others. 2003, ApJ, 596, 797

Garnett, D. R., Kennicutt, Jr., R. C., Chu, Y.-H., & Skillman, E. D. 1991, ApJ, 373, 458

Grandi, S. A. 1976, ApJ, 206, 658, doi: 10.1086/154424

Groves, B. A., Heckman, T. M., & Kauffmann, G. 2006, MNRAS, 371, 1559, doi: 10.1111/j.1365-2966.2006.10812.x

Gutkin, J., Charlot, S., & Bruzual, G. 2016, MNRAS, 462, 1757

Harikane, Y., Nakajima, K., Ouchi, M., et al. 2023, arXiv e-prints, arXiv:2304.06658, doi: 10.48550/arXiv.2304.06658

Hatano, S., Ouchi, M., Nakajima, K., et al. 2023, arXiv e-prints, arXiv:2304.03726, doi: 10.48550/arXiv.2304.03726

Heckman, T. M. 1980, A&A, 87, 152

Henry, R. B. C., Edmunds, M. G., & Köppen, J. 2000, ApJ, 541, 660

Hirschmann, M., Charlot, S., Feltre, A., et al. 2019, MNRAS, 487, 333, doi: 10.1093/mnras/stz1256

—. 2022, arXiv e-prints, arXiv:2212.02522. https://arxiv.org/abs/2212.02522

Ho, I. T., Kewley, L. J., Dopita, M. A., et al. 2014, MNRAS, 444, 3894, doi: 10.1093/mnras/stu1653

Hogarth, L., Amorín, R., Vílchez, J. M., et al. 2020, MNRAS, 494, 3541, doi: 10.1093/mnras/staa851

Izotov, Y. I., Chaffee, F. H., & Schaerer, D. 2001, A&A, 378, L45, doi: 10.1051/0004-6361:20011265

Izotov, Y. I., Schaerer, D., Blecha, A., et al. 2006, A&A, 459, 71, doi: 10.1051/0004-6361:20065622

Izotov, Y. I., & Thuan, T. X. 1999, ApJ, 511, 639

—. 2007, ApJ, 665, 1115

James, B. L., Pettini, M., Christensen, L., et al. 2014, MNRAS, 440, 1794

James, B. L., Tsamis, Y. G., Barlow, M. J., et al. 2009, MNRAS, 398, 2, doi: 10.1111/j.1365-2966.2009.15172.x

James, B. L., Berg, D. A., King, T., et al. 2022, arXiv e-prints, arXiv:2206.01224. https://arxiv.org/abs/2206.01224

Jaskot, A. E., & Ravindranath, S. 2016, ApJ, 833, 136

Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, MNRAS, 346, 1055, doi: 10.1111/j.1365-2966.2003.07154.x

Keel, W. C. 1983, ApJ, 269, 466, doi: 10.1086/161057

Kewley, L. J., & Dopita, M. A. 2002, ApJS, 142, 35

Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, ApJ, 556, 121

Kewley, L. J., & Ellison, S. L. 2008, ApJ, 681, 1183

Kewley, L. J., Groves, B., Kauffmann, G., & Heckman, T. 2006, MNRAS, 372, 961, doi: 10.1111/j.1365-2966.2006.10859.x

Kewley, L. J., Nicholls, D. C., & Sutherland, R. S. 2019, ARA&A, 57, 511, doi: 10.1146/annurev-astro-081817-051832

Komarova, L., Oey, M. S., Krumholz, M. R., et al. 2021, ApJL, 920, L46, doi: 10.3847/2041-8213/ac2c09

Korista, K. T., & Ferland, G. J. 1989, ApJ, 343, 678, doi: 10.1086/167739

Lamareille, F. 2010, A&A, 509, A53, doi: 10.1051/0004-6361/200913168

Lamareille, F., Mouhcine, M., Contini, T., Lewis, I., & Maddox, S. 2004, MNRAS, 350, 396, doi: 10.1111/j.1365-2966.2004.07697.x

Leitherer, C., Leão, J. R. S., Heckman, T. M., & others. 2001, ApJ, 550, 724

Luridiana, V., Morisset, C., & Shaw, R. A. 2015, A&A, 573, A42

Madau, P., & Haardt, F. 2015, ApJL, 813, L8

Mainali, R., Kollmeier, J. A., Stark, D. P., & others. 2017, ApJL, 836, L14

Mainali, R., Zitrin, A., Stark, D. P., et al. 2018, MNRAS, 479, 1180

Maiolino, R., Scholtz, J., Witstok, J., et al. 2023, arXiv e-prints, arXiv:2305.12492, doi: 10.48550/arXiv.2305.12492

Marasco, A., Belfiore, F., Cresci, G., et al. 2023, A&A, 670, A92, doi: 10.1051/0004-6361/202244895

Markwardt, C. B. 2009, in Astronomical Society of the Pacific Conference Series, Vol. 411, Astronomical Data Analysis Software and Systems XVIII, ed. D. A. Bohlender, D. Durand, & P. Dowler, 251

Maseda, M. V., Brinchmann, J., Franx, M., et al. 2017, A&A, 608, A4, doi: 10.1051/0004-6361/201730985

Matthee, J., Mackenzie, R., Simcoe, R. A., et al. 2022, arXiv e-prints, arXiv:2211.08255, doi: 10.48550/arXiv.2211.08255

Mayya, Y. D., Plat, A., Gómez-González, V. M. A., et al. 2023, MNRAS, 519, 5492, doi: 10.1093/mnras/stad017

McElroy, R., Croom, S. M., Pracy, M., et al. 2015, MNRAS, 446, 2186, doi: 10.1093/mnras/stu2224

Mingozzi, M., Berg, D. A., Chisholm, J., et al. 2022, ApJ

Mingozzi, M., Cresci, G., Venturi, G., et al. 2019, A&A, 622, A146, doi: 10.1051/0004-6361/201834372

Molina, M., Reines, A. E., Latimer, L. J., Baldassare, V., & Salehirad, S. 2021, ApJ, 922, 155, doi: 10.3847/1538-4357/ac1ffa

Nagao, T., Maiolino, R., & Marconi, A. 2006, A&A, 447, 863, doi: 10.1051/0004-6361:20054127

Nakajima, K., & Maiolino, R. 2022, MNRAS, 513, 5134, doi: 10.1093/mnras/stac1242

Nakajima, K., Schaerer, D., Le Fèvre, O., et al. 2018, A&A, 612, A94, doi: 10.1051/0004-6361/201731935

Nanayakkara, T., Brinchmann, J., Boogaard, L., et al. 2019, A&A, 624, A89, doi: 10.1051/0004-6361/201834565

Osterbrock, D. E., & Ferland, G. J. 2006, Astrophysics of gaseous nebulae and active galactic nuclei

Pascale, M., Dai, L., McKee, C. F., & Tsang, B. T. H. 2023, arXiv e-prints, arXiv:2301.10790, doi: 10.48550/arXiv.2301.10790

Pettini, M., & Pagel, B. E. J. 2004, MNRAS, 348, L59

Plat, A., Charlot, S., Bruzual, G., et al. 2019, MNRAS, 490, 978, doi: 10.1093/mnras/stz2616

Polimera, M. S., Kannappan, S. J., Richardson, C. T., et al. 2022, ApJ, 931, 44, doi: 10.3847/1538-4357/ac6595

Raiter, A., Fosbury, R. A. E., & Teimoorinia, H. 2010, A&A, 510, A109, doi: 10.1051/0004-6361/200912429

Ravindranath, S., Monroe, T., Jaskot, A., Ferguson, H. C., & Tumlinson, J. 2020, ApJ, 896, 170, doi: 10.3847/1538-4357/ab91a5

Reines, A. E., Condon, J. J., Darling, J., & Greene, J. E. 2020, ApJ, 888, 36, doi: 10.3847/1538-4357/ab4999

Rich, J. A., Dopita, M. A., Kewley, L. J., & Rupke, D. S. N. 2010, ApJ, 721, 505, doi: 10.1088/0004-637X/721/1/505

Rigby, J. R., Bayliss, M. B., Gladders, M. D., & others. 2015, ApJL, 814, L6

Rigby, J. R., Bayliss, M. B., Sharon, K., et al. 2018, AJ, 155, 104, doi: 10.3847/1538-3881/aaa2ff

Roberts-Borsani, G., Morishita, T., Treu, T., et al. 2022a, ApJL, 938, L13, doi: 10.3847/2041-8213/ac8e6e

Roberts-Borsani, G., Treu, T., Chen, W., et al. 2022b, arXiv e-prints, arXiv:2210.15639, doi: 10.48550/arXiv.2210.15639

Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, ApJL, 802, L19

Sartori, L. F., Schawinski, K., Treister, E., et al. 2015, MNRAS, 454, 3722, doi: 10.1093/mnras/stv2238

Schaerer, D., & Vacca, W. D. 1998, ApJ, 497, 618

Schaerer, D., Izotov, Y. I., Worseck, G., et al. 2022, arXiv e-prints, arXiv:2202.07768. https://arxiv.org/abs/2202.07768

Schmutz, W., Leitherer, C., & Gruenwald, R. 1992, PASP, 104, 1164, doi: 10.1086/133104

Senchyna, P., Chisholm, J., Berg, D. A., James, B. L., & others (in preparation). 2023a, ApJ

Senchyna, P., Plat, A., Stark, D. P., & Rudie, G. C. 2023b, arXiv e-prints, arXiv:2303.04179, doi: 10.48550/arXiv.2303.04179

Senchyna, P., Stark, D. P., Chevallard, J., et al. 2019, MNRAS, 488, 3492, doi: 10.1093/mnras/stz1907

Senchyna, P., Stark, D. P., Charlot, S., et al. 2021, arXiv e-prints, arXiv:2111.11508. https://arxiv.org/abs/2111.11508

Shirazi, M., & Brinchmann, J. 2012, MNRAS, 421, 1043

Sixtos, A., Wofford, A., Sander, A. A. C., & Peimbert, A. 2023, MNRAS, 519, 5656, doi: 10.1093/mnras/stac3732

Stanway, E. R., Eldridge, J. J., & Becker, G. D. 2016, MNRAS, 456, 485

Stark, D. P. 2016, ARA&A, 54, 761

Stark, D. P., Richard, J., Charlot, S., & others. 2015, MNRAS, 450, 1846

Stasińska, G., Izotov, Y., Morisset, C., & Guseva, N. 2015, A&A, 576, A83, doi: 10.1051/0004-6361/201425389

Steidel, C. C., Strom, A. L., Pettini, M., et al. 2016, ApJ, 826, 159

Stephenson, M. G., Arellano-Córdova, K. Z., Berg, D. A., Mingozzi, M., & James, B. L. 2023, Research Notes of the American Astronomical Society, 7, 31, doi: 10.3847/2515-5172/acbc12

Sutherland, R. S., & Dopita, M. A. 2017, ApJS, 229, 34, doi: 10.3847/1538-4365/aa6541

Tacchella, S., Eisenstein, D. J., Hainline, K., et al. 2023, arXiv e-prints, arXiv:2302.07234, doi: 10.48550/arXiv.2302.07234

Thuan, T. X., & Izotov, Y. I. 2005, ApJ, 161, 240

Tozzi, G., Maiolino, R., Cresci, G., et al. 2023, arXiv e-prints, arXiv:2302.04282, doi: 10.48550/arXiv.2302.04282

Treu, T., Roberts-Borsani, G., Bradac, M., et al. 2022, ApJ, 935, 110, doi: 10.3847/1538-4357/ac8158

Trump, J. R., Arrabal Haro, P., Simons, R. C., et al. 2023, ApJ, 945, 35, doi: 10.3847/1538-4357/acba8a

Vanzella, E., Grazian, A., Hayes, M., et al. 2010, A&A, 513, A20, doi: 10.1051/0004-6361/200913042

Veilleux, S., & Osterbrock, D. E. 1987, ApJS, 63, 295, doi: 10.1086/191166

Wise, J. H., Demchenko, V. G., Halicek, M. T., et al. 2014, MNRAS, 442, 2560

Wofford, A., Vidal-García, A., Feltre, A., Chevallard, J., et al. 2021, MNRAS, 500, 2908, doi: 10.1093/mnras/staa3365

Xiao, L., Stanway, E. R., & Eldridge, J. J. 2018, MNRAS, 477, 904, doi: 10.1093/mnras/sty646

Xu, X., Heckman, T., Henry, A., et al. 2022, ApJ, 933, 222, doi: 10.3847/1538-4357/ac6d56

—. 2023, arXiv e-prints, arXiv:2301.11498, doi: 10.48550/arXiv.2301.11498

APPENDIX

A. COMPARING UV AND OPTICAL KINEMATICS

Here we want to further discuss the emission line kinematics, also comparing optical and UV results. The left and right panels of Fig. 11 show the velocity offset of the UV stellar (left panel) and gas kinematics (right panel) with respect to the systemic velocity v_{sys} (derived from z_{lit}), consistent with the ionized gas velocity measured from optical emission lines, as a function of the galaxy stellar mass. In general, we find no systematic offsets between the ionizing stellar component velocity explored by fitting the UV spectra (magenta squares) and v_{sys}, with Δv mainly within $\pm \sim 50$ km/s and median value $< \Delta v > \sim 7$ km/s (dash-dotted magenta line). One implication is that the narrow components of the optical emission lines are really tracing non-outflowing gas (Rigby et al. 2018). The right panel of Fig. 11 instead shows the O III] $\lambda\lambda1661,6$, C III] $\lambda\lambda1907,9$, and C IV $\lambda\lambda1548,51$ velocities as red dots, blue triangles and yellow stars. There is a scatter between the UV and systemic velocities mainly within $\pm \sim 50$ km/s, and more enhanced for some C III] measurements, with three galaxies (J1359+5726, J0926+4427 and J0942+3547) showing a clear redshifted offset ($\Delta v \sim +100 - 200$ km/s) and three (I Zw 18, Mrk 996 and J1025+3622) showing a clear blueshifted offset ($\delta v \sim -150 - -200$ km/s). However, we do not see either any systematic offsets, as shown by the dashed red and dotted blue lines indicating the corresponding median values with respect to v_{sys}. Hence, we conclude it is unlikely that they are due to issues in the G185M grating calibration, covering the C III], discussed in detail in James et al. (2022).
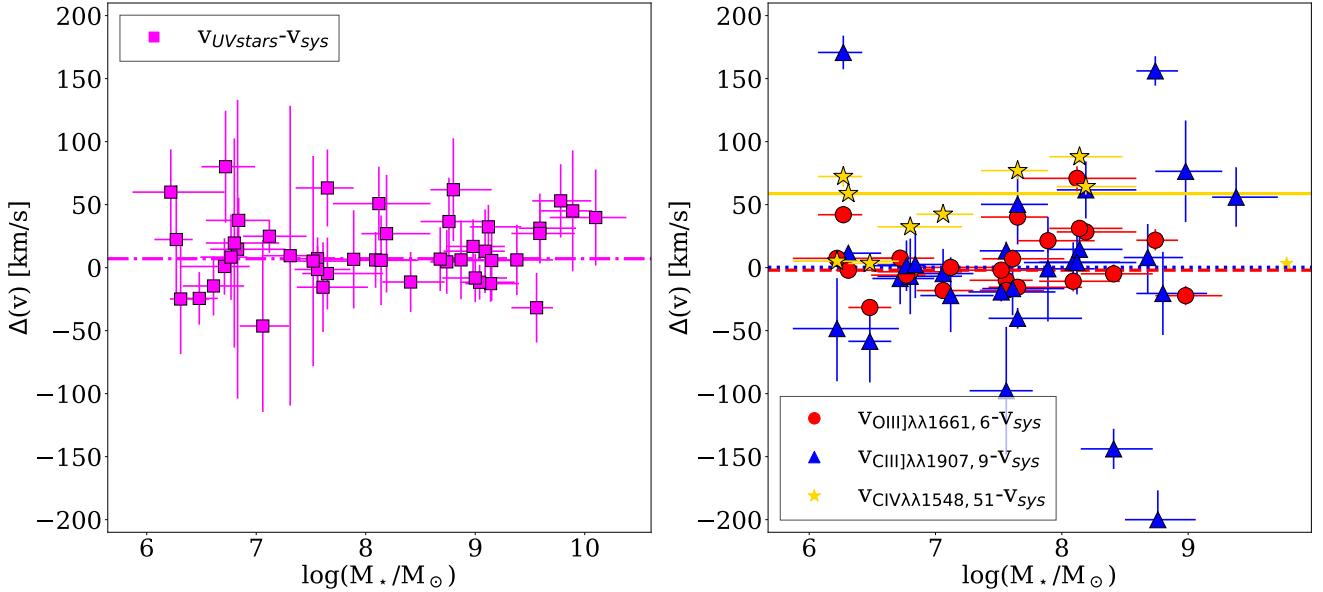

Figure 11. Velocity offset of the UV stellar (left panel, magenta squares) and gas kinematics (right panel) with respect to the systemic velocity v_{sys}, consistent with the ionized gas velocity measured from optical spectra, and as a function of the galaxy stellar mass. In the right panel, we indicate the O III] $\lambda\lambda1661,6$, C III] $\lambda\lambda1907,9$, and C IV $\lambda\lambda1548,51$ velocities as red dots, blue triangles and yellow stars, as reported in the legend.

Interestingly, we find a systematic offset between the v_{CIV} and v_{sys} with a median value of $< \Delta v > \sim +60$ km/s (solid gold line), implying that C IV emission is systematically redshifted with respect to the systemic velocity, consistently with resonant scattering/absorption (see also Berg et al. 2019; Wofford et al. 2021). For seven out of the nine galaxies for which we measure C IV nebular emission, this blueshifted offset increases with the stellar mass, going from ~ 0 km/s to ~ 90 km/s in the range $\log(M_\star/M_\odot)$ 6–8. The two exceptions are I Zw 18 and J1323-0132, with stellar masses of $\log(M_\star/M_\odot) \sim 6.3$, showing $\delta v \sim$ 65 km/s.

As discussed in Sec. 2.2 and in Fig, 1, optical emission lines of half of the CLASSY sample needed a narrow and a broad Gaussian component to be accurately reproduced. Looking at the single and narrow component σ values shown in the right panel of Fig, 1, we see that σ (and also its scatter) increases at larger stellar masses, with only objects with $M_\star \gtrsim 10^7$ M$_\odot$ needing

a 2-component fit. We highlight here that all the galaxies characterized by a broad component in optical emission lines (apart Mrk 996 and SBS 0335-052 E) were found to have an outflowing component in rest-UV absorption lines analyzed by Xu et al. (2022; 2023) (CLASSY Paper III, Paper VI). Recent studies have shown that low-mass systems ($M_\star < 10^9$ M$_\odot$) such as those of the CLASSY survey can have irregular velocity fields in the ionized gas studied through optical emission lines, indicating the presence of non-circular motions (e.g., Marasco et al. 2023). However, galactic winds, defined as gas at velocities larger than the galaxy escape speed, have found to be able to account only for a few percent of their observed fluxes both in the local universe (Marasco et al. 2023) and at $z \sim 1.2-2.6$ (Concas et al. 2022). Here we highlight that the lack of galactic wind detection in small systems could be also an issue of S/N, but in this work we do not explore this topic any further, leaving it to a future study using optical integral-field spectroscopy data, that can provide both a spatial and kinematical information of our galaxies, and thus a more complete picture that the integrated optical spectra analyzed here cannot provide.

B. ALTERNATIVE DIAGNOSTIC DIAGRAMS

Fig. 12 shows an alternative optical diagnostic diagram to discriminate the excitation properties, using [O II] $\lambda\lambda3727$/Hβ versus [O III] $\lambda5007$/Hβ line ratios (Lamareille et al. 2004; Lamareille 2010). The dotted, solid and dashed lines indicate the empirical dividing line between star-forming galaxies and AGN identified by Lamareille et al. (2004), while the dash-dotted line identifies the empirical locus of LI(N)ER and composite galaxies, found by Lamareille (2010). The advantage of this diagram is that it can be used to explore JWST targets at redshifts up to $z \sim 8.5$ in the observed NIR (with NIRISS or NIRspec), for which redder wavelengths are unavailable (e.g., Curti et al. 2022). However, this diagram should be taken with caution since, as shown in Table 2, it does not always disentangle ionizing mechanisms in our sample, with many galaxies (classified as star-forming according to the classical BPT) lying in the boundary region between SF and Seyferts and thus classified as composite. This shows the limit of [O II]/Hβ in distinguishing among ionization mechanisms.

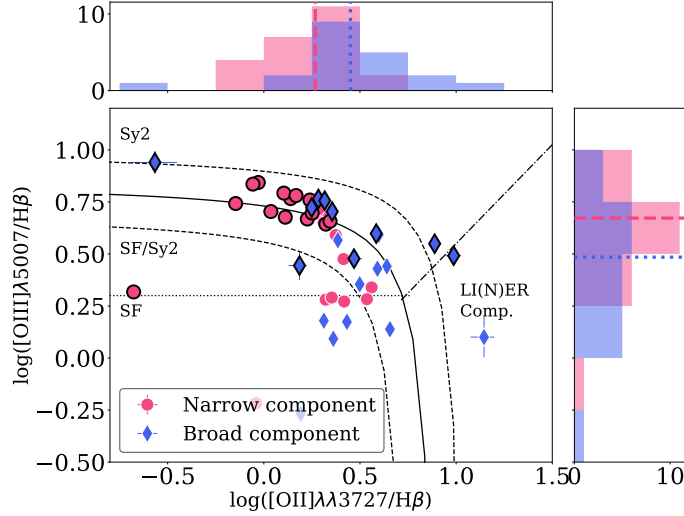**Figure 12.** Analogously to Fig. 2, [O II] $\lambda\lambda3727$/Hβ versus [O III] $\lambda5007$/Hβ diagnostic diagram for the narrow and broad components of the CLASSY galaxies, shown as red dots and blue diamonds, respectively. Only the measurements with $S/N > 3$ for all the lines involved are reported. The dotted and solid lines indicate the empirical dividing line between star-forming galaxies and AGN identified by Lamareille et al. (2004), the dashed ones show the uncertainty region, while the dash-dotted line identifies the empirical locus of LI(N)ER and Seyfert galaxies, found by Lamareille (2010). Unfortunately, this diagnostic plot does not represent a good tool to discriminate among the different ionizing mechanisms of our systems.

C. OTHER UV DIAGNOSTIC DIAGRAMS

Fig. 13 and 14 show the C3He2-C3O3 and C3He2-C4C3 diagrams, respectively, proposed together with the C3He2-O3He2 diagram (Fig. 6) by Feltre et al. (2016). C3He2-C3O3 behaves similarly to C3He2-O3He2 (Fig. 6), and thus, at sub-solar metallicities, it can clearly distinguish among SF, AGN and shocks, as shown by the Shock-SF separator
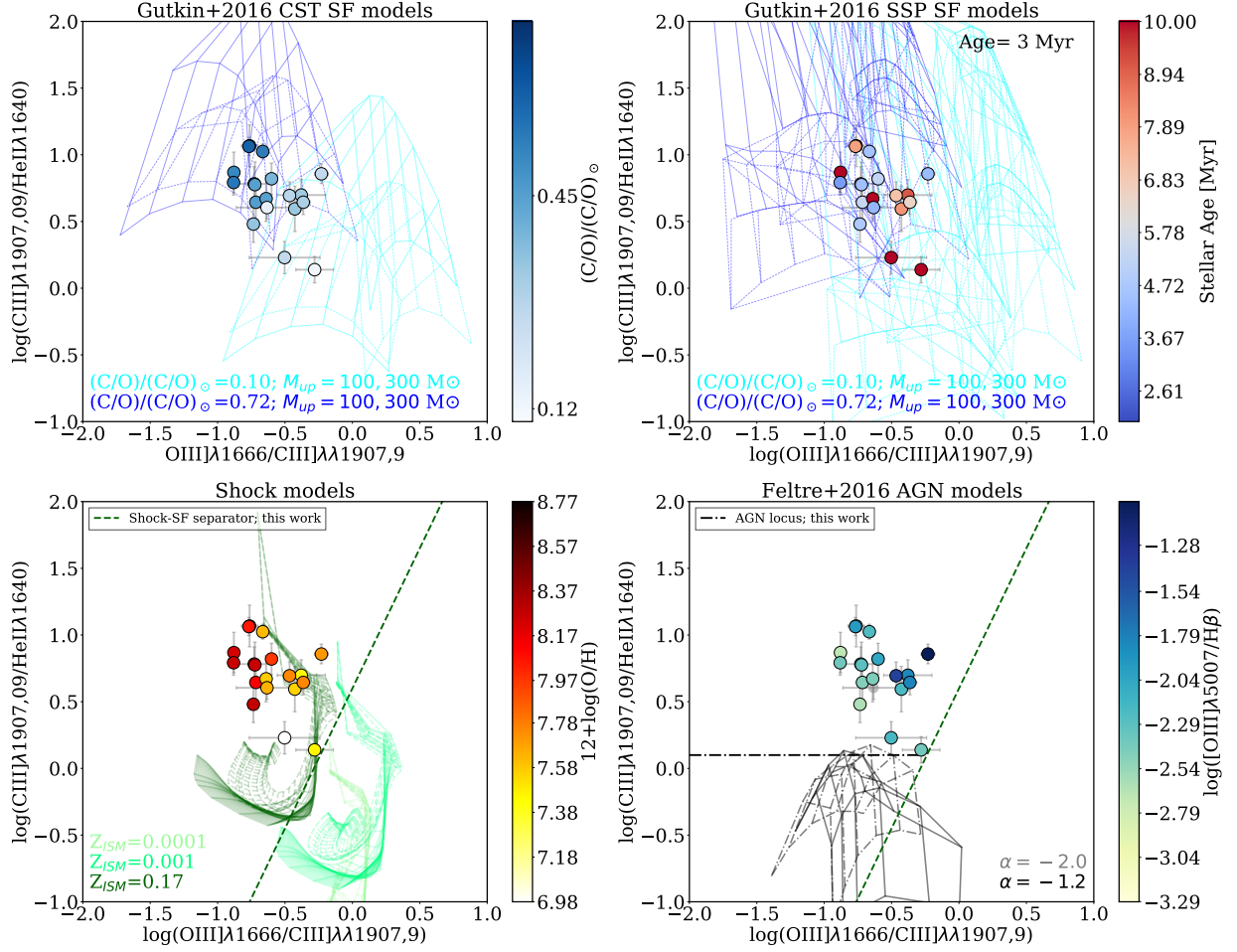
$$y = 2.1\,x + 0.6 \tag{C1}$$

Figure 13. C3He2-O3C3 diagram: C III] $\lambda\lambda$1907,9/He II λ1640 vs O III] λ1666/C III] diagnostic plot proposed by F16 to separate SF and AGN activities. The models superimposed are as explained in Fig. 5. This diagram clearly distinguishes among SF, AGN and shocks at sub-solar metallicities (i.e., excluding the dark-green shock grid), as C3He2-O3He2 (Fig. 6), as indicated by the dashed dark green and dash-dotted black lines.

and the AGN-SF separator lines

$$y = 0.1 \quad and \quad x < -0.25 \tag{C2}$$

with $x = \log([O\ III]\lambda1666/C\ III]\lambda\lambda1907,9)$ and $y = \log(C\ III]\lambda\lambda1907,9/He\ II\lambda1640)$. At sub-solar metallicities (i.e., excluding the dark-green shock grid), shocks can reach higher [O III]/C III] line ratios than AGN models, while SF grids are located at higher C III]/He II values.

C3He2-C4C3 instead behaves similarly to C3He2-C4He2 diagram (Fig. 5) and are a very similar version of the C4C3He2-C4C3 diagram (Fig. 7). There is not a clear difference in C IV/C III] line ratios between the different models, and thus the difference between SF and other mechanisms is mainly traced by the different C III]/He II (higher for SF grids, lower for AGN/shocks), as indicated by the Shock-SF separator

$$y \geq 0.1 \quad and \quad x \geq 0.1 \tag{C3}$$

and the AGN-SF separator lines

$$y = 0.1 \quad and \quad x < -0.25 \tag{C4}$$

with $x = \log(C\ IV\lambda\lambda1548,51/C\ III]\lambda\lambda1907,9)$ and $y = \log(C\ III]\lambda\lambda1907,9/He\ II\lambda1640)$ Also, AGN and shocks show similar C IV/C III], and thus this diagnostic diagram cannot distinguish them.
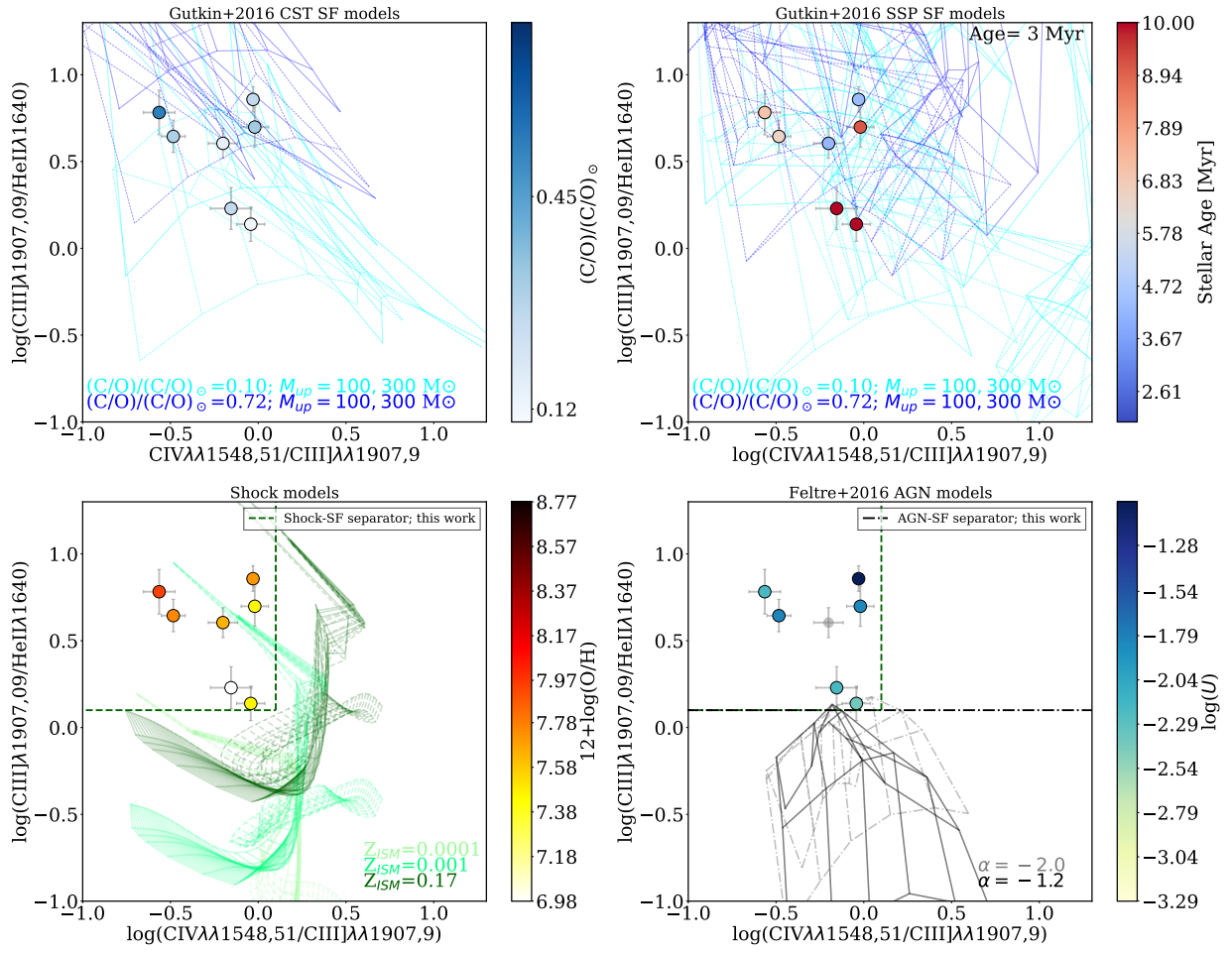
Figure 14. C3He2-C4C3 diagram: C III] λλ1907,9/He II λ1640 vs C IV λλ1548,51/C III] diagnostic plot proposed by F16 to separate SF and AGN activities. There is not a clear difference in C IV/C III] line ratios between the different models, and thus the difference between SF and other mechanisms is mainly traced by the different C III]/He II (higher for SF grids, lower for AGN/shocks), as indicated by the dashed dark green and dash-dotted black lines.