Diagnostics for PopIII galaxies and Direct Collapse Black Holes in the early universe

K. Nakajima and R. Maiolino

1 National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
2 Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK
3 Cavendish Laboratory Astrophysics Group, University of Cambridge, 19 JJ Thomson Avenue, Cambridge, CB3 0HE, UK
4 Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

Accepted for publication in MNRAS on Apr 25, 2022

ABSTRACT

Forthcoming observational facilities will make the exploration of the early universe routine, likely probing large populations of galaxies at very low metallicities. It will therefore be important to have diagnostics that can solidly identify and distinguish different classes of objects in such low metallicity regimes. We use new photoionisation models to develop diagnostic diagrams involving various nebular lines. We show that combinations of these diagrams allow the identification and discrimination of the following classes of objects in the early universe: PopIII and Direct Collapse Black Holes (DCBH) in pristine environments, PopIII and DCBH embedded in slightly enriched ISM (Z ∼ 10⁻⁵ − 10⁻⁴), (metal poor) PopII and AGN in enriched ISM. Diagnostics involving rest-frame optical lines (that will be accessible by JWST) have a better discriminatory power, but also rest-frame UV diagnostics can provide very useful information. Interestingly, we find that metal lines such as [O iii]λ5007 and C iv/λ1549 can remain relatively strong (about a factor of 0.1–1 relative H β and Heii/λ1640, respectively), even in extremely metal poor environments (Z ∼ 10⁻⁵ − 10⁻⁴), which could be embedding PopIII galaxies and DCBH.

Key words: galaxies: formation – galaxies: evolution – galaxies: high-redshift – galaxies: nuclei – galaxies: active

1 INTRODUCTION

Detecting galaxies hosting the very first generation of stars, formed out of the primordial, pristine gas, and often referred to as PopIII stars, is one of the primary, ambitious goals of modern astrophysics. These galaxies are expected to be found between the end of the “dark ages” (z ∼ 20–30) and the epoch of main reionisation (z ∼ 7) (e.g. Abel et al. 2002; Bromm et al. 2002; Schneider et al. 2006; Jeon et al. 2015; Jaacks et al. 2018), although some models expect that clumps of unprocessed gas may lead to formation of PopIII even at z ∼ 3 (Liu & Bromm 2020). PopIII galaxies are expected to be very faint (low mass) and rare, since the chemical enrichment (hence the transition to PopII) proceeds very quickly (e.g. Tornatore et al. 2007; Jaacks et al. 2018).

Some claims of PopIII detection (Sobral et al. 2015) in a luminous Lyα emitter, CR7, at z = 6.6 were not confirmed by later analyses (Bowler et al. 2017; Shibuya et al. 2018). The expectation is that PopIII galaxies might be identified with the forthcoming generation of optical/near-IR observatories, especially the James Webb Space Telescope (JWST) and the Extremely Large Telescopes (ELTs), although their detectability strongly depends on the specific model and cosmological simulations and their assumptions (e.g. Stiavelli & Trenti 2010; Pawlik et al. 2011; Inayoshi et al. 2018; Visbal et al. 2020; Vikaeus et al. 2021)

Another important class of primeval objects are the seeds of supermassive black holes at the center of primeval galaxies. The finding of supermassive black holes (M BH > 10⁶M☉) already in place by z ∼ 7.5, has triggered the development of numerous models and scenario about their formation and the nature of their progenitors, including massive stellar (PopIII) remnants accreting at super-Eddington rates, nuclear stellar clusters and the so-called Direct Collapse Black Holes (DCBHs), i.e. black holes with masses in the range of 10⁵ − 10⁶ M☉ resulting from the putative direct collapse of massive gas clouds under specific conditions (e.g. Volonteri 2012; Valiante et al. 2016; Inayoshi et al. 2020; Beckmann et al. 2019). Also in this case, some models expect that DCBH or other forms of supermassive black holes seeds may be detectable with the forthcoming generation of observing facilities (Volonteri et al. 2017; Valiante et al. 2018; Natarajan et al. 2017; Barrow et al. 2018; Habouzit et al. 2019; Whalen et al. 2020 and references therein). Some potential candidates of DCBHs have actually been suggested in the past based on broad-band photometry, but which still need spectroscopic confirmation (Pacucci et al. 2016). An interpretation of the nature of CR7 as primeval DCBH (Pacucci et al. 2017; Hartwig et al. 2016) has subsequently been questioned by subsequent analyses of this source (Bowler et al. 2017; Shibuya et al. 2018).

Various past studies have investigated what would be the properties of these classes of primeval objects and what would be their observational signatures. As a consequence of their higher temperature, PopIII stars are expected to have slopes bluer and ionising continua harder than later generation of (PopII) stars. Along with the lack of metals, these properties are expected to result into strong hydrogen and helium line emission and, in particular, strong He i emission lines. Specifically, the presence of the He ii/λ1640 line, along with
the absence of other metal lines, has been considered a signature of ISM photoionised by PopIII stars (e.g. Schaerer 2003; Raiter et al. 2010; Inoue 2011). CR7 was a notable example (Sobral et al. 2015), although the detection of HeII1640 and the claim of PopIII was not confirmed by subsequent studies (Bowler et al. 2017; Shibuya et al. 2018; see also Sobral et al. 2019).

The spectroscopic properties of accreting DCBH or other supermassive black hole seeds have been investigated by various studies in terms of broad photometric properties (Volonteri et al. 2017; Valiante et al. 2018; Habouzit et al. 2019), however fewer studies have investigated their observational properties in terms of nebular emission lines (Pacucci et al. 2016, 2017; Hartwig et al. 2016). Generally this class of objects is expected to have properties similar to PopIII, i.e. blue continua, strong nebular lines, including strong HeII emission.

Despite the various efforts discussed above, past studies had several limitations. Many of them focused on the rest-frame ultra-violet (UV) diagnostics, which are those observable from ground at high-redshift; however, JWST will provide additional key information also on the rest-frame optical diagnostics. Previous studies have explored a limited range of physical parameters (e.g. in terms of ionisation parameter or shape of the ionising continuum). Moreover, past studies have generally provided the expected strength of the nebular lines in the case of PopIII stars, but have not really explored the possible range of line ratios that can be safely associated with PopIII, i.e. it is not yet clear, even in the case of detection of HeII, what are the limits on the other lines that can unambiguously identify PopIII stars. Previous studies have only explored the case of PopIII with no metals in the ISM. However, recent theoretical works expect that pockets of pristine gas might survive amid enriched ISM (Liu & Bromm 2020), which might result into the formation of PopIII stars photoionising gas that emits some metal lines. Finally, given the similarity between PopIII stars and accreting DCBH in terms of ionising spectra, generally past studies have not attempted to identify diagnostics that might potentially discriminate between the two classes of sources.

In this study we make an effort to tackle these previous limitations. We develop new photoionisation models using different PopIII ionising continua, corresponding to different IMFs, and different physical conditions and explore what are the nebular emission lines, both in the rest-frame UV and in the rest-frame optical, and explore how different emission line ratios and equivalent widths (EWs) can distinguish among these different cases and also how safely they can separate PopIII galaxies. We also explore the new scenario of (mildly) enriched ISM photoionised by PopIII stars. Moreover, we consider the case of gas photoionised by accreting DCBH and investigate whether this can be distinguished from the case of PopIII stars. We provide diagnostic diagrams that will be useful to disentangle these various cases, especially through spectroscopic follow-up of candidate PopIII and DCBH with the forthcoming new observational facilities.

2 PHOTOIONISATION MODELS

We perform photoionisation models calculations by using CLOUDY (version 13.05: Ferland et al. 1998, 2013) to obtain spectroscopic predictions for primordial objects. We primarily model sources dominated by PopIII stellar populations and DCBH to find any difference in their spectroscopic features, which would allow distinguishing between the two populations. Moreover, we also consider evolved systems such as PopII galaxies as well as accreting supermassive black holes (AGNs) embedded in metal-enriched ISM, so to provide solid recipes for the identification of extremely metal-deficient systems and their discrimination from metal enriched systems.

We adopt similar procedures as detailed in Nakajima et al. (2018) to run the CLOUDY calculations. Briefly as for the gas properties, we assume constant-density gas clouds with a plane-parallel geometry. We do not include any dust in the ionised cloud and no elements are considered depleted onto dust grains except for the models of PopII galaxies (see §2.2) and unless otherwise specified. At the zero metallicity, we adopt the primordial helium abundance given by Hsyu et al. (2020). In non-zero metallicities, all elements except nitrogen, carbon, and helium are taken to be primary nucleosynthesis elements. For carbon and nitrogen, we use forms given by Dopita et al. (2006) and López-Sánchez et al. (2012), respectively, to take into account their secondary nucleosynthesis production in the high-metallicity range. These prescriptions are based on observations of nearby HII regions, and known to work for z ~ 2 – 3 galaxies, on average (e.g., Steidel et al. 2016; Amorín et al. 2017; Hayden-Pawson et al. 2021). For helium, we use a form in Dopita et al. (2006) but slightly updated to be consistent with the latest primordial value (Hsyu et al. 2020). The ionisation state is regulated by ionisation parameter (log U), whose definition in CLOUDY is the number of hydrogen-ionising photons to the total hydrogen density in a dimensionless form at the radius of the illuminated face of the plane-parallel cloud. The ionisation parameter log U is varied from ~3.0 up to an extreme value of ~0.5 with a step of 0.5 dex on the logarithmic scale. A gas electron density of 10² cm⁻³ is used as a fiducial value, but 10² and 10⁴ cm⁻³ are also explored and presented in Appendix §A. We choose this range of gas density because (i) it is observed in ionised HII regions of actively star-forming galaxies at z ~ 0 – 2 (Shirazi et al. 2014; Sanders et al. 2016), and (ii) a high gas density of ~ 10⁵ cm⁻³ is tenuously inferred in a z = 11 galaxy (Oesch et al. 2016; Jiang et al. 2021). We adopt the same fiducial value and range of gas densities for DCBHs and AGNs, as a comparable densities are observed in the narrow-line regions of typical AGNs (e.g., Nagao et al. 2006; Dors et al. 2014; see also Kewley et al. 2013). Since there are not yet direct determinations of the gas electron density in the ionised regions surrounding DCBHs, we treat the gas density for the DCBH models in the same way as for the AGN models.

Three types of incident radiation field are considered here: PopIII galaxies, PopII galaxies, and DCBHs. Their ionising spectrum shapes are illustrated in Fig. 1, and their main parameters in CLOUDY are summarised in Table 1. We give a detailed explanation for each of the models in the followings.

We note that single-cloud models, hence adopting a single value of the ionisation parameter and gas density, are only very rough approximations of the actual distribution of clouds properties in galaxies. More realistically, even within a single HII region, although chemically homogeneous, emitting gas clouds span a broad range of ionisation parameter and density, each of them emitting optimally a subset of nebular lines. This is even more true when considering the integrated emission from an entire galaxy, which includes different kind of HII regions. An accurate modelling would require combining photo-ionisation models with clouds spanning a broad range of ionisation parameter and density, and combining them to reproduce the observed nebular line ratios. Failure to do so results often in inferring metallicities which can be a factor of 2–3 different from the actual values (Marconi et al., in prep.). However, while we are aware of the limitations of single-cloud, single-ionisation parameter models, in this paper we are mostly interested in the overall trend of the diagnostics ratios over a range of metallicities spanning order of magnitudes, hence the effect of considering the contribution of multiple clouds with different ionisation parameters is secondary.
**Table 1.** Main parameters of CLOUDY photoionisation models explored in this paper.

| Models | PopIII galaxies | PopII galaxies | DCBHs |
|--------|-----------------|----------------|-------|
|        | SED shape       | SED shape      | SED shape |
|        | Graph           | Fig. 1(a)      | Fig. 1(b) | Fig. 1(c) |
| Parameters | Salpeter IMF, 1–100 $M_\odot$, 1–500 $M_\odot$, 50–500 $M_\odot$ | BPASS, Binary, Kroupa IMF, Upper mass: 100 $M_\odot$ and 300 $M_\odot$ | Power-law + Big Bump, $\alpha = -1.2, -1.6, -2.0$ | $Z_{\text{gas}} = Z_\star$ |
| Zero-age | 1 Myr and 10 Myr (continuous) | $T_{\text{bb}}$ (K) = 5e4, 1e5, 2e5 |

**Gas properties**

| $Z_{\text{gas}}$ | 0, 1e-5, 1e-4 (higher metallicities up to 0.014 (= 1 $Z_\odot$)) | 0, 1e-5, 1e-4, 1e-3 (higher metallicities up to 0.028 (= 2 $Z_\odot$)) |
|------------------|-------------------------|-------------------------|
| log $U$          | From −0.5 to −3.0 with a step 0.5 dex (down to −3.5 for evolved galaxies) | |
| $n_e$ (cm$^{-3}$) | $10^2$, $10^3$, and $10^4$ | |

**Figure 1.** Spectral shapes of the incident radiation fields used in this paper. (a) PopIII galaxies. (b) PopII galaxies. (c) DCBHs. The cumulative number of photons emitted above the energy $E$ is illustrated as a function of $E$. In panel (b), 10 Myr continuous star-formation as well as 1 Myr bursty SEDs are presented. Only models for IMF with $M_{\text{up}} = 300$ $M_\odot$ are displayed. Differences with respect to models with IMF having $M_{\text{up}} \leq 100$ $M_\odot$ are small. In panel (c), different power-law indices are indicated with different colors ($T_{\text{bb}} = 1e5$ K), and a different $T_{\text{bb}}$ are shown with different line-styles ($\alpha = -1.6$), as detailed in the legend.
moreover, at a given metallicity, we will show trends for different ionisation parameters, which may give an indication of the effects resulting from the contribution of different populations of clouds. In Appendix §A we also show photoionisation model results assuming gas densities different from the fiducial value. The results illustrate that different densities, and hence a potential variation of density in a system, do not alter the overall selection criteria and metallicity indicators that are derived based on the constant, fiducial density models as detailed below.

2.1 PopIII galaxies

We take three zero-metal stellar population models from Schaerer (2003) and Raiter et al. (2010) as SEDs of PopIII galaxies. They are calculated at zero-age with a Salpeter (1955) IMF, considering quite different mass ranges defined by the lower- and upper-mass cut-offs of IMF: 1–100 \( M_{\odot} \), 1–500 \( M_{\odot} \), and 50–500 \( M_{\odot} \). Their hard ionising spectra even above the He\(^+\)-ionising potential \( (E > 54 \text{ eV}) \) are presented in Fig. 1(a). As illustrated by Raiter et al. (2010), the most extreme model with the mass range from 50 to 500 \( M_{\odot} \) reproduces well the most top-heavy IMFs by comparison with other existing models (e.g., Tumlinson 2006). Since recent theoretical studies indicate that extremely metal-poor stellar populations should be characterised by a top-heavy IMF (e.g., Hirano et al. 2015; Chon et al. 2021), more top-heavy than the present-day IMF such as Salpeter (1955) with 1–100 \( M_{\odot} \), our models would cover a reasonable range of IMF variations expected for PopIII galaxies. A broader range of top-heavy stellar IMF for PopIII galaxies (with different analytical forms) will be explored in a forthcoming, separate paper.

We consider the zero-metal ISM as well as slightly metal-enriched cases with \( Z_{\text{gas}} = 10^{-3} \) and \( 10^{-4} \). The model calculations are stopped at the edge of the ionised nebular cloud, i.e., when the electron fraction, defined as the ratio of the number density of electrons to that of total hydrogen, falls below \( 10^{-5} \).

2.2 PopII galaxies

To identify sources with a PopIII stellar population and understand differences between more chemically-enriched systems, we need to also model PopII galaxies. We adopt SEDs of BPASS (v2.2.1; Eldridge et al. 2017; Stanway & Eldridge 2018) for modelling the ISM ionised by such PopII galaxies. BPASS models might be more appropriate for moderate/lower metallicity systems and less appropriate for high metallicity systems. However, here we are exploring differences between the extremely low metallicity regimes (akin PopIII, near-pristine systems) and low/moderate metallicity (PopII) systems, so we believe that our choice of using the BPASS SEDs is appropriate. We use publicly available fiducial SEDs with binary evolution which are based on a Kroupa (2001) IMF with the two upper-mass cuts of 100 \( M_{\odot} \) and 300 \( M_{\odot} \). We adopt a stellar age of 10 Myr as fiducial models assuming continuous star formation to take into account contributions of hard ionising photons produced by massive Wolf-Rayet (WR) stars. In addition, the youngest stellar age of 1 Myr is also considered as commonly adopted to model optical emission lines in H\( ii \) regions (e.g., Kewley & Dopita 2002). We adopt two different stellar ages primarily to discuss the EWs of emission lines. Usually EWs present the largest values when the stellar population is youngest (e.g., Jaskot & Ravindranath 2016; Nakajima et al. 2018), but it would not always be true for the H\( ii \) emission due to the contributions of WR stars at \( \sim 4 \text{ Myr} \) after the onset of star formation. The shapes of ionising spectrum for the PopII models at 10 Myr as well as those at 1 Myr are shown in Fig. 1(b) with different stellar metallicities. For each metallicity, the SEDs at 1 and 10 Myr are almost identical at \( E < 50 \text{ eV} \), but an enhanced production of high energy ionising photons \( (E \gtrsim 50 \text{ eV}) \) for the 10 Myr model. The model calculations are terminated at the edge of the ionised cloud.

For each of the PopII models, we assume the stellar metallicity is matched to the gas-phase metallicity. This is a fair assumption as the enrichment of oxygen and other \( \alpha \) elements proceeds quickly, within less than 10 Myr (Maiolino & Mannucci 2019). Although such a quick metal enrichment results in \( \alpha \)-enhanced abundance patterns with little iron enrichment, and hence a lower opacity in young stars for their oxygen abundance, as observationally inferred in star-forming galaxies at \( z = 2–3 \) (Steidel et al. 2016; Cullen et al. 2019; Harikane et al. 2020), this would not significantly enhance the nebular emission strengths for the PopII models as compared to the PopIII model results. Anyway, we recall that the BPASS models adopted by us (and adequate for low metallicity systems) are already characterised by hotter atmospheres and a higher yield of ionising photons than classical, metal rich models. Moreover, we include dust physics in the PopII models as detailed in Nakajima et al. (2018), assuming the dust model and the depletion factors are unchanged but the dust abundance is scaled linearly with the gas-phase metallicity.

2.3 DCBHs

Our mock accreting DCBH incident radiation fields are generated by the CLOUDY “AGN” continuum command. The function primarily consists of two components. One is the so-called a Big Bump component, and another is a power-law component in the high energy range reaching the X-ray band. We have decided to adopt this continuum template as the spectrum resulting from the accretion on primeval black holes is expected not to be much different from ‘standard’ AGN, though possibly with a broader range of the two components, as detailed below (Valiante et al. 2018, and M. Volonteri, priv. comm.). The Big Bump feature is known to appear in the UV to optical region and would trace thermal emission from an optically thick accretion disk, which feeds a massive black hole. It is parameterised by the temperature of the bump, \( T_{\text{bb}} \), which is varied between \( 5 \times 10^{4} \text{ K} \) and \( 1 \times 10^{6} \text{ K} \), and \( 2 \times 10^{4} \text{ K} \). The choice is based on the maximum temperature in the innermost region of the disc, \( T_{\text{max}} \sim 5 \times 10^{5} \text{ K} \) for black holes with masses of \( 10^{5} - 10^{6} M_{\odot} \). Their hard ionising photons than classical, metal rich models. Moreover, we include dust physics in the PopII models as detailed in Nakajima et al. (2018), assuming the dust model and the depletion factors are unchanged but the dust abundance is scaled linearly with the gas-phase metallicity.

Because the temperature of the disc decreases from inside out such that \( R^{-0.75} \), we adopt a wide parameter range to cover an effective temperature that would be expected in such an accreting disk around black holes with \( M_{\text{BH}} \sim 10^{5} - 10^{6} M_{\odot} \). It is still unclear though whether the lowest temperature \( T_{\text{bb}} = 5 \times 10^{4} \text{ K} \) can be achieved in DCBHs/AGNs (see §3). We note that the black hole mass range explored in this paper covers fairly well the theoretically suggested mass range of intermediate-mass black hole seeds, \( 10^{4.75} - 10^{6.25} M_{\odot} \) (Ferrara et al. 2014). The 0.25 dex lower (higher) mass would increase (decrease) the maximum temperature by a factor of only 1.15 (as \( T_{\text{max}} = M_{\text{BH}}^{-0.25} \)), which does not significantly impact the resulting diagnostics for the DCBH models.

We adopt the slope of the X-ray component with the default value, but change the power-law energy slope between the optical and X-ray bands, \( \alpha_{X} \) (Zamorani et al. 1981). The parameter corresponds to the power-law index \( \alpha \), where \( f_{\nu} \propto \nu^{-\alpha} \), determined in the range of a few to a few thousand eV, and hereafter called \( \alpha \) for simplicity. We vary the value between \( \alpha = -1.2, -1.6, \) and \( -2.0 \) (see Elvis et al. 2002). The ionising spectrum shapes for the DCBH models are shown in Fig. 1(c).
The model calculations are stopped when reaching a neutral gas (Z = 0) at 0.028 = 2 Z⊙ (red), and larger symbol sizes for higher ionisation parameters (Left: log U = −0.5 to −2.0 for PopIII and PopII galaxies, −0.5 to −3.0 for DCBHs; Right: log U = −1.0 to −3.5 for evolved galaxies, −0.5 to −3.0 for AGNs). A fiducial gas density of 10³ cm⁻³ is assumed. The black dotted lines denote the demarcation (Eq. (1)) between PopIII and the other populations.

We consider cases where these DCBHs are surrounded by pristine gas (zero metallicity) as well as mildly metal-enhanced gas with Z = 10⁻⁵, 10⁻⁴, 10⁻³, and much higher metallicities up to Z = 0.028 = 2 Z⊙.¹ The model calculations are stopped when reaching a neutral column density of N(HI) = 10²¹ cm⁻², in accordance with the AGN models (Kewley et al. 2013).

We note that the direct emission of the accretion disc may be obscured along our line of sight by the presence of a dusty torus, i.e. a ‘type-2’ AGN configuration. While it is unlikely this happen in extremely metal poor systems, whose dust content is greatly reduced, we note this might be a caveat for the EW of the nebular lines, not their line ratios.

3 RESULTS

3.1 Optical Diagnostics

3.1.1 Diagnostics involving only He and H optical lines

The rest-frame optical wavelength range will be observable for early galaxies with JWST. This section presents the photoionisation model predictions for some diagrams using the optical emission line fluxes and EWs that will be useful to identify PopIII galaxies and DCBHs.

One of the key emission lines in the optical wavelength is Hei4686 which is sensitive to the production of high energy ionising photons (E > 54 eV) and thus proposed to be useful to identify PopIII galaxies. Particularly interesting is the Hei/Hβ ratio, as both of them are recombination lines with the same dependence on density and temperature, and their ratio is essentially traces the ratio of ionising photons at E > 54 eV and at E > 13.6 eV, hence the shape of the ionising spectrum, nearly independently of other quantities such as gas metallicity, ionisation parameter and density. The EW(Hei) has a similar diagnostic power, by giving an indication of the fraction of highly energetic photons relative to the optical, non-ionising photons, and hence a probe of the overall shape of the spectrum of the ionising sources.

Fig. 2 presents a diagram using the key Hei observable quantities: EW(Hei) vs. Hei/Hβ. The left panel show the primitive objects cases, i.e. zero or very low metallicity. Different colors show the metallicity level (with dark blue being zero metallicity). Solid symbols are PopIII with different IMFs. Empty symbols are PopII with different IMF and ages. Starred and cross symbols are for DCBHs. Symbols of increasing size correspond to larger ionisation parameters, both for ionisation by stellar populations and in the case of DCBHs. The legend gives details of the different symbols and colors coding. The most interesting feature of this figure is that PopIII models are remarkably and distinctly positioned as having the strongest Hei emission, both in terms of EW and line ratios with respect to Hβ, covering a completely different and disjoint region in the diagram as compared to the PopII models. The PopII models generally show a larger EW(Hei) at 1 Myr than 10 Myr but the difference becomes small at low-metallicity, possibly as a result of the the increase of continuum level and the increase of Hei emission at 10 Myr, partly cancelling out each other. The different positions between the PopIII and PopII models on the diagram are clearly seen irrespective of the

¹ A solar chemical composition of Asplund et al. (2009) is assumed.
choice of stellar age for the PopIII models. The DCBH models having the hardest ionising spectrum, with the largest power-law slope $\alpha$ and the highest $T_{\text{phot}}$, present HeI/$H\beta$ ratios as large as those found in the PopIII models. However, a clear difference is seen in terms of EW(HeI), where PopIII models show a larger EW than the DCBH models. Although PopIII galaxies with an older stellar population present a decreased EW than illustrated in Fig. 2, the decrease would be within a factor of $\sim 2$ if the stellar age is younger than a few Myr (e.g., Schäerer 2003), and hence the diagram will remain useful to identify the primitive PopIII objects. We propose a selection criteria for identifying PopIII galaxies:

$$\text{EW}(\text{He} \, \text{I}) > 20 \, \AA \quad \text{and} \quad \text{He} \, \text{I}/H\beta > 0.03$$

(1)

These limits are indicated by the dotted lines in Fig. 2. This diagram has some potential also in discriminating between different PopIII IMFs, with more top heavy IMFs being characterised by higher EW(HeI) and higher HeI/$H\beta$. Appendix §A shows the same diagrams for different gas density models and illustrating that results and selection boundaries do not change significantly.

The right panel of Fig. 2 presents the same diagram for more chemically-evolved galaxies and AGNs. Symbols are the same as in the left panel and the color-coding shows different metallicities also in this case, although the metallicity range is shifted to much higher values than in the left panel. This panel illustrates that there is no contamination in the region of the PopIII galaxies on the diagram particularly for those with a top-heavy IMFs. Clearly this diagram can separate PopIII from PopII and from DCBH and evolved AGNs. However, it cannot distinguish between pr imeval DCBHs (in a zero/low metallicity environment) from more evolved AGNs living in a more metal enriched environment; disentangling these two classes of objects will require the inclusion of metal lines, as we will discuss later on.

It should be noted that various studies have found evidence for significant HeI nebular emission in populations of metal poor, dwarf star forming galaxies (e.g. Schäerer et al. 2019; Umeda et al. 2022). The observed equivalent width is about 1 Å, similar to what expected by our models in low metallicity regimes (especially if considering high ionisation parameters, which are indeed typically observed in low metallicity star forming galaxies). However, the HeI/$H\beta$ ratio observed in such low metallicity galaxies is often significantly higher ($\sim 0.01$) than predicted by our models. This is an issue that is faced also by other photoionisation models using stellar population templates (Senchyna et al. 2017, 2020; Chevallard et al. 2018; Umeda et al. 2022). Possible explanations that have been proposed include the contribution of X-ray binaries (e.g. Schäerer et al. 2019) and ultraluminous X-ray sources (Umeda et al. 2022), although other studies have argued against these explanations (Senchyna et al. 2020). Either of these scenarios implies the presence of an additional ionising spectral component associated with accretion on compact objects and, therefore, they are essentially equivalent to adding an AGN-like component, as indeed expected by our models. It should also be noted that the presence of AGN in a sizeable number of dwarf galaxies has been found by recent observations (e.g. Mezcua et al. 2016; Mezcua & Domínguez Sánchez 2020; Schutte & Reines 2022) and also suggested by recent theoretical work (Pacucci et al. 2021), so the high HeI/$H\beta$ ratio can actually be simply explained by the presence of AGN in many of them. However, regardless of the explanation, it remains true that the combination of high EW(HeI) and high HeI/$H\beta$ unambiguously identifies either PopII or DCBH with a power law index $\alpha \gtrsim -1.6$.

We finally note that these are predictions for the HeI nebular lines. The stellar HeI emission associated with WR stars is not considered here. However, the latter, if present, should be clearly distinguishable from its broad profile (typically of the order of 1000 km s$^{-1}$ or broader).

Other important emission lines are those associated with HeI, which should be observable even from zero-metallicity objects and can be useful to probe the ionising spectrum shape at the intermediate energy regime ($E \sim 24$ eV). Since these are primarily recombination lines, together with the HeI and hydrogen recombination lines, they would be in principle useful for determining the curvature of the ionising spectrum. However, the situation is not so simple in this case. Indeed, the relatively strong HeI recombination lines result from the cascades of HeI triplets, whose population/depopulation is affected by the collision effects (i.e., density and temperature) and the opacity properties of ISM (e.g., Izotov et al. 2014; Aver et al. 2015; Matsumoto et al. 2022). Yet, exploring the relative intensity of these lines is anyway useful.

Among several HeI recombination emission lines in the optical wavelength, we firstly test with HeI$\lambda$5876 as it is generally the strongest HeI line. We find, however, that the PopIII and DCBH models show a comparable HeI$\lambda$5876/$H\beta$ ratio. Rather, the line ratio depends more on the gas density in the ISM. Fig. 3 shows a diagram of HeI$\lambda$5876/$H\beta$ and HeI$\lambda$/$H\beta$ for the PopIII and PopII galaxies, varying the gas density from $10^2$ cm$^{-3}$ to $10^4$ cm$^{-3}$ in addition to the fiducial value of $10^3$ cm$^{-3}$. Clearly the HeI$\lambda$5876/$H\beta$ ratio increases with gas density (see also Aver et al. 2015). We therefore propose that this HeI$\lambda$5876 diagram can be used to probe the gas density for metal-deficient galaxies and even for DCBHs, such that

$$\log \text{He} \, \text{I}/H\beta = 0.07/(\log \text{He} \, \text{I}/H\beta + 0.98) - 0.5 \quad (n_0 = 10^2), \quad (2)$$

$$= 0.07/(\log \text{He} \, \text{I}/H\beta + 0.83) - 0.5 \quad (n_0 = 10^3), \quad (3)$$

$$= 0.07/(\log \text{He} \, \text{I}/H\beta + 0.73) - 0.5 \quad (n_0 = 10^4). \quad (4)$$

This diagram can substitute the density indicators using the metal collisional excitation lines such as [O I] $\lambda$13726/3729 and [S II] $\lambda$6717/6731 which will probably not be available for primitive, extremely metal-poor sources. A caveat is that this diagram
The diagrams discussed above, which use only the Hydrogen and Helium recombination lines, are useful to identify PopIII galaxies and DCBHs, but not sufficient to assess how much the surrounding ISM gas is chemically-enriched. Also, as already mentioned, DCBHs and more evolved AGNs, living in enriched ISM, would not be separable. To discriminate PopIII galaxies and DCBHs in a primitive environment from those embedded in a weakly, or even strongly, enriched ISM, optical metal lines are also needed. Fig. 5 presents three diagrams showing the line ratios of [O ii],[O iii],[S ii] vs. [N ii] for various models. The black dotted lines can be useful to disentangle some types of DCBHs.

\[ \text{[O ii]} /[\text{H} \beta] > 0.01 \text{ & } \text{[N ii]} /[\text{H} \beta] > 0.04. \] (5)

are most likely dominated by non-thermal hard ionising spectrum associated with accreting black holes. The separation is certainly less clear than in the diagram using the EW(Heii) (Fig. 2), but the diagram in Fig. 4, which uses only line ratios, is not affected by the potential issue of nuclear obscuring torus that might result into an artificially higher EW of the lines.

Once again, more evolved AGNs can also fall on the same region on this diagram (as shown in the right panel) and it is hard to assess the metal-enrichment only based on the Heii diagram. One would need other diagrams using the metal lines as discussed below.

### 3.1.2 Diagnostics involving metal optical lines

The diagrams discussed above, which use only the Hydrogen and Helium recombination lines, are useful to identify PopIII galaxies and DCBHs. However, they are not sufficient to assess how much the surrounding ISM gas is chemically-enriched. Also, as already mentioned, DCBHs and more evolved AGNs, living in enriched ISM, would not be separable. To discriminate PopIII galaxies and DCBHs in a primitive environment, metal optical lines are also needed. Fig. 5 presents three diagrams showing the line ratios of [O ii],[O iii],[S ii] vs. [N ii],[O ii],[N ii] for various models.

**Figure 4.** Distribution of the Cloudy models on the Heii4686/H\(\beta\) versus Heii4471/H\(\beta\) diagram, for primitive sources (Left) and for chemically-evolved systems (Right). Symbols as in Fig. 2. The black dotted lines can be useful to disentangle some types of DCBHs (Eq. 5).

The former two are, both widely used to separate AGNs from star-forming galaxies (e.g., Kewley et al. 2001; Kauffmann et al. 2003; Kewley et al. 2013). The left panel show the case for primeval, extremely metal-poor cases, the central panels show the case of evolved (significantly metal enriched) PopII galaxies, and the right panels show the case of chemically evolved AGNs.

In the central panels of the BPT-diagrams, we confirm that evolved star-forming galaxies follow or fall below the demarcation curves. We note that the BPT diagnostic line ratios are not significantly influenced by the choice of stellar age, and mostly governed by the youngest stellar population (see also Fig. 1b for the almost identical SEDs at 1 and 10 Myr for a given metallicity at \(E < 50\) eV). On the other hand as evidenced in the right panels, not all of the evolved AGN models are separable on the BPT-diagrams. AGNs with a gas-phase metallicity below \(Z < 0.5 Z_\odot\) can spill (and hence ‘contaminate’) into the star-forming galaxy region of the diagrams, as already indicated by Kewley et al. (2013). Moreover, the models with the least hard spectrum (\(\alpha = -2.0\)), if they really exist, could not be classified as AGNs with the classical BPT diagnostics even with a super-solar metallicity gas. The diagram in the 3rd row makes use of the [O i] and H\(\beta\) emission in place of [N ii], [S ii], and H\(\alpha\). This will become particularly useful when the wavelengths around H\(\alpha\) is not observationally available, e.g., for sources at \(z > 7\) with JWST/NIRSpec.

We note the evolved star-forming galaxy models in the middle panels of the [N ii] BPT-diagram shape a model grid consistent with Kewley et al. (2013) and other models results (e.g. Gutkin et al. 2016), and can reproduce the observed metallicity sequence of star-forming galaxies on the diagrams in the local universe (e.g., Maiolino et al. 2008; Curti et al. 2017, 2021), following the trend that the ionisation parameter typically gets larger in a more metal-deficient galaxy (e.g., Andrews & Martini 2013; Sanders et al. 2020).

\[ [\text{O ii}]_3727 \text{ denotes the sum of the doublet: } [\text{O ii}]_3726, 3729. \]
Figure 5. Distribution of the CLOUDY models on the optical metal-line diagnostics ([O iii] 5007/Hβ versus [N ii] 6584/Hα (1st row), [O iii] 5007/Hβ versus [S ii] 4661, 6731/Hα (second row), and [O iii] 4363/Hβ versus [O ii] 3727/Hβ (3rd row), for the primitive sources (Left) and for the chemically-evolved systems of galaxies (Middle) and AGNs (Right). Symbols as in Fig. 2. In the extremely low-metallicity regime (Z < 10^{-3}; left panels), all models follow a sequence for a given gas-phase metallicity irrespective of the different ionising sources. The dotted curves show the average sequences for the models with Z = 10^{-5} (Eqs. 6–8) and 10^{-4} (Eqs. 9–11). The conventional demarcation curves between AGNs and galaxies on the BPT-diagrams are depicted as introduced by Kewley et al. (2001) (dot-long dashed; “Kew01”) and Kauffmann et al. (2003) (short dash-long dashed; “Kau03”). The demarcation curve of Lamareille (2010) is plotted on the bottom diagram using [O iii] (dot-long dashed; “Lam10”). Clearly, these curves work for isolating AGNs if the metallicity is relatively high (Z ≥ 0.5 Z⊙) and the spectrum is hard (α ≥ −1.6).

Nakajima et al. 2021 submitted). A similar but offsetted sequence toward a larger [O iii]/Hβ ratio is reported for z = 2 − 3 galaxies in the [N ii] BPT-diagram (e.g., Steidel et al. 2014; Shapley et al. 2015; Strom et al. 2017). The z = 2 − 3 sequence is also compatible with our model grid, although the physical origins of the evolution remain under debate (e.g., Curti et al. 2021).

We also note that the sequence for evolved AGN found by us is also broadly consistent with other AGN photoionisation models (e.g. Groves et al. 2006; Feltre et al. 2016).

In the extremely low-metallicity regime in the left panels, all the models of PopIII galaxies, PopII galaxies, and DCBHs follow a similar sequence for a given gas-phase metallicity with the ionisation condition varied. Although these optical metal-line diagrams have, therefore, no power to distinguish between the objects powered by different ionising sources in such a low metallicity regime below
Z ≲ 10⁻³, they are very useful to infer the gas-phase metallicity, in a way similar to the strong line metallicity diagnostics (e.g., Maiolino & Mannucci 2019; Nakajima et al. 2021 submitted), but independently of the source of ionisation, and also allow the identification of sources that are not chemically-evolved at all.

Using the low-metallicity models found in the left panels, we define the average relationships as follows:

\[
\log \frac{[O \text{iii}]}{H\beta} = 0.61 / \log \frac{[N \text{ii}]}{H\alpha + 2.7} - 0.8 \\
= 0.72 / \log \frac{[S \text{ii}]}{H\alpha + 2.4} - 0.8 \\
= 0.65 / \log \frac{[O \text{ii}]}{H\beta + 1.2} - 0.8
\]

for a metallicity of Z = 10⁻⁵, and

\[
\log \frac{[O \text{iii}]}{H\beta} = 0.61 / \log \frac{[N \text{ii}]}{H\alpha + 1.7} + 0.2 \\
= 0.72 / \log \frac{[S \text{ii}]}{H\alpha + 1.4} + 0.2 \\
= 0.65 / \log \frac{[O \text{ii}]}{H\beta + 0.2} + 0.2
\]

for Z = 10⁻⁴.

Very interestingly, we note that the [O iii] λ5007 line remains relatively strong (at the level of one tenth of Hβ; see also Inoue 2011) even at extremely low metallicities, down to Z = 10⁻⁵, probably as a consequence of the fact that even a small amount of oxygen rapidly makes its collisionally excited transitions good coolant of the ISM. Objects that fall below the Z = 10⁻⁵ curves are good candidates of PopIII galaxies or DCBHs just formed in a totally primitive environment. As mentioned, even galaxies at Z = 10⁻⁵ or Z = 10⁻⁴ can actually be hosting PopIII galaxies or DCBHs embedded in slightly enriched ISM. The EW(Hεii) diagram (Fig.2) and the Hei diagram (Fig.4) will then be powerful to distinguish between the two ionising spectral shapes, i.e., PopIII stars vs. DCBHs.

3.2 UV Diagnostics

3.2.1 Diagnostics using only He and H UV lines

In this section we present the photoionisation model results for the UV emission lines. A major advantage of using the UV emission lines is that they are available for high-redshift galaxies even with the ground-based NIR instruments.

Fig. 6 focuses on the UV Hεii emission at 1640 Å, showing EW(Hεii) as a function of Hεii/Lyα. One caveat is that the Lyα strengths used here are simply predicted by our Cloudy models truncated at the edge of the ionised nebula (§2), and thus supposed to be the maximum value emitted from the cloud. Neutral hydrogen gas and dust in the systems would weaken the Lyα emission. Indeed, the dust-included PopIII models supplementarily plotted in Fig. 6 (hollow triangles, squares and pentagons) show a weakened Lyα and thus a higher Hεii/Lyα ratio, even higher by a factor of ~10 for the Z = 10⁻⁴ models. EW(Hεii) would also become enhanced if dust is included due to a weaker continuum level, but the effect is expected to be minor, by a factor of <2, as this is a secondary, differential effect. Lyα would also be absorbed by the inter-galactic medium (IGM) at high redshift, in particular for sources at, or beyond, the reionisation epoch (z > 6). The Hεii/Lyα ratios presented in Fig. 6 are therefore regarded as the lower-limits.

Using the predicted emission line strengths for the populations on this diagram, we propose the following selection criteria for selecting PopIII galaxies:

\[
\text{EW(He}\,$\,$\,$\,$\text{ii}|1640) > 40$ Å $\text{& He}\,$\,$\,$\text{ii}|1640/Lyα > 0.01
\]

which are traced by thick dotted lines in Fig. 6. These are based on the top-heavy IMF models reaching the maximum mass-cut of 500 $M_\odot$, which are located in a distinct region on this diagram. PopIII galaxies following such a top-heavy IMF, including the extreme one containing only massive stars of $50 - 500 M_\odot$, are separable from the other population on this UV Heii diagram. The PopIII models with the moderate (1 – 100 $M_\odot$, Salpeter) IMF, on the other hand,
present a maximum EW(Hei) of ~ 20 – 30 Å, which is comparable to the predicted EWs for the hardest DCBH/AGN models, and hence indistinguishable from each other on this diagram. The criteria of Eq. (12) is thus optimised to identify PopIII galaxies and minimise contaminations by DCBHs. In order to build a more complete sample of PopIII galaxies with a certain contamination by DCBHs allowed, the following, less strict criteria can also be useful:

$$\text{EW(He} \text{\textsc{i}1640)} > 20 \text{ Å} \ & \ \text{He} \text{\textsc{i}1640}/\text{Ly} \alpha > 0.006.$$  (13)

The choice of stellar age for the PopII models does not affect the PopIII selection criterion on the diagram, as it was the case for the optical EW(Hei) diagnostic (Fig. 2). In short, the UV Hei emission provides a good diagnostic to distinguish between PopIII galaxies and DCBHs, although the optical Hei has a better and more sharp discriminatory power.

### 3.2.2 Diagnostics using UV metal lines

To diagnose the chemical enrichment of the ISM (even in the case of mildly enriched gas photoionised by PopIII stars) using the UV emission lines alone, we propose to combine Hei\textsc{i}1640 with the UV metal lines of [C\textsc{ii}]+1907, [C\textsc{ii}]+1909 (hereafter their sum is referred to as C\textsc{ii}1909, or simply C\textsc{ii}) and C\textsc{iv}l1548, 1550 (hereafter C\textsc{iv}1549, or C\textsc{iv}). Fig. 7 displays the diagram correlating the ([C\textsc{ii}]+C\textsc{iv})/Hei ratio and the C\textsc{iv}/C\textsc{ii} ratio, each of which is named C34 and C43, respectively. In the left panel, we only plot PopIII galaxies and the DCBH models with the hardest ionising spectrum ($\alpha = -1.2$) that are to be selected with the above EW(Hei\textsc{i}1640) diagnostic (Eqs. 12 or 13). The C34-index is informative to probe the gas metallicity, although the secondary dependence on the ionisation condition is not negligible. The C43-index on the ordinate axis is then helpful to resolve the degeneracy and allows this diagram to infer the gas metallicity together with the ionisation parameter (see also, e.g., Nakajima et al. 2018).

The diagram indeed shows the power to diagnose the gas metallicity. Following the distributions of the models on the left diagram, we define two metallicity sequences for these primitive sources: Group A is for the PopIII galaxies with top-heavy IMF ($M_{\text{up}} = 500 M_\odot$), (pentagons and squares) and DCBHs with $T_{\text{bb}} = 5e4$ K (star symbols connected with a long-dashed line), and another group (Group B) for the other models. Each group shapes a relatively tight sequence on this diagram for a given metallicity, with Group A showing a systematically stronger Hei and/or a higher C\textsc{iv}/C\textsc{ii} ratio than Group B. The relationships are formulated as follows:

$$\log \text{C4C3} = 1.2 \log \text{C34} + 2.6 \ (A)$$  (14)

$$= 1.2 \log \text{C34} + 2.1 \ (B)$$  (15)

to trace sources with $Z = 10^{-5}$, and

$$\log \text{C4C3} = 1.2 \log \text{C34} + 1.5 \ (A)$$  (16)

$$= 1.2 \log \text{C34} + 1.0 \ (B)$$  (17)

for $Z = 10^{-4}$.

In the right panel of the same figure, metal-rich models of AGNs ($\alpha = -1.2$) and PopII galaxies are shown on the diagram. They present a larger C34 value, as expected, than the primitive sources for a given C43, and well apart from the $Z = 10^{-5}$ and $10^{-4}$ sequences above. The relationships thus provide a good measure of gas metallicity in such a low metallicity regime if all the necessary emission lines are available. Particularly, PopIII galaxies with top-heavy IMF would be robustly selected following the strict selection criterion (Eq. 12) and whose metallicity be inferred following the above relationships of Group A.

It is also expected that not all of the necessary lines are always available from sources in the early universe. To provide a substitute to probe the chemical enrichment of the ISM in an efficient way with limited spectroscopic measurements, we summarise in

---

**Figure 7.** Cloudy models of ([C\textsc{ii}]+C\textsc{iv})/Hei vs. C\textsc{iv}/C\textsc{ii} for the PopIII galaxies and DCBHs ($Z \leq 10^{-3}$) with the hardest ionising spectrum ($\alpha = -1.2$) that are supposed to be selected with the EW(Hei\textsc{i}1640) diagnostic (Left). Symbols as in Fig. 2. PopIII models with top-heavy IMF ($M_{\text{up}} = 500 M_\odot$) and DCBH models with $T_{\text{bb}} = 5 \times 10^4$ K (Group A) follow a well defined sequence for a given metallicity. The other models (Group B) fall on another metallicity sequence which is offset toward a weaker Hei and/or a lower C\textsc{iv}/C\textsc{ii} ratio. The sequences of $Z = 10^{-5}$ and $10^{-4}$ for both groups are illustrated with black dotted lines (Eqs. 14–15 and Eqs. 16–17, respectively). In the right panel, metal-rich models for AGNs ($\alpha = -1.2$) and PopII galaxies are shown on the same diagram, and are confirmed not to contaminate the low-metallicity sequences.
The considered metal lines include \( \text{O} \text{m} 1, \text{l} 1661, 1666 \) (hereafter \( \text{O} \text{m} \)) and \( \text{N} \text{v} \text{v} \text{v} 1239, 1243 \) (hereafter \( \text{N} \text{v} \text{v} \text{v} \) or \( \text{N} \text{v} \)). For each metal line, the primitive sources are displayed on the left-hand side, and the chemically-evolved systems on the right. Symbols as in Fig. 2. The dotted lines show the metallicity demarcations for PopIII galaxies (with the two EW(\( \text{H} \text{e} \text{v} \text{v} \text{v} \text{v} \text{v} \)) criteria; Eqs. 12–13) for \( Z = 10^{-5} \) (Eqs. 18–21) and \( = 10^{-4} \) (Eqs. 22–25).

Finally, we note that the C4C3-C34 diagram (Fig. 7) in Nakajima et al. (2018) is partly used to separate AGNs from star-forming galaxies. We reaffirm that the separation method works for galaxies with a gas metallicity higher than \( Z \geq 0.01 \) \( Z \text{☉} \), and that sources dominated by PopIII stars and AGNs are not distinguishable, as cautioned by Nakajima et al. (2018) and illustrated in Fig. 7 (left).

In Fig. 9, we also examine the other popular UV diagnostics combining \( \text{C} \text{m} 1 \), \( \text{O} \text{m} 1 \), \( \text{N} \text{v} \), and \( \text{He} \text{v} \) (Feltre et al. 2016; Gutkin et al. 2016; Laporte et al. 2017; Hirschmann et al. 2019; Saxena et al. 2020). As seen on the the C4C3-C34 diagram, the PopIII models are not well discriminated from other classes of metal-poor DCBHs and/or AGNs, and, more generally, discriminating different classes of targets and metallicities is difficult on these UV diagrams. Yet, it is confirmed that the \( \text{N} \text{v}/\text{He} \text{v} \text{v} \text{v} \text{v} \text{v} \text{v} \) ratio (right panels in Fig. 9) is powerful in discriminating AGN and PopII galaxies in the metal rich regime.

As an additional caveat of these diagrams, one should take into account that they depend on the assumptions of metal abundance ratios such as \( \text{C} \text{O}/\text{N} \text{O} \) and \( \text{C} \text{O}/\text{N} \text{O} \). Although our adopted prescriptions of the metal abundances are known to be generally valid for local \( \text{H} \text{ii} \) regions and \( z = 2–3 \) star-forming galaxies on average (Hayden-Pawson et al. 2021), they are subject to significant dispersion, which could be increased by some systematic uncertainties at high-redshift. It would be recommended to directly constrain the metal abundance ratios using appropriate sets of emission lines if all the necessary lines are available (e.g., Pérez-Montero & Amarín 2017; Hayden-Pawson et al. 2021) to correctly interpret observations. Yet, our assumptions provide a reasonable first step to explore the behaviours of these nebular emission lines, offering the diagnostics which will be useful even when limited sets of emission lines are observationally available.
4 DISCUSSION AND SUMMARY

We have presented CLOUDY photoionisation models predictions of the rest-frame optical and UV emission line strengths for PopIII galaxies and DCBHs, and we have proposed several diagnostic diagrams to select and distinguish between these early sources and more-evolved systems. The modelling reveals that it is possible to separate PopIII galaxies from other kind of primeval object, and in particular primeval black hole seeds (Direct Collapse Black Holes, DCBHs) by utilising the He\textsc{i} emission lines, both in the optical and UV, in conjunction with the other lines probing a wider range of ionising spectral shapes. Moreover, metal line diagrams allow the characterisation of the ISM properties, including gas metallicity and density in extreme conditions, including the exploration of scenarios in which PopIII or DCBH are embedded in slightly enriched ISM, and paving the way for an exploration of primitive sources (Z \(\lesssim\) 10\(^{-3}\)) in the early universe.

Among the most important diagnostics that we have identified, and interesting findings, we summarise the following:

- The EW of the He\textsc{i} optical line at 4686Å is the best discriminator of PopIII galaxies. A threshold of EW(He\textsc{i}4686) >20Å discriminates well PopIII from and from DCBH. Along with the He\textsc{i}/H\textsc{b} ratio, the EW(He\textsc{i}4686) can also discriminate among PopIII with different IMFs. The EW of the He\textsc{i} UV transition at 1640Å has a similar good diagnostic power, although not as sharp as the optical line.

- DCBHs can be distinguished well from PopIII and PopII on the EW(He\textsc{i}4686) vs. He\textsc{i}/H\textsc{b} diagram. The equivalent diagram in the UV, i.e. EW(He\textsc{i}1640) vs. He\textsc{i}1640 /Ly\text{x}a diagram, unfortunately does not have the same discriminatory power, although still useful to discriminate PopIII with top-heavy IMF.

- Very low metallicity systems, including PopIII and DCBH embedded in low metallicity ISM (Z \(\sim\) 10\(^{-5}\) – 10\(^{-3}\)), can be distinguished well from more evolved, moderate/high metallicity systems (PopII and AGN in metal-rich environments) through the optical diagrams [O\textsc{iii}]/H\text{b} vs. [N\textsc{ii}]/H\text{a}, or vs. [S\textsc{ii}]/H\text{z}, or vs. [O\textsc{ii}]/H\text{b}. In particular, candidate DCBH in a primeval ISM, identified through the previously suggested diagrams, can be distinguished well from more evolved AGN through the optical BPT diagrams. The trends of the UV lines are more complex, but we have identified tight sequences in the C\textsc{iv}1549 /C\textsc{iii}1909 vs. (C\textsc{iv}1549+O\textsc{iii}15007)/He\textsc{i}1640 diagram that can diagnose the ISM metallicities embedding PopIII and DCBHs.

- Interestingly, the [O\textsc{iii}]15007 optical transition remains strong even in very low metallicity regimes, by being as strong as H\text{b} at Z \(\sim\) 10\(^{-4}\) and only \(\sim\) 10 times fainter than H\text{b} at Z \(\sim\) 10\(^{-5}\). Similarly, the C\textsc{iv}1549 UV transition can be as strong as He\textsc{i}1640 even at metallicities below Z \(\sim\) 10\(^{-4}\), and, potentially, only \(\sim\)10 times fainter than He\textsc{i} at Z \(\sim\) 10\(^{-5}\). Therefore, galaxies with [O\textsc{iii}]15007/H\text{b} \(\lesssim\) 0.1 and C\textsc{iv}/He\textsc{i}1640 \(\lesssim\) 0.1 are good candidates for hosting PopIII or DCBH in a pristine or slightly enriched environment.

- Finally, the He\textsc{i}15876/H\text{b} ratio, together with the He\textsc{i}4686/H\text{b} ratio, is a good diagnostic of the gas electron density in the very low metallicity regime (where other density diagnostics associated with metal lines are not detectable).

While the focus of this work is to provide diagnostics for data that will be obtained in the early universe by forthcoming facilities (e.g. JWST and ELT), some of the currently available observations of local and high redshift galaxies can already be interpreted in the context of these diagnostics. Such a comparison will be presented in a forthcoming, parallel paper.

ACKNOWLEDGEMENTS

We thank the anonymous referee for helpful comments and discussions that improved our manuscript. KN acknowledges support from JSPS KAKENHI Grant JP20K25373. RM acknowledges ERC Advanced Grant 695671 QUENCH, and support from the UK Science and Technology Facilities Council (STFC). RM also acknowledges funding from a research professorship from the Royal Society.

DATA AVAILABILITY

The data presented in this article will be shared on reasonable request to the corresponding author.

REFERENCES

Abel T., Bryan G. L., Norman M. L., 2002, Science, 295, 93
Amorín R., et al., 2017, Nature Astronomy, 1, 0052
Andrews B. H., Martini P., 2013, ApJ, 765, 140
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Aver E., Olive K. A., Skillman E. D., 2015, J. Cosmology Astropart. Phys., 2015, 011
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Barrow K. S. S., Aykutalp A., Wise J. H., 2018, Nature Astronomy, 2, 987
Beckmann R. S., et al., 2019, A&A, 631, A60
Bowler R. A. A., McLure R. J., Dunlop J. S., McLeod D. J., Stanway E. R., Eldridge J. J., Jarvis M. J., 2017, MNRAS, 469, 448
In this appendix, we present the results with different gas densities for PopIII galaxies.

Fig. A1 presents the optical He\textit{\textsc{ii}} diagnostic for a gas density of 10^2 (left panels) and 10^4 cm^{-3} (right panels). This confirms that the recombination lines ratio He\textit{\textsc{ii}}/H\alpha is almost unchanged for any of the populations over the reasonable range of gas density from 10^2 to 10^4 cm^{-3}. EW(He\textit{\textsc{ii}}) shows a weak dependence on density, getting stronger by a factor of ~ 1.2 if a gas density higher by a factor of 10 is assumed. This dependence does not change the selection criteria for PopIII galaxies on this diagram (Eq. 1). We therefore reaffirm that the optical He\textit{\textsc{ii}} diagram is the best discriminator of PopIII galaxies.

Fig. A2 shows the UV He\textit{\textsc{i}} diagnostics for different gas densities.
A higher gas density leads to a stronger EW(\HeI/\LyA), by a factor of \approx 1.3 if assuming a gas density ten times higher, in a similar way as seen for the optical \HeII emission. The strict selection criteria (Eq. 12) of EW(\HeII/\LyA) > 40 Å is confirmed valid even for the low gas density case. Although some metal-rich AGNs with a hard ionising spectrum ($\alpha = -1.2$) and a high density ($\approx 10^4$ cm$^{-3}$) can present an EW(\HeII/\LyA) as strong as \approx 40 Å, such objects can be well-discriminated based on the diagnostics using the UV metal lines (§3.2.2). On the other hand, the relaxed selection criteria (Eq. 13) of EW(\HeII/\LyA) > 20 Å may miss some PopIII galaxies with a modest IMF if the gas density is as low as $\approx 10^2$ cm$^{-3}$. This brings us another caveat when using the UV \HeII diagnostic. The conservative lower-limits on \HeII/\LyA ratio remain valid irrespective of the gas density.

In Fig. A3, we show some key diagnostic emission line ratios using metal lines to assess their dependences on gas density. We only present the low-metallicity models for the [N\textsc{ii}] BPT diagram and C4C3-C34 diagram as representatives. The dotted curves tracing the $Z = 10^{-5}$ and $10^{-4}$ are as determined based on the fiducial gas density models (Eqs. 6–9, 14–17). Fig. A3 demonstrates that these dotted curves work well for the different gas density models, confirming that the gas density has little effect on the diagnostic emission line ratios, and hence on the metallicity estimations based on the diagrams. Although these metal lines are collisional excitation lines, the effect is minimal at the low metallicity regime, especially when we use line ratios.

We also emphasise that our primary purpose here is to explore variations of metallicity spanning orders of magnitude; compared with the effects of such large metallicity variations, the effects of gas density variations on the diagnostics is essentially unimportant (with the exception of the \HeI lines, as shown in Fig. 3). We can therefore conclude that the diagnostics developed in the main text work for different gas densities, and hence for a system with a possible internal variation of gas density, within the reasonable range of $10^2 - 10^4$ cm$^{-3}$, are solid for constraining (when possible) the nature of different classes of sources and their metallicities.