

Probing down to early cosmic epochs using limited redshift ($z \lesssim 1$) optical surveys with host galaxy age and lookback time analysis

S. Kasthurirangan,^{1*} M. Panchal¹, A. Joshi¹, G. Pawar²

¹*Department of Physics, University of Mumbai, Vidyanagari, Mumbai - 400098, India*

²*Nicolaus Copernicus Astronomical Center, Polish Academy of Sciences, ul. Rabinowska 8, 87-100 Toruń, Poland*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We present a detailed analysis of AGN identification diagnostics and host galaxy evolution using optical spectral diagnostics using ESO-GOODS-S data. We employ traditional Baldwin-Phillips-Terlevich (BPT) diagrams along with their modern extensions, the Mass-Excitation (MEx) and Colour-Excitation (CEx) diagrams, to classify AGNs from among up to 600+ candidates. We extract the spectral properties utilising an indigenous spectral fitting code. In addition, the code also incorporates inputs from a state-of-the-art stellar population synthesis model (Bruzual and Charlot, 2003, MNRAS, 344, 1000), extracting the host galaxy ages and metallicities as model parameters. Redshift values for all the studied spectra are obtained from the literature. Using these redshift values, we are able to assign a ‘lookback time’ (using standard Λ CDM cosmological parameters). When this lookback time is combined with the galaxy ages (from spectral fitting), it enables us to study galaxy evolution even up to early cosmological epochs (~ 13 Gyr), despite restricting the studied field of galaxies to $z \sim 1$ (lookback time of ~ 8 Gyr). We conclude that galaxies ultimately evolving into AGN hosts originate earlier (~ 8 Gyr ago) than non-AGN host galaxies (~ 6 Gyr ago). We relate this fact, that AGN hosts date back to an earlier epoch than non-AGN galaxies, as being likely due to higher cosmic temperatures at that time, probably enhancing black hole accretion. We emphasise that host galaxy age is an additional crucial parameter (apart from cosmological redshift) in defining our understanding of AGN and star formation co-evolution across cosmological time scales.

Key words: galaxies: active – galaxies: evolution – galaxies: formation – galaxies: nuclei – techniques: spectroscopic – methods: data analysis

1 INTRODUCTION

Active Galactic Nuclei (AGNs) are some of the most energetic and luminous systems in the universe, driven by the process of accretion onto supermassive black holes (SMBH) at the centers of galaxies (Rees 1984; Urry & Padovani 1995; Krolik 1999). The identification and characterisation of AGNs presents significant challenges, because no single observational indicator can fully capture the diverse and complex nature of black hole accretion and the accompanying wide-spectrum emission (Antonucci 1993a; Padovani et al. 2017). AGNs manifest with a wide range of observable signatures, depending on factors such as orientation, obscuration, and the surrounding environment (Urry & Padovani 1995; Netzer 2015; Hopkins et al. 2006). Many AGNs exhibit varying degrees of activity across different wavelengths, and their evolutionary processes can be influenced by multiple factors, including interactions with their host galaxies, cosmic environment, and merger history (Hopkins et al. 2006; Alexander & Hickox 2012; Kalfountzou et al. 2017).

Studies in different spectral wavebands probe distinct physical processes that contribute to the AGN phenomenon. For instance, optical emission lines reveal the ionized gas in the narrow- and broad-line

regions (Baldwin et al. 1981), while X-ray and UV wavelengths trace the highly energetic radiation from the innermost regions near the black hole (Silverman et al. 2005). Infrared observations detect reprocessed radiation from dust heated by the AGN, especially in obscured systems (Stern et al. 2005). Meanwhile, radio observations can detect relativistic jets and synchrotron emission, providing insights into the AGN’s interaction with its host galaxy and surrounding medium (Begelman et al. 1984).

Optical and infrared diagnostics have been essential in differentiating AGNs from star-forming regions in galaxies. One of the most widely used methods is the classical Baldwin-Phillips-Terlevich (BPT) diagram (Baldwin et al. 1981), which distinguishes AGNs from star-forming regions based on emission line ratios. This method is especially suited for dealing with wide field survey data, where the emphasis is on statistical trends rather than specific physical properties of the individual objects or regions under consideration. This diagram compares specific optical line ratios, such as $[\text{O III}]\lambda 5007/\text{H}\beta$ against $[\text{N II}]\lambda 6584/\text{H}\alpha$, to reveal the ionisation mechanisms at play within a galaxy. AGNs, which generate harder ionising radiation than star-forming regions (Osterbrock & Pogge 1985), occupy a distinct region of the BPT diagram, providing a seemingly clear diagnostic to identify their presence. More recent updates to the BPT method (Kewley et al. 2006; Stasińska et al. 2006) further refine these diag-

* E-mail: s.kasthurirangan@physics.mu.ac.in (SK)

nostics, extending their applicability across a wider range of object types (e.g. Low Ionisation Nuclear Emission Line Regions (LINERs) vs Seyfert galaxies etc.). Studying the possibility of modifying this classification scheme across different redshift regimes has also been discussed in the literature (Kewley et al. 2013; Trump et al. 2012; Coil et al. 2015).

Beyond the BPT diagram, newer methods like the Mass-Excitation (MEx) and Color-Excitation (CEX) diagrams offer complementary approaches to differentiate AGNs from star-forming regions, particularly at low-to-intermediate redshifts ($z \lesssim 1$). The MEx diagram, introduced by Juneau et al. (2011), incorporates the stellar mass of the galaxy as the ordinate, while retaining the $[\text{O III}]\lambda 5007/\text{H}\beta$ line ratio as the abscissa. This enables the MEx diagram to capitalise on the observed correlation between mass and ionisation properties. Similarly, Yan et al. (2011) studies the same line ratio against the (U-B) color index of the galaxy. This approach is especially effective in identifying AGNs in dusty, obscured environments where optical diagnostics alone may fall short.

One of the aims of this paper is to present a detailed study of the Great Observatories Origins Deep Survey South (GOODS-S) field data using spectral analysis to better understand the properties of galaxies, with a particular focus on AGNs. In order to do that, we use an indigenously developed spectral code (Panchal et al. 2024), which addresses key spectral features, including the host galaxy contribution, AGN emission continuum, and broad and narrow spectral lines, etc. Using the spectral fitting parameters obtained from the code, we construct various AGN diagnostic diagrams, like traditional BPT diagrams as well as the modified ones, like the MEx and CEX. A second major aim of this paper is to introduce one of the key features of the code, which is the integration of template spectra from the Bruzual & Charlot (2003) stellar population synthesis model (hereafter, BC03) into our fitting process, enabling us to provide estimates of the ages of the host galaxies, as model fitting parameters. Introducing host galaxy age as an independent parameter takes the AGN diagnostics beyond its traditional paradigm, making it possible to study the temporal evolution of the sources over different cosmological epochs. Such a holistic approach could eventually lead to a better understanding of the role of AGNs in the evolution of the universe. Such approaches are seen in the existent literature (Kewley et al. 2013; Goulding et al. 2014), but none of these take into account the age of the host galaxy as an independent parameter. This is a unique feature of this study.

The paper is organised as follows. In section 2, we describe the GOODS-S data used in our analysis. We also detail the selection criteria in terms of the availability of redshift data, and also in terms of redshift limits of the method used. In section 3, we outline the spectral fitting process, including the integration of template spectra from the BC03 stellar population synthesis models into the same. The key flux ratios which form the basis for constructing traditional BPT diagnostic diagrams (e.g., $[\text{O III}]/\text{H}\beta$, $[\text{N II}]/\text{H}\alpha$, $[\text{S II}]/\text{H}\alpha$), are obtained from the fitting parameters. In section 4, we present the results, including the classification of sources based on the traditional BPT diagrams ($[\text{N II}]$, $[\text{S II}]$, $[\text{O II}]$; see subsection 4.1) and more recent MEx and CEX diagrams (see subsection 4.2). Further, the galaxy ages estimated from the host galaxy fitting is then used to study trends in normal and active galaxies (see subsection 4.3). This section also explores the correlation between the derived ages of the sources and the lookback time (based on redshift), offering insights into how AGNs and star-forming galaxies evolve over different epochs. The broader conclusions from our findings are discussed in section 5, which also outlines the future directions for this line of research.

2 DATA

The GOODS-S field, centered at RA $3^{\text{h}} 32^{\text{m}} 32.74^{\text{s}}$ and DEC $-27^{\circ} 47' 20.40''$, spans 7.47 arcminutes and is characterized by its low hydrogen column density, minimizing galactic reddening and enabling clear extragalactic observations (Giavalisco et al. 2003). This field has been the focus of extensive multiwavelength surveys, starting with the *Chandra Deep Field South* (CDFS), which provides the deepest X-ray imaging to date, complemented by wider-field *XMM-Newton* observations (Luo et al. 2016; Xue et al. 2011). UV data have been contributed by *GALEX* and *Hubble's* UV Legacy Survey, probing young stellar populations and star formation in early galaxies (Martin et al. 2007; Teplitz et al. 2018). Optical observations include deep *HST* imaging with the ACS and WFC3 instruments, as part of the GOODS program, and spectroscopic studies from ground-based facilities like the VLT and Subaru, offering morphological classifications and redshift measurements (Giavalisco et al. 2004; Cooper et al. 2020). Infrared observations from *Spitzer* and *Herschel* have uncovered obscured star formation and thermal dust emission, while *JWST* now delivers unprecedented near- and mid-infrared data, resolving galaxies at $z > 10$ (Lutz et al. 2011; Finkelstein et al. 2023). Radio observations with the VLA at 1.4 GHz and submillimeter data from ALMA and ATCA have provided insights into star formation rates, molecular gas reservoirs, and dust content in galaxies (Morrisson et al. 2008; Dunlop et al. 2016). These multiwavelength datasets collectively make GOODS-S an unparalleled field for investigating galaxy formation, AGN activity, and the cosmic history of baryonic matter.

We have utilized optical spectral data from the ESO GOODS-S survey for our analysis. The data were obtained from the ESO data archive, observed using the FORS1 (Appenzeller et al. 1998a), FORS2 (Cesarsky 2014a), and VIMOS (Cesarsky 2014b) instruments mounted on the VLT at Paranal Observatory in Chile. The data sample analyzed in this study includes 8864 spectra covering the optical and near-optical wavelength regions, spanning from approximately 3600 Å to 10,000 Å (Appenzeller et al. 1998b). One of the aims of this study is to evaluate the various optical line-ratio-based classification schemes for AGNs, hence we primarily focus on optical spectroscopy data.

In addition to this primary dataset obtained from ESO, the redshift values of many of these sources were obtained separately from various surveys available for the same field, which are listed in Table 1 to complete our dataset. The spectra were mapped to the redshift surveys using a methodology available in the Python libraries such as *Astropy* and *TOPCAT* integrations (Astropy Collaboration et al. 2022), which uses RA and DEC values to match the sources across different surveys. This mapping was implemented using a Python-based code. From the available surveys, we mapped redshifts for 3534 unique sources.

In order to classify AGNs and star-forming galaxies, we have utilized traditional BPT diagrams (Baldwin et al. 1981), such as the $[\text{S II}]$, $[\text{N II}]$, and $[\text{O II}]$ diagnostic diagrams, which are commonly used to classify AGNs and star-forming galaxies based on their emission line ratios. These diagrams provide a clear separation between different types of galaxies, including AGNs, H II regions, and composite galaxies. The $[\text{S II}] \lambda\lambda 6717, 6731$, $[\text{N II}] \lambda 6584$, and $[\text{O II}] \lambda 3727$ emission lines are sensitive to the ionization state of the gas, making them key tracers of ionization mechanisms.

The flux ratios of $[\text{O III}] \lambda 5007/\text{H}\beta$, $[\text{N II}] \lambda 6584/\text{H}\alpha$, and $[\text{S II}] \lambda\lambda 6717, 6731/\text{H}\alpha$ are used to distinguish between star-forming galaxies and AGNs, with the latter showing strong emission from the AGN ionizing source. This methodology follows the diagnostic criteria de-

veloped by Baldwin et al. (1981), Veilleux & Osterbrock (1987), and further refined by Kewley et al. (2006) to interpret the gas excitation and ionization mechanisms across a wide range of redshifts and galaxy types.

We have also utilized the MEx diagram (Juneau et al. 2014), which uses galactic stellar mass as one of its axes. The stellar mass, an important parameter for understanding galaxy evolution, represents the mass present in stars in the galaxy. For this analysis, we have mapped the stellar mass data from the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS) by Santini et al. (2015) to the objects whose spectroscopy data is considered by us.

We have also used the CEx diagram in this study (Yan et al. 2011), which requires the $(U - B)$ color index as one of the axes. In general, the color index measures the difference between two measurements of the magnitude of an object made at different bands, with the longer wavelength value subtracted from the shorter wavelength value. The lower the color index, the bluer (or hotter) the object is, while a larger color index indicates a redder (or cooler) object. We have mapped the $(U-B)$ color index data from the COMBO-17 survey by Wolf et al. (2004).

We have also mapped the optical spectra to X-ray sources from the CDFS by Luo et al. (2016). X-ray luminosity data is a crucial metric for determining obscured AGNs, as it provides a direct probe of high-energy processes often hidden from optical wavelengths by dust and gas (Gilli et al. 2001). X-ray AGN sources can also serve as a confirmation for optically identified AGNs, particularly in cases where optical diagnostics are ambiguous or where sources are heavily obscured in the optical. Additionally, the X-ray classifications from Luo et al. (2016) provide a means for identifying trends and relating them to our AGN classifications, allowing for cross-verification and improved characterization of the AGN population. These combined datasets enable robust analysis of AGN properties, such as their luminosity functions, obscuration levels, and host galaxy environments (Buchner et al. 2014).

In the study of AGNs, the gas clouds surrounding the central engine are pivotal for understanding their spectral features (Schneider 2006). These clouds, typically found in spiral galaxies hosting AGNs, absorb ultraviolet and X-ray emissions from the central engine, leading to the emission of characteristic spectral lines (Osterbrock & Pogge 1985). These lines include strong hydrogen (H) lines and emission lines like $[\text{N II}]\lambda 6584$ and $[\text{O III}]\lambda 5007$, similar to those found in H II regions of our galaxy (Schneider 2006; Kewley et al. 2013). These spectral features provide insights into the composition and density of the gas clouds, categorized into two regions: the broad-line region (BLR) with denser and warmer clouds and the narrow-line region (NLR) with lower-density and colder clouds, both influenced by the AGN's strong gravitational field (Netzer 2015; Krolik 1999). According to the AGN unified model (Antonucci 1993b; Padovani et al. 2017), these two regions are not found in fundamentally different classes of AGNs but rather reflect the observer's line of sight with respect to the dusty torus surrounding the nucleus. When the central engine and the BLR are viewed directly ("face-on", type 1 AGN), both broad and narrow emission lines are observed. However, if the line of sight is obscured by the torus ("side-on", type 2 AGN), only the narrow-line region is visible. This orientation-dependent framework explains the diversity in AGN spectra while reinforcing that BLR and NLR are intrinsic components of all AGNs, while their observability is dependent on the viewing geometry.

The process of selection of data finally used in our analysis is outlined in Table 2. Initially, we have 8864 ESO spectra (Stage 1). Reliable redshifts are available only for 4845 spectra (Stage 2, see Table 1). In stage 3, we map these 4845 spectra to 3534 unique sources.

Table 1. List of redshift surveys used

Redshift Survey	Number of Sources
Vanzella, E. et al. (2005) [V05]	234
Vanzella et al. (2006) [V06]	723
Ravikumar et al. (2007) [R07]	156
Vanzella et al. (2008) [V08]	886
Balestra et al. (2010) [B10]	2824
Hsu et al. (2014) [H14]	261
Morris et al. (2015) [M15]	305
Joshi et al. (2021) [J19]	154

Of these, the data is of sufficient quality to undertake fitting with our model only in 2937 cases (Stage 4). These include even those with only host galaxy fitting. In the remaining cases, the fitting algorithm was failing due to poor S/N ratio. Further, spectra of only 726 unique sources are found to be in the wavelength range suitable for use in the diagnostic diagrams used in this study, since $[\text{O III}]\lambda 5007/\text{H}\beta$ ratio must be available for this (Stage 5). This corresponds to a minimum corrected wavelength of $4861 \text{ \AA}$, (for $z = 0$) to a maximum of $10000 \text{ \AA}$ (corresponding to $z \approx 1$). Specifically, 666 sources are used in the CEx diagram, 471 in the MEx diagram, 548 in the $[\text{O II}]$ BPT diagram ('blue' diagram), 104 in the $[\text{N II}]$ BPT diagram and 106 in the $[\text{S II}]$ BPT diagram. Also, 100 of these 726 sources have matched x-ray counterparts from the list of Luo et al. (2016). All this information is summarized in Table 2 which provides an overview of the data selection and analysis workflow.

A comment on reddening effects is also important here. In our study, galactic reddening is negligible ($N_{\text{H}} \sim 8 \times 10^{19} \text{ cm}^{-2}$), as the GOODS-South field lies in a region of low galactic hydrogen column density and dust extinction (Schlegel et al. 1998; Giavalisco et al. 2004), making the effect of foreground reddening minimal. To assess the presence of intrinsic reddening (in the host), we analyze the Balmer decrement (Grupe et al. 2010; Dong et al. 2008), *i.e.* the $\text{H}\alpha/\text{H}\beta$ ratio. Following the criterion proposed by Dong et al. (2008) and also adopted by Grupe et al. (2010), sources with $\text{H}\alpha/\text{H}\beta < 3.06$ are considered to exhibit no significant intrinsic reddening. Out of the 726 sources used in our diagnostic diagrams, 125 lie at redshifts where both $\text{H}\alpha$ and $\text{H}\beta$ fall within the observed spectral range, allowing for flux measurements for both lines. Among these, 37 sources ($\sim 30\%$) do not show significant broad-line $\text{H}\alpha$ or $\text{H}\beta$ emission, likely due to weak BLR activity. Of the remaining sources, 82 ($\sim 65\%$) show $\text{H}\alpha/\text{H}\beta$ ratios below 3.06, consistent with little to no intrinsic reddening. Only 6 sources ($\sim 5\%$) have $\text{H}\alpha/\text{H}\beta > 3.06$, suggesting the presence of intrinsic reddening in their BLR and the need for reddening correction. This analysis indicates that, statistically, a vast majority ($\sim 95\%$) of the sources in the data considered by us would not exhibit significant intrinsic reddening. Although the number of sources with available Balmer line measurements is limited, the sample of 125 is statistically sufficient to infer overall trends. Thus, we opine that reddening effects would not play a significant role in the broader interpretation of statistical trends in the present GOODS-south dataset.

3 SPECTRAL FITTING METHODOLOGY

There are a few spectral fitting codes for galaxy / AGN spectra already mentioned in the literature (Shen et al. 2011; Calderone et al. 2017; Ilić et al. 2023). Early tools like the SPECFIT package (Kriss 1994) provided foundational capabilities for continuum fitting and

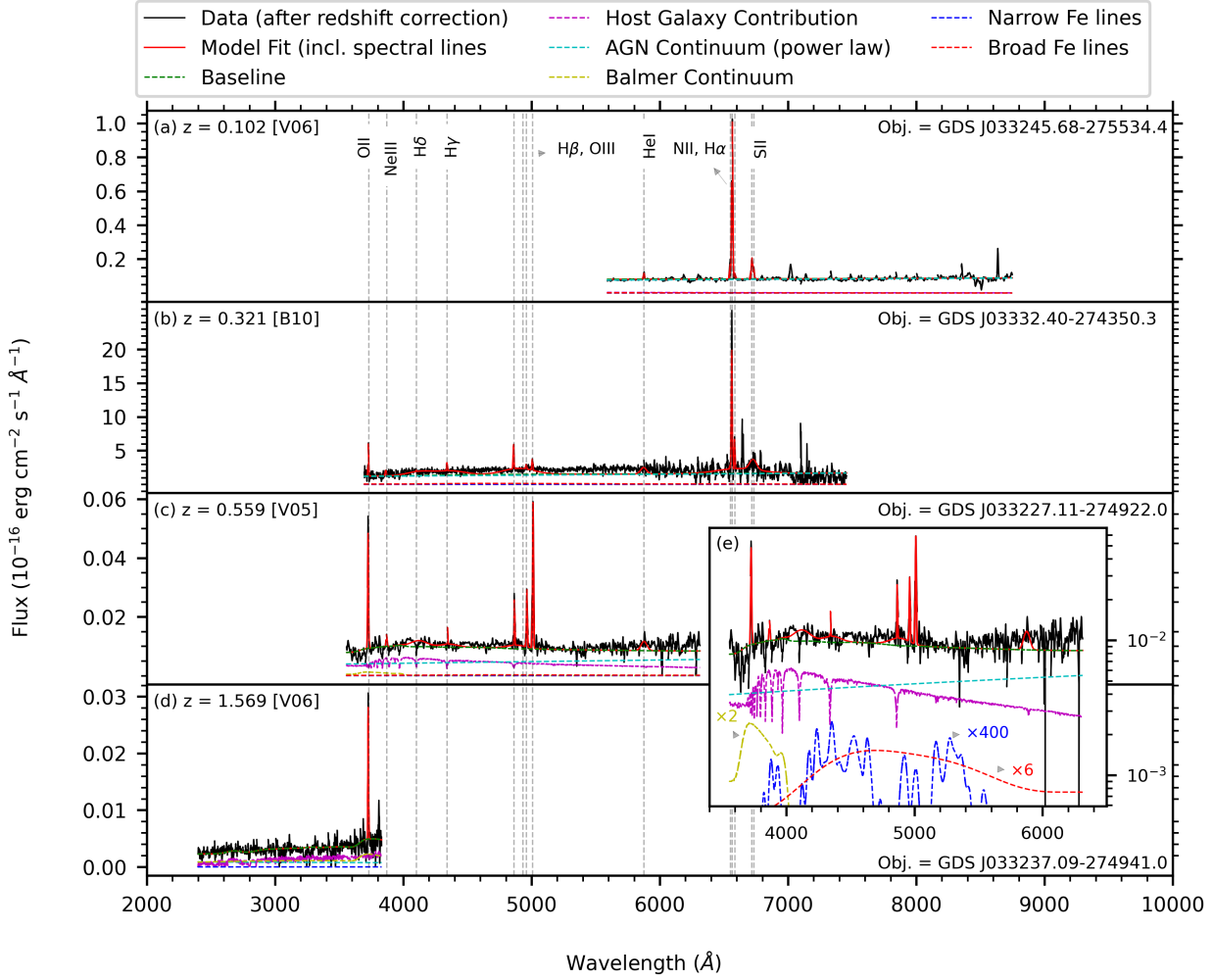


Figure 1. Spectral model fitting results for four sources: (a) *GDS J033245.68-275534.4* ($z = 0.102$), (b) *GDS J03332.40-274350.3* ($z = 0.321$), (c) *GDS J033227.11-274922.0* ($z = 0.559$), (d) highlights *GDS J033237.09-274941.0* ($z = 1.569$). Across the subfigures, we have shown the redshift-corrected spectra (black curve) and the overall model fit (red curve, including discrete emission lines), as well as baseline contributions, *i.e.* host galaxy contribution (magenta), AGN power-law continuum (cyan), and narrow and broad Fe lines (blue and red dashed). Subfigure (e): log-scaled version of (c), emphasizing all the spectral components for clarity.

Table 2. Data Reduction Stages and Source Statistics

Stage	Description	Count
Stage 1	Total ESO Spectra	8864
Stage 2	Total Redshifts Available	4845
Stage 3	Total Sources with Unique Redshifts	3534
Stage 4	Total Fitted Spectra (Includes sources with only Host Galaxy Fitting also)	2937
Stage 5	Total Unique Sources used in diagnostic diagrams (Considering sources with at least (O III/HB) ratio)	726
	Sources in CEx Diagram	666
	Sources in MEx Diagram	471
	Sources in [O II] BPT Diagram	548
	Sources in [N II] BPT Diagram	104
	Sources in [S II] BPT Diagram	106
	Matched X-ray Sources (from (Luo et al. 2016))	100

basic emission line modeling but lacked comprehensive integration of multi-component spectral features. A major leap was made by [Shen et al. \(2011\)](#), who performed a systematic analysis of over 100,000 SDSS quasar spectra, including detailed measurements of major emission lines and continua. Subsequent efforts, such as QSFIT by [Calderone et al. \(2017\)](#), introduced frameworks for global fitting of continuum, absorption, and emission features. More recently, the FANTASY code ([Ilić et al. 2023](#)) demonstrated high-fidelity multi-component fitting for AGN spectra, employing broken power-law continua and refined Fe II models derived from eigenspectra. However, there are also constraints associated with them, *e.g.* some spectral features are not included, or they are usable only for datasets from specific observatories, etc. Building upon these insights, we have indigenously developed a Python-based spectral fitting code ([Panchal et al. 2024](#)), and utilised it for the spectral analysis of GOODS-S optical spectra. The detailed algorithm is described in the Appendix. Here, we focus on more physical aspects of the code.

Firstly, the continuum emission from the galaxy is considered to be comprised primarily of four distinct components of different physical origins. The continuum emission from the (active) galactic nucleus is

modelled as a broken power law (e.g. see [Calderone et al. \(2017\)](#)). A novel feature of our code is the inclusion of a host galaxy contribution model, achieved through templates from the CB2019 library developed by Bruzual and Charlot ([Bruzual & Charlot 2003](#); [Coelho et al. 2019](#)), which provides spectral evolution data for stellar populations ranging from 0.1 Myr to 14 Gyr across a metallicity range of 0.0001 to 0.06. The host galaxy fitting process involves including the galaxy templates as a part of the baseline, and applying a least-squares fitting algorithm by iterating across all templates in the library to identify the best-fitting template. In the baseline continuum, we also consider the Balmer jump, or Balmer discontinuity, which occurs near the Balmer limit (3646 Å), with the Balmer continuum modelled using a blackbody function ([Wills et al. 1985](#); [Storey & Hummer 1995](#); [Calderone et al. 2017](#)). The model represents the continuum effectively below the Balmer edge wavelength, while suitably broadened Gaussian functions are used for the Balmer lines above it. Additionally, the model incorporates a pseudo-continuum due to both broad and narrow Fe lines, which must be included in estimating the overall continuum and addressing contamination (e.g. from lines like H β). This pseudo-continuum is generated by applying suitable Gaussian broadening to the iron lines in the UV ([Vestergaard & Wilkes 2001](#)) and optical ([Véron-Cetty, M.-P. et al. 2004](#)) wavelength ranges. Thermal broadening is a critical parameter in this process, with specific bounds established for narrow and broad lines in both regions, ensuring the physical validity and reliability of the fitting.

The discrete emission lines rising above the aforementioned continuum (baseline) typically include O II, Ne III, H δ , H γ , H β , H α , O III, He I, N II, and S II lines. The fitting of these lines is accomplished using Gaussian functions that model these lines based on parameters such as amplitude, FWHM, and velocity offset. Specifically, FWHM values for narrow lines range from 100-2000 km s⁻¹ and for broad lines from 1000-15000 km s⁻¹, while velocity offsets are constrained to ± 1000 km s⁻¹ for narrow lines and ± 3000 km s⁻¹ for broad lines. These bounds have been successfully used in the past for galaxy / AGN spectral analysis (see e.g. [Calderone et al. \(2017\)](#)). Overall, our model attempts to reproduce all the observed spectral features based on physically motivated model components.

4 RESULTS AND DISCUSSION

[Figure 1](#) shows the fitting results for four sources across different wavelength ranges (after redshift correction), showcasing many of the emission lines typically considered in this study. [Figure 1](#) (a) displays the object *GDS J033245.68-275534.4* with a redshift of 0.102 [V06]. The solid black curve represents the data after redshift correction, while the solid red curve indicates the model fit, including spectral lines. Contributions from the host galaxy, AGN continuum (broken power law), and narrow and broad Fe lines are shown as dashed magenta, cyan, blue and red lines respectively, forming an overall baseline (green dashed line), upon which prominent emission lines, such as [O III] λ 5007, H β , H α , etc. are fitted. Similarly, subfigures (b), (c) and (d) show the objects *GDS J033332.40-274350.3*, *GDS J033227.11-274922.0* and *GDS J033237.09-274941.0*, with progressively increasing redshifts of 0.321 [B10], 0.559 [V05] and 1.569 [V06]. The inset in [Figure 1](#) (e) provides a log-scaled view of the same spectrum as in [Figure 1](#) (c), emphasising specific spectral contributions. The same spectral fitting algorithm is used to fit all 4845 sources with known redshift values, achieving a decomposition of different components contributing to the spectral features in these sources [Panchal et al. \(2024\)](#).

4.1 [O II], [N II], and [S II] BPT diagrams

The classification of extragalactic sources into AGNs and SFs can be effectively achieved using various BPT diagrams, which are plotted together in [Figure 2](#), for the GOODS-S data considered in this paper.

[Figure 2](#) (a) presents the [O II] BPT diagram (the so-called ‘blue diagram’ ([Lamareille 2010](#))), which compares the ratios [O III] λ 5007/H β and [O II] λ 3727/H β . Most sources are located in the bottom-left region, identified as SF galaxies (blue dots). The upper and right regions are dominated by AGNs, including Seyfert galaxies (red dots) and LINERs (green dots). There is a region where the star-forming and Seyfert populations overlap, and the classification for these sources is uncertain. This is the so-called composite region (cyan dots). The empirical lines shown are adapted from [Lamareille \(2010\)](#).

Along the same lines, [Figure 2](#) (b) shows the [N II] BPT diagram which plots [O III] λ 5007/H β against [N II] λ 6584/H α . This diagnostic separates SF galaxies (blue dots) from AGNs (red and green dots for Seyferts and LINERs, respectively) and includes a composite region (cyan dots) for sources with ambiguous classifications. Square markers (blue, cyan, red, and green) indicate classifications from the [O II] diagram, revealing some discrepancies between the two methods, such as SF sources classified as AGNs and vice versa. The dividing lines are based on [Kauffmann et al. \(2003\)](#) and [Schawinski et al. \(2007\)](#).

Finally, [Figure 2](#) (c) displays the [S II] BPT diagram, which compares [O III] λ 5007/H β and [S II] λ 6717,6731/H α . It effectively distinguishes SF galaxies (blue dots) from AGNs (Seyferts and LINERs in red and green dots, respectively). Similar to [Figure 2](#) (b), square markers indicate the [O II] classification. The dividing lines are derived from [Kewley et al. \(2006\)](#). Across all subfigures, the majority of sources are concentrated in the SF region, consistent with the general consensus on the global AGN population, as well as previous reports, e.g. by [Kewley et al. \(2006\)](#) and [Juneau et al. \(2011\)](#).

Our dataset also contains some sources for which multiple spectra are available, acquired at different epochs (see [section 2](#)). For these selected sources, we have the opportunity to quantify the long-term time-variability in the emission lines, which may reflect changes in physical conditions, ionisation states, or other temporal factors in the sources ([Peterson 1988, 1993](#)). To represent this variability on the BPT diagram, we plot the observed emission-line flux ratios for each source, with the marker position indicating the mean flux ratio across multiple observations, and the error bars denoting the range of flux ratios observed at different times. These error bars are intended as a measure of the temporal variability, highlighting the dynamic nature of the emission-line regions, and the impact of this variability on the corresponding classification scheme. While AGNs exhibit variability due to fluctuations in their accretion rates and surrounding environments, and SF galaxies may undergo changes in their star formation activity, it is observed that the BPT classification remains largely unaffected. This is because the emission lines used in these diagrams originate from ionized gas regions that respond to long-term radiation fields rather than short-term fluctuations. Thus, irrespective of variability in AGNs or star-forming galaxies, BPT diagrams remain effective and reliable for distinguishing between different galaxy types.

Across the different BPT diagnostic diagrams, the classification statistics reveal consistent and well-separated source populations. In the [O II] BPT, a total of 459 sources are classified as SF (~84%), 41 as composite (~7%), and 48 as AGN (~9%). Similarly, the [S II] diagram identifies 84 SF and 22 AGN sources (~79% and ~21%,

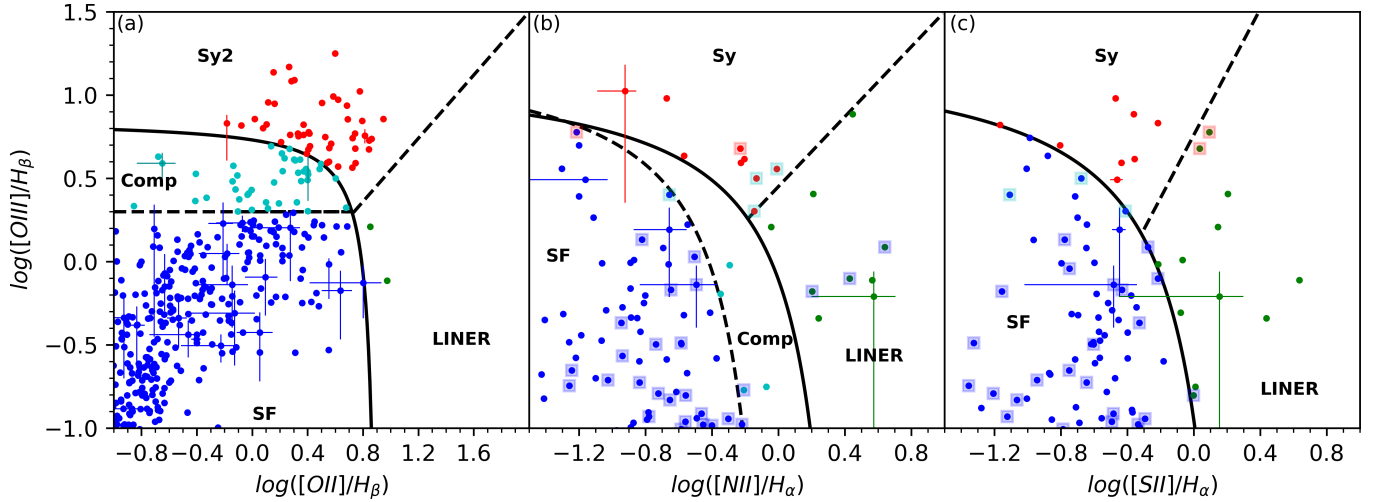


Figure 2. (a) The [O II] BPT diagram (‘blue diagram’), (b) The [N II] BPT diagram, (c) The [S II] BPT diagram. Across subfigures, blue dots represent star-forming galaxies (SFs), cyan dots indicate composite galaxies, red dots correspond to Seyferts (AGN), and green dots denote LINERs (AGN). Square markers (blue, cyan, red, and green) indicate classifications from the [O II] BPT diagram, providing a comparison across subfigures.

respectively); while the [N II] diagram yields 77 SF, 8 composite, and 19 AGN sources ($\sim 74\%$, $\sim 8\%$ and $\sim 18\%$, respectively).

When comparing [O II] classifications with those from [S II] and [N II], there is a strong level of agreement. Among the 459 SF sources in [O II], 28 are also present in the [S II] diagram, of which 27 ($\sim 96\%$) are classified as [S II]-SF. The corresponding numbers for [N II]-SF are 28 out of 33 ($\sim 85\%$). In the case of [O II]-identified composite and AGN sources, only 2 sources each are present in the [S II] and [N II] diagrams, and hence are statistically insufficient to define any concrete trends. Overall, the [O II] BPT is found to be quite consistent with [S II] and [N II], particularly for the SF population, and is able to cover a much larger number of sources, due to its applicability up to larger redshifts.

As discussed in section 3, the host galaxy contribution is integrally included in our spectral fitting procedure. This uses the stellar population synthesis models of Bruzual & Charlot (2003), which predict the spectral evolution of stellar populations over a wide range of ages (10^5 to 2×10^{10} years) and metallicities. Their high-resolution model incorporates updated stellar evolution tracks and a modern empirical spectral library, enabling the generation of synthetic spectra at $3 \text{ \AA}$ resolution in the optical range ($3200\text{--}9500 \text{ \AA}$). These models are widely used in galaxy evolution studies for interpreting spectra and deriving star formation histories, metallicities, and dust content (Kauffmann et al. 2003; Gallazzi et al. 2005; Popescu et al. 2011).

We use this library in our code to incorporate the host galaxy component, with the best-fit template providing the galaxy’s age. Figure 3 presents some results from this analysis. Figure 3 (a) shows the same [O II]-BPT diagram as Figure 2, but restricted to sources in two age bins: $0.5\text{--}1 \text{ Gyr}$ (cyan stars) and $>8 \text{ Gyr}$ (magenta squares), representing relatively young and old systems, respectively. Despite the limited number of sources, a clear pattern emerges: the older population of galaxies are well-represented across different regions of the diagram, while younger galaxies are virtually confined to the [O II]-SF and [O II]-composite regions. This is as expected for these age-wise populations, and also yields confidence in our incorporation of galaxy age in our model.

Further, Figure 3 (b) summarizes the age distribution of [O II]-classified SFs and AGNs across discrete host galaxy age bins. Each bin corresponds to the intervals used in Figure 3 (a), marked for

consistency. For each age bin, the relative fraction of SFs and AGNs is shown, allowing a direct comparison. A clear trend is observed: the AGN fraction increases with host galaxy age, suggesting a possible evolutionary link between star formation activity and AGN triggering over time.

In each host galaxy age bin, The two highlighted age regions from Figure 3 (a) are marked in Figure 3 (b) for reference. It is clearly seen that the fraction of AGNs out of all sources at a given age increases with host galaxy age. In fact, over the complete dataset, if we consider all [O II]-SF sources, they are uniformly distributed across the age bins ($\sim 6.5\%$), whereas the population of all [O II]-AGNs is significantly skewed towards higher HG ages, rising from $\sim 3\%$ at $\leq 1 \text{ Gyr}$ to $\sim 12\%$ at $\geq 8 \text{ Gyr}$ (not shown in figure).

This consistency between spectral classification and host galaxy age strengthens confidence in our fitted age values. The observed trend underlines an evolutionary connection between star formation and AGN activity: as galaxies age, AGN activity becomes more dominant, potentially due to gas depletion, black hole growth, and central fueling mechanisms (Hopkins et al. 2008; Heckman & Best 2014; Suh et al. 2020). Meanwhile, the wide age spread of SF galaxies indicates sustained star formation across a large fraction of cosmic time, likely regulated by gas accretion and feedback processes (Madau & Dickinson 2014). The increased fraction of AGN galaxies at older ages is also understandable in terms of AGN feedback quenching star formation in evolved systems (Fabian 2012; Kormendy & Ho 2013).

4.2 MEx and CEx diagrams

In addition to the original BPT diagnostic diagrams discussed in Figure 2, there have been recent efforts to use the same framework, modified by the inclusion of some other physical parameters. Examples of such efforts in the literature are the MEx (Juneau et al. 2011) and CEx (Yan et al. 2011) diagrams, which utilise the stellar mass M^* and the $(U - B)$ colour index respectively.

The MEx diagram, originally developed by Juneau et al. (2011), plots the stellar mass of the galaxies ($\log(M_*/M_\odot)$), against the [O III]/H β flux ratio, which reflects the ionization state of the gas. The diagram is divided into regions corresponding to different galaxy types: SF, composite, and AGN-dominated galaxies, separated by

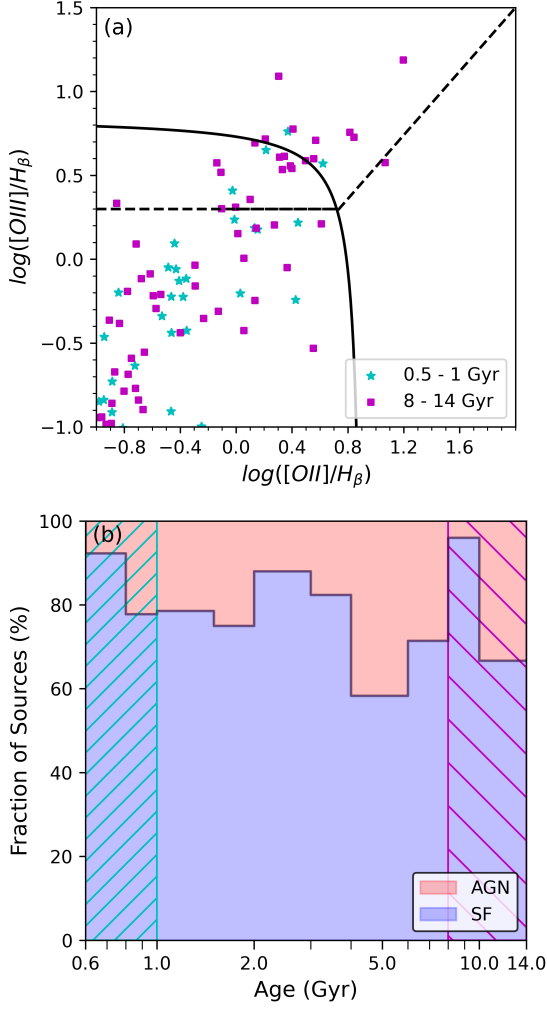


Figure 3. (a): [O II]-BPT diagram of Figure 2(a), limited to sources in two fitted age ranges — 0.5–1 Gyr (cyan stars) and >8 Gyr (magenta squares) — representing relatively young and old stellar populations, respectively. While AGNs are predominantly found in the older age bin, a small number of AGNs also appear in the younger age range (Heckman & Best 2014; Kauffmann et al. 2003). (b): Fractional distribution of [O II]-classified star-forming galaxies (blue) and AGNs (red) as a function of fitted host galaxy age. A clear trend is seen where the fraction of AGNs increases with stellar age. The two age bins shown in (a) are indicated by shaded boxes for reference.

solid and dashed empirical lines adapted from Juneau et al. (2014). The original demarcations in (Juneau et al. 2011) were updated by Juneau et al. (2014) to account for emission line detection limits, broadening the diagram’s applicability across redshifts.

To obtain the stellar masses (M^*) we have used the work of Santini et al. (2015), which summarises the efforts by 10 different teams within the CANDELS collaboration, who compute stellar masses using the CANDELS-GOODS-S dataset. For our analysis, we adopt stellar masses computed by the mass estimation method $2d_+$, which assumes a Salpeter Initial Mass Function (IMF). This choice is mainly to ensure consistency with the AGN classification criteria in the MEx diagram developed by Juneau et al. (2011, 2014).

The physical motivation and relevance of the MEx diagram has been emphasised in the literature (Juneau et al. 2011; Vitale et al. 2013), and is summarised below. Galactic stellar mass is an important parameter as it correlates with the overall gravitational poten-

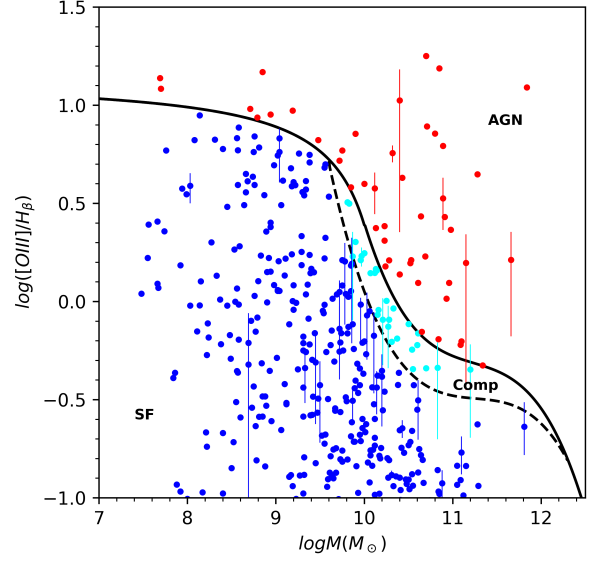


Figure 4. MEx diagram for the present data. The bottom-left region indicates star-forming galaxies (SF, blue dots), the composite region is located between the solid and dashed lines (Comp., cyan dots), and the top-right region corresponds to AGN activity (AGN, red dots). The empirical lines are adapted from Juneau et al. (2014).

tial and metallicity of the galaxy, influencing the ionisation state of the gas. Similarly, the [O III] λ 5007/H β ratio is sensitive to both the ionisation state and the metallicity. AGNs typically exhibit higher [O III] λ 5007/H β ratios due to the harder ionising radiation from the AGN compared to the softer radiation from young stars in star-forming galaxies. Galaxies with higher stellar masses are more likely to host AGNs, as these are often found in more massive galaxies with older stellar populations and higher central concentrations, which favor AGN activity (Vitale et al. 2013).

In the MEx diagram for the data considered by us (see Figure 4), 400 sources are classified as SF (~85%). These are located in the bottom-left region and are represented by blue dots. The middle (composite) region contains 26 galaxies (~5%), whose classification is ambiguous, marked by cyan dots. The top-right region comprises 45 sources classified as AGNs (~10%), indicated with red dots. Towards the end of this section, we present a comparison of the MEx results with our earlier results from subsection 4.1.

Apart from the MEx diagram, the CEx diagram is another diagnostic tool used to classify galaxies based on their spectral and photometric properties, particularly for distinguishing between SF galaxies, composite systems, and AGN-dominated sources. Developed by Yan et al. (2011), the diagram employs the [O III] λ 5007/H β flux ratio and the rest-frame ($U - B$) color index to distinguish between these classes, with empirical boundaries (black solid and dotted lines) separating different regions. Star-forming galaxies occupy the region with low [O III] λ 5007/H β values and bluer ($U - B$) colors, due to ionization dominated by young, massive stars. Composite galaxies lie in an intermediate zone with moderate values of both parameters, indicating contributions from both star formation and AGN activity. AGN-host galaxies, with high [O III] λ 5007/H β and redder ($U - B$), are located in the upper-right region, where ionization is primarily driven by accretion onto supermassive black holes.

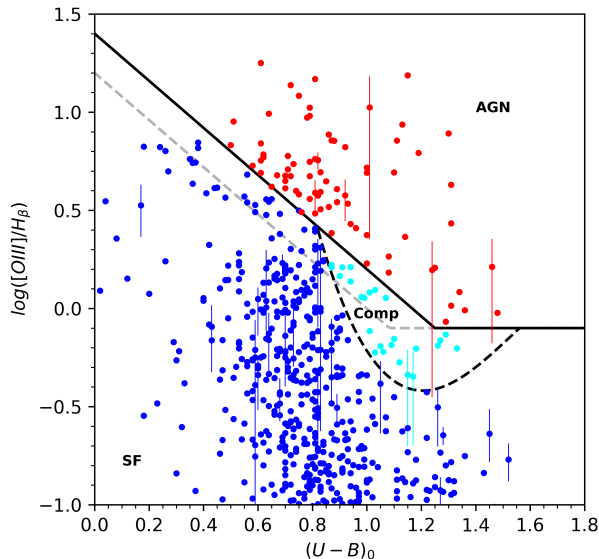


Figure 5. CEx diagram for the present data. Star-forming galaxies (SF, blue dots) occupy the region with low $[\text{O III}]\lambda 5007/H\beta$ ratios and bluer $U - B$ colors, composite galaxies (Comp., cyan dots) are located in the intermediate region, and AGN-host galaxies (AGN, red dots) are found in the region with high $[\text{O III}]\lambda 5007/H\beta$ ratios and redder $U - B$ colors. The empirical boundaries are adapted from Yan et al. (2011).

Trump et al. (2012) later refined the AGN/SF division (grey dotted line) by recalibrating the empirical boundaries using a larger and more diverse galaxy sample, including sources at $z > 1$. This makes the CEx diagram particularly effective in cases where traditional BPT diagrams are limited, such as for galaxies lacking all necessary emission lines.

Physically, the diagram is founded on the observation that nearly all BPT-identified AGNs reside in red or green galaxies, while very few are found in blue ones (Yan et al. 2011). Blue galaxies typically have lower stellar masses and bulge-to-disk ratios, which correlate with smaller central black holes (BHs) (Häring & Rix 2004; Jahnke & Macciò 2011). Consequently, AGNs in such galaxies tend to be less luminous and more easily overwhelmed by star formation, making them harder to detect using standard BPT diagnostics (Trump et al. 2012; Juneau et al. 2014). The CEx method attempts to correct for this bias by incorporating galaxy color, allowing for more reliable AGN identification across the full color sequence.

In our analysis, we use $(U - B)_0$ values from the COMBO-17 survey (Wolf et al. 2004), as described in section 2. Since the COMBO-17 magnitudes are in the Vega system and the CEx diagnostic was calibrated using AB magnitudes (Yan et al. 2011), we converted the $(U - B)$ colors to the AB system using the correction provided by Willmer et al. (2006).

In the CEx diagram (see Figure 5) for the data considered by us, 569 sources are classified as star-forming (SF; $\sim 85\%$). These galaxies, represented by blue dots, predominantly belong to the blue, low-mass population, where ionization is primarily attributed to young, massive stars. The diagram also identifies 23 composite galaxies ($\sim 4\%$, marked by cyan dots, whose classification is ambiguous (e.g. see Yan et al. (2011)). Finally, 74 sources ($\sim 11\%$) are classified as AGNs and are shown in red. These galaxies typically lie on the redder end of the color sequence and host more massive black holes, where

ionization is expected to be dominated by AGN accretion activity. The CEx diagram thus offers a valuable classification scheme, especially for galaxies where the BPT diagnostics are limited or unavailable.

In Figure 6, we examine the behavior of sources across various diagnostic diagrams, viz., in the sequence shown in the figure, the MEx (a), CEx (b), and $[\text{O II}]$ BPT (c) diagnostic diagrams. The subfigures are colour-coded such that classification of galaxies into SF, composite, and AGN-dominated categories is indicated by the colour of the markers (dots). In subfigures (b) and (c), the MEx classification is additionally shown in the form of shaded squares. Furthermore, sources identified as X-ray AGNs by Luo et al. (2016) from the CDF-S source catalog are also shown as black triangles.

In Figure 6 (a), the MEx diagram shown is the same as in Figure 4, and classifies sources into SFs (blue dots), composites (cyan dots), and AGNs (red dots). Additionally, X-ray AGNs are marked as black triangles, showing that 11 out of 32 X-ray AGNs are correctly identified as MEx-AGNs, while 20 are classified as MEx-SFs, and 1 as a MEx-Composite.

Figure 6 (b) shows the CEx diagram (same as Figure 5), with the same colour-coding of markers (dots). In addition, the MEx classification is also overlaid in the form of blue squares (MEx-SFs), cyan squares (MEx-Composites), and red squares (MEx-AGNs), for comparison. In the CEx classification, out of 35 X-ray AGNs, 7 are identified as CEx-AGNs, 25 as CEx-SFs, and 3 as CEx-Composites (accuracy $\sim 20\%$).

Figure 6 (c) displays the $[\text{O II}]$ BPT diagram (see Figure 2(a)), where blue, cyan, and red squares represent O II -SFs, O II -Composites, and O II -AGNs, respectively, and the MEx classification is overlaid for comparison similar to Figure 6 (b). Out of 30 X-ray AGNs, only 3 are classified as O II -AGNs, with the majority falling into the O II -SF and O II -Composite categories.

In broader terms, misclassifications of X-ray AGNs into SF or composite categories highlight the limitations of all these diagnostic methods. To summarise our results, the MEx diagram provides better AGN identification among X-ray-identified AGNs compared to other diagrams, correctly classifying $\sim 35\%$ (compared to $\sim 20\%$ and $\sim 10\%$ for the CEx and $[\text{O II}]$ -BPT, respectively). This observation is consistent with the analysis by Ricci et al. (2017a,b); Hickox & Alexander (2018), in which they conclude that $\sim 70\%$ of AGNs exhibit obscuration to varying degrees, making them easier detected in the X-ray spectrum than in the optical. These results emphasize the need for complementary multi-wavelength data to improve AGN detection.

Using the MEx diagram as a reference, we also assess the consistency of source classifications from the $[\text{O II}]$ and CEx diagrams. Among the 276 sources classified as star-forming in MEx, 264 ($\sim 88\%$) are also identified as SF in the $[\text{O II}]$ diagram, and 361 ($\sim 93\%$) are consistent with the SF classification in the CEx diagram. In contrast, for the 45 sources classified as AGN/composite in MEx, only 25 ($\sim 68\%$) match the AGN/composite category in $[\text{O II}]$, whereas 39 ($\sim 89\%$) show consistent AGN/composite classification in CEx.

These results suggest that the CEx diagram shows stronger correlation with the MEx classification, especially for both SF and AGN/composite sources, while the $[\text{O II}]$ diagram exhibits slightly lower consistency, particularly in the AGN/composite regime.

4.3 Temporal Evolution: Age vs Lookback time

In this section, we present a novel use of the host galaxy ages determined as part of the spectral fitting. We combine the derived host galaxy ages with the redshift values (z) for the respective galaxies.

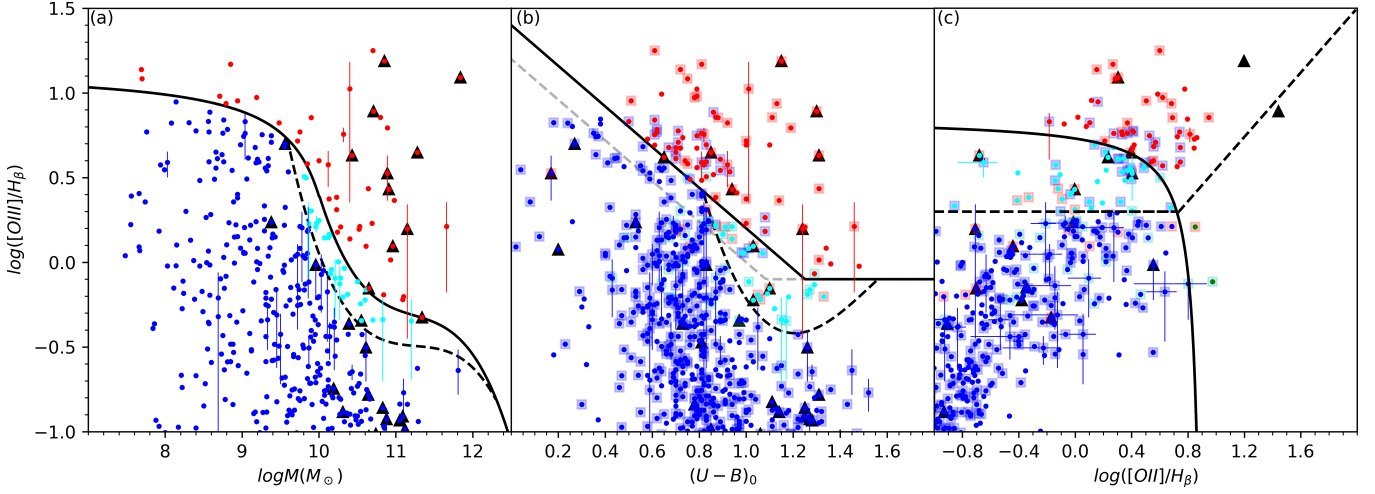


Figure 6. The figure presents source classifications across the MEx, CEx, and [O II] BPT diagnostic diagrams. Subfigures are color-coded to indicate star-forming galaxies (blue), composites (cyan), and AGNs (red). (a) shows the MEx diagram, where AGN identification is compared to X-ray-selected sources. (b) displays the CEx diagram, with classifications also represented as shaded squares for comparison. (c) presents the [O II] BPT diagram, illustrating classification variations and the challenges of distinguishing AGNs from star-forming and composite galaxies. Sources identified as X-ray AGNs from the Luo et al. (2016) CDF-S catalog are marked as black triangles, providing insight into how different diagnostics classify these objects.

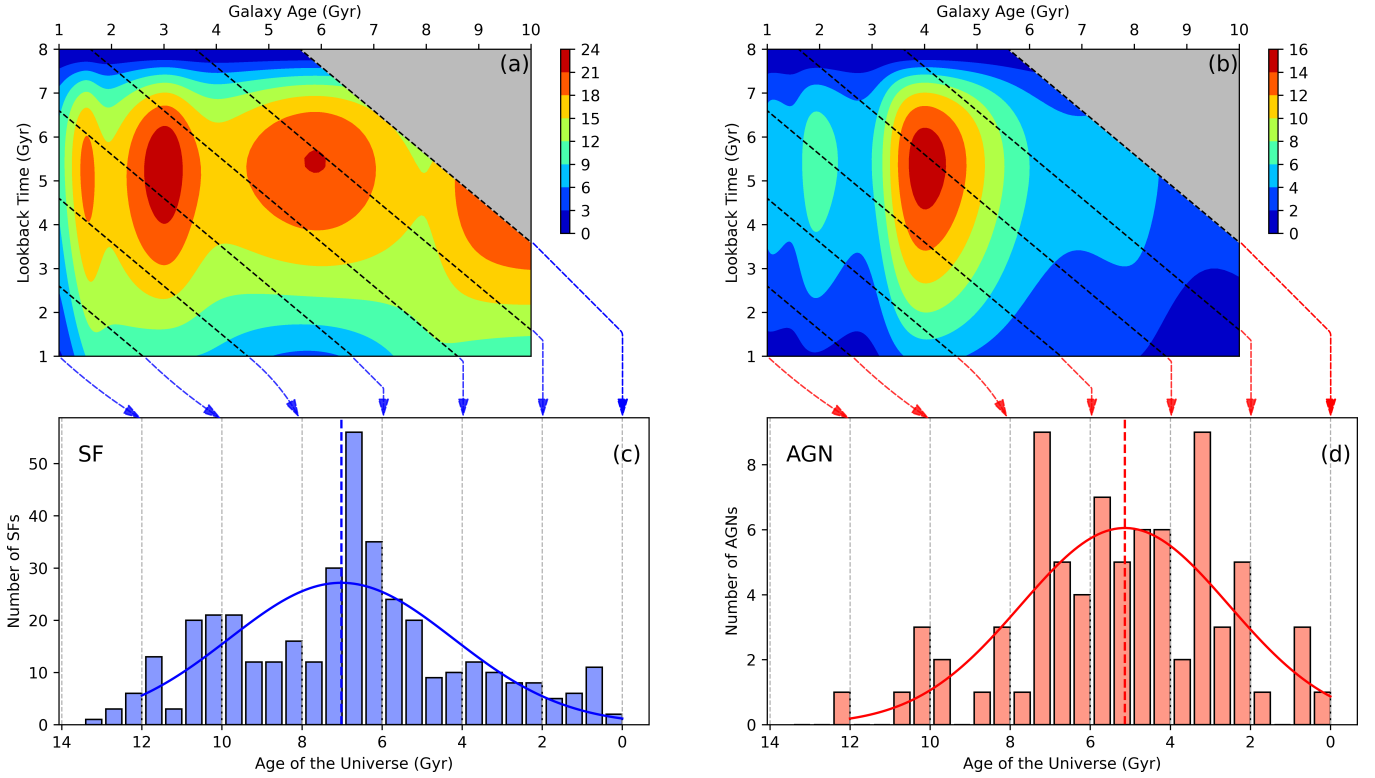


Figure 7. The Lookback Time vs. Host Galaxy Age plot reveals distinct evolutionary trends for star-forming galaxies and AGNs. The peak of AGN activity occurs at earlier cosmic times, suggesting that black hole growth was more significant when the universe was hotter, likely due to enhanced gas accretion. In contrast, star-forming galaxies are observed across all epochs, with their peak activity occurring later, indicating sustained star formation over cosmic time. The distribution of sources shows a clear shift, with AGNs being more prevalent in older host galaxies, while star formation remains widespread. This suggests a transition in galaxy evolution, where AGN feedback may have played a role in regulating star formation in massive systems.

The redshift provides us with the ‘lookback time’ (τ_{lookback}), which indicates how far back in time we are observing the galaxy. The host galaxy age (τ_{age}), as derived from the spectra, represents the age of the galaxy at the time of light emission. The sum of these two quantities ($\tau_{\text{tot}} = \tau_{\text{lookback}} + \tau_{\text{age}}$) provides the total time elapsed since the

epoch when the respective galaxy originated. This, in turn, acts as a measure of the age of the Universe at that epoch ($13.6 \text{ Gyr} - \tau_{\text{tot}}$).

In simple terms, τ_{lookback} tells us ‘how long ago’ the light was emitted, while τ_{age} tells us ‘how old’ was the galaxy at the time of emission. The lookback time (τ_{lookback}) is derived from the redshift

using cosmological parameters from the Planck 2018 results (Planck Collaboration et al. 2020), adopting $\Omega_m = 0.315$, $\Omega_\Lambda = 0.685$, and $H_0 = 67.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Using these values, the current-epoch age of galaxies is computed as the sum of Lookback Time and Galaxy Age ($\tau_{\text{tot}} = \tau_{\text{lookback}} + \tau_{\text{age}}$). The uniqueness of our methodology is that both time parameters (τ_{age} and τ_{lookback}) are derived entirely from the optical spectrum of the given galaxy. Furthermore, τ_{age} is derived from the continuum emission spectrum (host galaxy), while τ_{lookback} is derived from the discrete spectral lines (redshift), and are hence, in principle, independent of each other. Treating them as two independent parameters, one can study the distribution of galaxies across this 2-dimensional data space.

In Figure 7 (a and b), such a contour map is displayed, showing of the number distribution of galaxies over cosmic time. The horizontal axis represents Galaxy Age τ_{age} (in Gyr), as determined from spectral fitting. The vertical axis represents lookback time τ_{lookback} (in Gyr), as determined from the redshift. By our previous arguments, the diagonal dotted lines are lines of equal universe age ($\tau_{\text{tot}} = \tau_{\text{lookback}} + \tau_{\text{age}}$), and are spaced at intervals of 2 Gyr, starting from 13.6 Gyr on the extreme right, and decreasing towards the bottom left. The coloured contours indicate regions of equal number density of galaxies. The left and right upper subfigures (Figure 7 (a) and (b)) show this distribution for MEx-SFs and MEx-AGNs respectively.

The lower subfigures (Figure 7 (c) and (d)) are histograms generated by diagonally summing the data (*i.e.* summing at equal τ_{tot}) from the upper contour plots, as indicated by the dotted-line arrows, but with a bin-width of 0.5 Gyr. Since the diagonal strips represent lines of equal age of the universe, the histogram shows the fraction of galaxies originating at each cosmic epoch. The horizontal scale spans the age of the universe (*i.e.* 13.6 Gyr - τ_{tot} , shown increasing from right to left).

The plots in Figure 7 also provide a deeper perspective on the evolutionary timeline of galaxies and their connection to cosmic conditions, which we proceed to discuss below. Our results show that the formation of galaxies which ended up as MEx-SFs peaked when the age of the universe was 7.02 ± 2.80 Gyr, whereas MEx-AGNs peaked earlier at 5.14 ± 2.61 Gyr. This suggests that AGN formation-conducive activity would have been more prominent at earlier cosmic times. This conclusion is also consistent with previous studies indicating that SMBH growth and AGN activity peaked at earlier cosmic times, likely driven by an abundant cold gas supply and higher merger rates in the early universe (Hopkins et al. 2008; Heckman & Best 2014). Furthermore, using standard cosmological parameters, the estimated CMB temperature (Noterdaeme, P. et al. 2011) at universe age 5.14 Gyr (*i.e.* the peak of AGN-candidate formation in this study) was 5.8 K, while at 7.02 Gyr (the peak of SF-candidate formation), it was 4.7 K. This suggests that a higher cosmic temperature increases the probability of a particular galaxy to evolve into an AGN, possibly by enhancing gas accretion onto central black holes, thereby triggering AGN-driven feedback (see *e.g.* Weinberger et al. (2017)). Conversely, a lower background temperature at later epochs appears to be more conducive to star formation (Madau & Dickinson 2014).

This trend reinforces the idea that AGN activity and star formation are tightly linked to the physical conditions of the Universe at each epoch. The early peak in AGN activity reported by us may have contributed to shaping galaxy evolution, regulating star formation through feedback mechanisms (see *e.g.* Fabian (2012); Kormendy & Ho (2013)). Meanwhile, the sustained presence of SF (*i.e.* non-AGN) galaxies throughout cosmic history highlights the prolonged role of secular evolution, gas accretion, and internal galaxy dynamics in sustaining star formation (Tacconi et al. 2020).

5 SUMMARY

In this study, we have used the traditional Baldwin-Phillips-Terlevich (BPT) diagram framework ([N II], [S II], [O II] BPT diagrams) to classify AGNs and star-forming galaxies using spectral data from the GOODS-South field. We have also used more modern extensions of the BPT framework (MEx, CEx diagrams) in a detailed comparative study. Our analysis utilized an indigenously developed spectral fitting code, which incorporates host galaxy templates based on the stellar population synthesis models of Bruzual & Charlot (2003). This inclusion allows us to extract the underlying host galaxy continuum from the overall AGN background, and estimate host galaxy ages. This in turn enables a more refined study of AGN and star-forming galaxy evolution.

Our key findings are as follows:

- (i) Traditional BPT diagrams ([N II], [S II], [O II]) remain a relevant classification tool, even in the presence of spectral variability, with the [O II] BPT diagram serving as a useful alternative when the H α region is unavailable (beyond $z \sim 0.4$).
- (ii) The Mass-Excitation (MEx) diagram and the Colour-Excitation (CEx) diagram provide more viable alternatives with the inclusion of stellar mass and (U-B) colour index respectively. In the present study of the GOODS-S field, the MEx diagram outperforms the other alternative diagnostics, *e.g.* correctly identifying a larger fraction of X-ray-selected AGNs compared to the CEx and [O II] BPT methods. However, the results from both the MEx and CEx diagrams are highly correlated, and prove effective in identifying AGNs across a broader parameter space, especially at higher redshifts.
- (iii) Cross-validation with X-ray-selected AGNs from the Chandra Deep Field South (CDFS) survey shows that optical diagnostics alone miss a significant fraction of AGNs ($\sim 65\%$ - 70%), reinforcing the necessity of multi-wavelength approaches.
- (iv) Host galaxy age estimates indicate that AGNs preferentially reside in older host galaxies, supporting the scenario in which black hole growth becomes more dominant as galaxies evolve.
- (v) We have presented a methodology whereby the age as well as lookback time of a galaxy can be estimated from its optical spectrum. If utilised correctly, this provides an immensely valuable tool to understand the evolution of galaxies, and hence of the universe, on cosmic time scales.
- (vi) Figure 7 presents a novel approach to studying the co-evolution of star-forming galaxies and AGNs. By analyzing lookback time in combination with host galaxy age, we can directly probe how AGN activity and star formation have evolved over cosmic time. This temporal evolution analysis reveals that AGN-candidate formation peaks earlier (~ 8 Gyr ago), compared to the average SF-galaxy formation (~ 6 Gyr ago), although star formation undoubtedly remains widespread across all epochs.
- (vii) Furthermore, we can conclude that AGN activity peaked in an era of higher cosmic temperature ($T_{\text{CMBR}} = 5.8$ K at $\tau_{\text{tot}} \sim 5$ Gyr), likely driven by enhanced gas accretion.
- (viii) By combining lookback time with galaxy ages derived from spectral fitting, we can probe galaxy evolution all the way to early cosmic epochs (up to ~ 13 Gyr ago), even though the observed sample is limited to galaxies at $z \sim 1$ (with lookback times of only ~ 8 Gyr).

In conclusion, our findings emphasize the importance of integrating host galaxy modeling with multi-wavelength data to improve AGN classification. By extracting host galaxy age from optical spectra, we add an important — and hitherto overlooked — dimension to understanding galaxy evolution. While redshift indicates when we observe a galaxy, stellar age reveals how long it had been evolving

before that point. Together with lookback time, this enables us to estimate when a galaxy likely formed, offering a more complete view of its life history and the physical conditions that shaped it. Our spectral fitting framework, combined with both traditional and extended AGN diagnostics, shows that AGNs are preferentially found in older galaxies and likely formed earlier in cosmic history than star-forming systems. This work demonstrates that even limited-redshift ($z \lesssim 1$) optical surveys can yield meaningful insights into galaxy evolution when host galaxy age and lookback time are jointly analyzed. The observed trends can serve to inform future studies on the role of physical conditions — such as cosmic background temperature — in shaping galaxy evolution.

DATA AVAILABILITY STATEMENT

This research made use of data obtained from the European Southern Observatory (ESO) Science Archive Facility, in accordance with the ESO Data Access Policy. Observations were originally carried out using the FORS1 and FORS2 instruments (Appenzeller et al. 1998b), and the VIMOS instrument (Le Fèvre et al. 2003). We made use of the publicly available FORS2 GOODS dataset (DOI: 10.18727/archive/28) and the VIMOS GOODS dataset (DOI: 10.18727/archive/31), obtained via the ESO Science Archive Facility. A complete list of ESO archival DOIs is available at <https://archive.eso.org/wdb/wdb/doi/collections/query>.

ACKNOWLEDGEMENTS

This work also made extensive use of the Python programming language and several open-source scientific libraries, including NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), Pandas (Wes McKinney 2010), Matplotlib (Hunter 2007), and Astropy, a community-developed core Python package for astronomy (Astropy Collaboration et al. 2022). We gratefully acknowledge the developers and maintainers of these packages for making their tools freely available to the scientific community.

GP acknowledges support provided by the Polish National Science Center (NCN) through grants no. 2023/49/B/ST9/01671, and 2024/53/N/ST9/03885.

The authors also thank Prof. M. N. Vahia, Dr. A. N. Raghav and Mr. B. Palit for their valuable comments and suggestions during the preparation of this work.

REFERENCES

Alexander D. M., Hickox R. C., 2012, *New Astron. Rev.*, **56**, 93
 Antonucci R., 1993a, *ARA&A*, **31**, 473
 Antonucci R., 1993b, *ARA&A*, **31**, 473
 Appenzeller I., et al., 1998a, *The Messenger*, **94**, 1
 Appenzeller I., et al., 1998b, *The Messenger*, **94**, 1
 Astropy Collaboration et al., 2022, *ApJ*, **935**, 167
 Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, **93**, 5
 Balestra I., et al., 2010, *A&A*, **512**, A12
 Begelman M. C., Blandford R. D., Rees M. J., 1984, *Rev. Mod. Phys.*, **56**, 255
 Bruzual G., Charlot S., 2003, *MNRAS*, **344**, 1000
 Buchner J., et al., 2014, *A&A*, **564**, A125
 Calderone G., Nicastro L., Ghisellini G., Dotti M., Sbarbato T., Shankar F., Colpi M., 2017, *MNRAS*, **472**, 4051

Cesarsky C., 2014a, The Great Observatories Origins Deep Survey - VLT/FORS2 Spectroscopy, doi:10.18727/ARCHIVE/28, <https://doi.eso.org/10.18727/archive/28>
 Cesarsky C., 2014b, The Great Observatories Origins Deep Survey - VLT/VIMOS spectroscopy, doi:10.18727/ARCHIVE/31, <https://doi.eso.org/10.18727/archive/31>
 Coelho P. R. T., Bruzual G., Charlot S., 2019, *Monthly Notices of the Royal Astronomical Society*, **491**, 2025
 Coil A. L., et al., 2015, *The Astrophysical Journal*, **801**, 35
 Cooper M. C., Newman J. A., Davis M., et al., 2020, *The Astrophysical Journal Supplement Series*, **247**, 38
 Dietrich M., Hamann F., Shields J. C., Constantin A., Vestergaard M., Chaffee F., Foltz C. B., Junkkarinen V. T., 2002, *ApJ*, **581**, 912
 Dong X., Wang T., Wang J., Yuan W., Zhou H., Dai H., Zhang K., 2008, *MNRAS*, **383**, 581
 Dunlop J. S., McLure R. J., Biggs A. D., et al., 2016, *Monthly Notices of the Royal Astronomical Society*, **466**, 861
 Fabian A. C., 2012, *ARA&A*, **50**, 455
 Finkelstein S. L., Bagley M. B., Haro P. A., et al., 2023, *The Astrophysical Journal Letters*, **946**, L13
 Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, *MNRAS*, **362**, 41
 Gialalisco M., Ferguson H. C., Koekemoer A. M., et al., 2003, *The Astrophysical Journal Letters*, **600**, L93
 Gialalisco M., et al., 2004, *ApJ*, **600**, L93
 Gilli R., Salvati M., Hasinger G., 2001, *A&A*, **366**, 407
 Goulding A. D., et al., 2014, *The Astrophysical Journal*, **783**, 40
 Grandi S. A., 1982, *ApJ*, **255**, 25
 Grupe D., Komossa S., Leighly K. M., Page K. L., 2010, *The Astrophysical Journal Supplement Series*, **187**, 64
 Häring N., Rix H.-W., 2004, *ApJ*, **604**, L89
 Harris C. R., et al., 2020, *Nature*, **585**, 357
 Heckman T. M., Best P. N., 2014, *ARA&A*, **52**, 589
 Hickox R. C., Alexander D. M., 2018, *ARA&A*, **56**, 625
 Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006, *ApJS*, **163**, 1
 Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, *ApJS*, **175**, 356
 Hsu L.-Y., Stockton A., Shih H.-Y., 2014, *The Astrophysical Journal*, **796**, 92
 Hunter J. D., 2007, *Computing in Science and Engineering*, **9**, 90
 Ilić D., Rakić N., Popović L. Č., 2023, *The Astrophysical Journal Supplement Series*, **267**, 19
 Jahnke K., Macciò A. V., 2011, *ApJ*, **734**, 92
 Joshi B. A., Cohen S., Windhorst R. A., Jansen R., Pirzkal N., Hathi N. P., 2021, *VizieR Online Data Catalog*, p. J/ApJ/883/157
 Juneau S., Dickinson M., Alexander D. M., Salim S., 2011, *ApJ*, **736**, 104
 Juneau S., et al., 2014, *ApJ*, **788**, 88
 Kalfountzou E., et al., 2017, *Monthly Notices of the Royal Astronomical Society*, **471**, 28
 Kauffmann G., et al., 2003, *MNRAS*, **346**, 1055
 Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, *MNRAS*, **372**, 961
 Kewley L. J., Maier C., Yabe K., Ohta K., Akiyama M., Dopita M. A., Yuan T., 2013, *The Astrophysical Journal Letters*, **774**, L10
 Kormendy J., Ho L. C., 2013, *ARA&A*, **51**, 511
 Kriss G., 1994, in Crabtree D. R., Hanisch R. J., Barnes J., eds, *Astronomical Society of the Pacific Conference Series Vol. 61, Astronomical Data Analysis Software and Systems III*. p. 437
 Krolik J. H., 1999, *Active galactic nuclei : from the central black hole to the galactic environment*
 Lamareille F., 2010, *A&A*, **509**, A53
 Le Fèvre O., et al., 2003, in Iye M., Moorwood A. F. M., eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes*. pp 1670–1681, doi:10.1117/12.460959
 Luo B., et al., 2016, *The Astrophysical Journal Supplement Series*, **228**, 2
 Lutz D., Poglitsch A., Altieri B., et al., 2011, *Astronomy & Astrophysics*, **532**, A90

- Madau P., Dickinson M., 2014, *Annual Review of Astronomy and Astrophysics*, 52, 415–486
- Martin D. C., Wyder T. K., Schiminovich D., et al., 2007, *The Astrophysical Journal Supplement Series*, 173, 342
- Morris A. M., et al., 2015, *AJ*, 149, 178
- Morrison G. E., Owen F. N., Dickinson M., et al., 2008, *The Astrophysical Journal Supplement Series*, 175, 423
- Netzer H., 2015, *ARA&A*, 53, 365
- Noterdaeme, P. Petitjean, P. Srianand, R. Ledoux, C. López, S. 2011, *A&A*, 526, L7
- Osterbrock D. E., Pogge R. W., 1985, *ApJ*, 297, 166
- Padovani P., et al., 2017, *A&ARv*, 25, 2
- Panchal M., Joshi A., Sawant A., Sheth M., Kasthurirangan S., 2024, in 42nd meeting of the Astronomical Society of India (ASI). p. P53, <https://ui.adsabs.harvard.edu/abs/2024asi.confP..53P>
- Peterson B. M., 1988, *PASP*, 100, 18
- Peterson B. M., 1993, *PASP*, 105, 247
- Planck Collaboration et al., 2020, *A&A*, 641, A6
- Popescu C. C., Tuffs R. J., Dopita M. A., Fischera J., Kylafis N. D., Madore B. F., 2011, *A&A*, 527, A109
- Ravikumar C. D., et al., 2007, *A&A*, 465, 1099
- Rees M. J., 1984, *ARA&A*, 22, 471
- Ricci C., et al., 2017a, *ApJS*, 233, 17
- Ricci C., et al., 2017b, *Nature*, 549, 488
- Santini P., et al., 2015, *ApJ*, 801, 97
- Schawinski K., Thomas D., Sarzi M., Maraston C., Kaviraj S., Joo S.-J., Yi S. K., Silk J., 2007, *MNRAS*, 382, 1415
- Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *ApJ*, 500, 525
- Schneider P., 2006, *Extragalactic Astronomy and Cosmology*. Springer Berlin Heidelberg, Berlin, Heidelberg, doi:10.1007/978-3-540-33175-9, <http://dx.doi.org/10.1007/978-3-540-33175-9>
- Shen Y., et al., 2011, *ApJS*, 194, 45
- Silva L., Granato G. L., Bressan A., Danese L., 1998, *ApJ*, 509, 103
- Silverman J. D., et al., 2005, *The Astrophysical Journal*, 618, 123
- Stasińska G., Cid Fernandes R., Mateus A., Sodré L., Asari N. V., 2006, *MNRAS*, 371, 972
- Stern D., et al., 2005, *The Astrophysical Journal*, 631, 163
- Storey P. J., Hummer D. G., 1995, *Monthly Notices of the Royal Astronomical Society*, 272, 41
- Suh H., Civano F., Trakhtenbrot B., Shankar F., Hasinger G., Sanders D. B., Allevalo V., 2020, *ApJ*, 889, 32
- Tacconi L. J., Genzel R., Sternberg A., 2020, *ARA&A*, 58, 157
- Teplitz H. I., Rafelski M., Siana B. D., et al., 2018, *The Astronomical Journal*, 156, 257
- Tran H., Hartmann J.-M., 2008, *Journal of Geophysical Research: Atmospheres*, 113
- Trump J. R., et al., 2012, *The Astrophysical Journal Letters*, 763, L6
- Urry C. M., Padovani P., 1995, *Publications of the Astronomical Society of the Pacific*, 107, 803
- Vanzella, E. et al., 2005, *A&A*, 434, 53
- Vanzella E., et al., 2006, *A&A*, 454, 423
- Vanzella E., et al., 2008, *A&A*, 478, 83
- Veilleux S., Osterbrock D. E., 1987, *The Astrophysical Journal*, 163, L99
- Vestergaard M., Wilkes B. J., 2001, *ApJS*, 134, 1
- Virtanen P., et al., 2020, *Nature Methods*, 17, 261
- Vitale M., et al., 2013, *Astronomy & Astrophysics*, 556, A11
- Véron-Cetty, M.-P. Joly, M. Véron, P. 2004, *A&A*, 417, 515
- Weinberger R., et al., 2017, *MNRAS*, 465, 3291
- Wes McKinney 2010, in Stéfan van der Walt Jarrod Millman eds, *Proceedings of the 9th Python in Science Conference*. pp 56 – 61, doi:10.25080/Majors-92bf1922-00a
- Willmer C. N. A., et al., 2006, *The Astrophysical Journal*, 647, 853
- Wills B. J., Netzer H., Wills D., 1985, *ApJ*, 288, 94
- Wolf C., et al., 2004, *A&A*, 421, 913
- Xue Y. Q., Luo B., Brandt W. N., et al., 2011, *The Astrophysical Journal Supplement Series*, 195, 10
- Yan R., et al., 2011, *The Astrophysical Journal*, 728, 38

APPENDIX A: SPECTRAL FITTING ALGORITHM

In this appendix, we describe the spectral fitting methodology adopted in this paper. The current status of the scientific literature in this field has been summarized in the main body of the paper (see section 3). We focus here only on the methodology used.

The spectral fitting process in our code is primarily divided into two main stages: baseline fitting, and spectral peak fitting. In addition, certain data pre-processing is necessary, which constitutes a preliminary stage. The process begins with correction for atmospheric absorption features such as the telluric oxygen dip at 765 nm using methods described by [Tran & Hartmann \(2008\)](#). The spectra are then redshift-corrected using redshift values from literature (see Sec. 2 and Table 1). Once the spectrum is shifted to the rest frame, we fit the continuum baseline. For this, the characteristic emission peaks are masked to prevent them from biasing the modeling of the underlying continuum.

In the baseline fitting stage, the continuum is modeled as a combination of multiple physical components, each contributing to the observed spectrum. The continuum model includes a broken power-law component (see e.g. [Calderone et al. \(2017\)](#)) to represent the non-thermal AGN emission. This power-law is characterized by the luminosity density and spectral slope, capturing the broadband behavior of the AGN continuum. Previous studies have adopted varied strategies to model the host galaxy contribution in AGN spectra. For instance, [Shen et al. \(2011\)](#) employed a geometric mean of composite spectra, while [Calderone et al. \(2017\)](#) modeled the host using a 5 Gyr elliptical galaxy template from [Silva et al. \(1998\)](#). More recently, [Ilić et al. \(2023\)](#) utilized a linear combination of normal galaxy and quasar eigenspectra. In our code, we have incorporated a state-of-the-art host galaxy stellar contribution model, using templates from the BC03 library ([Bruzual & Charlot 2003](#)), which includes a wide range of stellar population ages (0.1 Myr to 14 Gyr) and metallicities ($Z = 0.0001$ to 0.06). During the fitting of these templates the component's parameters are varied until the differences between the data and the model are minimized, following a least squares minimization algorithm. The goodness of fit is assessed using the R-squared statistic.

In addition, the baseline model in our code also includes the Balmer series continuum. The continuum above the Balmer series limit (3646 Å), and the closely spaced discrete Balmer lines constituting a pseudo-continuum are modelled based on [Wills et al. \(1985\)](#); [Grandi \(1982\)](#); [Dietrich et al. \(2002\)](#). Lines up to principal quantum number $n = 40$ are included.

Yet another (pseudo) continuum consists of broad and narrow Fe lines, which may especially overlap with the $H\beta$ region. We incorporate these Fe II line templates. Fe II lines in the UV region are taken from the [Vestergaard & Wilkes \(2001\)](#) template, while optical Fe II features are from [Véron-Cetty, M.-P. et al. \(2004\)](#). These templates are convoluted with Gaussian profiles whose widths are constrained to physically meaningful ranges: 100–2000 km s^{−1} for narrow components and 1000–15,000 km s^{−1} for broad ones.

The construction of the baseline model proceeds by sequentially fitting and summing each component in a physically motivated order. First, the stellar continuum is modeled using the BC03 templates, establishing a baseline that reflects the host galaxy's stellar light. Next, the AGN continuum is added via the broken power-law, capturing the broad-band non-thermal contribution. The Balmer continuum is then incorporated to account for the pseudo-continuum near 3646 Å. Finally, Fe II templates are fitted and added, fine-tuning the model particularly around the $H\beta$ and UV regions. Throughout this process, masking is applied to exclude prominent emission lines to avoid bi-

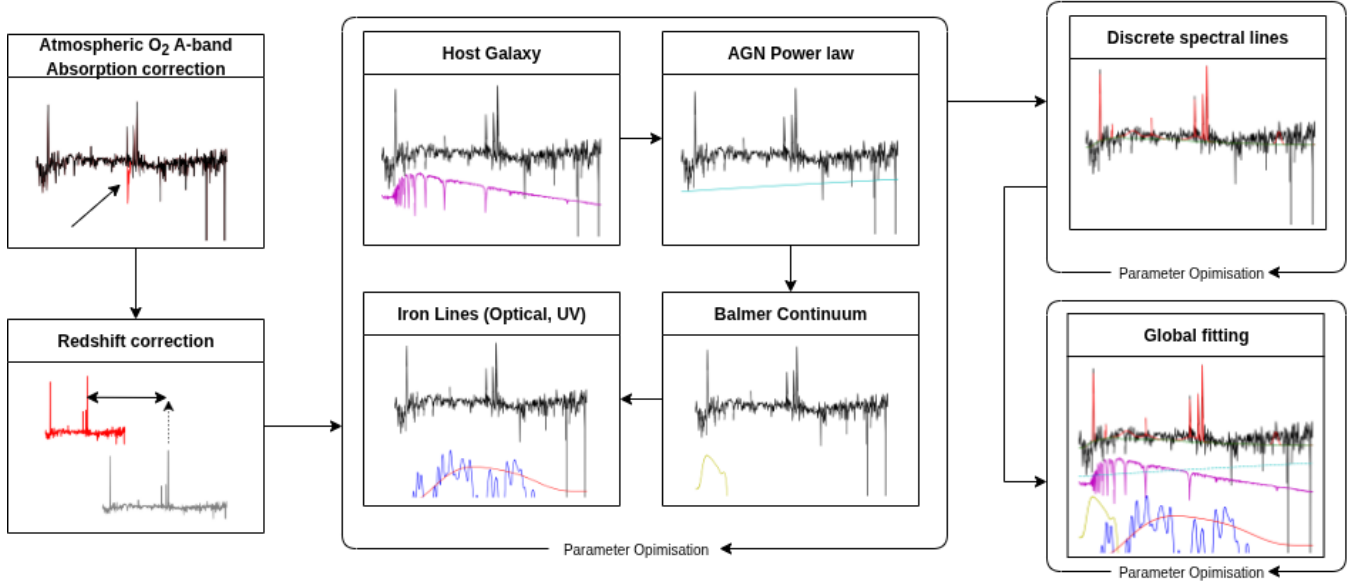


Figure A1. Flowchart illustrating the spectral fitting pipeline used in this work. The process begins with data preprocessing steps, including telluric correction and redshift shifting to the rest frame. The baseline fitting stage models the stellar continuum using BC03 templates, AGN emission with a broken power-law, Balmer continuum around the 3646 Å break, and Fe II pseudo-continuum features in the optical and UV regions. Emission lines such as H α , H β , [O III] λ 5007, [N II] λ 6584, and others are then fitted using Gaussian profiles with physically motivated constraints on width and velocity offset. A final global fit jointly refines both continuum and line parameters. The code outputs all derived quantities and fitted spectra, enabling detailed analysis of stellar populations, AGN activity, and gas properties.

using the continuum fitting. This multi-step fitting ensures that each physical component is isolated and constrained before the full model is synthesized.

Once the baseline continuum has been fitted to the masked regions of the spectrum, the final stage involves fitting the full (unmasked) spectrum, including discrete emission lines. These are modeled within localized wavelength windows using Gaussian profiles. The emission lines include [O II], [Ne III], H δ , H γ , H β , [O III], He I, H α , [N II], and [S II]. Each line is characterized by its amplitude, full-width-at-half-maximum (FWHM), and velocity offset. FWHM values are constrained between 100–2000 km s^{−1} for narrow components and 1000–15,000 km s^{−1} for broad components, while velocity offsets are limited to ± 1000 km s^{−1} (narrow) and ± 3000 km s^{−1} (broad). Subsequently, the continuum and emission-line parameters are jointly optimized in a global fitting procedure.

The code also tabulates all derived parameters and stores the fitted models and spectral plots, facilitating further analysis of the gas composition, density, and dynamics in galaxies and AGNs. Such a global fitting approach allows us to decompose complex spectra into their constituent physical components and extract key astrophysical insights.

This paper has been typeset from a T_EX/L^AT_EX file prepared by the author.