ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN ASTROFISICA

CICLO 37

Settore Concorsuale: 02/C1 - ASTRONOMIA, ASTROFISICA, FISICA DELLA TERRA E DEI PIANETI

Settore Scientifico Disciplinare: FIS/05 - ASTRONOMIA E ASTROFISICA

Hidden Supermassive Black Holes in the Deep Universe

Presentata da: Giovanni Mazzolari

Coordinatore Dottorato
Prof. Andrea Miglio

Supervisore
Prof.ssa Marcella Brusa

Co-supervisore
Dr. Roberto Gilli
Dr. Marco Mignoli

Esame finale anno 2025

A Elena

Abstract

The existence of a large population of hidden active galactic nuclei (AGN), especially at high redshift, raises fundamental questions about their origins, their co-evolution with host galaxies, and their influence on the growth of supermassive black holes (SMBH) across cosmic time. A thorough understanding of the AGN population at all redshifts is essential for testing SMBH formation models, explaining the SMBH-galaxy relations observed in the local universe, and unraveling the complexities of accretion physics at different epochs. This PhD thesis explores three key aspects of obscured AGN, with particular emphasis on high redshift: selection methods, demographic characterization, and their physical properties. I investigated these topics using a multiwavelength approach, incorporating radio, X-ray, optical, and infrared data, as well as using both photometry and spectroscopy, which is crucial for identifying and understanding these elusive sources. In the first part of this thesis, I focus on how radio observations can be used to identify obscured AGN, especially the highly obscured Compton-thick (CTK) population. I test the effectiveness of radio emission in selecting these sources both from an analytical and observational point of view, finding that with deep continuum radio observations it is possible to unveil much more heavily obscured AGN than what it is possible to do relying only on the X-rays. The thesis further investigates new spectroscopic methods for identifying obscured AGN at high redshift, where the *James Webb* Space Telescope (JWST) is opening new incredible opportunities for the investigation of these sources. I present three new diagnostic diagrams based on the [OIII]λ4363 auroral line, enabling the identification of AGN even in low-metallicity environments typical of the high-redshift universe, and providing a more effective tool for characterizing AGN in the JWST era. Finally, I describe the selection and the detailed multiwavelength analysis of narrow-line AGN (NLAGN) selected from the JWST CEERS survey, covering the redshifts range $2 < z < 9$. I will explore their spectral properties, the effectiveness of the different emission line diagnostics diagrams, and I will provide insights on their X-ray and radio characteristics. This thesis aims to improve our understanding of obscured AGN, especially at high redshifts, emphasizing the use of a comprehensive and multiwavelength approach that, together with the potential of new facilities like JWST, offers the best way to identify hidden AGN and explore their role in cosmic evolution.

Contents

List of Abbreviations

AGN	Active Galactic Nucleus/Nuclei
BH	Black Hole(s)
BHARD	Black Hole Accretion Rate Density
BLAGN	Broad-line Active Galactic Nucleus/Nuclei
BLR	Narrow-Line Region
CTK	Compton Thick AGN
CXB	Cosmic X-ray Background
EM	Electromagnetic
FIR	Far Infrared
IR	Infrared
MIR	Mid Infrared
NIR	Near Infrared
NLAGN	Narrow-line Active Galactic Nucleus/Nuclei
NLR	Narrow-Line Region
RL	Radio-Loud (AGN)
RQ	Radio-Quiet (AGN)
SED	Spectral Energy Distribution
SF	Star Formation
SFG	Star Forming Galaxies
SMBH	Supermassive Black Hole(s)

Chapter 1

Introduction

Active Galactic Nuclei (AGN) represent one of the most energetic and enigmatic phenomena in the universe, with their intense luminosities driven by the accretion of matter onto supermassive black holes (SMBHs) located at the centers of galaxies. These compact regions exhibit extraordinary energy outputs, often outshining the entire stellar population of the host galaxy. The energy emitted by AGN spans the entire electromagnetic spectrum, from radio waves to gamma rays, with bolometric luminosities that can exceed 10^{46} erg/s. The fact that SMBHs are expected to be ubiquitous in galaxies and that their mass (ranging from $10^6 - 10^{10} M_\odot$) is a consequence of the accretion is now well established but was far from being trivial in the late 80'. A key piece of evidence supporting the presence of SMBH in most galaxies came from the Soltan argument (Soltan, 1982). By assuming that the radiation emitted by AGN is powered by the accretion of matter onto SMBHs and knowing the efficiency of this accretion process, Soltan showed that the total mass density accreted during Cosmic time is of the same order as the total observed SMBH mass density in the local universe. This implies that the SMBHs observed in nearby galaxies were largely built through accretion, with AGN representing a phase in their evolutionary history. The Soltan argument provides a crucial connection between past AGN activity and the current mass density of SMBHs, supporting the notion that AGN are a significant growth phase in the life of a SMBH and of a galaxy.

Initially, AGN were classified into a wide variety of categories based on the distinct features observed in different bands of the electromagnetic spectrum, resulting in a complex "zoo" of AGN types. However, in the early 1990s, a unification model was proposed to explain this apparent diversity (Antonucci, 1993; Urry & Padovani, 1995). According to this model, the differences among AGN could be attributed primarily to two factors: the presence of an obscuring structure (often referred to as a torus) surrounding the central SMBH, and the orientation of the AGN relative to our line of sight. This simple yet powerful model suggested that much of the observed variation was due to our viewing angle rather than intrinsic differences between AGN.

More recently, advances in observational capabilities and refined theoretical models have led to a revision of this unification model. While the obscuring torus and orientation effects remain important, it is now understood that AGN properties arise from a more complex interplay of physical processes occurring over various spatial and temporal scales. Indeed, AGN are transient phenomena and can turn "on" and "off" over time in what is referred to as a duty cycle. Thus, rather than being a distinct class of astrophysical objects, AGN represent

a phase in the lifecycle of a galaxy, where the central SMBH becomes actively accreting (Kormendy & Ho, 2013; Merloni & Heinz, 2013). The effect of the AGN energetic and kinetic output on the host galaxies, representing the so-called AGN feedback, was proved to play a crucial role in shaping the galaxies' physical properties and also the large-scale cosmic structure's formation (Harrison & Ramos Almeida, 2024, for a recent review).
Therefore, It is clear that the study of AGN is critical for several reasons. By tracing the demographics of AGN we gain valuable insights into the processes that drive black hole accretion, AGN-galaxy coevolution, and large-scale structure formation (Merloni & Heinz, 2008; Silverman et al., 2009). However, detecting AGN and accurately characterizing their populations remains a challenge due to their diverse observational signatures, which can vary dramatically across the electromagnetic spectrum. AGN can be obscured by dust, diluted by host galaxy light, or exhibit weak emission in certain bands, making them difficult to detect with a single observational method. Therefore, developing new AGN selection techniques, particularly those that combine data from multiple wavelengths (e.g., optical, infrared, X-ray, and radio), is essential. A comprehensive understanding of the AGN population is crucial for building accurate models of galaxy evolution and constraining the role of AGN in the broader context of cosmic history, in a complete and unbiased way.
In the present thesis, we will address these issues, focusing in particular on the questions related to AGN demography and selection.

In the following section, we start describing the key components of the AGN structure responsible for their emission across the electromagnetic spectrum.

1.1 AGN Structure and emission

1.1.1 Supermassive Black Hole

Theoretical models describing the physical properties of SMBH suggest that black holes can be described using only three key parameters: mass, spin, and electric charge, the last being the most poorly constrained and usually assumed to be zero because of plasma neutralization (Zajaček et al., 2018).
Different techniques have been developed historically to measure the SMBH masses. Already in the '70s, in the hypothesis that the broad emission lines observed in AGN spectra (see Sect. 1.1.5) were due to photoionization driven from the accretion onto a SMBH, it was suggested that the observed lags between AGN continuum and broad-line variation were due to the light travel time from the ionizing source to the broad line region (BLR) (Bahcall et al., 1972). Therefore, the measure of the lag allowed the measure of the BLR size (the so-called reverberation mapping technique Blandford & McKee, 1982), and, assuming that the motion of the BLR clouds was dominated by the gravitational field of the central SMBH, it was possible to obtain the first measurements of SMBH masses. Reverberation mapping measurements also revealed a correlation between the size of the BLR and the AGN optical luminosity (the so-called $R - L$ relation Kaspi et al., 2000, 2005; Bentz et al., 2009), which provides a much less expensive way to estimate the size of the BLR based on the luminosity of the AGN. Subsequently, this relation has been used to develop the so-called "single-epoch virial black hole mass estimators", that allow the measurement of the SMBH mass using the AGN luminosity

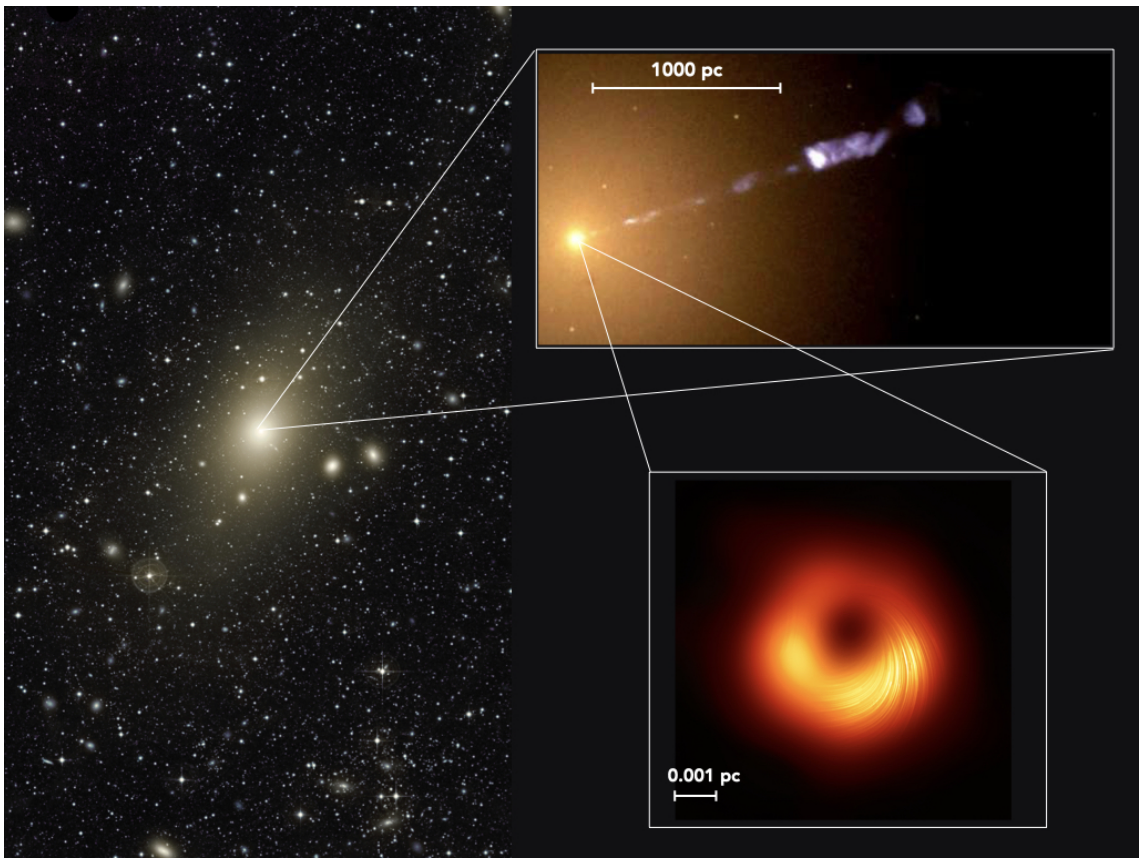

FIGURE 1.1: Multiscale representation of the massive elliptical galaxy M87. On the left the optical image of the huge halo around giant elliptical galaxy M87 (credits:ESO), on the top right Hubble imaging of the region of the radio jet (credits:STScI), on the bottom right the image of the SMBH at the center of M87 taken from Event Horizon Telescope Collaboration et al. (2021).

and the velocity dispersion of the broad emission lines (Greene & Ho, 2005; Reines et al., 2013; Reines & Volonteri, 2015, and many others). Recently, efforts from the GRAVITY collaboration (Abuter et al., 2024; GRAVITY Collaboration et al., 2024) also led to the first measure of a SMBH mass at $z \sim 2$ directly from the study of the dynamics of resolved broad line region clouds' photocenters, finding results in agreement with the predictions from virial relations. Another approach that can be pursued in the absence of spectroscopic information (or when virial measurements are not feasible) is to use SMBH scaling relations, such as the $M - \sigma$ relation (Gebhardt et al., 2000) and the Magorrian relation (Magorrian et al., 1998; Gültekin et al., 2009). These relations leverage empirical correlations between SMBH mass and the properties of their host galaxies, such as the stellar velocity dispersion or the luminosity of the bulge, and were calibrated in local SMBH samples with mass measurements coming from stellar dynamics or from reverberation mapping.
The size of a SMBH is defined by its event horizon (EH). The EH of a Schwarzschild (non-rotating) SMBH of mass M is defined as $R_{EH} = R_S = 2R_G = 2GM/c^2 \sim 1.5 \times 10^{13}(M/10^8 M_\odot)$cm. If the SMBH is also characterized by a rotation, then the space-time around it is described by the Kerr metric solution, where the dimensionless spin parameter ($0 < a < 1$) accounting for rotation, is introduced. In the case of a maximally spinning SMBH ($a = 1$), the event horizon is restricted to $R_{EH} = R_G$. Measuring the SMBH spins has been so far extremely challenging, and the methods explored (High-Frequency Quasi-Periodic Oscillations, X-ray polarimetry, study or the line profile of the Iron line, or inner disk reflection modeling) generally returned high values of a ($a > 0.5$, Brenneman, 2013), but with large uncertainties.
In 2017, the Event Horizon Telescope (EHT) collaboration (Collaboration, 2019) published the first and sharpest observation of a SMBH, the $6.5 \times 10^9 M_\odot$ SMBH at the center of the giant elliptical galaxy M87 in the Virgo cluster (see Fig. 1.1), directly catching the synchrotron emission coming from electrons at a few Schwarzschild radii from it. A detailed analysis of these observations revealed SMBH properties that are extremely consistent with those predicted by general relativity.

1.1.2 Accretion disk

What really powers AGN is the radiation emitted by the accretion of gas onto an SMBH. This process is the most efficient known process for converting mass into energy (Peterson, 1997). The luminosity associated with SMBH accretion can be written in the following form:

$$L = \eta \dot{M} c^2, \tag{1.1}$$

where $\dot{M}$ is the mass accretion rate and η is the radiative efficiency of the accretion process, i.e., the efficiency in transforming mass into radiation, and typically $\eta \simeq 0.1$. Using this value in Eq. 1.1 we find that to power fairly high-luminosity AGN of $L = 10^{46}$erg/s it needs only a $\dot{M} \sim 2 M_\odot/yr$. However, the value of the efficiency is highly uncertain and depends on different parameters and physical processes connected to the accretion mechanism.
If we assume a simple spherical accretion model of a completely ionized hydrogen gas, in order to avoid disintegration of the accreting structure, the outward force generated by the radiation pressure produced by the accretion has to be counterbalanced by the inwards gravitational force. This simple argument leads to the definition of the Eddington limit to the

luminosity of a source spherically (and isotropically) accreting:

$$L_E = \frac{4\pi G c m_p}{\sigma_T} M_{BH} \simeq 1.26 \times 10^{38} \frac{M_{BH}}{M_\odot} \text{ erg s}^{-1}, \tag{1.2}$$

where G is the gravitational constant, m_p is the proton mass and σ_T is the Thompson scattering cross-section. Equivalently, it is possible to also define an Eddington limit for the accretion rate using Eq. 1.1.
However, the assumption of spherical accretion is generally not physically supported, while the simplest physically motivated accretion disk model considers an optically thick and geometrically thin structure rotating at keplerian velocities around the central SMBH (Shakura & Sunyaev, 1973) at typical distances of $10^{-6} - 10^{-3}$pc from the central SMBH (see Fig. 1.2). Assuming the virial theorem, half of the gravitational energy lost by the gas spiraling toward the center goes into radiation, and considering that the disk is optically thick, it is possible to derive its temperature profile, which has the form: $T(r) \propto (r^{-3/4})$. Following (Bonning et al., 2007), the peak temperature of an accretion disk can be expressed in terms of the bolometric luminosity and the mass of the SMBH in the form:

$$T_{peak} = c \times 10^{5.54} M_{BH}^{-1/4} L_{bol}^{1/4}, \tag{1.3}$$

where c is a factor depending of the SMBH spin, L_{bol} is the bolometric luminosity of the AGN in erg/s, and M_{BH} is the SMBH mass in solar masses. The geometrically thin disk structure implies that the rate at which energy is advected inward is negligible compared to the fraction that is reradiated in the vertical direction (radiation-dominated accretion), determining the full accretion disk spectrum to be a composite of optically thick thermal emission spectra. The peak frequency of this multicolor black body generally resides in the range between the UV ($\sim$ 100Å) and soft X-ray ($\sim$ 100eV) and gives origin the so-called big-blue bump clearly observed in unobscured AGN (Sanders et al., 1989).
At high accretion rates, approaching the Eddington limit, the accreting gas can prevent part of the upward-flowing radiation from escaping, allowing the disk to extend in the vertical direction, becoming geometrically thick but also, potentially, optically thin. In this condition, the radiation-driven cooling timescale becomes longer than the radial advection timescale, and the accretion becomes advection-dominated. These kinds of disks, which are radiatively inefficient, are usually well described by the so-called advection-dominated accretion flow (ADAF) models (Ichimaru, 1977; Blandford & Begelman, 1999). Due to the advection and to the radial energy transfer, these disks can be generally approximated as a single temperature blackbody, reaching temperatures of $\sim 10^4$K.

1.1.3 X-ray Corona

The corona is thought to be a magnetically-powered plasma, with electrons that are accelerated, kept hot, and maintained close to the central SMBH thanks to the magnetic field energy and its specific pattern (Fabian et al., 2015; Laor & Behar, 2008). In particular, electrons in the corona are expected to reach temperatures $T \sim 10^{8-9}$K, with number densities as high as $n \sim 10^8 \text{cm}^{-3}$ but restricted in a region of few $\sim R_G$ (De Marco et al., 2013; Chartas et al., 2016). The electrons and ions constituting the ionized plasma of the corona are

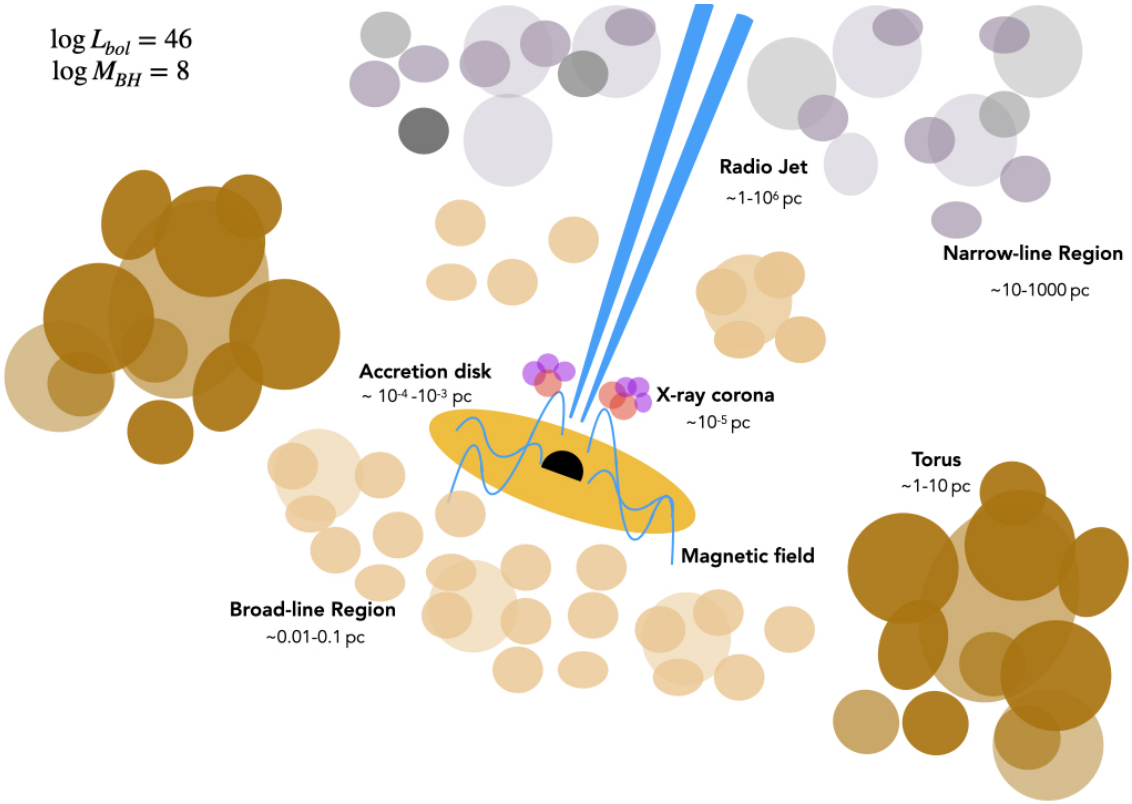

FIGURE 1.2: Skecth of the internal structure of an AGN (not in scale). Physical scales of the different components are computed for a reference AGN with bolometric luminosity $L_{bol} = 10^{46}$erg/s and SMBH of mass $M_{BH} = 10^8 M_\odot$.

thought to come from the accretion disk, where are stripped by magnetic forces. The electrons in the corona play a fundamental role in the production of the characteristic X-ray emission observed in AGN. Indeed, by means of inverse Compton scattering, the optical and UV photons radiated from the accretion disk are up-scattered by the corona electrons and gain energy (subtracting it from gas thermal energy), becoming X-ray photons (Haardt et al., 1994). The X-ray emission accounts for $\sim 10\%$ of the AGN bolometric luminosity (Duras et al., 2020), and follows a power-law spectrum of the form:

$$F_\nu = N_\nu \times h\nu \propto \nu^{1-\Gamma}, \tag{1.4}$$

where the first factor (N_ν) accounts for the distribution of the number of photons per unit of time and energy (depending on the frequency by means of the photon index Γ), while the second factor represents the energy of each photon. From the fit of the X-ray spectrum of an AGN, it is possible to retrieve the observed value of Γ, whose intrinsic range is generally between $1.7 < \Gamma < 2.1$ (Piconcelli et al., 2005; Tozzi et al., 2006), with some relevant exceptions, like Narrow-line Seyefert I (NLS1) that are generally characterized by $\Gamma > 2 - 3$ (Fabian et al., 2002). However, obscuration can significantly affect the observed AGN X-ray spectrum because absorption mostly affects soft X-ray photons (energies $< 2keV$), determining a flattening of the spectrum. X-ray photons are characterized by small optical depths compared, for example, with those of optical photons coming from the accretion disk, but can still be absorbed by metals via photo-electric absorption processes at energies $\lesssim$ 10keV (in particular by oxygen and iron species), and by Compton scattering at higher energies. It is common to parametrize the abundance of the metals absorbing the X-ray emission in terms of Hydrogen column density N_H assuming a given gas metallicity. The AGN X-ray selection and the effects of obscuration will be further described in Sect. 1.3.1.

1.1.4 Magnetic field and radio emission

The existence of a magnetic field in the innermost part of an AGN has been postulated for a long time (Blandford & Znajek, 1977). As presented in the previous paragraph, the X-ray corona is expected to be magnetically connected with the accretion disc and sustained by the magnetic energy. A magnetic field is also essential to explain AGN emission in the radio waveband, which can arise from various processes occurring at scales from the innermost regions near the SMBH to the Mpc scale, as we will see. The formation of a magnetic field in the region close to the central SMBH is expected, given the presence of plasma and the differential rotation of the accretion disc. The disk, acting as a dynamo, is expected to generate the seed magnetic field that is subsequently amplified by mechanisms such as buoyancy and magnetic instabilities (Di Matteo, 1998; Merloni & Fabian, 2002; Merloni et al., 2003).
Electrons accelerated at relativistic velocities in the presence of a magnetic field emit synchrotron radiation. Considering a relativistic electron of mass m_e and Lorenz factor γ in the presence of a magnetic field B, they emit photons at frequency $\nu \propto BE^2$, where $E = \gamma m_e c^2$. Assuming a power law energy distribution of the electrons it is possible to derive the typical power spectrum of synchrotron radiation:

$$F_\nu = \nu^{-\alpha}, \tag{1.5}$$

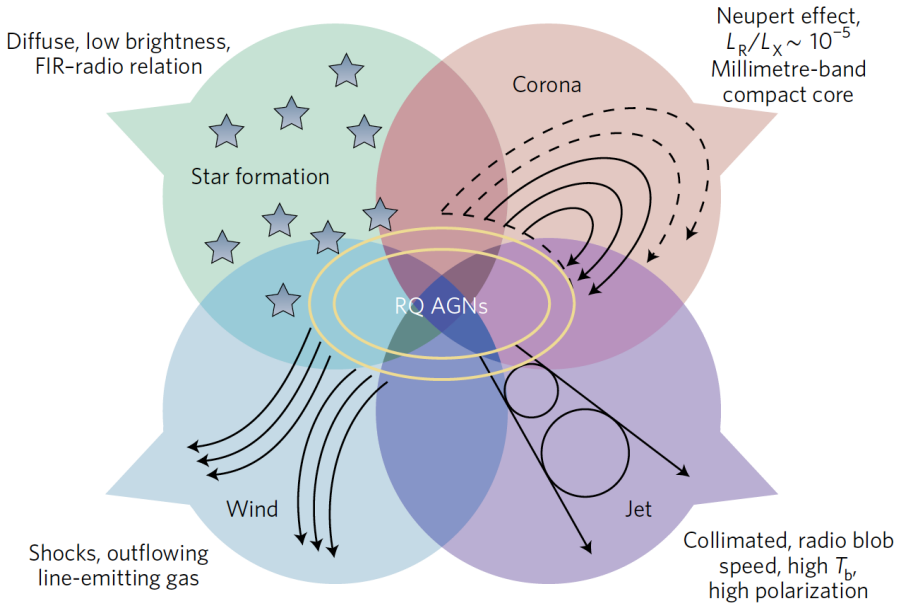

FIGURE 1.3: Taken from Panessa et al. (2019). Sketch of the possible physical processes that can contribute to the radio emission in RQ AGN.

where F is the flux, and the spectral index α typically assumes values $\alpha \simeq 0.7$ for extragalactic sources emitting optically thin synchrothron radiation (Longair, 2011).
The emission produced by synchrotron radiation is expected to be highly polarized, and by studying polarization it is possible to trace the underlying magnetic field configuration and the magnetized plasma properties, as was done by the EHT collaboration for M87 (Event Horizon Telescope Collaboration et al., 2021) and SgrA* (Event Horizon Telescope Collaboration et al., 2024).
The AGN synchrotron emission can originate from processes occurring at different scales. Radio jests, for example, are powerful radio structures that extend from close to the SMBH, possibly out to ∼Mpc scale (Rawlings & Saunders, 1991; Padovani et al., 2017). This large-scale collimated radio emission is characteristic of the population of so-called radio-loud (RL) AGN, whose emission in the radio band generally exceeds $10^{24} - 10^{25}$W/Hz (Padovani et al., 2015). On the contrary, AGN that do not show a powerful radio emission and extended radio structures are usually referred as radio-quiet (RQ) AGN. Another, but not completely overlapping, way to distinguish the two populations of radio AGN is among ´radio-jetted and ´non-radio-jetted´ sources (Padovani et al., 2017). The presence of strong radio jets is also generally associated with radiatively inefficient accretion modes that can produce a strong magnetic field, while non-jetted and RQ AGN are more frequently associated with radiatively efficient accretion disks (Heckman & Best, 2014). We will further discuss the typical features and a more precise classification of the populations of RQ and RL AGN in Sect. 1.3.1, while here we want to focus on the different physical processes giving origin to the radio emission in these sources.
While in RL AGN the expanding radio jets are expected to originate from the interaction of the electrons with the strong magnetic field amplified by the rotation of the accretion disk and of the SMBH (Blandford & Znajek, 1977), different mechanisms have been proposed in

the literature to explain the origin of the radio emission in RQ AGN (see Panessa et al., 2019, for a review, and see Fig. 1.3 for a sketch). Some high-resolution observations of RQ AGN revealed a scaled-down version of RL jets (Falcke & Biermann, 1995, on scales of a few tens of pc), where the difference between the two populations resides in a different efficiency in accelerating and collimating relativistic electrons on subparsec scales.
Another possible source of the RQ AGN radio emission can originate from inner shock winds. The observed correlation between radio luminosity and [O III] λ5007 luminosity and velocity dispersions in RQ AGN (Nims et al., 2015; de Vries et al., 2007; Zakamska & Greene, 2014) suggested that inner shock winds, driven by the accretion disk or by the AGN mechanical feedback, are able to accelerate relativistic electrons and produce synchrotron radiation on scales $>$ 100pc.
Albeit a number of works suggested that radio emission in RQ AGN can be simply explained with star-formation (SF) (Padovani, 2016; Bonzini et al., 2015), clear indications of radio emission in excess of what is expected from pure SF were observed in several samples of RQ AGN (Smolčić et al., 2017b; Delvecchio et al., 2021; Del Moro et al., 2013).
Another correlation found in different samples of radio AGN (both RQ and RL) is the X-ray - radio luminosity relation (Panessa et al., 2015; D'Amato et al., 2022; Fan & Bai, 2016, and many other works). As described in Sect. 1.1.3 the corona, at the origin of the AGN X-ray emission, is expected to be magnetically heated. The connection between X-ray and radio emission in AGN can arise if we assume, for example, that energetic electrons, emitting synchrotron radiation in the innermost AGN region, can heat the X-ray corona depositing energy that is later re-emitted in the form of X-ray emission. The $L_X - L_R$ luminosity relation was also studied in the wider context of correlation between the AGN radio emission, the AGN X-ray emission, and the SMBH mass, the so-called fundamental plane of black hole activity (Merloni et al., 2003). Many of these relations were derived not only considering SMBH, but also the accretion onto X-ray binaries and young stellar objects, generalizing the relations between the accretion properties and the emission at radio and X-ray wavelengths. The origin of the fundamental plane relation is still not fully understood, but its existence suggests that the physics characterizing the AGN accretion onto the central SMBH has an impact on both the X-ray and radio AGN emissions. The possibility of a coronal-magnetic origin of the radio (together with the X-ray) emission in RQ AGN is supported by some observational results (in particular in the mm waveband Inoue & Doi, 2014; Doi et al., 2016; Inoue & Doi, 2018) and theoretical computations (Raginski & Laor, 2016), and is expected to arise from a very compact region. High-frequency and high-resolution radio observation of AGN, reaching the mas resolution corresponding to sub-pc scales, revealed that nuclear emission shows an almost flat $\alpha \sim 0$ spectrum, while the spectral index increases moving towards external regions, reaching values around 0.7-0.8 (Chen et al., 2023, 2024). This is consistent with an optically thick (self-absorbed) synchrotron emission in the core and optically thin synchrotron emissions at larger scales, possibly associated with SF.

In Chapter 2, and Chapter 3 we will explore the potentialities of radio band in identifying AGN, in particular those most obscured, while in Chapter 5, presenting the radio analysis of a sample of high-z AGN, we will further investigate the possible mechanisms giving origin (or suppressing) the radio emission.

1.1.5 Broad and Narrow line region

The Broad-Line Region (BLR) and Narrow-Line Region (NLR) both consist of gas clouds illuminated by radiation from the inner accretion disk, but they are distributed on different spatial scales and exhibit distinct physical and kinematic properties.
The BLR clouds are distributed on scales that range $0.01 < R_{BLR} < 1$pc, are directly illuminated by the disk radiation, and are supposed to move with almost Keplerian velocities between $\sim 600 < v/\mathrm{km\ s^{-1}} < 10^4$ under the direct gravitational influence of the central SMBH. The GRAVITY collaboration, using near-infrared (NIR) interferometry at a Very Large Telescope Interferometer (GRAVITY Collaboration et al., 2024), almost spatially resolved, for the first time, the BLR photocenters in a sample of luminous local AGN, and detailed modeling revealed a disk structure in Keplerian motion, as expected from theory and reverberation mapping. As reported in Sect. 1.1.1, BLR sizes are used to compute the mass of SMBH assuming virialization and to determine the mass-luminosity relations ($R_{BLR} - L_{opt}$) over which the 'single epoch' mass measurements (based on the luminosity and FWHM of broad lines) were calibrated. The results presented by the GRAVITY collaboration were (somehow unexpectedly) consistent with previous estimates of SMBH mass and $R_{BLR} - L_{opt}$ relations.

BLR clouds are highly ionized by disk radiation and generate broad emission lines that provide crucial indirect information on the SMBH mass and intrinsic luminosity of the AGN (Stern & Laor, 2012; Reines et al., 2013; Reines & Volonteri, 2015). BLR clouds, whose size is estimated to be a $\sim 5 \times 10^{10}$cm, $\sim 0.7 R_\odot$, are extremely dense, and almost optically thick even to the X-ray radiation, reaching hydrogen column densities $> 10^{23-24}\mathrm{cm}^{-2}$ (Risaliti et al., 1999), but generally distributed with covering factors $\sim 10\%$ (Jin et al., 2012; Gaskell, 2009). BLR densities are also sufficiently high to prevent the emission of forbidden, collisionally excited lines, while their emission is typically detected as a broad component in the rest-UV and optical permitted lines like the Balmer series, C III] 1908, C IV 1557,1560, MgII] 2795,2802, Lyα 1215. AGN that show BLR emission components in their spectra are generally called Broad-line AGN (BLAGN) or Type 1, while those AGN in which the broad emission line components are not seen (because of obscuration) are also called Narrow-line AGN (NLAGN) or Type 2. The composite rest-frame optical spectra of Type 1 and Type 2 AGN are shown in Fig. 1.4.
The intensities of the broad emission lines have been proven to be directly related to the AGN bolometric luminosity, and several broad emission-line to bolometric luminosities relations have been developed over the years (Greene & Ho, 2005; Richards et al., 2006; Stern & Laor, 2012).

NLR clouds are instead distributed on larger scales compared to the BLR, potentially extending up to the Kpc scale, and constitute the most distant region directly and predominantly illuminated by the ionizing radiation from the central SMBH. For this reason, narrow emission lines can be used to identify AGN by means of narrow-line diagnostic diagrams, as we are going to present in Sect. 1.3.2.
The velocity of the NLR clouds is generally lower compared to the BLR, and typically $v_{NLR} \leq$ 650km/s. The lower densities characterizing the NLR allow the emission (and the detection) also of forbidden, collisionally excited, transitions (contrary to the BLR). Therefore, the rich information coming from NLR emission lines allows the direct investigation of the NLR gas

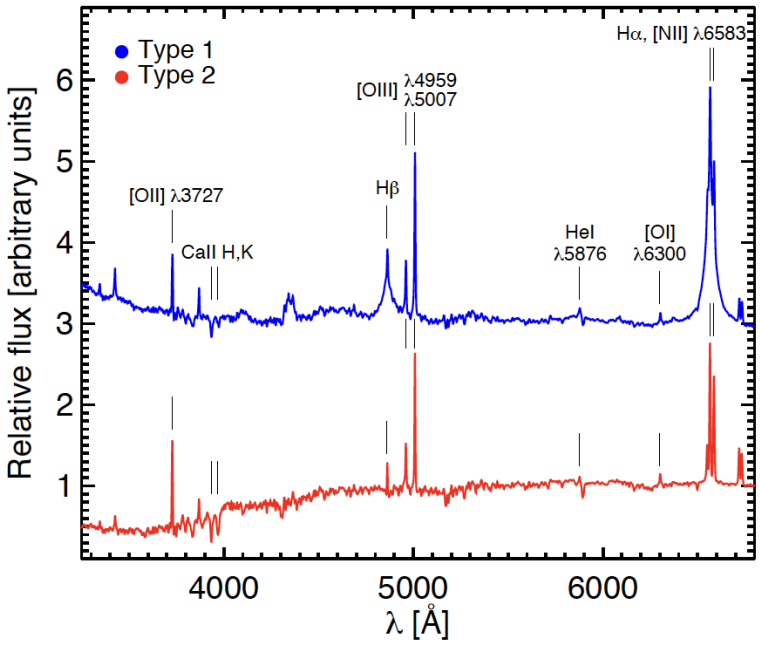

FIGURE 1.4: Adapted from DiPompeo et al. (2018). The plot shows the composite rest-frame optical spectrum of Type 1 (blue) and Type 2 (red) AGN from SDSS, with the prominent emission lines highlighted. The broad components in the Hα and Hβ emission lines are present only in the Type 1 spectrum, while narrow lines are present in both.

physical properties like temperature, densities, metallicities, and the ionization parameter (Osterbrock, 1989; Baldwin et al., 1981; Kewley et al., 2019; Maiolino & Mannucci, 2019).
The metallicity of a gas cloud, or in general of a galaxy and generally indicated with Z, is defined as the ratio between the mass of all the elements heavier than Helium to the total gas mass, and is usually expressed in terms of the solar metallicity, that is $Z_{\odot} = 0.014$. Instead, the ionization parameter (U, usually reported as $\log U$) is defined as the ratio between the photon number density of the incident radiation and the particle density. Determining these two quantities from observables is far from trivial, and several works have tried to define emission-line diagnostic diagrams or correlations able to provide U or Z based on easy-to-observe emission line ratios (Maiolino et al., 2008; Curti et al., 2017; Nakajima et al., 2022; Curti et al., 2023b; Sanders et al., 2023). In this sense, photoionization models play a fundamental role (see Sect. 1.3.2) in predicting the intensities of emission lines starting from known gas physical conditions.
NLR electron temperatures and densities are instead measured from emission lines that arise from a single ion species. In particular, temperatures are computed from collisionally excited emission lines that have very different excitation potentials so that the rate at which the different levels are populated by collision is highly temperature-dependent. The sets of lines generally involved in the electron temperature measurements are [O III] λ4363,4959,5007 and [N II] λ5755,6548,6583. In particular, we will exploit the property of [O III] λ4363 emission line in tracing the gas temperature in Chapter 4.
Electron densities are instead retrieved by measuring the intensity ratio of two lines from

a single ion, which arise in closely spaced upper states and decay to a common lower state. Common dublets used to estimate gas densities are [S II] λ6716,6731 and [O II] λ3726,3729. Since the NLR resides outside of the dust sublimation radius, another important effect to take into account when dealing with NLR emission lines is the effect of dust attenuation. Dust attenuation is not uniform across wavelengths, dimming the emission line fluxes more significantly at shorter wavelengths. We will further explore this topic in Sect. 1.3.3.
Different works investigated correlations between the AGN bolometric luminosity and the intrinsic (attenuation-corrected) luminosity of some emission lines coming from the NLR, assuming that all the emissions come from AGN-driven ionization. For example, Netzer (2009, 2019) found a correlation between the Hβ narrow-line emission and L_{bol}, while Lamastra et al. (2009) used the luminosity of the [O III] λ5007 line to derive AGN bolometric luminosity correction.

As already mentioned, in Chapter 4 and Chapter 5 we will make large use of the narrow AGN emission lines to define new NLAGN diagnostic diagrams but also to select and investigate in detail the population of high-z NLAGN.

1.1.6 Torus

Observational properties in the Mid-Infrared (MIR) part of the spectral energy distribution of AGN, as well as early AGN unification models, require the presence of a dusty structure close to the central SMBH. Recent studies show that the dusty torus is probably constituted by an evolving ensemble of dense dust clouds distributed on scales $0.1 < r < 10$pc. The inner boundary of this structure is thought to be set by the dust sublimation radius, while the outer one probably corresponds with the edge of the gravitational sphere of influence of the SMBH. The structure of the torus is not fixed, but it is thought to evolve with galaxy/AGN activity, depending on the AGN intrinsic luminosity and variability. The dust of the torus absorbs and reprocesses part of the emission radiated by the accretion disk and by the BLR, re-emitting in the 1-50μm wavelength range, where an infrared bump is typically observed in AGN Spectral Energy Distribution (SED) (Edelson & Malkan, 1986; Barvainis, 1987). The obscuration caused by the torus is important not only for the optical and UV radiation from the disk but also for the absorption of the X-ray emission coming from the X-ray corona.
A critical parameter defining the properties of the torus is the covering factor (CF), which is generally defined as the fraction of the whole solid angle covered by the obscuring structure ($\Omega/4\pi$, e.g. Hamann et al., 1993). Under the hypothesis of randomly oriented AGN in the sky, when results from X-ray surveys are considered, the fraction of obscured sources (those with $\log N_H > 22$) is generally used as a proxy of the mean CF at a given luminosity (Ueda et al., 2014; Hasinger, 2008; Merloni et al., 2014). Similarly, also the ratio between Type 2 and Type 1 AGN has been used to estimate the average CF of a population of AGN (Lawrence & Elvis, 1982).
The torus properties are thought to be dynamic, and therefore, also the CF. Indeed, some general trends of the CF with the AGN luminosity and redshift have been largely investigated in the literature. The CF is observed to decrease with AGN luminosity (Maiolino et al., 2007; Trefoloni et al., 2024), while the trend with redshift is still debated and less clear, with some works finding an increase of the CF with z (Treister & Urry, 2006; La Franca et al., 2005;

Iwasawa et al., 2012) and some other not (Lusso et al., 2013; Gilli et al., 2007, 2022). The decrease of the CF with luminosity is generally interpreted in the context of a ´receding torus model‘ (Lawrence, 1991), in which more luminous AGN progressively move the sublimation radius to outer radii and blows away more obscuring material (thanks to radiation pressure), decreasing the CF.
The existence of a clumpy and broadly axisymmetric structure around the central SMBH that obscured the AGN disk emission also has to be seen in the context of a complex environment of interactions between the SMBH and the surrounding inflowing and outflowing gas. Hydrodinamical simulations showed that the clumpy obscuring structure can potentially extend also to larger scales ($<$ 100pc), and in the polar direction (Wada et al., 2016; Hopkins et al., 2016), as confirmed by high-resolution observations of local AGN tori using MIR interferometry (Hönig et al., 2012; Tristram et al., 2014).

1.2 Obscured AGN

As described in Sect. 1.1 AGN are among the most luminous extragalactic sources known, but in some cases, their emission can be almost completely hidden at certain wavelengths due to the effect of obscuration. The definition of 'obscured AGN' depends on the wavelengths and selection technique considered, as we will further see in this section, but classically refers, from a spectroscopic point of view to the absence of detection of broad line components in AGN emission lines. Another way of defining obscured AGN is based on the absorption of their typical X-ray emission, and in particular by means of the absorbing gas column density estimated from the X-ray spectral analysis. Based on this quantity, AGN are generally classified into unobscured AGN ($N_H < 10^{22}\ cm^{-2}$), Compton-thin AGN ($10^{22}\ cm^{-2} < N_H < 10^{24}\ cm^{-2}$), and Compton-thick AGN ($N_H > 10^{24}\ cm^{-2}$).
Obscuration is thought to be not only a feature characterizing some AGN, but to represent a critical phase of the SMBH-host galaxy evolution history (Hickox & Alexander, 2018; Hopkins et al., 2008). The accretion onto the SMBH releases enormous amounts of energy through radiation, outflows and, if present, relativistic jets. This energy has a significant impact on the evolution of the host galaxy (Harrison & Ramos Almeida, 2024, for a recent review), and many theoretical models have shown the importance of AGN feedback in correctly reproducing the observed galaxy luminosity function, especially in the high-luminosity tail (Benson et al., 2003; Harrison, 2017). These models predict that most of the SMBH and galaxy growth occur when large gas (and consequently dust) reservoirs (that are needed to fuel both accretion and SF) are available, for example as a consequence of mergers, interactions, or instabilities (see Fig. 1.5). These conditions can naturally lead to a (possibly prolonged) phase of obscured AGN activity. Then, the radiative and kinetic feedback from the accretion strongly impacts on the host galaxy (Fabian, 2012), potentially suppressing SF and clearing out the SMBH obscuring material.

The identification of obscured AGN has traditionally been done using deep X-ray surveys because, as we described in Sect. 1.1.3, the medium is almost optically thin to the X-ray radiation up to CTK hydrogen column densities. Different results using both deep and wide X-ray observations show that there might be an increase in the fraction of obscured AGN with redshift. This is quite expected from a naif point of view, given that at high-z galaxies appear richer in gas and dust (Tacconi et al., 2018; Aravena et al., 2020) and also more compact

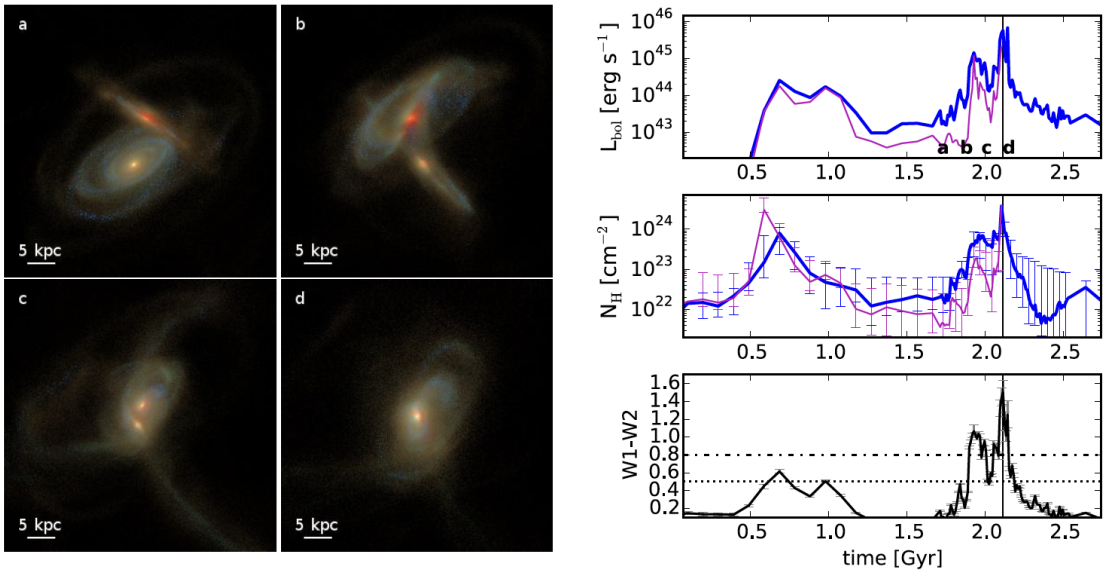

FIGURE 1.5: Taken from Blecha et al. (2016). The left panel shows the sequence of images taken from simulated galaxies at different stages of galaxy mergers. On the right are reported the panels showing the evolution with time of the AGN bolometric luminosity, obscuration, and WISE W1-W2 colors during the merger.

(van der Wel et al., 2014; Langeroodi & Hjorth, 2023), therefore denser. In particular Vito et al. (2018) showed that the fraction of AGN with X-ray luminosity $L_X > 10^{44}$ erg/s (i.e. bolometric luminosity $L_{\rm bol} > 10^{45}$ erg/s) and obscured by gas column densities of $N_H > 10^{23} cm^{-2}$ raises from $10-20\%$ in the local Universe up to $80-90\%$ at $z \sim 4$. An increasing fraction of obscured AGN was also observed by other works based on X-ray observations (Buchner et al., 2015; Aird et al., 2015; Signorini et al., 2023; Peca et al., 2022).

Cosmic X-ray Background (CXB) models (Comastri et al., 1995; Gilli et al., 2007; Ueda et al., 2014; Buchner et al., 2015; Ananna et al., 2019), by studying the integrated X-ray emission of faint extragalactic point-like sources, are able to provide an AGN census that includes also the most obscured AGN population, poorly sampled even in the deepest single X-ray surveys. The different CXB models almost agree in predicting that a consistent fraction of CTK AGN among the whole AGN population is needed to reproduce the peak of the X-ray background model at energies $E \sim 30$keV. In particular, CTK AGN are expected to be the $30 - 50\%$ of the whole AGN population (Gilli et al., 2007; Buchner et al., 2015; Aird et al., 2015; Ananna et al., 2019). However, by studying the CXB integrated emission is difficult to derive information on the evolution with redshift of the obscured AGN fraction.

As we reported in Sect. 1.1.6, a clear evolution with redshift of the torus CF has not been established yet, and some recent work suggested that the torus is actually not sufficient to justify the increase of the obscured AGN fraction observed with redshift (Trefoloni et al., 2024).

A possible driver of the evolution of the obscuration with redshift may be the interstellar medium (ISM) distribution, occurring on the galaxy (Kpc) scale. The galaxy ISM is made of

gas and dust, with a typical dust-to-gas ratio of 0.01 (Rémy-Ruyer et al., 2014). In the co-evolutionary picture of SMBH and galaxy growth, obscuration on the galaxy scale might play a significant role, in particular, in the earliest phase. Gilli et al. (2022), measuring the cosmic evolution of the physical properties of the ISM around AGN, found an ISM column density evolution with redshift $N_H(z) \sim (1+z)^{3.3}$ on the range $0 < z < 6$, which imply an AGN obscuration driven by the host-galaxy ISM 10 times larger at $z \sim 3$ than in the local universe, and approaching the CTK level at $z > 6$. Similar results were also outlined by Alonso-Tetilla et al. (2024), and also Silverman et al. (2023) using JWST/NIRCam images found that a significant contribution to the X-rays obscuration can come from host galaxy extinction already at $z \sim 2$. This scenario suggests that a large fraction of SMBH in the early universe might be obscured or heavily obscured, and may naturally explain the evolution with redshift of the obscured AGN fraction. Zoom-in simulations investigating the contribution of galaxy ISM to AGN obscuration showed that gas funneling toward the center during the active AGN phase can produce large obscuration, with most of the obscuring material distributed on scales of a few hundred parsecs (Wada et al., 2016; Blecha et al., 2018). Also the results from large cosmological simulations showed that the AGN host galaxy at high-z can be so dense and compact that the inferred column density can be as high as CTK (Ni et al., 2020).

1.2.1 The importance of obscured AGN

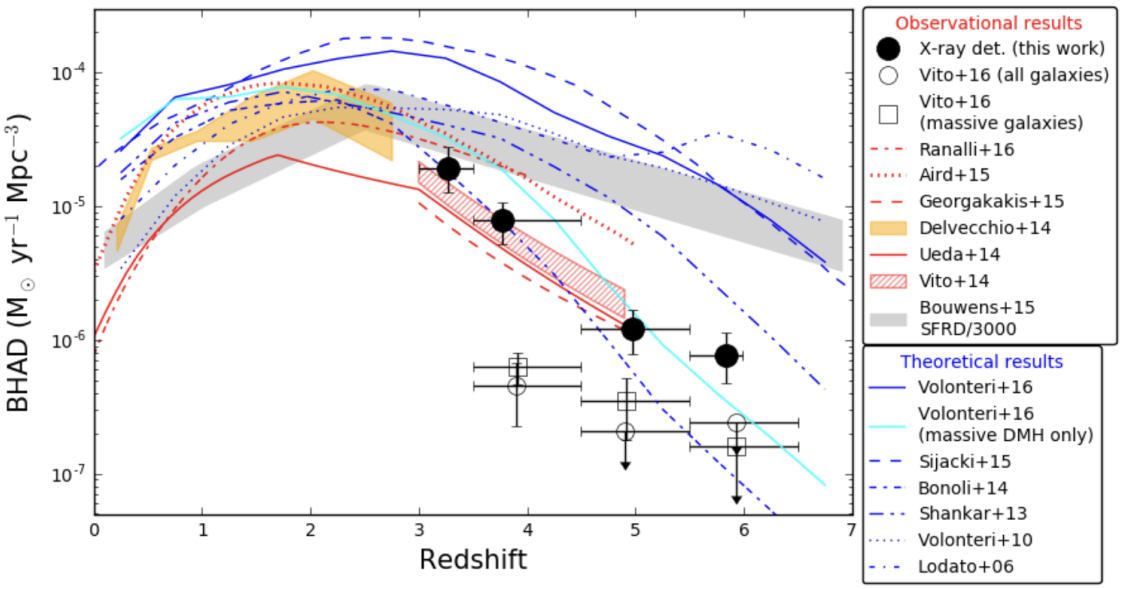

FIGURE 1.6: Adapted from Vito et al. (2018). Comparison between observational and theoretical BHARD functions, as reported in the legend on the right.

Obscured AGN represent a population of sources extremely important to investigate. Given their obscured nature, they are more difficult to identify and select, as we will further discuss in Sect. 1.3. However, efforts in the identification of obscured AGN are necessary

if we want to build up a complete census of the AGN population across cosmic times. Retrieving a complete AGN demography is fundamental both to characterize AGN properties across redshifts and for observational cosmology since we can derive information on the cosmological buildup of SMBH and on the average accretion efficiency of SMBH at different epochs. Moreover, as reported in the previous section, the phase of obscured AGN activity is invoked as a key phase in the host-galaxy SMBH coevolution framework. In this picture, most of the SMBH and host galaxy growth is expected to happen during a phase of obscured AGN activity, triggered, for example, by galaxy mergers or episodes of intense SF that made large amounts of gas available to be funneled toward the central region. Different works also showed that obscured sources are generally hosted by more massive halos, possibly explaining the obscuration with the larger availability of inflowing gas (DiPompeo et al., 2014). From the cosmological point of view, obscured AGN may also be interesting for the cosmological assembly of galaxies. Indeed Donoso et al. (2012); DiPompeo et al. (2015, 2017) have suggested that obscured sources present a higher clustering signal than unobscured sources, not driven by a host galaxy segregation bias but rather appearing as a genuine property of obscured AGN to inhabit denser environments than unobscured AGN. These findings also reveal the limitations of the traditional AGN unification models, where the obscuration is driven solely by orientation, while it might also be related to the large-scale galaxy assembly.
Investigating the population of high redshift obscured AGN can also be relevant for SMBH seeding models. For example, direct collapse seed black hole models generally predict an early obscured phase of SMBH growth, given that they require very gas-rich dark matter halos (Volonteri, 2010; Mayer et al., 2010). These models also predict the early stages of accretion to be characterized by a high-covering factor that can also imply CTK absorption.

Knowing how many AGN are accreting in a specific period of cosmic history can also provide us a clearer picture of the global AGN properties, such as the average efficiency of the accretion process. A typical quantity that is used to trace the global AGN radiation efficiency across redshift is the black hole accretion rate density function (BHARD). This function depends both on the AGN number density and on the accretion rate at which the SMBH grows. The common evolution of star-formation rate density (SFRD) and the BHAD at $z \lesssim 3$ (see Fig. 1.6) is generally considered to be a manifestation of the SMBH-galaxy coevolution. However, the X-ray BHARD declines by a factor > 10 between $z = 3 - 6$, with a steeper slope compared to the SFRD (Silverman et al., 2008; Aird et al., 2015). Different coevolution models and simulations (Shankar et al., 2014; Sijacki et al., 2015; Volonteri et al., 2016) derived BHARD functions that agree with the observational results coming from deep X-ray survey only until $z < 3$, then they predict an increasing excess of AGN with respect to what is found by deep X-ray surveys (see Fig. 1.6). If X-rays are really able to probe the complete AGN radiation output, it is necessary to scale down the high-z SMBH growth in the theoretical models, which, however, would then hardly reproduce the observed AGN number density in the local universe. One plausible solution to solve the tension between BHRD derived from simulations and X-ray observations is the existence of a heavily obscured AGN population at $z > 3$, where the effects of obscuration are expected to be stronger, which is almost completely missed by X-ray surveys (Barchiesi et al., 2021; Yang et al., 2023a; Akins et al., 2024). This possibility, implying a substantially higher SMBH mass density at early times, may not necessarily imply a significantly lower accretion efficiency. Indeed, Comastri et al. (2015) showed that the total radiation output from accreting SMBH could be increased

even by a factor of four without violating the X-ray or the infrared (IR) backgrounds prescriptions, if the additional accretion is related to heavily obscured sources ($\log N_H > 25$). Consequently, the existence of such a population of AGN at these redshifts would keep the global radiation efficiency close to the theoretically expected value $\eta \sim 0.1$ and at the same time solve the high-z tension between models and X-ray observations.

Recovering a complete census of the AGN demography across different epochs is fundamental to testing these scenarios and verifying the correctness of theoretical predictions, from the unification model to the seed black hole formation mechanisms. Therefore, it is necessary to investigate the selection techniques that have been developed so far to identify these sources across the EM spectrum.

1.3 AGN selection

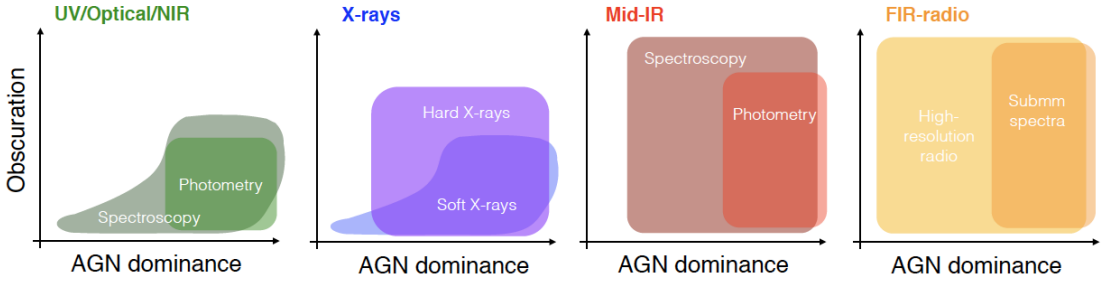

FIGURE 1.7: Taken from Hickox & Alexander (2018). Representation of the effectiveness of the AGN selection techniques developed in each band of the EM spectrum as a function of AGN obscuration and AGN dominance over the light emitted by the host galaxy.

The AGN accretion processes determine the Spectral energy distribution (SED) of an AGN to be characterized by peculiar features across the EM spectrum that generally allow us to distinguish these sources from normal galaxies. However, the effectiveness of the different selection techniques depends on the physical properties of the AGN, like the obscuration, the accretion rate, the luminosity, and also the metallicity of the gas surrounding the SMBH. In particular, obscured AGNs are more challenging to identify than unobscured AGNs for two main reasons. First, the obscuring material dims the AGN UV and optical emission, reducing the possibility of catching the typical AGN features in the SED. Second, the absorption of most of the AGN-related emission allows star formation processes in the host galaxy to more easily overwhelm and dilute the residual AGN contribution, making it difficult to distinguish from a normal SFG. We will further discuss this aspect in Sect. 1.3.
It is helpful to consider two distinct terminologies when discussing selection techniques: reliability and completeness. The first refers to the ability of a selection method to correctly identify AGN, an efficient technique implies a high purity of the selected AGN sample. The second refers to the ability to select a complete sample of sources; a highly complete selection means that it was possible to select nearly all the AGN present in the considered sample of sources. In Fig. 1.7 I show a sketch representing the effectiveness (combination of reliability

and completeness) of selection techniques in different bands of the EM spectrum as a function of AGN obscuration and AGN dominance over the host galaxy.

1.3.1 Photometric selections

Optical and MIR colors

Unobscured AGN are characterized in the rest-frame optical and UV band by emission dominated by the accretion disk, which generally outshines the emission coming from SF in the host galaxy. Different color-color diagrams based on the optical and UV bands have been traditionally developed to identify AGN (Sandage et al., 1965; Richards et al., 2002, 2004), and are based on the identification of those sources showing the Big-Blue Bump excess (see Sect. 1.1.2). A specific class of color-color diagrams was also developed to select sources at a given redshift (generally at $z > 3$) when spectroscopic redshifts are not available. These diagrams take advantage of the flux drop blueward of the rest frame wavelengths of the $Ly\alpha$ line due to the absorption by the increasing amount (with redshift) of neutral hydrogen in the intergalactic medium (Ono et al., 2018; Venemans et al., 2015; Bañados et al., 2016).
Constraining both the redshift and the AGN nature using sets of rest frame UV-optical diagnostics is feasible for unobscured sources, but they are not suited for the population of obscured AGN, whose AGN continuum emission in this wavelength range is generally largely suppressed by obscuration and overwhelmed by SF emission.
However, as reported in Sect. 1.1.6, in sources where the dusty torus obscures the disk radiation, the reprocessed radiation is partially re-emitted in the MIR, producing the AGN characteristic emission in the 5-40μm band. Different MIR color-color diagrams were developed to identify obscured AGN, like those reported in Stern et al. (2005); Donley et al. (2012); Assef et al. (2013). Given that the torus is expected to be present both in obscured and unobscured AGN, MIR selection allows for the selection of both obscured and unobscured AGN, contrary to optical-UV color diagrams. However, also MIR color selections have stringent limitations. Indeed, they are generally able to select the most luminous AGN and only up to $z \sim 2 - 3$ because, at higher redshifts, SFG colors become similar to those of AGN, and the selection is no more effective (Messias et al., 2014).

X-ray emission

As reported in Sect. 1.1.3, a typical tracer of the AGN emission are X-ray photons. X-ray emission from star formation processes (X-ray binaries) has been found to depend on the stellar mass and star formation rates of galaxies, and very rarely exceeds $L_X > 10^{42}$erg/s (Lehmer et al., 2016). Therefore, it became customary to consider that only AGN are able to justify X-ray luminosities above this threshold. As already discussed, AGN X-ray selection is generally highly reliable but not necessarily complete, especially when the population of the most obscured AGN is considered. The detection in X-ray spectra of a strong reflection component at $E > 10keV$ (and consequently an almost flat X-ray spectral index) and of prominent Kα emission lines, generally suggest signatures of CTK obscuring material.

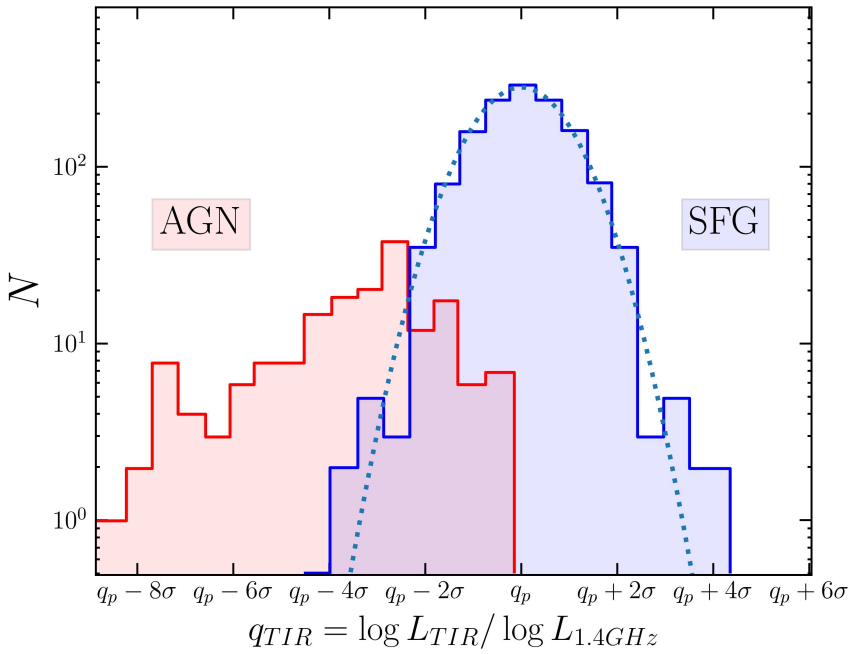

FIGURE 1.8: Adapted from Delvecchio et al. (2021). Distribution of the IR-to-radio luminosity ratio $q_{\rm TIR}$ for SFG (blue) and AGN (red). The value $q_p = 2.64$ corresponds to the peak of the distribution, dominated by SFGs, while $\sigma = 0.22$ is the dispersion of the Gaussian that fits the SFG population.

Radio emission

As already presented in Sect. 1.1.4, AGN are also characterized by a radio emission that is produced by the physical processes connected to the accretion mechanisms and/or by the large-scale relativistic jets. A great advantage of using radio emission to select AGN is that it is generally largely unaffected by obscuration (Hildebrand, 1983) and, therefore, extremely powerful in detecting the most heavily obscured sources.
The first radio surveys were mostly sensitive to the powerful ($L_{1.4\,\rm GHz} > 10^{25}$ W Hz^{-1}) population of RL AGN (White et al., 1997; Becker et al., 1995, 2001), for which a significant fraction of the power produced by the accretion processes is released in kinetic form through the formation of relativistic jets that may expand up to the Mpc scales. However, new-generation radio surveys, reaching sub μJy sensitivity, are deep enough to detect also the radio emission coming from faint and distant RQ AGN (Heywood et al., 2020; Alberts et al., 2020; van der Vlugt et al., 2021; Hale et al., 2023).
One concern for radio AGN selection is contamination from radio emission driven by SF in the host galaxy. At frequencies around a few GHz (those typically explored by current radio surveys), the radio emission is, both for AGN and SFG, mostly due to optically thin synchrotron radiation emitted by electrons accelerated at relativistic velocities by different acceleration mechanisms. In SFG, the acceleration is provided by supernova explosions in SF regions, while in AGN by the mechanisms described in Sect. 1.1.4. A way to distinguish between SF or AGN-related radio emission is provided, for example, by the well-known Far Infrared Radio Correlation (Novak et al., 2017; Delvecchio et al., 2021). This relation arises because the same population of massive stars that heats up dust, causing it to reradiate its

energy in the far-infrared (FIR), produces supernovae that generate relativistic particles emitting synchrotron radiation at radio frequencies. This correlation is quantified by a parameter

$$q_{\rm TIR} = \log(L_{TIR}) - \log(L_{1.4GHz}), \tag{1.6}$$

that is the ratio between the total IR luminosity integrated between 8-1000μm and the 1.4GHz luminosity (see Fig. 1.8). This correlation has been found to depend both on the redshift and the stellar mass (M_*) of the sources (Delhaize et al., 2017; Delvecchio et al., 2017, 2021). Galaxies with radio emission in excess of what is predicted by the FIR-radio correlation (therefore showing low values of $q_{\rm TIR}$) likely possess an AGN-driven radio component, independently of which exact mechanism is responsible for it. It is common to indicate these sources as radio-excess AGN. In Chapter 3 I will use the radio-excess selection to investigate the presence of obscured AGN candidates in a field covered by a deep radio observation.
It is important to note that selecting radio-excess AGN does not mean selecting RL AGN. Indeed, the former are simply AGN characterized by an excess in radio emission compared to the emission due to star formation, while RL AGN are classified comparing the radio emission to the AGN-related emission in another band, for instance, at optical or X-ray bands (see Kellermann et al. 1989; Terashima & Wilson 2003). Nevertheless, RL AGN are typically part of the radio-excess selected AGN, as they also satisfy the radio-excess definition.
There are also other parameters defined in the literature to quantify the radio excess of sources, and are based on the ratio between the radio-luminosity or radio flux density (or a directly related quantity, such as the radio SFR) and a quantity that should not be contaminated by the AGN emission, such as the 24μm flux density or the SFR inferred from NIR/FIR emission (Padovani et al., 2011; Bonzini et al., 2013; Arango-Toro et al., 2023).
There are also other criteria that can be used to classify AGN based on their radio properties. One, obvious, is the detection of the extended radio morphology typical of RL AGN lobes or a radio luminosity higher than 10^{25} W/Hz, as these powers can not be produced by SF. Furthermore, if radio emission at different frequencies is available, a flat radio spectral index generally indicates a synchrotron self-absorbed emission and, therefore, a compact source. This is highly indicative of the presence of an AGN given that star formation radio emission is diffuse, being produced at larger scales (Chen et al., 2023).

SED-fitting

A natural evolution of single-band photometric AGN selection is broadband SED-fitting. Several SED fitting codes have been developed to find the best model capable of minimizing the differences with the observed broadband photometry, which can range from the radio to the X-rays. Two main classes of SED-fitting codes have been traditionally developed: template-based and parametric or module-based SED fitting codes (Thorne et al., 2021; Pacifici et al., 2023; Newman & Gruen, 2022). More recently, machine learning-based SED fitting codes have also been designed (Davidzon et al., 2019; Simet et al., 2021). The codes in the first class consider sets of real or synthetic SED of galaxies and AGN (or their linear combination) and look for the solution that best fits the source photometry. Instead, the second approach is based on the reconstruction of the source SED considering the emission coming from each possible emission mechanism that can characterize that source, which is generally parametrized in the code by a grid of variables and models.

The SED fitting codes can be used for multiple scopes: deriving photometric redshift when no spectroscopy is available, estimating the physical properties of the sources (like obscuration, SFR, stellar masses, luminosity, accretion parameters, etc.), and identifying any possible AGN component in the fit. Among the SED fitting codes that were mainly developed to derive photometric redshift there are `EAZY` (Brammer et al., 2008), `Lephare` (Arnouts et al., 1999; Ilbert et al., 2006), and `Hyperz` (Bolzonella et al., 2000), which are all template-based. Instead, among the most common SED-fitting codes mainly used to derive galaxies' physical properties there are `CIGALE` (Boquien et al., 2019; Yang et al., 2020), `MAGPHYS` (da Cunha et al., 2008), `Prospector` (Leja et al., 2017), and many others. Different studies showed that when photometric catalogs cover enough wavebands, the results from different SED-fitting codes and fitting strategies are reasonably consistent, in particular for what concerns the main physical properties: photometric redshift, stellar mass, star formation rates, and obscurations (Dahlen et al., 2013; Pacifici et al., 2023).
The clear advantage of using SED fitting techniques to select AGN is that they can account for more photometric information than only a few broad-band fluxes, allowing us to obtain more robust results. Furthermore, if the available photometry covers a sufficiently broad range of the EM spectrum, SED-fitting techniques are able to catch both obscured and unobscured AGN emission features. However, the need for a wide photometric coverage can also limit the use of SED-fitting codes to select AGN since photometry across a large portion of the EM spectrum is not always available.

1.3.2 Spectroscopic selection

A generally highly reliable method for identifying AGN consists in the analysis of their spectral features. While spectroscopy is much more time-consuming than photometry, especially for large samples and with current technology, it is highly valuable for smaller samples. For instance, it can be used to validate the accuracy of photometric AGN selection techniques or to measure additional properties that cannot be inferred only by photometry.

Diagnostic diagrams

As we reported in Sect. 1.1.5, from the emission lines FWHM, fluxes, or the ratio between different emission line fluxes it is possible to derive information on the physical properties of galaxies, such as the velocity dispersion of the gas, the density, the ionization parameter or the temperature of the gas. One clear spectroscopic evidence for the presence of an AGN is given by the presence of a broad emission component in the permitted lines (for example, Hα Hβ) not detected in the forbidden lines (such as the [O III] λ5007), which suggests an emission directly coming from the AGN BLR (where collisionally excited lines are suppressed). This allows the selection of BLAGN (or Type 1). When no broad emission lines are detected, the spectrum can still hide an obscured, narrow line, AGN (NLAGN, also called Type2), that can be identified using specific ratios between line fluxes. In Baldwin et al. (1981) were firstly defined three effective narrow-line diagnostic diagrams able to separate the NLAGN population from SFG (see Fig. 1.9). These diagnostics are based on the ratio of ionized forbidden lines to the neutral permitted line fluxes. The harder (i.e., more energetic) ionizing radiation produced by the accretion disk in AGN generally determines larger ratios of these lines compared to SFG and allows the separation of the AGN population from the

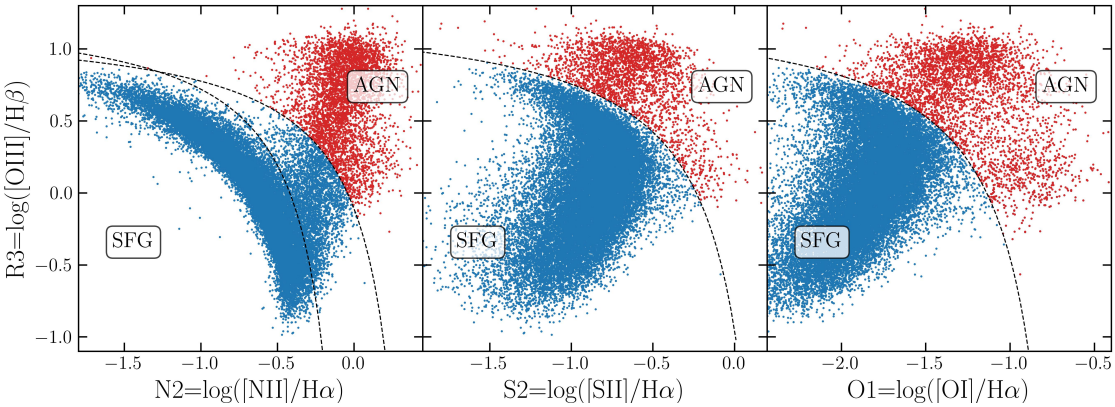

FIGURE 1.9: Optical 'BPTs' emission-line diagnostic diagrams used to spectroscopically separate AGN from SFG. They were first introduced by (Baldwin et al., 1981). The demarcation lines are those reported in Kewley et al. (2001).

other species by means of demarcation lines. The emission lines involved in the line ratios are usually chosen close to each other in order to reduce the effect of dust reddening, which has the effect of artificially increasing the ratio between far-away lines.
Thanks to the hundreds of thousands of spectra made available from the Sloan Digital Sky Survey (SDSS York et al., 2000), we are able to fully populate these diagnostic diagrams and even investigate the different positions occupied by the SFG and AGN populations when they are characterized by different physical conditions (temperature, densities, metallicities, etc.). Indeed, diagnostic diagrams are sensitive to these properties, and also the demarcation lines between the population of SFG and AGN can change significantly by varying the physical conditions of the sources, potentially up to the point of becoming useless in certain conditions (as we will see in Sect. 1.4.2). This topic will also be addressed in Chapter 4, where I will describe the definition of new NLAGN diagnostic diagrams based on the [O III] λ4363 emission line.

Photoionization models

Tracking the different positions of AGN and SFG in a diagnostic diagram according to the variation of gas parameters (temperature, densities, metallicities, ionizing radiation, etc.) is not an easy task. If we want to consider only observational samples, we need large spectroscopic surveys, and we have to derive these parameters by exploiting correlations calibrated for 'standard' conditions (Kewley et al., 2019). Therefore, the quickest approach to trace the evolution of the position of sources in diagnostic diagrams with the variation of their physical properties is to use photoionization models. These models are produced by codes such as `MAPPINGS` (Sutherland & Dopita, 1993; Allen et al., 2008; Dopita et al., 2013) and `Cloudy` (Ferland, 1993; Ferland et al., 2013), which, starting from a large grid of initial conditions of the gas and on the incident radiation (stellar or accretion driven), return the intensities of selected emission lines produced when the stopping conditions are reached. Once the results of photoionization models are calibrated and tested against observational samples, they can be used to study the position of specific types of astrophysical sources in diagnostic diagrams according to the variation in metallicities, ionization parameters, density, or temperature,

and even explore ranges of these parameters not already probed by the observational campaigns.
In Chapter 4 I will take great advantage of the distribution of photoionization models to define new AGN diagnostic diagrams based on the [O III] λ4363 line.

1.3.3 Obscuration from spectra

When spectroscopy allows the detection of multiple lines belonging to the Balmer series, a common approach to derive the average amount of obscuration affecting the AGN NLR or BLR or the galaxy ISM is to use the Balmer decrement technique (Calzetti et al., 1994, 2000, and many others). If these lines show both a broad and narrow component, it is necessary to disentangle the two components and apply the Balmer decrement separately to the broad and narrow lines. This technique takes advantage of the fact that the intrinsic ratios between Hα , Hβ , and Hγ fluxes are almost fixed, with only small variations with the temperature and density of the gas. Generally, CASE-B recombination (Osterbrock, 1989) is assumed that for typical electron temperatures $T_e = 10^4$K and electron densities $n_e = 10^2$cm^{-3}, and for a completely dust-free HII region, it implies Hα /Hβ =2.863 and Hγ /Hβ =0.47. Given that the attenuation of line fluxes due to obscuration is wavelength-dependent, with bluer wavelengths being more affected than the red ones, it is possible to correlate the deviation of the observed Hα /Hβ (or Hγ /Hβ) ratio to the average amount of obscuration A_V, once a specific attenuation curve is fixed. Given that the Balmer lines are in the rest-frame optical part of the spectrum, they are sensitive to moderate amounts of obscuration (up to $A_V \sim 4$), but for larger amounts the Hβ line can be completely suppressed, especially if the source is intrinsically faint. However, it is possible to use the same approach also with the lines from the Paschen series, which are at longer wavelengths and, therefore, less affected by obscurations.
It is possible to retrieve an indication of the amount of AGN obscuration also taking advantage of correlations that have been found between the observed luminosity of some forbidden lines, like for example the [O III] λ5007 that is generally quite bright even in obscured systems, and the observed X-ray luminosity (Maiolino et al., 1998). By assuming the AGN unified model, the ratio between the hard X-ray luminosity and an isotropic indicator of the intrinsic luminosity should provide indications on the amount of absorption affecting the nuclear X-ray source. The luminosity of the [O III] λ5007 line can be considered an isotropic indicator of the nuclear intrinsic luminosity, one corrected for the dust extinction deduced from the Balmer decrement. The effect of a high absorbing column density is to lower the $L_{2-10keV}/L_{[OIII]}$ ratio with respect to what is observed for unobscured systems.

1.4 New windows opened by JWST

The James Webb Space Telescope (JWST) was launched on December 25, 2021, and from the first scientific results published in late 2022, it became clear that we now have the opportunity to explore the early epochs with unprecedented sensitivity and resolution.
JWST is a large (6.6 m), cold (<50 K), infrared-optimized observatory (Gardner et al., 2023) with a hexagon segmented mirror design. JWST is located near the Lagrange point L2, offering a stable environment for its scientific operations. The telescope is equipped with four

advanced instruments: NIRCam, NIRSpec, NIRISS, and MIRI, allowing it to perform both photometry and spectroscopy across a broad wavelength range of approximately 1 to 25 microns. JWST's versatile capabilities are further enhanced by its ability to operate in five distinct observing modes. These include imaging with NIRCam, MIRI, and NIRISS; slit spectroscopy with NIRSpec and MIRI; slitless spectroscopy with NIRCam and NIRISS; integral field unit (IFU) spectroscopy using NIRSpec and MIRI, which provides spatially resolved spectral data; and coronographic imaging with NIRCam and MIRI, designed to block out the light from bright sources like stars, enabling detailed studies of nearby faint objects such as exoplanets. The combination of multiple instruments, observing modes, and wide wavelength coverage makes JWST exceptionally versatile and powerful. It is capable of addressing a wide range of scientific objectives, from probing the earliest galaxies in the universe to studying the atmospheres of nearby exoplanets in exquisite detail. Its imaging and spectroscopic capabilities represent a significant leap forward compared to previous observatories, enabling groundbreaking discoveries across many areas of astrophysics.
The extremely sharp resolution of JWST images (0.031"/pixel and 0.063"/pixel for NIRCam short and long wave filters, respectively, and a diffraction-limited PSF FWHM of ~ 2 pixels) allows us to investigate the composite structure of galaxies up to $z > 6$ (Witten et al., 2023; Huertas-Company et al., 2024) and to perform extremely detailed studies even on single star-forming clumps or dust structures in galaxies in the local universe ($\sim$ 10pc resolution Lee et al., 2023). At the same time, JWST spectroscopy, thanks to the NIRSpec 0.8-5.3μm coverage, allows the investigation with unprecedented sensitivity of the rest frame UV and optical spectroscopic properties of galaxies up to $z \sim 14$ (Carniani et al., 2024; Curtis-Lake et al., 2022; Robertson et al., 2022).

1.4.1 Properties of high-z AGN

Within this context, recent studies, exploiting both spectroscopic and imaging data from JWST, have revealed a large population of AGN at high redshift (Kocevski et al., 2023; Übler et al., 2023; Übler et al., 2024; Matthee et al., 2023; Maiolino et al., 2024b, 2023b; Greene et al., 2023; Bogdán et al., 2023; Goulding et al., 2023; Kokorev et al., 2023; Furtak et al., 2023; Juodžbalis et al., 2024a; Scholtz et al., 2023b; Chisholm et al., 2024), providing the unique opportunity to study the properties of SMBH and the AGN-galaxy coevolution, since very early times.
Different works, selecting AGN at high-z using JWST NIR and MIR photometry, highlighted that the AGN population at $z > 3$ might probably be larger than previously expected and dominated by obscured or heavily obscured sources. Yang et al. (2023a), taking advantage of the JWST-MIRI photometry of the Cosmic Evolution Early Release Science Survey (CEERS; Finkelstein et al., 2022), investigated the AGN population using SED modeling, and found a BHARD at $z > 3 \sim 0.5$ dex higher than what was expected from previous X-ray AGN studies (Vito et al., 2016, 2018), and more in line with theoretical models (see Fig. 1.10 to be compared with Fig.1.6). Similarly, Akins et al. (2024), by selecting AGN candidates with NIRCam and MIRI color-color diagnostics on the 0.54 deg^2 of the COSMOS-Web survey, derived an almost flat BHARD from $z \sim 2 - 9$ (even if galaxy contaminants can artificially increase the BHARD). Lyu et al. (2024) performing a multi-wavelength AGN selection also involving the JWST/MIRI data of the SMILES survey (de Graaff et al., 2024) selected a

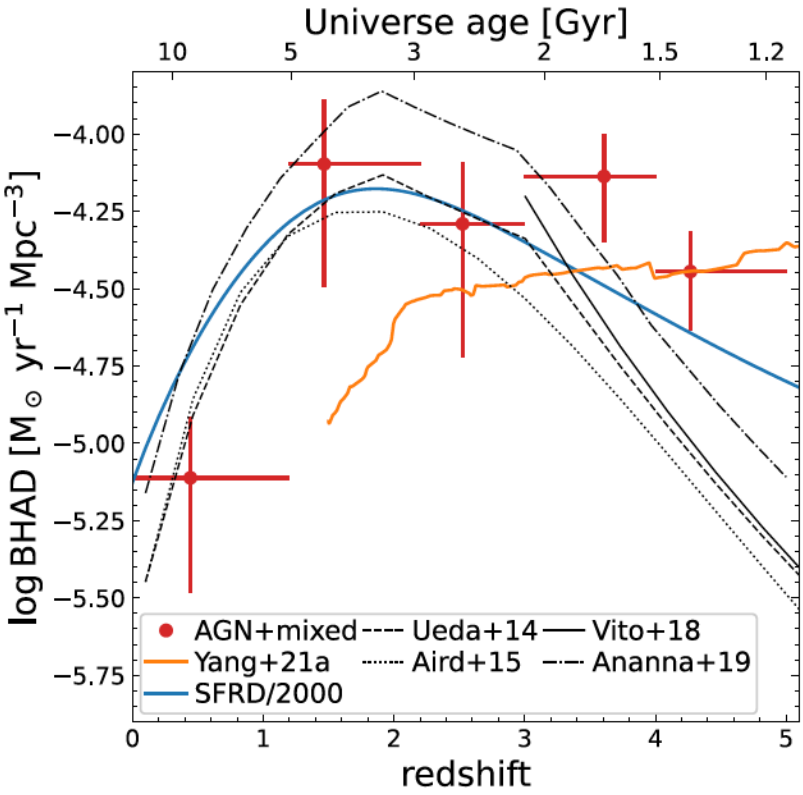

FIGURE 1.10: Taken from Yang et al. (2023a). BHARD as a function of redshift. The red data points are derived based on the MIRI-selected AGN on the CEERS survey. The orange curve represents the inferred BHAR for bulge-dominated galaxies and is taken from Yang et al. (2021). The black curves are from X-ray measurements with references marked in the legend. The blue curve represents SFRD (Madau & Dickinson, 2014) downscaled by a factor of 2000

remarkable fraction of AGN among the MIRI detected sources ($\sim 7\%$) and found a statistically significant increase of the obscured AGN fraction with redshift, confirming the results of other pre-JWST works (Signorini et al., 2023; Gilli et al., 2022; Buchner et al., 2015; Aird et al., 2015).
The existence of a larger-than-expected AGN population at early times was also shown by spectroscopic studies. For example, Maiolino et al. (2023b) and Harikane et al. (2023), by selecting BLAGN among JWST/NIRSpec spectra of the JADES and CEERS survey, found significant AGN excess at $z > 4$ with respect to the AGN luminosity functions previously derived using X-ray data (Giallongo et al., 2019).

JWST photometry and spectroscopy also enabled the identification of an apparently new type of sources, the so-called Little Red Dots (LRD). Even if a proper definition of this class of sources is not yet been established, these galaxies are generally unresolved by JWST PSF and characterized by a steep red continuum in the rest-frame optical, and by relatively blue colors in the rest-frame UV (Kocevski et al., 2023; Harikane et al., 2023; Matthee et al., 2023; Greene

et al., 2023; Killi et al., 2023), therefore featuring a sort of "V-shaped" SED. They also generally show large obscuration from Balmer decrement analysis ($A_V > 3$). The exact nature of these sources is not well established yet, and it is probable that this class contains both heavily-obscured AGN and dust-obscured starburst galaxies (Pérez-González et al., 2023). Some recent works (Greene et al., 2023; Kocevski et al., 2024), performing follow-up spectroscopy of some of these sources, demonstrated that over 80% of photometrically-selected LRDs show broad-line emission, indicative of an AGN nature, and that their number density peaks around $z \sim 8$, then experiencing a rapid decline below $z \sim 5$. The redshift distribution of these obscured nuclei is consistent with the coevolutional scenario, described in Sect. 1.2.1, in which the gas, fueling the SMBH, and the dust, that determines the obscuration, are initially tied to the dissipative gas collapse that drives the inside-out growth of the galaxy and the accretion onto the SMBH during early epochs. The subsequent decline of this population, instead, would be due to star formation moving to larger scales, in the ISM of the host galaxy, resulting in less dust being deposited near the central SMBH.

JWST data not only allow the identification of a higher fraction of AGN at high-z but also show some new and unexpected features in this early population of SMBHs. Considering the population of spectroscopically confirmed BLAGN discovered with JWST, they have bolometric luminosities $L_{bol} \sim 10^{44-45}$erg s^{-1}, which means 2-3 dex below that of bright quasars identified by ground-based surveys at similar redshifts (Willott et al., 2010; Jiang et al., 2016; Mazzucchelli et al., 2017; Matsuoka et al., 2019). Furthermore, they are generally powered by SMBHs with masses typically of 10^{6-7} M$_\odot$, making them among the least-massive SMBHs known in the early Universe. These faint quasars are probably more representative of the global population of SMBH at high redshifts and, given the importance of retrieving a complete census of the AGN population at different epochs (see Sect. 1.2.1), they can be crucial in constraining models of SMBH seeding (Pacucci & Loeb, 2022; Li et al., 2023), to investigate the contribution of AGN to hydrogen reionization (Dayal et al., 2020; Yung et al., 2021; Giallongo et al., 2019), and the early coevolution of galaxies and SMBHs (Habouzit et al., 2022; Inayoshi et al., 2022; Pacucci et al., 2023).

Different studies have also revealed that this population of early SMBHs detected with JWST is generally overmassive relative to the host galaxy stellar mass when compared with the local AGN distribution (Maiolino et al., 2023b; Bogdán et al., 2023; Furtak et al., 2023; Juodžbalis et al., 2024a), and considering the local AGN scaling relations (Reines & Volonteri, 2015). This might suggest that the early stages of the SMBH - host galaxy coevolution can be dominated by a first phase of SMBH growth followed by a rapid phase of star formation (see Fig. 1.11). Alternatively, this could also suggest that the preferred driving channel for the SMBH formation could be the so-called "Direct Collapse Black Hole" scenario, together with episodes of super-Eddington accretion (Scholtz et al., 2023a; Lupi et al., 2024). However, estimates of the host galaxies' stellar masses are still highly uncertain, given the difficulties in decomposing the AGN light from stellar contribution. Therefore, considering also the intrinsic scatter of the relation used to derive the SMBH masses, the uncertainties and degeneracies are large, even higher than 1 dex. By using spectroscopic data at high-spectral resolution to estimate dynamical masses, Maiolino et al. (2023b) pointed out that actually early BHs might be much closer to the local $M_{BH} - M_{dyn}$ relation (as compared to

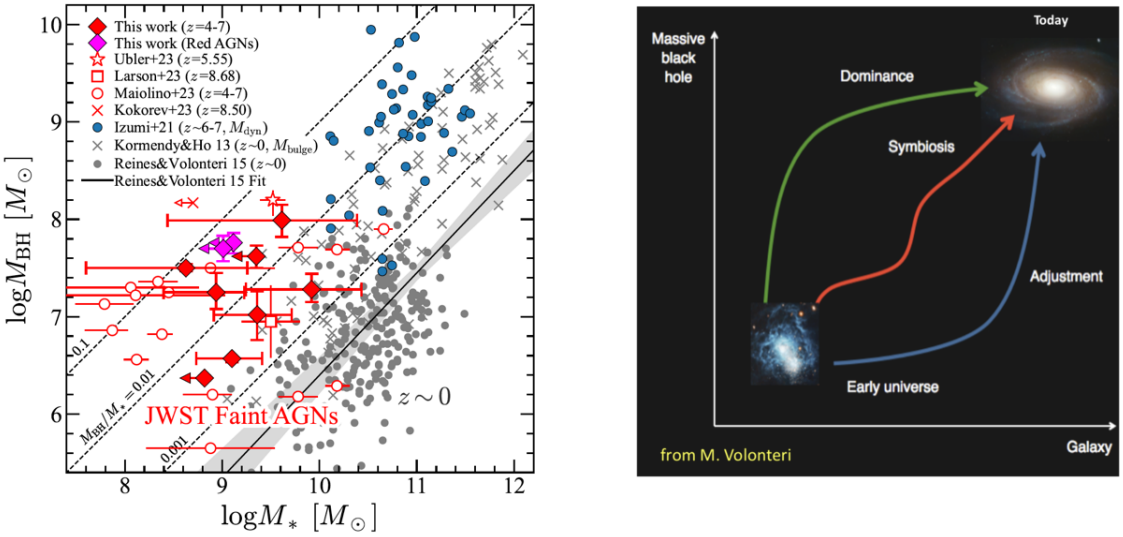

FIGURE 1.11: Taken from Harikane et al. (2023) (left) and from Volonteri (2010) (right). *Left:* SMBH mass versus host galaxy stellar mass for different samples of high-z JWST selected BLAGN, compared to sources in the local universe, as labeled. *Right:* Schematic representation of the possible different path of coevolution between SMBH and host galaxy at early stages.

$M_{BH} - M_{\odot}$). This may suggest that the host galaxies contain the adequate total baryonic mass but have been inefficient in converting gas into stars, possibly as a consequence of the photo-dissociative feedback of the AGN. In this case, it would be the galaxy's stellar mass to be under-massive relative to the SMBH. Yet, also the few dynamical masses of AGN host galaxies inferred by Maiolino et al. (2023b) are subject to high uncertainties being derived solely by the velocity dispersion of the narrow emission lines (and without taking into account the mass distribution).

Another remarkable feature of high-z AGN discovered by JWST is that a large fraction of them is undetected in the available deep X-ray images, suggesting a significant X-ray weakness compared to the low-redshift AGN population. Specifically, most of the newly selected AGNs, including BLAGN, lack any X-ray emission (Maiolino et al., 2024a; Yue et al., 2024; Ananna et al., 2024; Mazzolari et al., 2024b), even if located in fields covered by some of the deepest extragalactic X-ray observations ever performed, such as the Chandra Deep Field South (CDFS Luo et al., 2017) or Chandra Deep Field North (CDFN Xue et al., 2016). The X-ray undetection persists even when X-ray stacking is performed, leading to an observed X-ray weakness, derived by comparing the observed X-ray luminosity with the expected intrinsic one, of 2-3 dex (Maiolino et al. 2024a, see Fig. 1.12). Given the very few exceptions of X-ray AGN detections in the early Universe (Goulding et al., 2023; Kovacs et al., 2024; Maiolino et al., 2024a), it is possible that the X-ray weakness could be due to intrinsic properties of high-z AGN (Yue et al., 2024), as it was also suggested for some low-z AGN (Simmonds et al., 2018; Zhang et al., 2023). In this view, these AGN might be characterized by a different accretion-disk/coronal structure that can determine, for example, a larger ratio of the optical to X-ray emission (α_{OX}) due to a much lower efficiency of the corona in producing X-ray

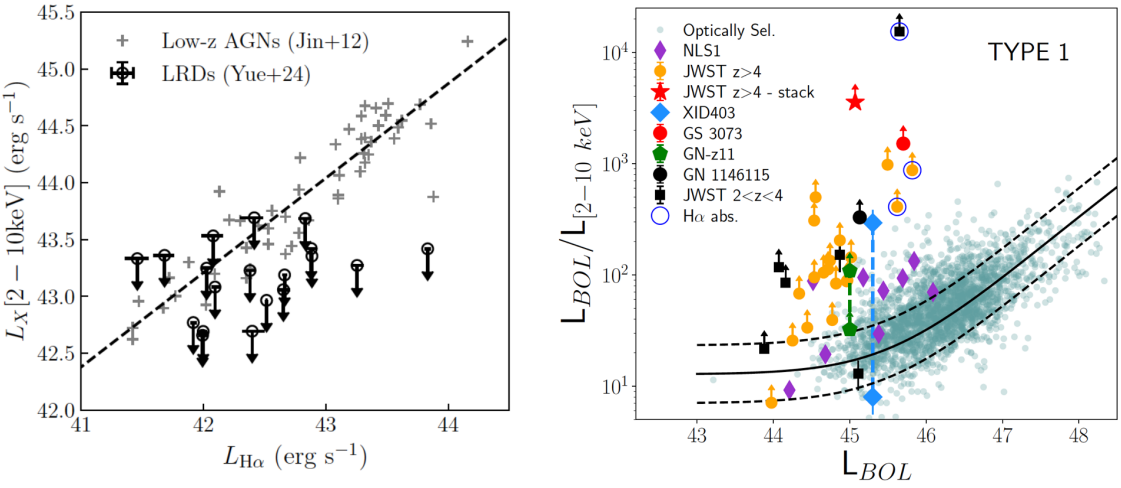

FIGURE 1.12: Taken from Yue et al. (2024) (left) and Maiolino et al. (2024a) (right). *Left*:Observed X-ray luminosity versus Hα luminosity for a sample of LRDs compared to the distribution of local BAGN. None of the sources is X-ray detected. *Right:* Ratio between the AGN bolometric luminosity and the Xray (2-10 keV) luminosity as a function of the bolometric luminosity for a large sample of high-z, JWST detected, BLAGN on the CDFN and CDFS. Light blue small points are from the low-redshift sample of Lusso et al. (2020), representative of "normal", optically/UV selected blue quasars. Other points represent peculiar sources with similar behavior to the high-z BLAGN.

photons (Proga, 2005), or simply a lack of the corona itself. Another possibility investigated by different recent works is that these sources are undergoing (or have undergone) a phase of rapid, super-Eddington accretion (see Sect. 1.1.2), collimating the X-ray emission and altering the X-ray corona structure (Pacucci & Narayan, 2024; Lambrides et al., 2024; Madau & Haardt, 2024).
Maiolino et al. (2024a), analyzing the lack of X-ray emission in a large sample of high-z sources unambiguously identified as AGN, suggested that their X-ray weakness could also be ascribed to the presence, in the inner region of the AGN, of a spherical distribution of clouds with CTK column densities and very low dust content, such as the BLR clouds. In this hypothesis, high-z BLAGN do not need to be characterized by any kind of intrinsic X-ray weakness, as their observed lack of X-ray emission would simply be caused by the inner CTK BLR gas distributed with large covering factors. This hypothesis is further sustained by the detection of a non-negligible fraction of the JWST selected high-z AGN with absorption features in the Balmer or HeI emission lines (Juodžbalis et al., 2024b), suggesting the presence of outflows of dense gas close to the central SMBH.

JWST spectroscopy offers unprecedented insights into the nature of high-redshift sources, which, as discussed, are transforming our understanding of galaxy and SMBH evolution in the early universe. For what concerns the AGN population, besides the detection of many BLAGN, which are likely to still represent only a small fraction of the total AGN population, it also enables the deep spectroscopic study of the more elusive NLAGN at very high redshifts. However, identifying and selecting these obscured AGN remains a more challenging task.

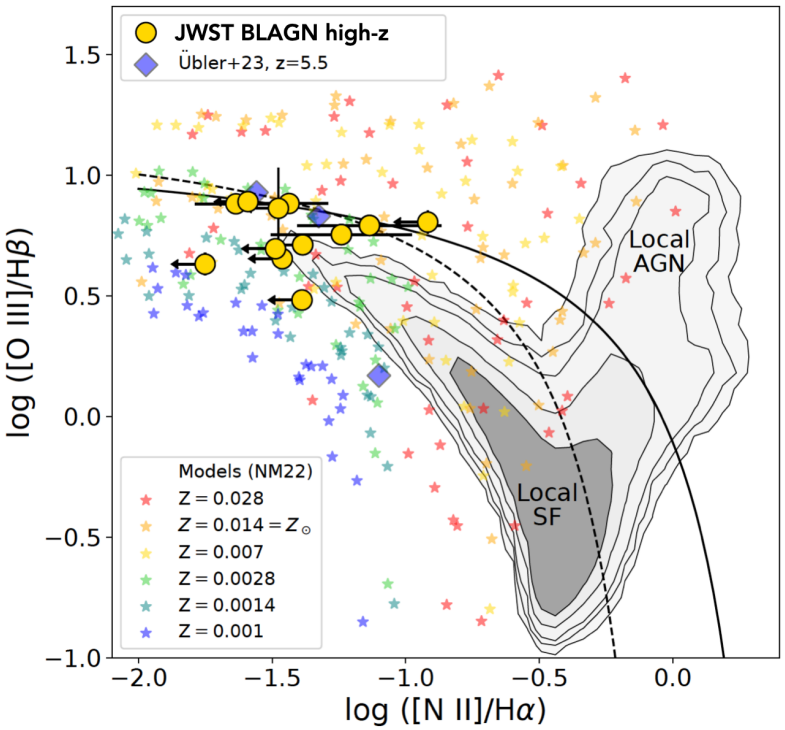

FIGURE 1.13: Taken from Maiolino et al. (2023b). Position on the BPT diagram (Baldwin et al., 1981), of the high-z BLAGN discovered by JWST in the JADES survey. For the line ratios only the narrow component of the emission lines is considered. The black contours mark the distribution of AGN and SFG from the SDSS survey. Stars represent photoionization models from Nakajima & Maiolino (2022) color-coded by metallicity, as labeled.

1.4.2 High-z AGN spectroscopic selection

So far, the detection of a broad component in the permitted lines and not in the forbidden lines (such as the [O III] λ5007) with $FWHM > 1000$km/s is still considered a safe indication of the presence of an AGN also at high-z (see Sect. 1.1.5), because the kinetic properties of the BLR clouds are not expected to evolve with redshift. On the contrary, a significant difference between high-z and low-z AGN is represented by the remarkably different physical environments in which they are set, i.e. their host galaxies.

Early galaxies are systematically more metal-poor and characterized by younger stellar populations (Curti et al., 2023b) that, emitting much more UV photons, increase the ionization parameter of the gas in the ISM (Cameron et al., 2023; Curti et al., 2023b). The so-called stellar mass-metallicity relation (M_*-Z, MZR Tremonti et al., 2004; Schaye et al., 2015) indicates that the metallicities of galaxies increase with increasing stellar masses. At the same time, the fundamental metallicity relation (FMR Mannucci et al., 2010) suggests that the galaxies with higher star formation rates tend to have lower metallicities for a given stellar mass. Taking advantage of JWST photometry and spectroscopy, it is possible to prove the population of low mass SFG ($M_* < 10^{7-8} M_\odot$) at $z > 4$, and different works showed that a consistent fraction of these high-z galaxies are compact and frequently observed in a bursty and above main sequence star-forming state (Endsley et al., 2023; Rinaldi et al., 2023; Fukushima et al.,

2023; Tang et al., 2023; Langeroodi & Hjorth, 2023; Hopkins et al., 2023). The effects of a lower metallicity and a larger ionization parameter in the host galaxies (and therefore also in AGN narrow line region clouds) have a significant impact on the selection of NLAGN using traditional optical emission-line diagnostic diagrams, like those reported in Fig. 1.9. Considering the BPTs diagrams, the larger ionization parameter, on one hand, makes SFGs move towards the AGN locus on the BPT; on the other hand, the lower metallicity of the NLR of AGN makes their line ratios move towards and overlap with the SFG locus on these diagnostic diagrams, as shown in Fig. 1.13 (Übler et al., 2023; Kocevski et al., 2023; Maiolino et al., 2023b; Scholtz et al., 2023b; Mazzolari et al., 2024a).
The detection of high-ionization emission lines (like N v λ1242, [Ne iv] λ2424, [Ne v] λ3426) can still be considered as a safe tracer for the presence of an AGN (Calabrò et al., 2023; Cleri et al., 2023; Brinchmann, 2023; Chisholm et al., 2024), but their detection at high-redshift is still difficult even in the deepest JWST spectroscopic surveys (Bunker et al., 2023a; Scholtz et al., 2023b). Furthermore, given the peculiar conditions that might characterize star formation in galaxies at high-z, some works also discussed the possibility that these lines could originate even from extreme star-forming processes (Cleri et al., 2023; Topping et al., 2024).

The spectroscopic selection and physical characterization of high-redshift AGN using JWST data will be the main topic of Chapter 4 and Chapter 5.

1.5 Thesis Overview

The work presented in this thesis focuses on the study of the most obscured AGN population, especially at high-z. In particular, I worked on the identification of new selection techniques, able to unveil the most elusive, but probably dominant, population of SMBH across Cosmic Time. The successful application of these new selections allowed me to deal with statistical samples of heavily obscured AGN and to study in detail their physical properties, using a multiwavelength approach spanning from radio to X-rays. Furthermore, I took advantage of these selections to explore their distribution and abundance across different redshifts.
The work presented in this thesis was done with a fruitful collaboration with many other people across different institutes and is presented in four chapters, as described below:

- In Chapter 2, I present an analytical model developed to investigate the effectiveness of heavily obscured AGN selection using the radio band and focusing particularly on CTK AGN. We first translate the AGN X-ray luminosity functions into the radio band (1.4GHz) assuming the population synthesis model of the CXB, and we used it to predict the number of detectable AGN over a radio image. We derive the 1.4GHz luminosity function, and radio number count, comparing the results with those in the literature. We then apply the model to the major extragalactic fields covered by X-ray and radio observations, discussing for which type of AGN the 1.4GHz emission performs better than the X-ray one in selecting AGN. Finally, we also present the expectations for the Square Kilometer Array Observatory (SKAO) surveys that will be done in the next future.

- In Chapter 3, I report the selection and analysis of the population of radio excess AGN on the field around the quasar SDSS J1030+0524. In particular, we looked for those sources that are expected to be CTK and at high-redshift ($z > 1.5$) to observationally test the effectiveness of the radio selection of heavily obscured AGN and to measure their abundance for the first time from a radio perspective. We defined a radio excess parameter considering the available photometry on the field, and we then selected all the radio-excess sources without a counterpart in the deep X-ray image, suggesting heavy obscuration. We thus performed a detailed X-ray stacking analysis of the selected sources, and we finally computed their number density. Then, the results of the radio selection of heavily obscured AGN are compared with results for the same population retrieved from the X-rays.

- In Chapter 4, I present the selection of high-z obscured AGN using spectroscopic techniques. In particular, given the above-mentioned difficulties in using the traditional diagnostic diagrams to select NLAGN at high-z, I show three new NLAGN diagnostic diagrams based on the [O III] λ4363 auroral line, that are also effective in selecting AGN at high-z. Using both known observational samples and photoionization models of AGN and SFG, we demonstrated that these diagnostics can separate the two populations, allowing the selection of AGN even in the low-metallicity conditions typically found at high-redshift.

- In Chapter 5, I present the NLAGN selection performed among the JWST spectra of the CEERS survey and their detailed physical characterization. In particular, we

compare the effectiveness of the different emission line AGN diagnostic diagrams in selecting NLAGN across different redshifts, and then we study the AGN and host galaxy properties using spectral and SED-fitting. Lastly, we explore the average X-ray and radio properties of the selected AGN using a stacking analysis. This is among the first works performing this kind of analysis on a statistical sample of AGN spanning $2 < z < 9$.

Finally, in Chapter 6, I will present the conclusions of this PhD thesis, summarizing the results obtained and discussing the possible future perspectives opened by this work.

Here and for the rest of the Thesis we assume a flat ΛCDM universe with $H_0 = 70 \mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$, $\Omega_m = 0.3, \Omega_\Lambda = 0.7, \Omega_k = 0$.
If not stated, we assume AGN radio spectra of the form $S_\nu \propto \nu^{-\alpha}$, with $\alpha = 0.7$, which is the typical spectral slope considered for extragalactic synchrotron emission (Smolčić et al., 2017a; Novak et al., 2017). When $\mathrm{L}_{1.4\mathrm{GHz}}$ is reported we refer to $\nu \mathrm{L}_\nu$ (with $\nu = 1.4\mathrm{GHz}$), in units of $\mathrm{erg\,s}^{-1}$.

Chapter 2

Heavily obscured AGN in the radio band: an analytic approach

2.1 Introduction

As described in Sect. 1.2.1, retrieving a complete census of the AGN population across different redshift is crucial to test the predictions of coevolutional models, SMBH seed models, and also to derive information on the overall SMBH accretion efficiency along cosmic time. The fact that we are potentially missing a large fraction of the SMBH population due to obscuration effects, in particular at high-redshift, calls for the identification of novel obscured AGN selection techniques that can overcome the limitations of the X-rays and MIR selections for heavily obscured sources (see Sect. 1.3.1).
In Sect. 1.3.1 we also presented the advantages of radio emission in selecting AGN. At radio frequencies it is possible to uncover low-luminosity AGN that do not show the strong signatures of high-luminosity AGN; furthermore radio emission is largely unaffected by AGN obscuration since both gas and dust opacities are almost negligible at typical radio frequencies, as described in Sect. 1.1.4.
In this Chapter, we investigate the effectiveness of AGN selection methods based on radio and X-ray emission considering AGN at different levels of obscuration and at different intrinsic luminosities and redshift, focusing particularly on CTK AGN. The aim is to investigate for which AGN and in which conditions radio emission can be (or not) more useful than X-ray emission in unveiling the most obscured AGN population.
To do this in Sect. 2.2 we translate the AGN X-ray luminosity functions into the radio band (1.4GHz). In Sect. 2.3 we present the analytically derived 1.4GHz luminosity function (Sect. 2.2.2), radio number counts (Sect. 2.3.2) and AGN predictions (Sect. 2.3.3), comparing the results with those in the literature. In Sect. 2.4 we apply our model to the major extragalactic fields covered by X-ray and radio observations predicting the number of detectable AGN over a radio field, given its depth, area, and completeness corrections, and then we discuss for which type of AGN the 1.4GHz emission performs better than the X-ray one. In Sect. 2.4.3, we focus on the predictions for the high-redshift Universe ($z > 3$), while in Sect. 2.4.4 we present the expectations for the Square Kilometer Array Observatory (SKAO) surveys that will be done in the next future.
The results presented in this Chapter have been published in Mazzolari et al. (2024c).

One might argue that using X-ray results as a starting point would not allow us to eventually unveil the missing AGN population (see Sect. 1.2.1). However, with the present model, we want to investigate the differences in the AGN detection and selection rates in the two bands and not to discuss whether X-ray models fail or not in recovering the whole AGN census. This point, instead, will be the main topic of Chapter 3.

2.2 Methods

Cosmic X-ray Background (CXB) models (Gilli et al., 2007; Ueda et al., 2014; Buchner et al., 2015; Ananna et al., 2019), by studying the integrated X-ray emission of faint extragalactic point-like sources, can provide an AGN census for any obscuration level, including also the most CTK AGN population, poorly sampled even in the deepest X-ray surveys. The different CXB models almost agree in predicting a consistent fraction of CTK AGN among the whole AGN population, that is between $30 - 50\%$.
To derive the AGN radio luminosity function (RLF) we followed the approach described in the following sections which is based on the CXB model of Gilli et al. (2007) (G07 hereafter) and considers the AGN radio-hard-X luminosity relation ($L_{1.4GHz} - L_{HX}$) derived by D'Amato et al. (2022).
A similar approach was also followed in Ballantyne (2009) and in La Franca et al. (2010), the latter to investigate the global AGN kinetic energy release in the context of AGN feedback processes.

2.2.1 Hard-X luminosity function

The CXB model in G07 considers for the unobscured AGN population the soft-X-ray luminosity function (SXLF, energy range: 0.5-2 keV) derived by Hasinger et al. (2005), of the form:

$$\frac{d\Phi(L_X, z)}{dL_X} = A\left[\left(\frac{L_X}{L_*}\right)^{\gamma_1} + \left(\frac{L_X}{L_*}\right)^{\gamma_2}\right]^{-1} \cdot e_z(z, L_X) \cdot e_{\rm decl}(z), \tag{2.1}$$

where the values of the normalization A, of the characteristic luminosity L_*, and of γ_1, γ_2 are taken from Tab. 5 of Hasinger et al. (2005). The evolution factor e_z was derived in Hasinger et al. (2005) considering a luminosity-dependent evolution model. The term $e_{\rm decl}$:

$$e_{\rm decl}(z) = \begin{cases} 1 & \text{for} \quad z < 2.7 \\ 10^{[-0.43\cdot(z-2.7)]} & \text{for} \quad z \geq 2.7 \end{cases} \tag{2.2}$$

is introduced to reproduce the steep decline in the density of AGN observed at high-z (e.g. see Fig. 8 in Brusa et al., 2009). Following the same approach as in G07, we obtain the corresponding Hard-X luminosity function (HXLF, energy range: 2-10 keV) assuming a Gaussian distribution of X-ray spectral indices centered at $\Gamma = 1.9$ (Piconcelli et al., 2005). Then, using the obscured-to-unobscured AGN ratio in G07, we derive the HXLF of the different subpopulations of AGN in terms of their level of obscuration. In particular, we derive the HXLF of obscured Compton-thin and CTK AGN, the latter assumed to have the same number density of the obscured Compton-thin AGN. The orange solid line in Fig. 2.1 corresponds

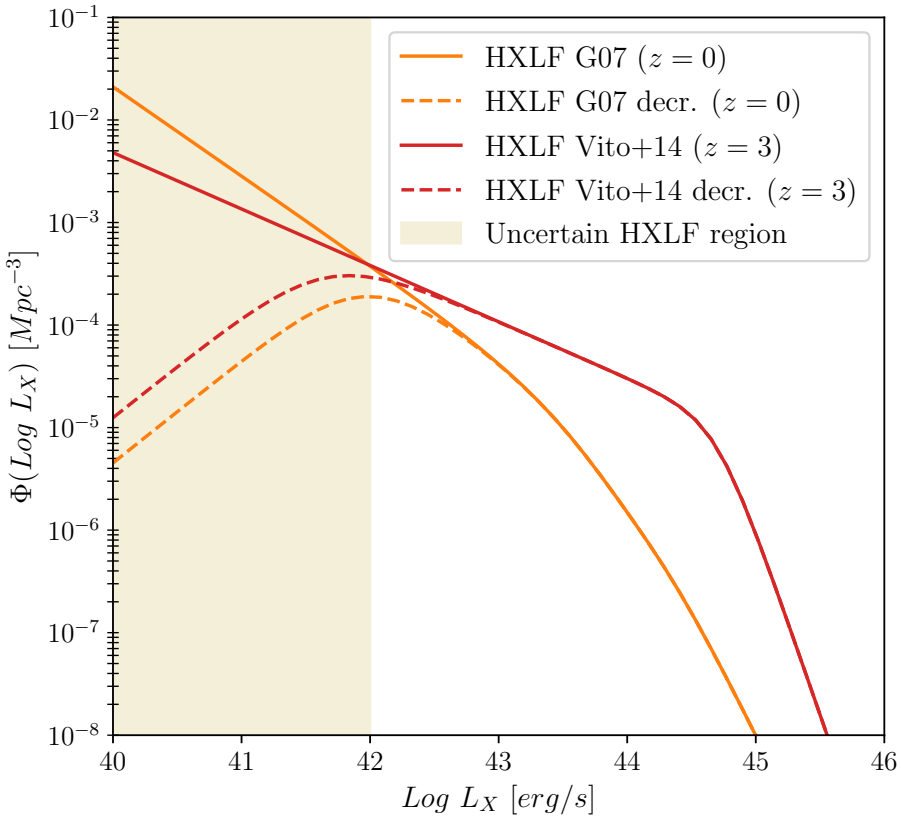

FIGURE 2.1: HXLF of all AGN obtained following the prescription of the CXB model in G07 and computed at $z = 0$ (solid orange line). The dashed orange line is the same HXLF but introducing a cutoff at $\log L_X < 42$, where the HXLF measurements are most uncertain. The red curve represents the HXLF as derived by Vito et al. (2014), computed at $z = 3$, and implemented as the baseline model for $z > 3$. The dashed red line is the Vito et al. (2014) HXLF with the same cutoff described above.

to the total HXLF at $z = 0$, computed by summing the luminosity functions of all the sub-populations of obscured and unobscured AGN.
The faint end of the XLF is poorly observationally constrained for values of $\log L_X \leq 42$, since at these X-ray luminosities it is difficult to distinguish between the X-ray emission coming from AGN and that coming from SF processes in normal galaxies (Ranalli et al., 2003; Lehmer et al., 2016). Also recent results deriving the XLF combining wide and deep X-ray observations conservatively takes as lower limit $\log L_X = 42$ for the definition of the XLF (Peca et al., 2023). To derive information in the faintest AGN luminosity regime ($\log L_X \leq 42$), X-ray luminosity functions are usually extrapolated using the steep slopes determined at brighter luminosities, but the possibility that they remain constant or even decline cannot be excluded. Consequently, we introduced a cut-off in the Hasinger et al. (2005) SXLF at $\log L_X < 42$:

$$\frac{d\Phi(L_X, z)}{dL_X} = A\left[\left(\frac{L_X}{L_*}\right)^{\gamma_1} + \left(\frac{L_X}{L_*}\right)^{\gamma_2} + \left(\frac{L_X}{L_1}\right)^{\gamma_3}\right]^{-1} \cdot e_z(z, L_X) \cdot e_{\rm decl}(z), \qquad (2.3)$$

where $\mathrm{L_1} = 10^{40}$ erg s^{-1} and $\gamma_3 = -1$ (see orange dashed line in Fig. 2.1). The region spanned at $\log L_X < 42$ by the HXLF with and without the declining factor defines the uncertainty on the HXLF at these X-ray luminosities. To check the goodness of the chosen decline factor we converted the [O III] λ5007-luminosity function derived for low-luminosity optically selected obscured AGN from zCOSMOS (Bongiorno et al., 2010) into an HXLF using the conversion reported in Heckman et al. 2005. At $\log L_{[\rm OIII]5007} \sim 39.5 - 40.5$ (corresponding to $\log L_{HX} \sim 41 - 42$) we found $\Phi_{\rm low-lum.\ AGN} \simeq 10^{-3} - 10^{-4}$, that is perfectly inside the range spanned by the HXLF with and without the cut-off, demonstrating the reliability of our approach.
Additionally, since the X-ray luminosity function of Hasinger et al. (2005) was largely derived from AGN samples at $z < 3$, we also considered the high-z X-ray luminosity function presented in Vito et al. (2014). In this work the authors, assembling a sample of 141 AGN at $3 < z < 5$ from X-ray surveys of different sizes and depths, built an HXLF specific for high-z AGN. In our work, we used the Vito et al. (2014) results as baseline HXLF at $z > 3$. In Vito et al. (2014) HXLF we assumed a constant obscured AGN fraction with luminosity and a number ratios between unobscured, obscured Compton-thin, and obscured Compton-thick AGNs of 1:4:4. Both these assumptions are in agreement with the observational results reported in Vito et al. (2018). The total HXLF is presented in red in Fig. 2.1 at $z = 3$ (red solid line) and has the following analytic expression:

$$\frac{d\Phi(L_X, z)}{dL_X} = A\left[\left(\frac{L_X}{L_s}\right)^{\gamma_a} + \left(\frac{L_X}{L_s}\right)^{\gamma_b}\right]^{-1} \cdot \bar{e}_z(z, L_X), \qquad (2.4)$$

where the values of the parameters L_s, γ_a, γ_b, are taken from Table 5 in Vito et al. (2014) and the redshift evolution factor, $\bar{e}_z(z, L_X)$ is given by Eq. 7 of the same work. The HXLF derived in Vito et al. (2014) provides a very good description of the high-z obscured AGN fraction and AGN space density measured in the 2Ms Chandra Deep Field North (CDFN Xue et al., 2016) and 7Ms Chandra Deep Field South (CDFS Liu et al., 2017) (see Vito et al., 2018).
In conclusion, our baseline model uses the HXLF derived by the CXB model in G07 for $z < 3$

and of the HXLF of Vito et al. (2014) one for $z > 3$. To account for the uncertainty on the X-ray luminosity function faint end also at high-z, we applied to the Vito et al. (2014) HXLF the same cut-off ($(L_X/L_1)^{\gamma 3}$) reported in Eq. 2.3, (see red dashed line in Fig. 2.1).

2.2.2 Predicted AGN radio luminosity function

To convert the HXLF into a RLF requires a relation between AGN luminosities in the two bands (e.g. Panessa et al., 2015; Merloni et al., 2003; Dong et al., 2021). In this work, we assume the relation of D'Amato et al. (2022), that was derived using the deep radio and X-ray data available in the $\sim 0.2\mathrm{deg}^2$ field centered on the quasar SDSS J1030+0524 (hereafter J1030):

$$\log L_{1.4\mathrm{GHz}} = 0.83 \log L_{HX} + 3.17. \tag{2.5}$$

This relation was computed from a sample of X-ray selected AGN and early-type galaxies (ETG), spectroscopically confirmed up to $z \sim 3$, spanning a wide range of luminosities. The 2-10 keV luminosities cover the range $\log \mathrm{L_{HX}} \in [40.5, 45]$, while the radio luminosities extend over $\log L_{1.4\mathrm{GHz}} \in [37, 43]$. Radio upper limits were taken into account using survival analysis. Since the relation was computed using X-ray selected sources, and since we started from an X-ray luminosity function, this relation is perfectly suited for our aim. Furthermore, since the sample of D'Amato et al. (2022) includes sources up to $z \sim 3$, their relation extends well beyond the local universe, as opposed to other $\mathrm{L_{1.4GHz}} - \mathrm{L_{HX}}$ relations in the literature. The relation of D'Amato et al. (2022) does not consider any prior selection in the radio loudness parameter of the sources, but 83% of the sample (87% accounting radio upper limits) is radio quiet according to the radio loudness threshold $R_X = \log(L_{1.4\mathrm{GHz}}/L_{HX}) < -3.5$, originally defined in Terashima & Wilson (2003). The relation derived from D'Amato et al. (2022) differs significantly from those found in the literature for RL AGN (e.g. Fan & Bai, 2016). This means that the RL AGN population cannot be described by this relation and they will not be included in our predictions (see Sect. 2.3).
To transform the HXLF into a RLF we considered a Gaussian probability distribution function that returns the probability P that an AGN with a 2-10 keV luminosity $\mathrm{L_{HX}}$ has a radio luminosity $\mathrm{L_{1.4GHz}}$ according to the D'Amato et al. (2022) $\mathrm{L_{1.4GHz}} - \mathrm{L_{HX}}$ relation:

$$\mathrm{P}(\mathrm{L_{1.4GHz}}|L_{HX}, \sigma_\mathrm{R}) = \tag{2.6}$$

$$= \frac{\exp(-(0.83 \log L_{HX} + 3.17 - \log L_{1.4GHz})^2/2\sigma_\mathrm{R}^2)}{\sqrt{2\pi}\sigma_\mathrm{R}}, \tag{2.7}$$

where the Gaussian dispersion $\sigma_\mathrm{R} = 0.5$ is given by the intrinsic dispersion of the $\mathrm{L_{1.4GHz}} - \mathrm{L_{HX}}$ relation of D'Amato et al. (2022). Then, we computed the RLF by weighing the HXLF by the probability distribution computed above:

$$\Phi(\mathrm{L_{1.4GHz}}, \mathrm{z}) = \int_{\mathrm{L_{HX_{min}}}}^{\mathrm{L_{HX_{max}}}} \Phi(\mathrm{L_{HX}}, \mathrm{z}) \cdot \mathrm{P}(\mathrm{L_{1.4GHz}}|\mathrm{L_{HX}}, \sigma_\mathrm{R})\, \mathrm{dL_{HX}}, \tag{2.8}$$

where $L_{HX_{max}} = 10^{47}\mathrm{erg\ s^{-1}}$ and $L_{HX_{min}} = 10^{40}\mathrm{erg\ s^{-1}}$. The AGN RLF computed in Eq. 2.8 is shown in Fig. 2.2 for different redshift bins.

The shape of the RLF mainly depends on two parameters: the shape of the HXLF, in particular at its faint end, and the value of the dispersion $\sigma_{\rm R}$ of the $L_{\rm 1.4GHz} - L_{HX}$ relation.
Our baseline model adopts the HXLF extrapolated to low luminosities using its original slope, combined with $\sigma_{\rm R} = 0.5$ (i.e. the 1σ scatter of the X-ray-radio relation derived by D'Amato et al. (2022)). The shaded area in Fig. 2.2 represents the assumed uncertainties on the RLF. The lower limit of the shaded region corresponds to the RLF computed starting from the HXLF with the cut-off at $\log {\rm L_X} < 42$ (Eq. 2.3), and taking a slightly smaller dispersion $\sigma_{\rm R} = 0.4$. The upper limit of the green shaded region is derived from the baseline model, but assuming a slightly larger dispersion $\sigma_{\rm R} = 0.65$.
We note that a larger (lower) value of σ_R with respect to our baseline model increases (decreases) the slope of the RLF, in particular at high ${\rm L_{1.4GHz}}$. The two values of the dispersion $\sigma_{\rm R}$ for the upper and lower boundaries are chosen to provide conservative estimates of the RLFs. The value $\sigma_{\rm R} = 0.65$ is the one obtained from the D'Amato et al. (2022) sample when excluding radio upper limits; $\sigma_{\rm R} = 0.4$ allows us to obtain a steeper slope of the RLF at high radio luminosities.

2.2.3 Radio number counts

Starting from the RLF we computed the cumulative number counts, namely the number of AGN above a given 1.4GHz flux S_{lim} and in a given redshift range in units of ${\rm sr^{-1}}$ or ${\rm deg^{-2}}$. Taking the comoving volume element:

$$\frac{dV}{dz} = \frac{4\pi c}{H_0} \frac{D_{\rm M}^2}{\left(\Omega_{\rm m}(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda\right)^{1/2}}, \tag{2.9}$$

where $D_{\rm M}$ is the comoving distance, the number of AGN per steradian with a flux ${\rm S} > {\rm S_{lim}}$ is given by:

$$N(> S_{lim}) = \int \frac{dN}{ds} ds \tag{2.10}$$

$$= \frac{1}{4\pi} \int_0^{z_{max}} \int_{max[l_{lim}, L_{min}]}^{L_{max}} \Phi(L, z) \cdot \frac{dV}{dz}\, dL\, dz, \tag{2.11}$$

where $l_{lim} = L_{1.4GHz}(S_{lim}, z)$, and we fixed $z_{\rm max} = 10$, $L_{\rm max} = 10^{43} {\rm erg\ s^{-1}}$ and $L_{\rm min} = 10^{37} {\rm erg\ s^{-1}}$.

2.3 Results

2.3.1 Luminosity function comparison

In Fig. 2.2 we present the RLF obtained with our model compared to that measured in the COSMOS field for radio-excess AGN at different redshifts (Smolčić et al., 2017c).
The COSMOS radio-excess AGN are selected as radio sources exhibiting a $>3\sigma$ radio emission excess with respect to the one expected from their hosts (IR-based) SFR (Smolčić et al.,

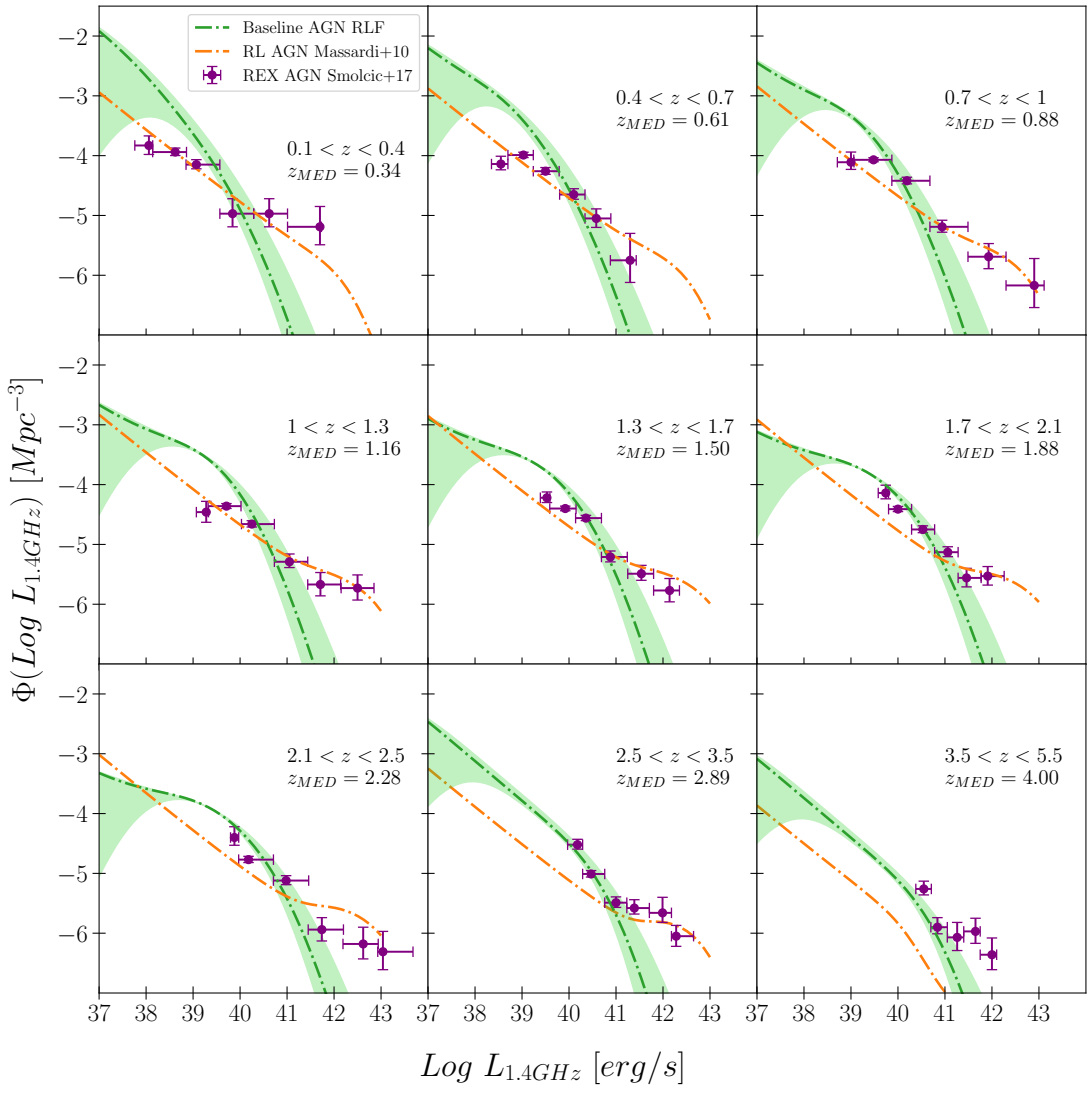

FIGURE 2.2: Baseline RLF of all AGN derived from our model (green line) compared to the radio-excess AGN RLF (purple data points) measured by Smolčić et al. (2017c). Different redshift ranges are investigated, as labeled. The boundaries of the shaded area represent the uncertainty region of the RLF as described in Sect. 2.2.2. In all panels we also report in orange the RL luminosity function empirically derived by Massardi et al. (2010).

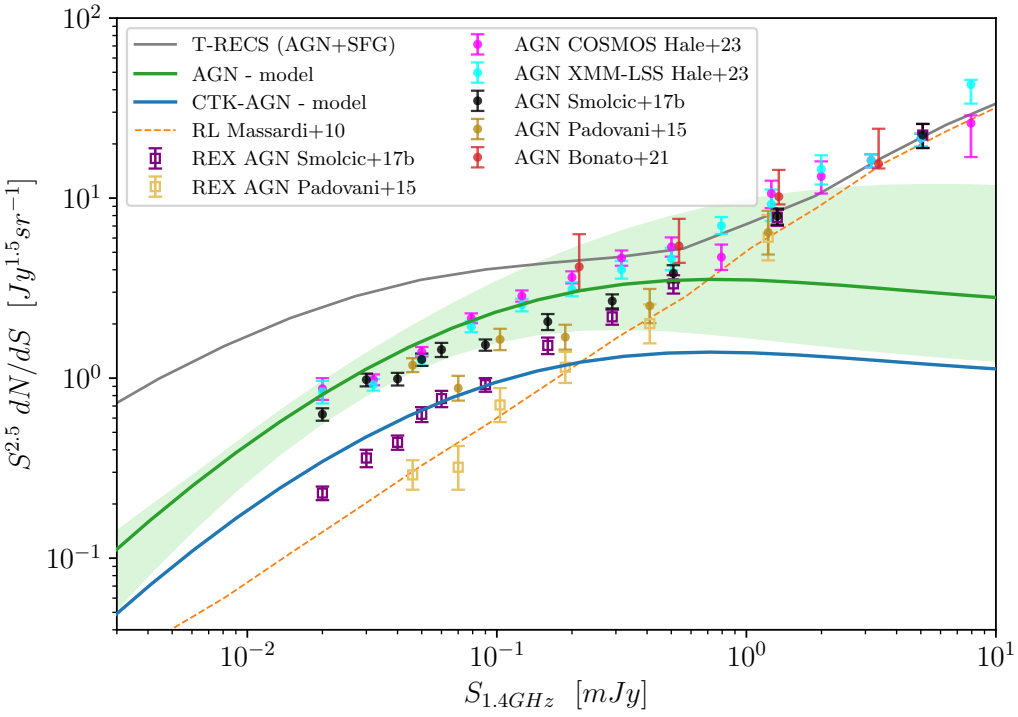

FIGURE 2.3: Differential AGN number counts derived with our model for all AGN (green line and shaded area) and CTK AGN (blue curve), compared with AGN number counts derived in the literature from different radio fields, as labeled. The uncertainties on the CTK number counts are of the same order as those of the whole AGN population but are not reported for clarity. Number counts measured for the whole radio AGN population in the different works are shown as filled points, while the radio-excess AGN only are shown as empty squares. The gray curve shows the number counts of AGN plus SFG predicted by the T-RECS simulations, while the orange curve shows the radio-loud AGN number counts derived from Massardi et al. (2010).

2017c). This criterion ensures that at least 80% of the radio emission in these sources is associated with an AGN, resulting in a highly reliable sample.
Fig. 2.2 shows that our model generally predicts a higher AGN number density at low radio luminosities with respect to Smolčić et al. (2017c), while it predicts lower AGN number densities at high radio luminosities. On the other hand, the modeled RLFs are consistent with the data (within the error bars) in the intermediate luminosity range. The fact that our model cannot reproduce the high-radio luminosity AGN number densities can be explained by the fact that the assumed $L_{1.4\mathrm{GHz}} - L_{HX}$ relation (D'Amato et al., 2022) is not valid for the RL AGN population, which is therefore not included in our model (see Sect. 2.2.2). To highlight the RL AGN component missed by our model, we show in Fig. 2.2 the empirical RL AGN luminosity function derived by Massardi et al. (2010) by summing three different populations of RL AGN: flat-spectrum radio quasars, BL Lacs and a steep spectrum RL AGN population. The three populations are described considering different evolutionary properties (see Massardi et al. (2010) for details). The RL AGN luminosity functions are in good agreement with the Smolčić et al. (2017c) datapoints at all luminosities (at least up to $z \sim 2.5$), and particularly at high radio luminosities, where we expect that the radio-excess and RL AGN classifications trace the same population.
The excess of the modeled AGN number densities at low radio luminosities ($\log L_{1.4\mathrm{GHz}} < 39.5$) can be explained as follows. AGN with low radio luminosities are more difficult to select via the statistical AGN radio-excess selection techniques since their radio emission can be easily overwhelmed by the radio emission coming from the SF of the host galaxy. Indeed, from a statistical point of view, the SFG population start to dominate over the RQ AGN one at $\log L_{1.4\mathrm{GHz}} \sim 39$ (Mancuso et al., 2017). Therefore, the radio-excess AGN sample of Smolčić et al. (2017c) is probably incomplete in this low-luminosity regime. This incompleteness almost disappears when combining multiple AGN selection techniques (see Sect. 2.3.2).

2.3.2 Number counts comparison

In Fig. 2.3 we present the Euclidean normalized differential radio number count obtained from our model compared with other models and measurements in the literature. The green shaded area around the AGN number count, obtained using Eq. 2.10, reflects the assumed RLF uncertainties (see Sect. 2.2.2). The CTK number counts follow a trend similar to that of the total AGN population, and hence their fraction is almost constant (∼40%) with radio flux. This stems from the prescriptions of the CXB model, where $\sim 40\%$ of all AGN are assumed to be CTK (once averaged over the whole luminosity range) and from the fact that radio emission is unaffected by absorption, probing the assumed intrinsic CTK AGN fraction.
It is useful to compare our AGN model with the global population of radio sources, represented by the T-RECS simulated number count (Bonaldi et al., 2019, gray line). The T-RECS simulation includes both RL AGN and SFG, dominating respectively at the highest and lowest radio fluxes. As discussed earlier on, our model does not include RL AGN. We decided not to forcibly include the RL AGN population in our model for two main reasons. First, RL AGNs are expected to constitute a small fraction of the overall AGN population, $\sim 10\%$ of all the AGN, as reported by different works studying RL AGN at different redshift (Williams & Röttgering, 2015; Liu et al., 2021). Second, they dominate the number count at

$S_{1.4GHz} > 0.5 - 1$mJy, as it is possible to see also from Fig. 2.3 (see the orange dashed line derived from the RL AGN luminosity function of Massardi et al., 2010). Therefore, they are not the main focus of this work, where we are primarily interested in quantifying the contribution of AGN at the faintest radio fluxes. In this respect, it is worth noting that the $L_{1.4GHz} - L_{HX}$ we used for our model was derived considering an X-ray selected spectroscopically confirmed AGN+ETG sample, which is mostly composed of RQ AGN, and where no preselection based on radio excess was done (D'Amato et al., 2022). This allows us to compare our models with radio-selected samples in which AGN are not only selected based on their radio excess, but also through selection criteria based on AGN signatures at other bands (such as those discussed in Sect. 2.1). Thus, our model is able to account for a larger population of low-luminosity radio-selected AGN. In Fig. 2.3 we compare our AGN model with AGN samples extracted in some of the major deep radio surveys and fields: the 3GHz JVLA-COSMOS (Smolčić et al., 2017a,b,c), the Extended Chandra Deep Field South (ECDFS) (Miller et al., 2013; Padovani et al., 2015), the Lockman-Hole (LH) (Prandoni et al., 2018; Bonato et al., 2021) and the MIGHTEE radio survey (Whittam et al., 2022). The AGN number count shown in Fig. 2.3 as filled points are extracted from the aforementioned surveys by selecting AGN using all the available selection techniques (radio excess, X-ray emission, MIR colors, SED fitting, and spectroscopy). Instead, the squares refer to radio-excess AGN only in the COSMOS and in the ECDFS fields. In particular, the COSMOS radio-excess data points are derived from the same radio-excess AGN sample that was used to build the RLF presented in Fig 2.2.

For fluxes $S_{1.4GHz} < 1$mJy, the number counts derived by selecting radio AGN using multiple criteria, are in very good agreement with our AGN model predictions, down to the faintest fluxes $\sim 20\mu$Jy. Instead, the data from radio-excess AGN are below our model, with a discrepancy that increases going to fainter radio fluxes. On the contrary, the radio-excess AGN number counts seem to be in good agreement with the RL AGN number count (orange dashed line) derived from Massardi et al. (2010). This result supports the scenario discussed in Sec 2.3.1: the radio-excess AGN RLF is incomplete, in particular at low radio luminosities (and consequently low fluxes), and it mostly accounts for the brighter RL AGN population.

As a final remark, we notice that we made some sanity checks on our AGN model. First, we compared the number count derived from our AGN model with those derived by the model of La Franca et al. (2010). We found that the two radio counts do not differ significantly. In particular, the number count derived in La Franca et al. (2010) predicts slightly lower AGN number densities, but still within our uncertainty region. Second, we derived the number count considering a flatter AGN radio spectral index, namely $\alpha = -0.4$, finding no significant deviation from those derived with $\alpha = -0.7$, as again they lie within our uncertainty region.

2.3.3 Expectations from the main radio and X-ray deep fields

We are interested in assessing the relative effectiveness of deep X-ray and radio surveys to detect the elusive obscured AGN populations and particularly CTK AGN. We use the number count derived from our radio model and the X-ray population synthesis model of the CXB, to respectively predict the number of radio-detected and X-ray detected AGN and CTK AGN

Field	Radio (1.4GHz)			X-ray (0.5-2 keV)		
	Area [deg^{-2}]	flux limit [μJy]	T_{OBS} [hrs]	Area [deg^{-2}]	flux limit 10^{-16}[erg s^{-1} cm^{-2}]	T_{OBS} [hrs]
J1030	0.18	10.6	30	0.09	0.57	140
COSMOS	2.4	18	385	2.15	1	1278
LH-XMM	0.31	22	75	0.19	1.9	322
CDFS	0.32	16	40	0.13	0.08	1944
ECDFS	0.32	16	40	0.31	0.74	69
XMM-LSS	5	32	48	5.36	2.7	361
BOOTES	6.7	80	192	9.26	1.5	944
ELAIS-S1	2.7	83	410	3.2	4.7	278
CDFN	0.07	12	39	0.125	6	556
LHN	0.4	15	140	0.5	3.2	175

TABLE 2.1: Summary of the area covered by radio and X-ray observations, flux limits, and observing time for each of the considered fields. The X-ray and radio flux limits correspond to the flux of the faintest sources in the $0.5-2$keV X-ray catalogs and in the 1.4GHz catalogs, respectively. We note that for the CDFS and ECDFS we considered the same radio survey. The fields are listed with the references for the radio observation followed by the X-ray one: J1030 (D'Amato et al., 2022; Nanni et al., 2020), COSMOS (Smolčić et al., 2017a; Civano et al., 2016), LH-XMM (Biggs & Ivison, 2006; Brunner et al., 2008), ECDFS (Miller et al., 2013; Xue et al., 2016), CDFS (Miller et al., 2013; Liu et al., 2017), XMM-LSS (Heywood et al., 2020; Chen et al., 2018), BOOTES (de Vries et al., 2002; Masini et al., 2020), ELAIS-S1 (Franzen et al., 2015; Ni et al., 2021), CDFN (Owen, 2018; Xue et al., 2016), LHN (Lockman-Hole North Owen & Morrison, 2008; Trouille et al., 2008)

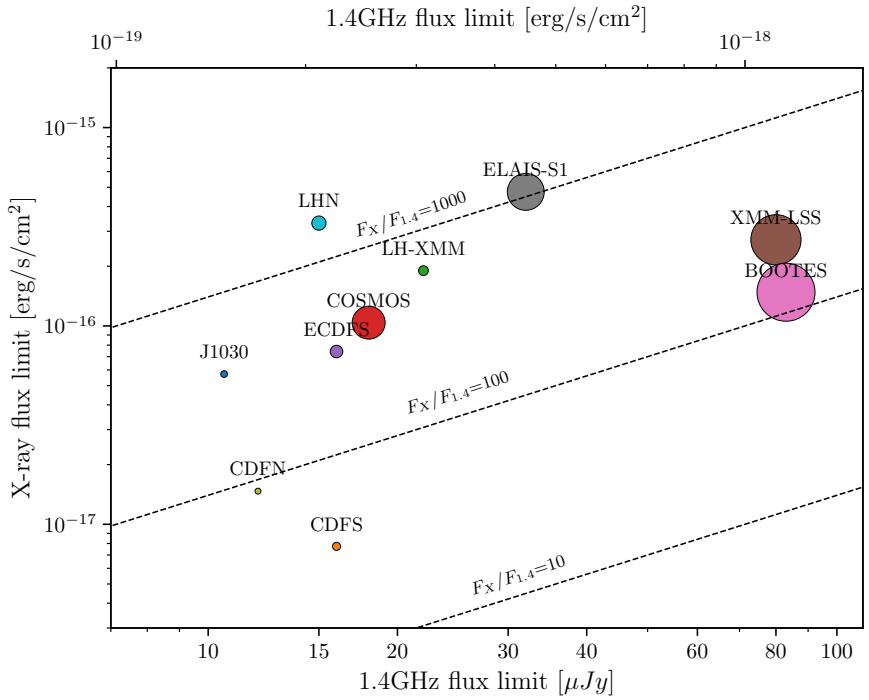

FIGURE 2.4: Distribution of the $0.5-2$keV versus 1.4GHz flux limit for the ten radio and X-ray fields considered in this work. The reported flux limits correspond to the faintest sources' flux in the $0.5-2$keV and in the 1.4GHz catalogs, respectively. Each field marker size is proportional to the area of the overlapping region between the radio and X-ray observations. The three dashed lines correspond to level curves for different ratios between the X-ray and radio fluxes.

over some of the major deep extragalactic radio fields, also covered by X-ray observations. We report the areas, sensitivities, and exposure times of the considered fields in Tab. 2.1. For the COSMOS field, the original 3GHz fluxes and limits were translated to 1.4GHz assuming a spectral shape $S_\nu \propto \nu^\alpha$, with $\alpha = -0.7$.
Fig. 2.4 shows the X-ray soft band (SB; 0.5-2 keV) versus 1.4GHz flux limit plane for each of the 10 considered fields. Each field marker size is proportional to the area of the overlapping region between the observations in the two bands.

Number counts corrections

Our radio model forecasts the number of AGN per deg^{-2} with fluxes larger than a given S_{lim} over an ideal radio field of constant sensitivity. However, in every radio pointing the sensitivity decreases from the center to its outskirt, as described by the so-called visibility function. Moreover, other issues affect the completeness of a radio catalog, such as background noise, Eddington bias (Eddington, 1913), and resolution bias (Prandoni et al., 2001). The first two effects affect the incompleteness of the catalog, particularly at its faint end, while the resolution bias causes a possible loss of extended sources. To get realistic radio AGN predictions, we computed for each field a total "completeness" function (or factor) that takes into account all the aforementioned effects, and we apply it to the model forecasts. In practice, we

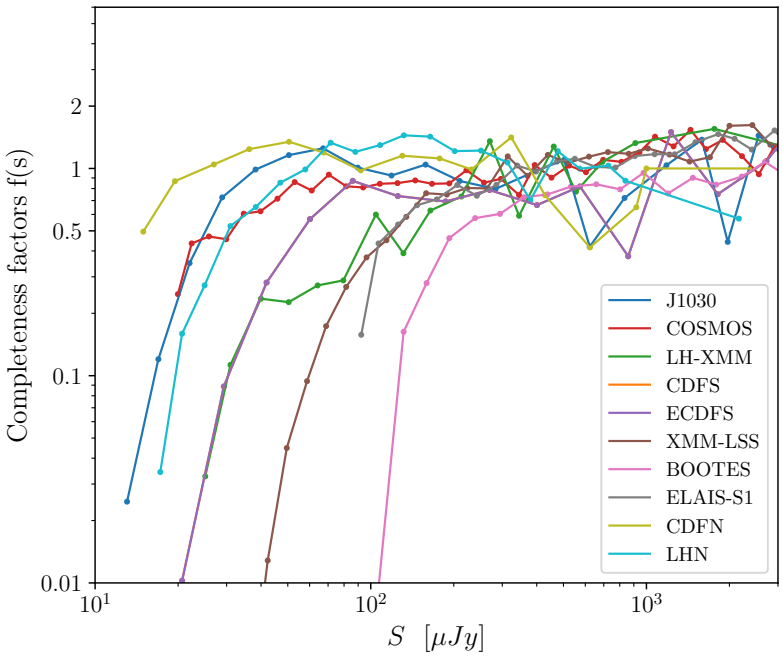

FIGURE 2.5: Completeness factors of each radio catalog derived as discussed in Sect. 2.3.3. Each color represents a different radio field, as labeled. The dots represent the center of the flux bins used to compute the differential number count corrections (see text for details).

proceeded as follows. For every field we first obtained the differential number count from the sources listed in the corresponding radio catalog, namely $dN_{catalog}/ds$. These should be considered as 'raw' number counts, as no attempt to correct for the aforementioned incompleteness effects is made in this case. Then we compared $dN_{catalog}/ds$ with the differential number count expected from an ideal (fully complete) catalog, namely dN_{ideal}/ds. The ratio between these two quantities defines the completeness factor:

$$f(\bar{s}) = \frac{dN_{catalog}}{ds}(\bar{s}) \Big/ \frac{dN_{ideal}}{ds}(\bar{s}), \tag{2.12}$$

where $\bar{s}$ is the central flux of each flux bin. We chose the $25deg^2$ mock catalog released as part of the T-RECS simulations (Bonaldi et al., 2019) as ideal catalog, as it provides a fairly good representation of the observed 1.4GHz radio source counts, both in the bright and in the faint flux regime (Bonaldi et al., 2019; D'Amato et al., 2022). As mentioned in Sect. 2.3.2, the T-RECS simulations include SFG and AGN, the two main extra-galactic radio source populations detected in deep radio fields.

The completeness factors of the different fields considered in this work are presented in Fig. 2.5. The radio catalogs are largely incomplete at their faint end, while they are virtually complete ($f(\bar{s}) \rightarrow 1$) at bright fluxes, as expected. The steep decrease of the completeness functions at faint fluxes is mainly driven by the visibility function, which represents the fraction of the survey area over which sources above a given flux density can be detected. Generally, faint sources can be detected only over a limited area, close to the center of the

radio image, where the local noise is lowest, while the brightest sources can be detected everywhere, even at the outskirt of the radio images where the local noise is highest. This is reflected in a rapid decrease of the visibility function going to faint fluxes. In addition, the background noise introduces errors in the measurements of source fluxes that mostly affect the completeness at low signal-to-noise ratios and so mainly at low radio fluxes.
We caveat that the correction factors shown in Fig. 2.5 were obtained considering both AGN and SFGs. Assuming that these corrections hold independently of the specific type of radio source considered, we can also apply them to the radio AGN-only model we built. This assumption can be justified by the fact that the faint AGN detected in deep radio fields tend to occupy the same flux density regime as SFG, and tend to be unresolved, like SFG, at the typical (arcsec-scale) angular resolution of these surveys (see e.g. Bonzini et al., 2013). This implies that SFG and AGN are similarly affected by incompleteness effects. Also Hale et al. (2023), studying the incompleteness of MIGHTEE radio images, found that the completeness corrections for the two populations are the same for a vast range of fluxes, showing only small differences for $S < 10\mu Jy$, a flux density limit not crossed by none of the considered radio catalogs. The second assumption we make is that the corrections derived in Eq. 2.12 hold independently of the level of AGN obscuration. Since the 1.4GHz emission is unaffected by obscuration, this assumption appears to be fine.

Field	Radio (1.4GHz)					X-ray (0.5-2 keV)				
	#sources catalog	#AGN model	#CTK model	Σ_{AGN} model	Σ_{CTK} model	#sources catalog	#AGN model	#CTK model	Σ_{AGN} model	Σ_{CTK} model
J1030	1489	533	222	0.87	0.36	256	184	8	0.57	0.025
COSMOS	10830	4893	2024	0.57	0.24	4016	3320	67	0.43	0.01
LH-XMM	506	207	87	0.19	0.08	409	394	15	0.58	0.022
CDFS	883	361	149	0.31	0.13	1008	696	111	1.49	0.24
ECDFS	883	361	149	0.31	0.13	1003	955	58	0.86	0.052
XMM-LSS	5762	2747	1124	0.15	0.06	5242	6052	80	0.31	0.004
BOOTES	3172	1695	690	0.07	0.03	6891	5892	64	0.18	0.002
ELAIS-S1	2084	1095	446	0.11	0.046	2630	2410	20	0.21	0.002
CDFN	795	270	112	1.07	0.44	683	545	66	1.21	0.147
LHN	2056	775	304	0.54	0.21	761	695	14	0.39	0.008

TABLE 2.2: Expected number of AGN (total and CTK) compared with the total number of sources observed in each field. Radio and X-ray predictions are shown in the left and right side of the table, respectively. The total and CTK AGN densities (Σ) are computed by dividing the number of predicted AGN by the area of the radio or X-ray image (see Tab. 2.1) and are reported in units of arcmin^{-2}.
1.4GHz expectations for the ECDFS and CDFS are the same since we considered the same radio survey for both.

Radio predictions

To quantify the number of AGN predicted in the different radio catalogs we applied the following equation:

$$\mathrm{N_{AGN}} = \mathrm{A} \int_0^\infty \frac{\mathrm{dN_{AGN\ model}}}{\mathrm{ds}}(\mathrm{s}) \cdot \mathrm{f(s)\ ds}, \tag{2.13}$$

where A is the area of the radio field, $\mathrm{S_{min}}$ is the minimum radio flux density of the radio catalog and $f(s)$ is the completeness function defined in Eq. 2.12. $\mathrm{dN_{AGN\ model}/ds}$ are the differential number count derived by our model (Eq. 2.10) considering $\log \mathrm{L_{1.4GHz\ min}} = 37$, $\log \mathrm{L_{1.4GHz\ max}} = 43$, $z_{min} = 0$, $z_{max} = 10$. We assigned $\log \mathrm{L_{1.4GHz\ min}} = 37$ since it is an acceptable threshold for the lower limit of the AGN 1.4GHz luminosity distribution, as reported by some of the deepest radio observations performed so far (Algera et al., 2020; Alberts et al., 2020). The number of AGN and CTK AGN predicted by our model at 1.4GHz for each field is presented on the left side of Tab. 2.2. The predictions are based on our baseline model. By using the uncertainty boundary regions described in Sect. 2.2.2, the prediction would get on average 30% higher (upper boundary) or 40% lower (lower boundary).
The top panel of Fig. 2.6 shows the fraction of AGN and CTK AGN expected by our model obtained by dividing the predicted number of AGN by the whole number of radio sources in each radio catalog. The fields are sorted in ascending order of radio flux limit. The expected fraction of AGN ranges between 36% in the J1030 field, the deepest radio field in our sample, up to 53% in the BOOTES and ELAIS-S1 fields, which are among the shallowest ones. The top panel of Fig. 2.6 shows a clear trend where the AGN fraction in radio catalogs decreases as the sensitivity of the survey increases. This is in accordance with the evidence, presented by several works in literature and also shown in Fig. 2.3, that AGN dominates over SFGs in the bright-fluxes regime of the global radio number counts (gray line in Fig. 2.3). On the contrary, the lower the radio flux, the larger the relative contribution of SFG to the global radio number count (see again Fig. 2.3), and consequently also the fraction of AGN gets reduced.
All our predictions, particularly those for the shallower fields, have to be considered conservative. Indeed, the given AGN fractions do not include, by construction, the most powerful radio AGN (the RL population) that dominate the bright end of the number count. About $\sim 40\%$ of all AGN, that is about $\sim 20\%$ of the radio sources in each catalog, are expected to be CTK nuclei, as reported also in Sect. 2.3.2.

X-ray predictions

To compare radio versus X-ray AGN yields in the considered fields, we proceeded as follows. We computed the AGN differential number counts based on the CXB model, taking $\log \mathrm{L_{X\ min}} = 41$, $\log \mathrm{L_{X\ max}} = 47$, $z_{min} = 0$, $z_{max} = 10$. The X-ray lower limit $\log \mathrm{L_{X\ min}} = 41$ corresponds to the radio lower limit $\log \mathrm{L_{1.4GHz\ min}} = 37$ used in Eq. 2.13, when assuming the $L_{1.4\mathrm{GHz}} - L_{HX}$ relation of our model. This value is also a reasonable lower limit to the known AGN X-ray luminosity distribution, and we checked that this choice does not violate the total CXB flux constraint. Using the sky coverage $A(s)$ and the CXB differential number count, $\mathrm{dN_{AGN\ CXB}/ds}$, the number of AGN over a certain X-ray field is predicted as follows:

$$\mathrm{N_{AGN}} = \int_0^\infty \frac{\mathrm{dN_{AGN\ CXB}}}{\mathrm{ds}}(\mathrm{s}) \cdot \mathrm{A(s)\ ds}, \tag{2.14}$$

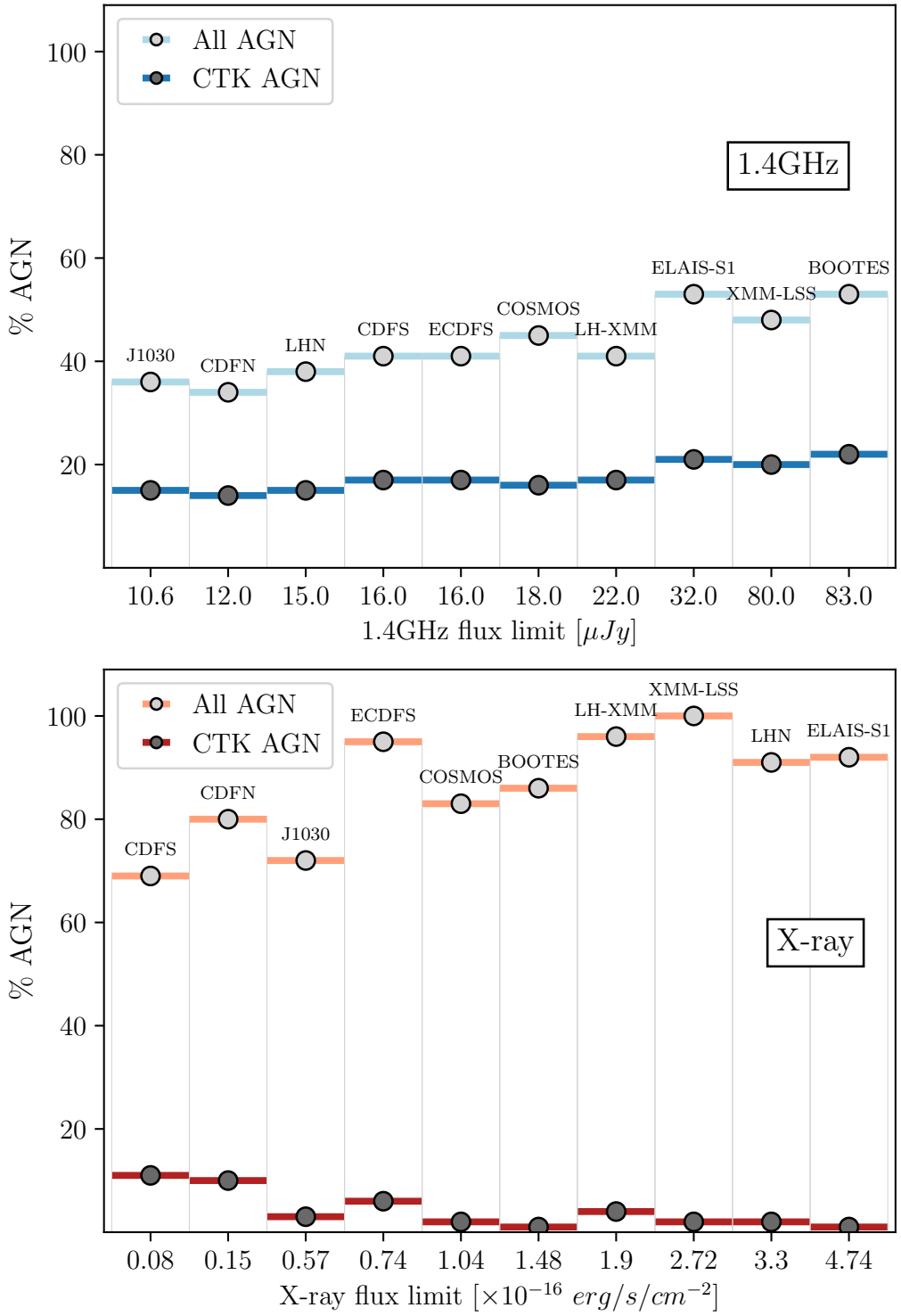

FIGURE 2.6: *Upper panel:* predicted fractions of AGN (light blue) and CTK AGN (dark blue) as a function of radio survey flux limit. *Lower panel:* same as in the upper panel but for the X-ray predictions. In both panels the surveys are sorted on the x-axis by increasing flux limit (faintest source in the catalog).

where S_{min} is the minimum SB X-ray flux of the respective catalog and $A(s)$ represents the sky coverage, namely the observed area as a function of the sensitivity of the survey. The function $A(s)$ is generally retrieved by performing completeness simulations, already implementing any possible completeness correction factor. The results are reported on the right side of Tab. 2.2.

We show only the predictions in the SB for two reasons. Firstly, the numbers of AGN and CTK AGN predicted by the CXB model in the SB are larger than in the hard band (HB, $2-10$keV), for every field. Indeed, for all the considered fields, the flux limit in the SB is always lower by a factor$\sim 5-10$ than the one in the HB, according to the fact that for all the X-ray catalogs the number of sources detected in the SB is larger than in the HB[1]. Secondly, the expected fraction of CTK AGN predicted by the CXB model is larger in the SB than in the HB for $z > 0$. This happens because of the deeper SB flux limit and because of the presence of a minor (but not negligible) SB reflected component in the CTK AGN rest-frame spectrum implemented in the CXB model (Fig.1 in Gilli et al., 2007). We finally notice that in the case of the XMM-LSS field, the number of AGN predicted by the X-ray model exceeds the number of X-ray sources in the related catalog by a fraction $\sim 15\%$. This excess can be justified by field-to-field variation of the number counts or by an overestimate of the CXB model in the SB already observed in a previous work (Marchesi et al., 2020).

In the lower panel of Fig. 2.6 we report the fraction of X-ray AGN and CTK AGN predicted over the different X-ray fields sorted by increasing X-ray flux limit. In the X-ray band AGN are the dominant population: in all cases the fraction of AGN is $\geq 70\%$, with again an increasing trend for shallower flux limits. Indeed, the X-ray emission of SFGs starts to be detected only at very faint fluxes, and are rarely detected in the shallower surveys. The increasing trend of the AGN fraction going to higher flux limits in the X-ray panel is more scattered than in the radio panel. This is mainly due to the shapes of the sky coverage of the different X-ray observations that can change significantly from field to field. In general, X-ray surveys allow the detection of a larger fraction of AGN with respect to radio ones, where the contamination from SFG can be large, especially at faint fluxes. However, the large majority of the AGN detected in the X-rays are not heavily obscured: the fraction of CTK AGN in the X-rays is always between 1%-11% of all the X-ray sources and between the 1%-15% of all the X-ray AGN predicted by the CXB model, the larger fraction reached in the deepest survey, namely the CDFS. Indeed, the fraction of detectable CTK AGN shows a remarkable dependence on the X-ray flux limit: the lower the X-ray flux limit the higher the fraction of CTK AGN detected. This trend is justified by the fact that heavily obscured sources are usually very faint in the X-rays and so their detection requires very deep surveys. On the contrary, in the radio band, the CTK AGN are between $14-22\%$ of all the radio sources and $\sim 40\%$ of all the radio AGN predicted by our model, and show an opposite dependence on radio flux limit.

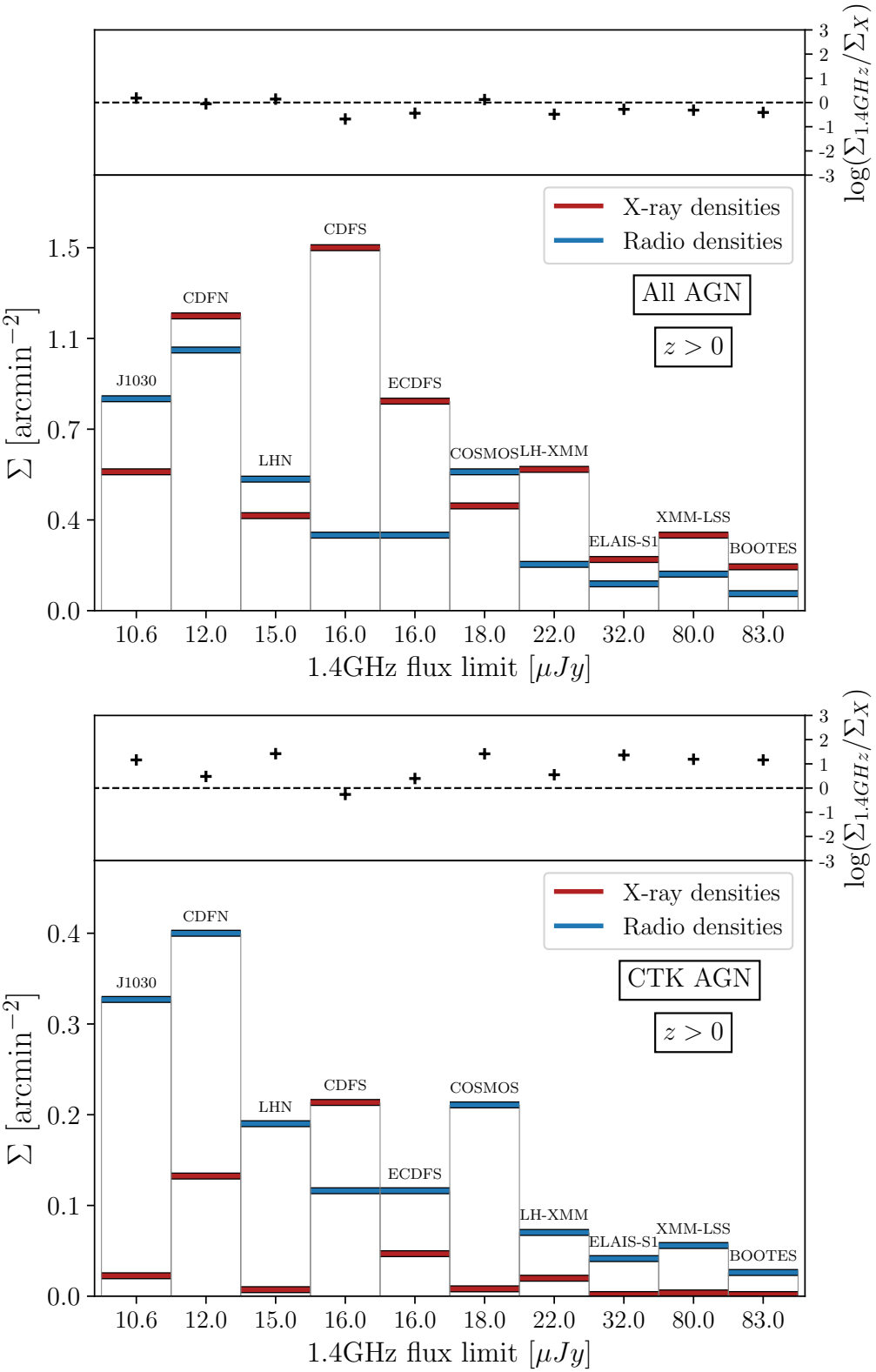

FIGURE 2.7: *Upper panel:* AGN densities in X-ray (red) and radio (blue) surveys as a function of 1.4GHz survey flux limit, as reported in Tab. 2.2. The ratio between the 1.4GHz and X-ray AGN densities is shown in the top position of the plot. *Lower panel:* same as in the upper panel but for CTK AGN. In both panels the surveys are sorted on the x-axis by increasing radio flux limit (faintest source in the catalog).

2.3.4 AGN densities

In Fig. 2.7 we plot the radio and X-ray densities predicted by the two models for both total and CTK AGN. These densities, also reported in Tab. 2.2, are computed by dividing the number of AGN predicted by the area of the radio or X-ray observations. In the upper panel, we note that the X-ray observations generally return a larger density of AGN per arcmin^{-2}, except for the J1030, LHN and COSMOS fields. However, the differences between the densities predicted in the two bands are generally less than one order of magnitude. When only the CTK population is taken into account (lower panel) the situation is the opposite: for every field (except the CDFS) the CTK AGN density is much larger in the radio than in the X-rays, by even more than one order of magnitude. Furthermore, as one would expect, the 1.4GHz AGN and CTK AGN densities increase for deeper radio observations.
These two plots clearly show that X-ray observations are on average more effective than radio ones in detecting AGN when the global AGN population is considered, but if we want to detect the most obscured AGN the radio selection is more effective than the X-ray one (by on average a factor of 10 in term of surface densities). In Sect. 2.4.3 we show that this result holds even at high redshift.

2.4 Discussion

2.4.1 Radio catalogs: AGN predictions versus observations

Since the radio model we built is not set to predict only the radio-excess AGN population but, as presented in Sect. 2.3.2, it can be used to forecast the whole population of radio-quiet AGN we compared our results with those found in the literature using multiple selection techniques.
In the case of the COSMOS field, as mentioned in Sect. 2.3.1, Smolčić et al. (2017b) distinguished radio sources into pure SFG, medium-luminosity AGN (MLAGN), and high-luminosity AGN (HLAGN). Using color selection, X-ray counterpart analysis, SED fitting decomposition, and radio excess, they identified 1623 HLAGN and 1648 MLAGN among all the radio sources with a counterpart in the COSMOS2015 photometric catalog (Laigle et al., 2016). All the AGNs represent around the 42% of the COSMOS radio sources in the reference catalog, a fraction very close to the 45% of AGN our model predicts for COSMOS. However, our predictions have to be considered conservative since our model does not take into account the classical RL AGN, which are expected to be around 10% of the whole AGN population. In the same work, Smolčić et al. (2017b) made also an AGN selection based only on the radio information, selecting only those AGN showing a ($> 3\sigma$) radio excess with respect to what is expected from a pure SFG, to ensure that at least the 80% of their radio emission is due to the AGN component. Their final radio-excess AGN sample is constituted by 1846 sources, that is only 26% of all the radio sources in COSMOS. This shows that the

[1]The only exception is the J1030 field, where the number of X-ray sources reported in Nanni et al. (2020) in the SB (HB) is 193 (208). However, the Chandra observation of the J1030 field is the latest one among those considered in this work and suffers from the SB Chandra sensitivity deterioration (Peca et al., 2021)

radio excess technique alone is not able to provide complete AGN samples, and multiple selection methods are needed to this end[2].
Recent MIGHTEE observations of the COSMOS field (Whittam et al., 2022; Zhu et al., 2023) obtained results similar to those of Smolčić et al. (2017b). The 1.4GHz MIGHTEE data in the COSMOS field reach almost the same sensitivity as the 3GHz VLA COSMOS image, but was obtained on a smaller portion of the field $\sim 1.6\mathrm{deg}^2$. Using the radio excess criterion, following the more recent prescriptions of Delvecchio et al. (2021), Whittam et al. (2022) found 1332 radio excess AGN over a sample of 5223 sources, corresponding to $\sim 25\%$. On the contrary, when they account for all the other selection techniques, they retrieve a fraction of AGN around 35%. Whittam et al. (2022) also present a consistent fraction of unclassified sources; if these are split into AGN and SFG based on the flux density ratio of the classified sources, the fraction of the whole AGN gets to $\sim 40\%$, close to our predictions.
Another example comes from the radio observation of the CDFS (Miller et al., 2013), where source classification is reported in Bonzini et al. (2013) and Padovani et al. (2015). Among the 883 radio sources in their catalog, they identified 381 AGN (43%) using IRAC colors, X-ray counterparts, and the radio excess parameter. Our model predicts a similar number of AGN (361) that has to be taken as a conservative estimation since RL AGN are excluded. Bonzini et al. (2013) selected as radio-excess AGN only 173 sources, only 19% of the whole radio sources, again suggesting that when multiple, deep, multiband analyses are adopted the completeness of the radio AGN sample largely increases and the number of AGN get very similar to what is predicted by our model.
The radio selection of heavily obscured AGN has been only marginally investigated in the literature. The works mentioned above did not investigate the radio properties of AGN in terms of their obscuration levels. Instead Andonie et al. (2022), starting from an IR-based AGN selection in the COSMOS field (and limiting to $z < 3$), investigated the obscuration properties of their sample by means of X-ray and radio data. They found that 73% of IR-selected AGN were in the VLA 1.4GHz or 3GHz catalogs, and in particular that 63% of these radio detected AGN are obscured ($\log \mathrm{N_H} > 22$).
Despite the broad agreement with the total number of AGN observed in radio catalogs, we remark that the number of heavily obscured CTK AGN expected by our model is driven by the assumptions made in the CXB modeling, and is therefore affected by the uncertainties discussed below. The population synthesis model of the CXB considers X-ray data that extend beyond the traditional 0.5-10 keV band (the peak of the CXB is around 20-30keV), allowing one to constrain the population of heavily obscured AGN better than what is possible with Chandra and XMM observations alone. However, the derived total abundance of CTK AGN, and in particular that of the most heavily obscured, reflection-dominated AGN ($\log \mathrm{N_H} > 25$), is degenerate with the reflection efficiency, and hence overall normalization, assumed for their X-ray spectra. In principle, by decreasing the reflection efficiency by a factor of ~ 2, which would still be consistent with broad-band X-ray observations of local CTK AGN, would increase the number of reflection-dominated CTK AGN by an equal factor. An additional uncertainty is related to the scatter of the measured CXB intensities. For example, Comastri et al. (2015) showed that by assuming a space density of reflection-dominated CTK AGN 4x larger than what was assumed by Gilli et al. (2007), the produced CXB flux would

[2]Considering only the X-ray AGN with a radio counterpart (335 sources) we found that only the $\sim 30\%$ of them are selected as radio-excess AGN.

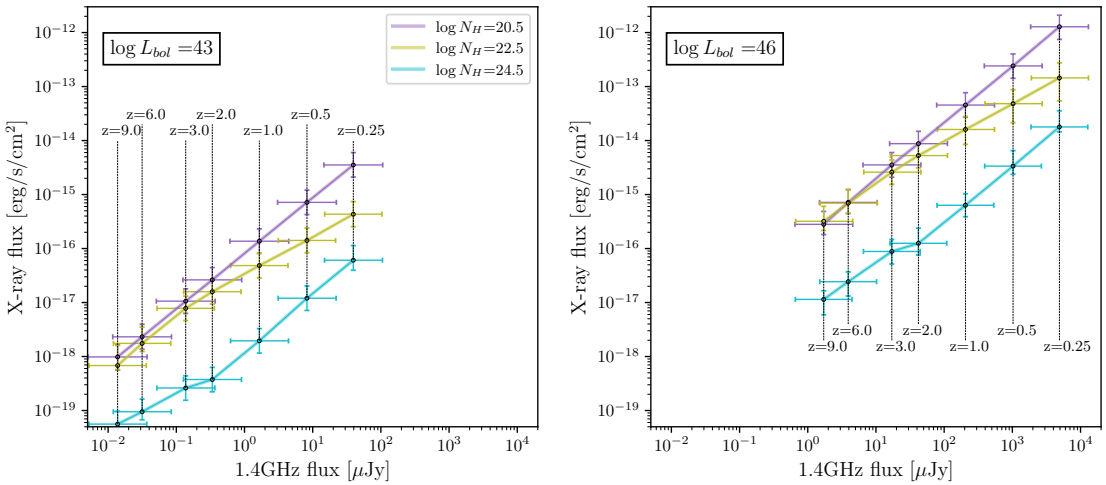

FIGURE 2.8: Typical SB X-ray flux versus 1.4GHz flux at different redshifts computed for an AGN with $\log L_{bol} = 43$ (upper panel) and $\log L_{bol} = 46$ (lower panel). Error bars show the 1σ distribution values due to the scatter in the $L_{bol} - L_{HX}$ relation (vertical error bar), and in the $L_{1.4GHz} - L_{HX}$ relation (horizontal error bar). The dashed lines indicate the redshift at which the fluxes are computed.

still be in agreement with the CXB measurements within their scatter. However, we do not expect the fraction of CTK AGN to be severely higher (or lower) than what we assumed, as suggested by a number of works in the literature that searched for CTK AGN by means of multiband tracers that are in principle not affected by obscuration and can be used as proxies of the intrinsic AGN nuclear power. For example, Daddi et al. (2007) and Fiore et al. (2009) selected CTK AGN by means of their mid-IR excess (hot dust) emission as compared to their X-ray emission, finding space densities of CTK AGN at $z \sim 1.5 - 2.0$ consistent with the CXB model expectations. A similar agreement was found at $z \sim 0.8$ by Vignali et al. (2014) who selected obscured AGN through their narrow [Ne v] λ3426 Å emission line. The uncertainties related to the high-redshift predictions of our model will be discussed in Sect. 2.4.3.

2.4.2 Radio versus X-ray AGN detection

For a given combination of radio and X-ray flux limits the relative efficiency of the two bands in detecting an AGN depends on the shape of the AGN spectral energy distribution, which in turn depends on a number of physical properties like the SMBH's accretion parameters, the presence of radio jets, the duty cycle, the bolometric luminosity and the obscuration level. Here we want to investigate the detection efficiency of 1.4GHz and X-ray selections by varying three main AGN parameters: redshift, bolometric luminosity and obscuration level. Starting from a fixed bolometric luminosity, we computed the HB X-ray AGN luminosity using the bolometric correction derived in Duras et al. (2020), which are valid both for type 1 and type 2 AGN. Then we converted the HB X-ray luminosity into the corresponding SB luminosity assuming a power-law spectrum with photon index $\Gamma = 1.9$ (Piconcelli et al., 2005; Gilli et al., 2007). To derive the X-ray flux we used one of the X-ray source mock catalogs

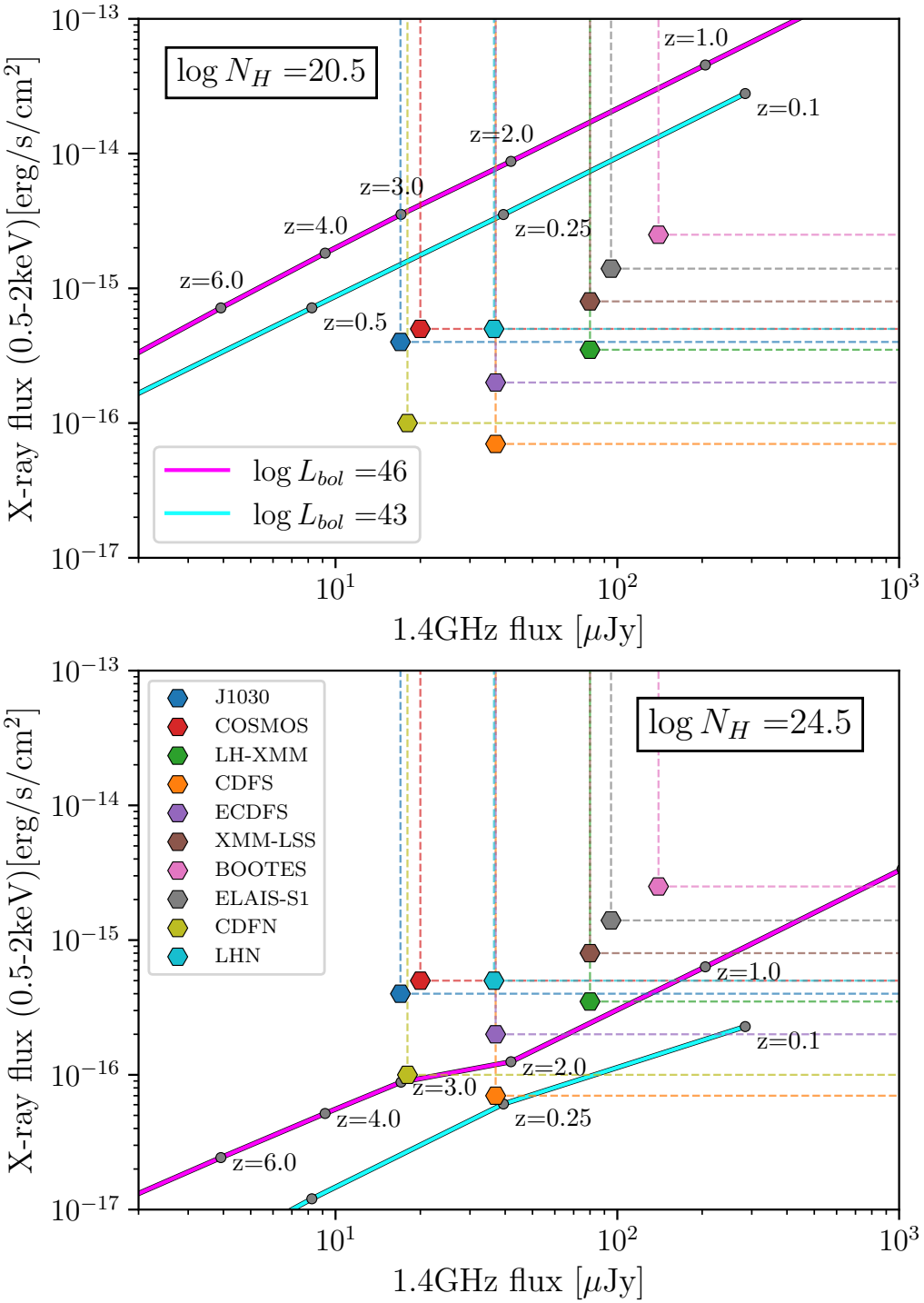

FIGURE 2.9: Median SB X-ray and 1.4GHz detection fluxes for each of the fields color-coded as labeled in the right panel. The two panels consider tracks for two different levels of AGN obscuration: $\log N_H = 20.5$ (left) and $\log N_H = 24.5$ (right). In cyan and in magenta we show the typical $0.5 - 2$keV versus 1.4GHz flux tracks, computed at different redshifts, for an AGN with $\log L_{bol} = 43$ and $\log L_{bol} = 46$, respectively.

presented in Marchesi et al. (2020). The considered catalog contains 5.4M sources simulated down to very faint X-ray fluxes, 10^{-20}erg cm^{-2} s^{-1}, over an area of 100deg^2 and with different obscuring hydrogen column densities, following the prescription of the CXB model in G07. The catalog provides the SB flux and the corresponding SB X-ray luminosity for each source at a given redshift and obscuration level. Therefore, from given combination of SB X-ray luminosity and obscuration, we derived X-ray flux versus redshift dependences.
The intrinsic X-ray luminosity was then converted into 1.4GHz luminosity using Eq. 2.5. Finally, by considering the 1.4GHz luminosity-flux relation:

$$\frac{\mathrm{L_{1.4GHz}}}{\nu_{\mathrm{1.4GHz}}} = 4\pi \mathrm{D_L^2} \frac{\mathrm{S_{1.4GHz}}}{(1+\mathrm{z})^{1+\alpha}}, \tag{2.15}$$

we obtain the corresponding radio flux density ($\mathrm{S_{1.4GHz}}$). In Fig 2.8 we report the AGN X-ray versus radio flux density for redshifts between $z = 0$ and $z = 9$, considering two different bolometric luminosities and three different levels of AGN obscuration. We also report as error bars the 1σ intervals produced by the scatter in the $\mathrm{L_{bol}} - \mathrm{L_{HX}}$ relation ($\sigma = 0.27$ dex, vertical direction), and in the $L_{1.4\mathrm{GHz}} - L_{HX}$ relation ($\sigma = 0.5$ dex, horizontal direction).
At any given redshift, the 1.4GHz AGN flux density is the same for the three levels of AGN obscuration, since the radio emission is not affected by obscuration. On the contrary, the X-ray fluxes at the same redshift differ significantly and decrease with increasing obscuration. We notice that the X-ray flux of an obscured AGN ($\log \mathrm{N_H} = 22.5$) at a given bolometric luminosity tends to approximate the flux of unobscured AGN ($\log \mathrm{N_H} = 20.5$) for $z > 3$. This is due to the effect of the K-correction that shifts absorption out of the X-ray bandpass for AGN at high-z. Considering the two values of $\log \mathrm{L_{bol}}$ the trend of the AGN fluxes are similar for corresponding values of N_H, but with a shift of $\sim$2 orders of magnitude in radio and of 2.5 dex in the X-ray.

To understand which AGN are detectable in the radio and/or X-ray fields considered in this work, we compared these curves with the fields' X-ray and radio median sensitivities. In each panel of Fig. 2.9 we plot the 1.4GHz median flux limit ($5\sigma_{med}$) versus the X-ray sensitivity at 50% of the field area (that can be seen as a median X-ray flux limit). These are compared with the tracks showing X-ray versus radio fluxes for AGN at two given obscuration levels and two different bolometric luminosities.
As shown in Fig 2.9, the tracks for unobscured AGN ($\log \mathrm{N_H} = 20.5$; left panel), intercept the vertical projection of all the fields' positions before intersecting the horizontal ones. This means that, when typical unobscured AGN reaches the median radio sensitivity of the fields, its X-ray flux is largely above the median X-ray sensitivity of the same fields. Therefore, when an unobscured AGN is detected in the radio band, it is already detected also in the X-ray image of the same field. This justifies why, as we see in Sect. 2.3.3, unobscured AGN are preferentially detected in the X-rays for all the considered fields.
A very different behavior is observed for CTK AGN, with $\log \mathrm{N_H} = 24.5$ (right panel). In this case, for most of the fields, the AGN tracks intercept the horizontal projections before the vertical ones, meaning that when one of these AGN becomes detectable in the X-ray it is already detected in the radio image of the field. The only field for which this is not true for both values of $\log \mathrm{L_{bol}}$ is the 7Ms CDFS, where, due to the deep Chandra imaging, X-ray selection is more effective than radio selection in identifying AGN even in the presence of

Results $z > 3$								
Field	Radio (1.4GHz)				X-ray (0.5-2 keV)			
	#AGN model	#CTK model	Σ_{AGN} model	Σ_{CTK} model	#AGN model	#CTK model	Σ_{AGN} model	Σ_{CTK} model
J1030	8.0	3.5	0.013	5.8×10^{-3}	8.8	0.15	0.027	4.7×10^{-4}
COSMOS	68	30	0.008	0.35×10^{-3}	135	0.06	0.017	0.1×10^{-4}
LH-XMM	2.6	1.1	0.002	1×10^{-3}	19	0.19	0.027	2.7×10^{-4}
CDFS	4.6	2	0.004	1.8×10^{-3}	53	4.4	0.113	9.3×10^{-3}
ECDFS	4.6	2	0.004	1.8×10^{-3}	51	1.3	0.046	1.1×10^{-3}
XMM-LSS	28	12	0.002	7×10^{-4}	222	0.005	0.011	1×10^{-5}
BOOTES	11	5.1	0.0005	2×10^{-4}	168	0.11	0.005	1×10^{-5}
ELAIS-S1	7.4	3.3	0.0008	3×10^{-4}	68	1×10^{-4}	0.006	1×10^{-5}
CDFN	4.3	1.9	0.017	7.6×10^{-3}	36	2.3	0.08	5.2×10^{-3}
LHN	11	4.8	0.008	3.3×10^{-3}	31	0.01	0.017	1×10^{-5}

TABLE 2.3: Expected number and surface densities of $z > 3$ AGN (total and CTK) in each field both in the radio (left) and X-ray band (right).

heavy obscuration. In the 2Ms CDFN, the detection efficiency of luminous ($\log L_{bol} = 46$) CTK AGN is the same for both radio and X-ray observations: the tracks in Fig. 2.9 (right) exactly crosses the CDFN datapoints. Less luminous CTK AGN are instead preferentially detected in the radio band.

2.4.3 High redshift predictions

In Tab. 2.3 we report the predictions of our model for AGN and CTK AGN at $z > 3$. For consistency with the assumptions in our baseline model for radio AGN, we use for the X-rays the expectations of Vito et al. (2014) as implemented in the corresponding mock catalog (see Marchesi et al., 2020). The procedure followed to forecast the number of detectable high-z AGN is the same as reported in Sect. 2.3.3 and Sect. 2.3.3.

The change in the starting HXLF also allows us to introduce an intrinsic increment in the fraction of heavily obscured AGN at $z > 3$, that has been observed or predicted by different works in the literature (Aird et al., 2015; Ananna et al., 2019; Gilli et al., 2022; Ni et al., 2020; Lapi et al., 2020).

In the CXB model the fraction of CTK AGN depends on the luminosity and reaches 4/9 only for the low luminosity regime or, equivalently, at the faintest X-ray fluxes (see Eq. 4 in Gilli et al., 2007). On the contrary, in Vito et al. (2014) HXLF, we assumed (according to Vito et al., 2018; Marchesi et al., 2020) a luminosity-independent and constant fraction of CTK AGN equal to 4/9 of the whole AGN population.

The number of X-ray AGN reported in Tab. 2.3 are obtained considering the $0.5 - 2$keV X-ray band, for which the total number of AGN detected at $z > 3$ is larger than in the HB. We verified that using the HB, we would have obtained larger numbers of CTK AGN (within a factor of 2), but we kept the SB information for consistency with the previous sections.

In Fig. 2.10 we show the densities of AGN (upper panel) and CTK AGN (lower panel) both

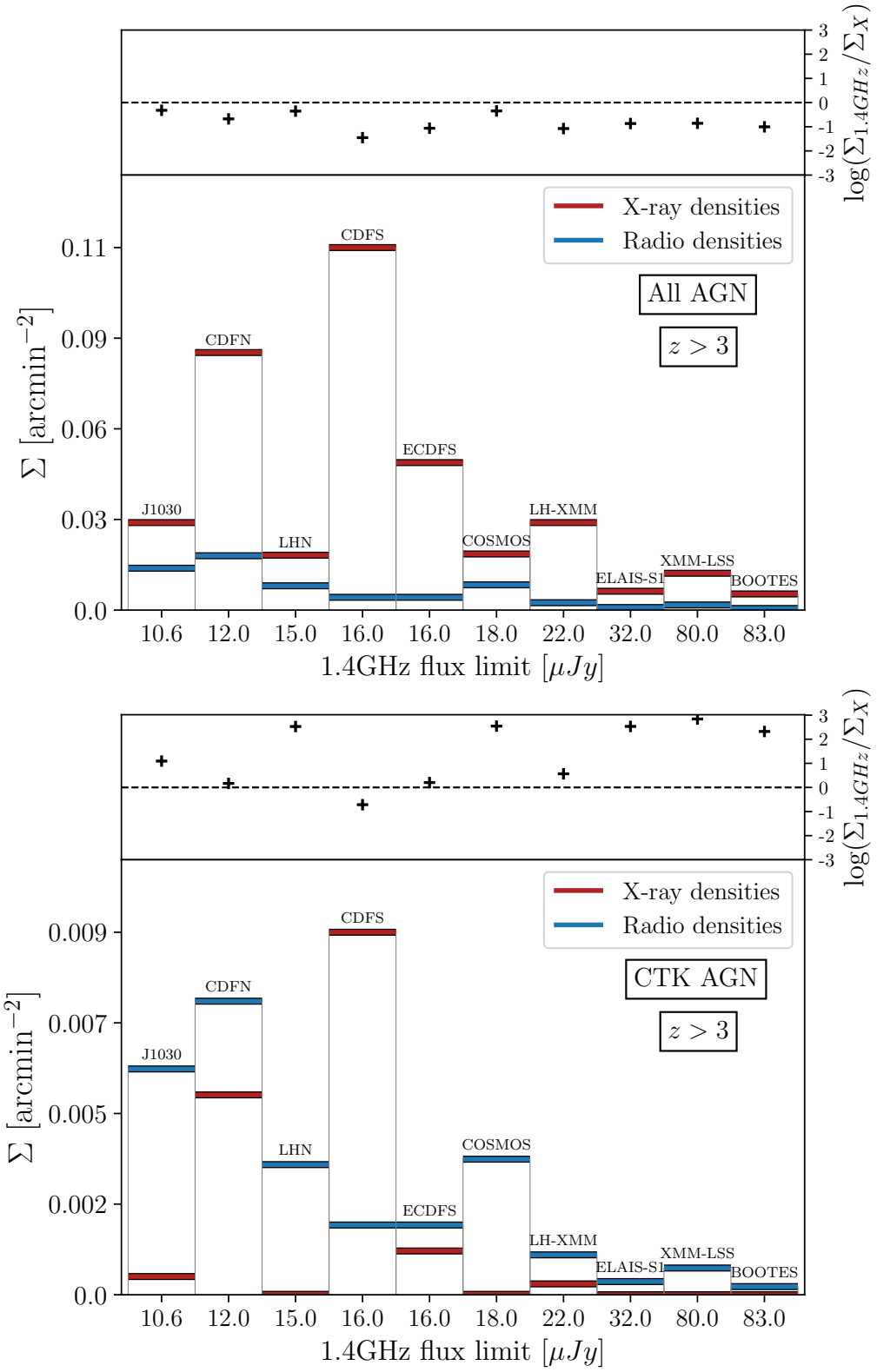

FIGURE 2.10: Same as Fig. 2.7, but for $z > 3$.

in the radio and X-ray band. The densities of $z > 3$ AGN are, for both the bands, ~ 1.5 dex lower than what was predicted for $z > 0$, remarking the difficulty of detecting such high-redshift sources with surveys with the current facilities. Even at $z > 3$, X-ray observations are more efficient in detecting unobscured or mildly obscured AGN (predicting AGN surface densities on average 10 times larger than the radio ones), whereas the radio model predicts much larger CTK AGN densities for all the fields (except the CDFS). At $z > 3$ the surface density of CTK AGN radio detected is generally more than 10 times larger than the X-ray one and more than 100 times larger for half of the fields. This strongly supports the radio selection's effectiveness in detecting heavily obscured AGN, even at high redshift.

2.4.4 SKA predictions

By the end of this decade, the Square Kilometer Array Observatory (SKAO) will become fully operational. As reported in Prandoni & Seymour (2015), SKAO will have among its primary science drivers the investigation of the SMBH-galaxy coevolution and the astrophysics related to the accretion processes. Prandoni & Seymour (2015) have proposed three 1.4GHz radio continuum surveys, optimized to address AGN and galaxy evolution science cases. These surveys are organized in tiers, following a nested wedding cake strategy; the Ultra-Deep, the Deep, and the Wide surveys with the following area – 1σ sensitivity combinations: 0.05μJy over 1 deg^2, 0.2μJy over 10-30 deg^2, 1μJy over 10^3 deg^2, respectively.
Using the number count derived from our 1.4GHz model we computed the number of AGN and CTK AGN detectable in the three different survey tiers for high-redshift ($z > 3$) and very high-redshift ($z > 6$, $z > 10$) ranges. The results for the three tiers are reported in Tab 2.4. All the predictions are computed assuming a flat sensitivity over the survey area. With the wide SKA survey thousands of AGN are expected to be detected at $z > 6$ and a consistent fraction of them is expected to be CTK (at least $> 45\%$). Also for $z > 10$, a redshift range that since the advent of JWST was almost completely unexplored, SKA will be able to detect several tens of AGN.
The $z > 6$ universe is still very unconstrained, particularly for the AGN X-ray and radio LFs. Without better information, we decided to use our baseline radio model and to extrapolate it to higher redshifts. However, we are aware that the results should be taken cautiously. For example, we assumed that the radio-to-X-ray relation measured by D'Amato et al. (2022) holds even at $z >> 3$, but we did not consider whether the increased ISM density at high-z (Gilli et al., 2022) may provide stronger free-free absorption and depress radio emission. Similarly, we did not consider whether the effects of self-synchrotron absorption (SSA) may depress radio emission as well, as one would require on average stronger magnetic fields to overcome energy losses by Comptonization of the enhanced CMB photon field. We defer a detailed discussion of these effects to future work entirely dedicated to SKA forecasts. Another uncertain aspect is the fraction of CTK AGN at these redshifts. While our $z > 3$ baseline model assumes a constant fraction of CTK AGN, the large majority of the AGN above $z > 6$ are expected to be CTK (Gilli et al., 2022; Ni et al., 2020; Lapi et al., 2020; Lambrides et al., 2020; Lusso et al., 2023).
At $z = 6$ the 5σ sensitivities of the three survey correspond to 1.4GHz luminosities of 7.8×10^{38}erg s^{-1}, 3.1×10^{39}erg s^{-1}, 1.5×10^{40}erg s^{-1} respectively for the Ultra-Deep,

Deep and Wide surveys. Converting the 1.4GHz luminosity into the corresponding HB X-ray intrinsic luminosity using Eq. 2.5, and taking the HX to bolometric luminosity correction factor of Duras et al. (2020), we derived the minimum $z > 6$ AGN bolometric luminosity that SKA would detect in each of the three surveys. These bolometric luminosity lower limits corresponds to $L_{bol} \sim 2 \times 10^{44}$erg s^{-1} for the Ultra-Deep, $L_{bol} \sim 10^{45}$erg s^{-1} for the Deep and $L_{bol} \sim 8 \times 10^{45}$erg s^{-1} for the Wide survey.

To investigate the population of AGN that SKA will detect in the three surveys, we used the X-ray AGN mock catalog generated using the Vito et al. (2014) HXLF and provided in Marchesi et al. (2020). This 100 deg^2 catalog includes $\sim 2.5 \times 10^6$ AGN in the redshift range $3 < z < 20$ down to a 0.5-2keV luminsoity $\log L_{SB} = 40$, reaching fluxes below $\sim 2 \times 10^{-20}$erg s^{-1} cm^{-2}. Using the same relations as in Sect. 2.4.2, we derived the bolometric luminosity, the 1.4GHz luminosity and the flux density of each source.

The panels of Fig. 2.11 show the distribution in $\log L_{1.4GHz}$ and $\log L_{bol}$ of the $z > 6$ sources with a radio flux density larger than the sensitivity threshold of the three tiers.

Since the area of the mock catalog is 1/10 of the area of the Wide SKA survey, the number of sources in the Wide Survey panel is almost 1/10 of the AGN predicted for the respective survey at $z > 6$ and reported in Tab. 2.3. The median bolometric luminosities of AGN detected in the three surveys, at $z > 3$, z>6, and $z > 10$ are reported in Tab. 2.4.

We remind that the $L_{1.4GHz} - L_{HX}$ relation used to derive the radio luminosities from the mock X-ray catalog does not include the most powerful RL AGN that should populate a fainter region of the $\log L_{bol}$ distribution. Furthermore, the upper limits of the distributions of the bolometric and radio luminosities presented in Fig 2.11 are affected by the shape of the HXLF of Vito et al. (2014). These HXLF are partially incomplete in their brightest part, since they were computed on deep pencil-beam X-ray fields that miss the most luminous X-ray sources. Taking into account a larger fraction of bright X-ray sources the $\log L_{bol}$ distribution in Fig. 2.11 would extend to larger values.

SKA high-z predictions

Survey	Area	Sensitivity (5σ)	$z > 3$			$z > 6$			$z > 10$		
			#AGN	#CTK	$\log L_{bol}$	#AGN	#CTK	$\log L_{bol}$	#AGN	#CTK	$\log L_{bol}$
Ultra-Deep	1	0.25	1870	831	43.8	34	15	44.4	2	1	44.8
Deep	20	1	13780	6120	44.4	220	98	44.9	6	3	45.6
Wide	1000	5	182000	81000	45.1	1980	880	45.4	35	16	46.6

TABLE 2.4: Expected number of high-z AGN in the three 1.4GHz continuum surveys planned with SKAO. Area and sensitivities of the three tiers are reported in units of $\deg^2$ and μJy, respectively. The values of $\log L_{bol}$ refer to the median values in the simulated samples described in Sect. 2.4.4.

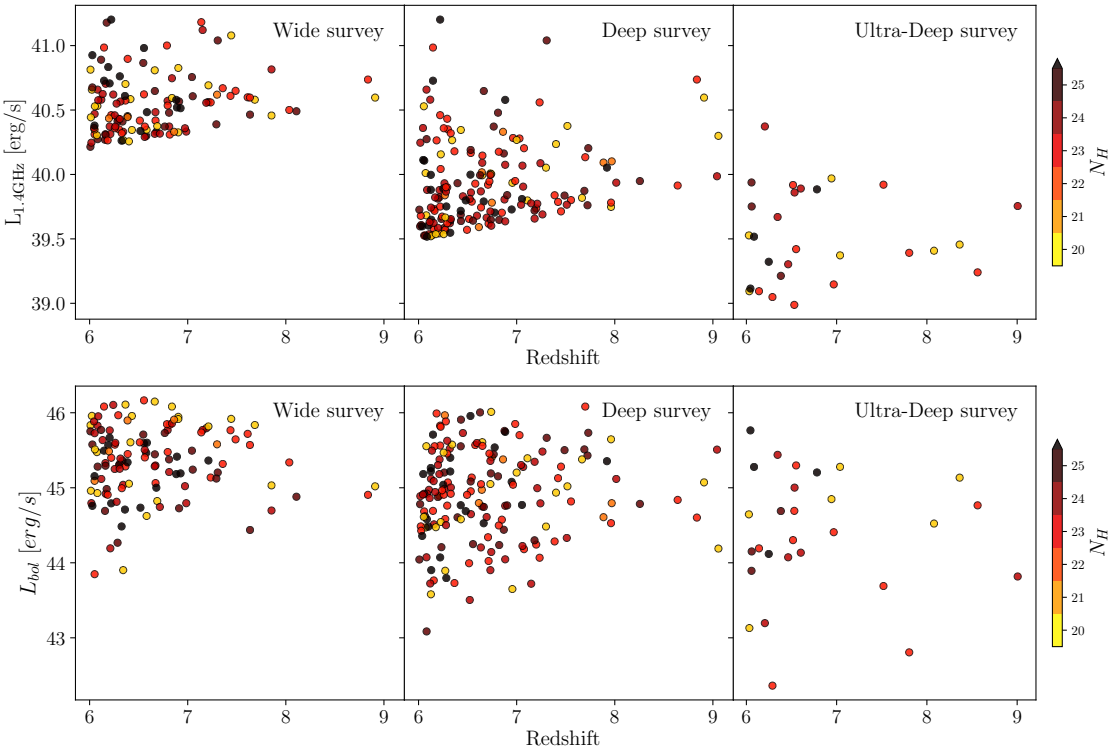

FIGURE 2.11: Simulated $z > 6$ AGN to be detected by SKAO in the three tiers of its continuum 1.4GHz surveys. Each point is color-coded according to its obscuration level. The top three panels show the distribution of $\log L_{1.4GHz}$ while the bottom panels show the distribution of $\log L_{bol}$. The number of simulated sources in the Wide survey panels is almost 1/10 of the number of AGN reported in Tab. 2.4 (see text for details).

The possibility to identify thousands of AGN and CTK AGN at $z > 6$ will transform our understanding of the co-evolution of MBHs with their galaxies in the first Gyr of the cosmic time.

The combination of radio data from SKA with multiband data from current and next-generation facilities will be crucial to detect distant AGN and separate them from SFGs. Indeed SKA surveys are planned to overlap with forthcoming wide-area, sensitive optical and NIR-surveys, like those by LSST and Euclid (see Prandoni & Seymour 2015 for a detailed list of the multiwavelength survey synergies). Therefore, it will be possible to apply radio-excess selection techniques similar to those mentioned earlier in this work (see e.g. Smolčić et al. 2017c) to very large statistical samples, and especially at high redshifts. In addition, in combination with ALMA follow-up data, SKA will prove whether the far-infrared-radio correlation normally used to separate AGN and SFGs is still valid at high-z, and also for low stellar masses and SFR. As an example, the sensitivity of the SKA deep surveys will allow detection of SFR $\sim 10 M_\odot$/yr in galaxies up to $z \sim 3-4$, and SFR $\geq 50 M_\odot$/yr at $z \sim 6-7$ (McAlpine et al., 2015). Observations of well-known extra-galactic fields where dense multiband information, from UV to far-IR, is available will in particular increase the identification efficiency of CTK AGN. JWST and ELT observations will also allow for spectroscopic follow-up of radio-selected, high-z AGN candidates to determine their redshift and ultimately measure

the radio AGN luminosity function at $z > 6$. The combination of the SKA radio observations with future X-ray data from telescopes like STAR-X[3], Athena (Nandra et al., 2013), and AXIS (Mushotzky et al., 2019) will enable improved AGN identification and measurement of their column densities.
In the absence of other multiband AGN diagnostics, the identification of AGN and CTK AGN using SKA data may rely on either multifrequency radio information, or high-resolution (SKA 10GHz observations will provide angular resolution of $0.05" - 0.1"$ compared to the $0.4" - 0.5"$ of 1.4GHz observations) and/or multi-epoch follow-ups. All these diagnostics are amply used in the literature. Flat or convex radio spectral indices would point to the presence of compact AGN cores (O'Dea & Saikia, 2021). High-resolution follow-ups can pinpoint AGN through the measurement of high ($T > 10^{5-6}$ K) brightness temperatures (see Morabito et al., 2022), while multi-epoch observations may identify the tiny fraction of variable AGN at μJy flux density levels (Radcliffe et al., 2019).

2.5 Conclusions

In this chapter, we developed an analytical model to compute the 1.4GHz luminosity function and associated number counts for the whole AGN population and for subpopulations of AGN with different obscuration levels. In particular, we converted into the radio band the prescriptions of the population synthesis model of the CXB using an X-ray to radio luminosity relation derived for faint X-ray sources.
We applied our model to some of the major extragalactic fields covered by deep radio and X-ray observations to predict the number of AGN and CTK AGN detectable in each band. The main results can be summarized as follows:

1. We found a very good agreement between the number of radio-detected AGN predicted by our model in the fields and the observed number of AGN identified via multiple selection techniques. This means that our model is able to give an almost complete census of the radio AGN population, excluding the minority population of RL AGN, that is missed by construction.

2. On average X-rays are able to detect a larger number of AGN, but most of the X-ray detected AGN are unobscured. The CTK AGN surface density detected at 1.4GHz is on average 10 times higher than the X-ray one.
   This result also stands at high redshift ($z > 3$), where the surface density of CTK AGN expected in radio surveys becomes, for some fields, even 10^3 times larger than the X-ray density.

3. Our model predicts the existence of thousands of CTK AGN already detected in the radio fields investigated in this work that are largely missed by the corresponding X-ray observations. Both our model and the results coming from the literature suggest that radio emission may provide an unbiased picture (in terms of obscuration) of the AGN demography. However, multiple selection techniques employing different multiwavelength indicators of nuclear activity are required to identify AGN among SFGs.

[3]http://star-x.xraydeep.org/

The investigation of the CTK AGN selection among the sources in the available radio catalogs will be the focus of the next-coming work.

4. In the future the SKAO will be able to detect in its three-tier surveys show that it will detect more than 2000 AGN at $z > 6$ (down to $\log L_{bol} \sim 43$) and some tens at $z > 10$, opening new windows in exploration of the AGN parameter space at these redshifts.

We demonstrated that radio emission can be a powerful tool to detect the elusive population of heavily obscured, Compton-thick AGN. In the future, the synergies between deep continuum surveys performed by SKA and multiband information provided by present and future telescopes and observatiories, will allow for an extensive testing of radio-based AGN selection tecniques on large statistical samples, and for a detailed exploration of the radio emission properties of these objects in the high-z Universe.

In the next chapter, we will apply the radio-excess selection to the radio sources of the J1030 field, trying to uncover the most obscured AGN population. Taking advantage of deep radio observation, we will demonstrate that radio emission is not only able to unveil a larger population of heavily obscured AGN than corresponding X-ray observations but, at $z \gtrsim 3$, even to return more CTK AGN than what is predicted by the X-ray background models.

Chapter 3

Heavily obscured AGN in the radio band: an observational approach

3.1 Introduction

As we discussed in Sect. 1.2.1, comparing the black hole accretion rate density (BHARD) derived from deep X-ray surveys with those derived from theoretical models, they are in good agreement up to $z < 3$, whereas at larger redshifts the models seem to overpredict the BHARD derived from X-ray surveys. While this issue might be due to an overprediction of SMBH in simulations (Vito et al., 2016), one of the most widely proposed solutions to solve this tension is the existence of a highly obscured AGN population at high redshifts, which has been missed by X-ray surveys (Hickox & Alexander, 2018; Barchiesi et al., 2021).
This possibility is further supported by some recent results obtained by JWST (Yang et al., 2023a), that taking advantage of the JWST photometry or spectroscopy are unveiling a larger population than expected at high-z ($z > 4$, see Sect. 1.4.1), with the large majority of these AGN lacking any X-ray emission. It is therefore necessary to look at this population of heavily obscured AGN using multiple and complementary approaches to possibly retrieve a complete AGN census, in particular at early times.

In the previous Chapter, we investigated the potentialities of radio emission (compared to the X-ray one) in selecting the most obscured AGN thanks to the fact that radio waves are almost unaffected by obscuration. In this Chapter, we want to face the same issue but from an observational point of view.

As presented in Sect. 1.1.4 AGN radio emission is a tracer of the AGN activity, also in RQ AGN. Therefore different radio-excess AGN selection techniques have been developed to separate the radio AGN population from SFG (see Sect. 1.3.1), relying on the fact that AGN (even if RQ) are expected to have a stronger radio emission with respect to normal SFG with the same stellar mass and at the same redshift.
In this chapter, we will investigate the population of radio excess AGN on the field around the quasar SDSS J1030+0524 (the J1030 field, already mentioned in Chapter 2), focusing in particular on those that are expected to be CTK and at high-redshift ($z \gtrsim 2$). In this way, we will test the efficiency of deep radio surveys in selecting the most obscured AGN from an observational point of view, and, in turn, we will measure their abundance for the first time

from a radio perspective.

In Sect. 3.2 we provide a description of the large photometric dataset of the J1030 field, ranging from radio to X-rays, presenting the main scientific cases explored so far with these data, and focusing particularly on the photometric redshifts derived for all the K-band detected sources on the field. In Sect. 3.3, we described the methods used to compute the main physical quantities needed for the selection of radio-excess AGN, and we defined our radio-excess parameter. In Sect. 3.4 we present the sample of radio-excess selected AGN candidates and we compute the expected obscuration values for those that are not X-ray detected. In Sect. 3.5 we discuss the possible contaminants of our AGN selection and we perform a detailed X-ray stacking analysis to prove both the AGN nature of the radio-excess sources and also the high-level of obscuration of those that are not X-ray detected. Finally in Sect. 3.5.5 we compute the radio selected CTK AGN number density and we compare it with results coming from X-ray observations and models. Then in Sect. 3.6 we draw our conclusions and we present a summary of the work presented in this chapter. In Sect. 3.7 we attach the supplementary material of this chapter.

The work presented in this Chapter is part of two papers that will be submitted shortly, the first describing the new J1030 photometric catalog with associated photometric redshift (Mignoli, Mazzolari, et al. in prep.), the second exploiting the heavily obscured AGN radio-excess selection (Mazzolari et al. in prep.). The work associated with this second part has to be considered still in progress, and the results shown as preliminary.

3.2 The J1030 field

Starting from the last decade, the J1030 collaboration began an optical imaging campaign to investigate the existence of galaxy overdensities around four luminous QSOs at $z \sim 6$, using the Large Binocular Camera (LBC) on the Large Binocular Telescope (LBT). The LBT observations were performed in four different bands (g, r, i, z bands) and the LBC's wide field of view (25×25 arcmin^2) enabled the search for $z \sim 6$ Lyman Break Galaxy (LBG) candidates across regions up to 4 Mpc in projection from the quasars. By combining data from all four fields, they identified strong evidence of large-scale LBG overdensities around these early quasars (Morselli et al., 2014). In particular, the region surrounding the SDSS quasar J1030+0524 at z=6.31 (RA: 157.6129317, DEC: 5.4153177), powered by a SMBH of $\sim 10^9 M_\odot$, displayed the highest overdensity of LBG candidates. Further deep near-infrared imaging of the J1030 field with WIRCAM at CFHT (Y, J, K NIR bands) and archival Spitzer/IRAC data (Channels 1,2,3,4 and MIPS Annunziatella et al., 2018) allowed the detection of $z \sim 6$ galaxy candidates down to fainter magnitudes, refined photometric redshifts, and improved contaminant exclusion, further solidifying the large-scale overdensity (Balmaverde et al., 2017). Then, the spectroscopic confirmation of six members of this overdensity determined the first spectroscopic identification of a galaxy large-scale structure around a SMBH in the first billion years of the Universe (Mignoli et al., 2020).
In the meanwhile, the J1030 field gathered a considerable amount of follow-up observations in different bands of the EM spectrum, ranging from radio to X-rays. Besides being part of

the MUSYC survey (UBVRIzJHK bands), it was observed in the mm domain by AzTEC (Zeballos et al., 2018), while the inner square arcmin around the quasar has been observed by HST-ACS and HST-WCF3 (Stiavelli et al., 2005), VLT-MUSE (Mignoli et al., 2020), ALMA (D'Amato et al., 2020; Decarli et al., 2018).
The central region of the J1030 field has also been observed by a JWST NIRCam mosaic (imaging and slitless spectroscopy) as part of the EIGER program (Kashino et al., 2023), and the recently released data will be analyzed soon in Loiacono et al. in prep.
From the spectroscopic point of view the J1030 field has been targeted by multiple optical, near-infrared and mm follow-up spectroscopic observations performed by: LBT/MODS (Mignoli et al., 2020; Marchesi et al., 2021, 2023), Keck/DEIMOS (Decarli et al., 2019; Marchesi et al., 2021; Meyer et al., 2020), VLT/MUSE (Mignoli et al., 2020; Meyer et al., 2020), VLT/FORS2 (Mignoli et al., 2020; Marchesi et al., 2021), LBT/LUCI (Gilli et al., 2019), and NOEMA (Decarli et al., 2019).

The central 17×17 arcmin^2 of the J1030 field was also observed with 500ks of Chandra time in Cycle 17 (PI R.Gilli), with the aim of obtaining the best X-ray spectrum of a $z > 6$ quasar and investigating the presence of moderate-luminosity AGN within the candidate overdensity around it. As presented in Fig. 2.4 of Chapter 2, the Chandra survey of the J1030 field is among the deepest X-ray images performed so far. It reaches a flux limit of $\sim 3, 0.6, 2 \times 10^{-17}$ erg cm^{-2} s^{-1} in the innermost region in the 0.5-7 keV (full), 0.5-2 keV (soft) and 2-7 keV (hard) bands, respectively. The X-ray source catalog, containing 256 X-ray sources was presented in Nanni et al. (2020). On a subsample of 54 obscured AGN candidates was then performed an X-ray spectral analysis with XSPEC to derive both a redshift estimate and the obscuring column densities (Peca et al., 2021). In Signorini et al. (2023), the X-ray spectral analysis was extended to the whole sample of X-ray-detected sources to derive the fraction of obscured AGN (at both $\log N_H > 22$ and $\log N_H > 23$) as a function of redshift and X-ray luminosity.
The sample of X-ray-detected sources on the J1030 field was followed up by multiple spectroscopic programs, with a total of 123 X-ray sources associated with a spectroscopic redshift. In Marchesi et al. (2021), the authors presented the photometric and spectroscopic redshift measurements of the J1030 X-ray sources, showing the effectiveness of X-ray selected AGN to track large-scale structures over physical scales up to several Mpc. Indeed, after the spectroscopic confirmation of a galaxy overdensity around the quasar at $z \sim 6$, two other large-scale structures were detected on the J1030 field, one around the FRII radio galaxy at the center of the field at $z \sim 1.7$ (Gilli et al., 2019) and the other at $z \sim 2.8$ (Marchesi et al., 2021).

The central part of the J1030 field was also observed with VLA at 1.4 GHz, reaching a sensitivity of $\sim 20\mu$Jy (Petric et al., 2003). New radio continuum JVLA imaging was obtained at L-band (1-2 GHz) in 2018. The L-band half primary beam width (HPBW $\sim$27 arcmin) nicely matches the entire region of interest and, in particular, covers the whole Chandra field of view. The $\sim$ 30hrs of JVLA observation at 1.4GHz reached a depth of $\sim 2\mu$Jy of rms in the central region, comparable with that of the deepest extragalactic radio surveys to date (Mazzolari et al., 2024c). The beam of the radio image, defining the angular resolution of the observation, has a major (minor) axis of 1.31 (1.05) arcsec, allowing the best compromise between resolution and sensitivity. The JVLA radio observation of the J1030 field and its related catalog were presented and described in D'Amato et al. (2022).

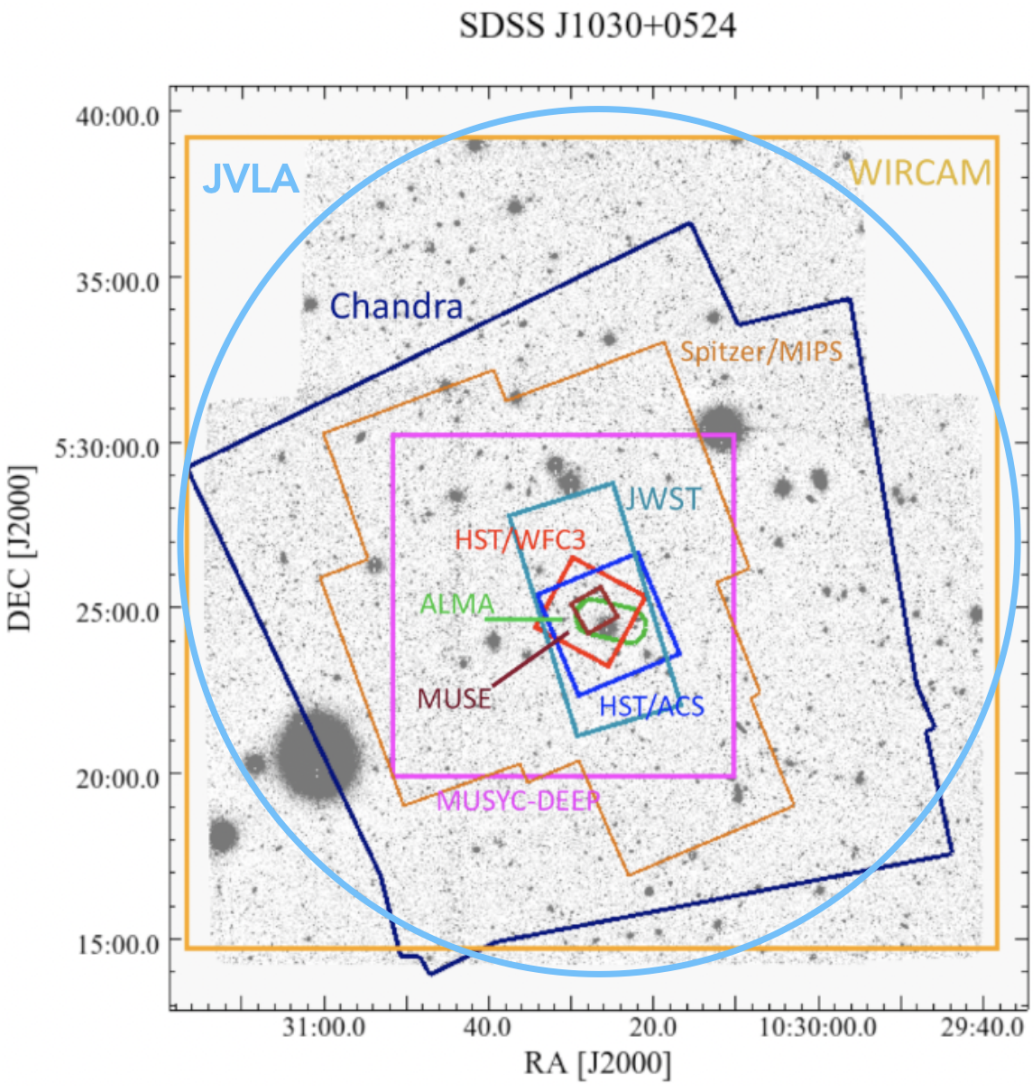

FIGURE 3.1: LBT/LBC z-band image of the field. The coverage of different surveys/instruments is shown. The entire field is also covered by XMM, VLA, MUSYC wide, and Spitzer/IRAC.

A cutout of the J1030 field with the field of view of the different observations performed on it is presented in Fig. 3.1.

3.2.1 Photometric redshift of the J1030 field

The J1030 multiband photometric catalog, contains 14631 sources detected in the deep CHFT/WIRCAM Ks-band (limit AB magnitude $\sim$ 24) and with associated aperture photometry in 11 bands, from MUSYC/U band ($\sim$ 350nm) to IRAC/CH2 ($\sim$ 4500nm). The details on the photometry extraction are described in a work in preparation (Mignoli, Mazzolari et al. in prep). During my PhD, I worked on the photometric redshift catalog associated with these 14631 sources. We decided to use three different SED-fitting codes to derive photometric redshifts (mentioned also in Sect. 1.3.1), and then combine the probability distributions returned by

each code or set of templates to define the final redshift solution. In particular, we used `eazy-py` (Brammer et al., 2008), `LePhare` (Ilbert et al., 2009), and `Hyperz` (Bolzonella et al., 2000). In the following sections, I will briefly describe the settings of the three codes and the procedure that I followed to derive the photometric redshift.
The basic algorithm of these three codes is similar, and essentially steps through a user-defined grid of redshift and, at each redshift, finds the best fitting template spectrum (or linear combination of template spectra) by minimizing the χ^2:

$$\chi^2_{z,i} = \sum_{j=1}^{N_{filt}} \frac{(T_{z,i,j} - F_j)^2}{\delta F_j^2} \tag{3.1}$$

where the sum is done on the number of photometric points available, $T_{z,i,j}$ is the synthetic flux of template i in filter j and for redshift z, F_j is the observed flux in the filter j, and δF_j the uncertainty in F_j. Then the redshift-template combination with the lowest χ^2 is returned as best solution.
The J1030 photometric catalog was converted from AB magnitudes into fluxes (in different units depending on the SED-fitting code). In case of nondetection of a source in a given band, we assumed $F_j = \delta F_j = F(m_{UL})/2$, where $F(m_{UL})$ is the flux corresponding to the upper limit on the magnitude of the j-th band.

`eazy-py`

This is the `Python` version of the C++ photometric redshift code `EAZY` described in Brammer et al. (2008). It is widely used in the literature to derive photometric redshift for wide and deep multiband surveys (such as those on the COSMOS or CANDELS fields Dahlen et al., 2013; Weaver et al., 2022) but also used for JWST surveys (Finkelstein et al., 2023; Castellano et al., 2022).
For the fit of the J1030 catalog, we used two different sets of `eazy-py` templates. The first set is called "tweak fsps QSF v12 v3", and is constituted by 12 templates based on the Flexible Stellar Population Synthesis code (FSPS Conroy & Gunn, 2010), including dust attenuation and a more representative set of templates (inclusive of emission lines) than the original `EAZY` templates. This template set has further been corrected (or "tweaked") for systematic offsets observed between data and the models. The second set of templates is referred in `eazy-py` as "sfhz" and is made of 13 templates with redshift-dependent star formation histories (SFHs) that disfavor SFHs that start earlier than the age of the universe at a given epoch. While fitting, `eazy-py` linearly combines the templates of a given set to reproduce at best the photometry of the source.
The photometric redshift grid explored by this code (as well as by the other codes) ranges from $z = 0$ to $z = 9$ in steps of $\Delta z = 0.01$.
We also used `eazy-py` to derive zero-point correction to the photometry of the J1030 catalog. These are computed by `eazy-py` by performing a fitting of the galaxies with a spectroscopic redshift, iteratively fitting their SED, and computing the median fractional residual between photometry and templates in every filter. After five iterations, the correction routine is stopped because the updates on the photometry correction become generally negligible. We found zero-point corrections of the order of a few percent for most of the bands, except for the U, g, and i filters, for which the flux corrections were of the order of $\sim 20\%$.

We tested the robustness of the zero-point corrections using different sets of templates, and then we applied these corrections directly to the photometry of our catalog. To improve the fit, `eazy-py` can also take an apparent magnitude prior. This prior was introduced to mitigate the degeneracies between template colors and redshift, helping to choose the correct photometric redshift solution when multiple peaks in the redshift probability distribution are present. However, we decided not to include the prior into the final `eazy-py` run because we noted that it produced a larger scatter in the z_{phot} versus z_{spec} plane and also in the K_{AB} versus z relation derived for radio galaxies (Willott et al., 2003).
The code `eazy-py`, besides performing the SED-fitting with galaxy templates, also allows to perform a fit with a library of stellar templates to identify stars among the catalog sources.

`LePhare`

The `LePhare` SED-fitting code was originally described in Arnouts et al. (2002); Ilbert et al. (2006). This code allows the fit with galaxy and AGN templates separately. In particular, the galaxy templates include the elliptical and spiral galaxy models from Polletta et al. (2007) (19 templates) and 12 blue star-forming galaxy models from Bruzual & Charlot (2003). The AGN templates include models from Salvato et al. (2009, 2011). For these AGN templates in some cases, the AGN is fully dominant, while in others the template is a hybrid where both the AGN and the host contribute to the overall SED, with the relative contribution of the AGN component varying in the range 70–90%. Also in `LePhare` are available stellar templates that allow the identification of sources that are best fitted by star SED. We include in the fitting also the extinction, considering different attenuation curves (Calzetti et al., 2000; Prevot et al., 1984, and two modified Calzetti law with a bump at 2175 Å) and allowing $E(B-V)$ to vary between 0 and 2.

`Hyperz`

The `Hyperz` was originally presented in Bolzonella et al. (2000), and with this code we followed the SED-fitting procedure described in Marchesi et al. (2021). The fitting is performed twice for each object, the first using galaxy templates and the second using AGN templates. The first set comprises 75 templates dominated by stellar emission. More in detail, we used the same galaxy templates used for `LePhare` plus 44 additional templates based on the Bruzual & Charlot (2003) stellar population synthesis models (Ilbert et al., 2013). For the fit with galaxy templates, we allowed the absolute, non-absorption corrected magnitude of the sources to vary in the range $r_{AB} = [-30 : -16]$. This range is fairly conservative, particularly at the high-luminosity end. We then fit our SEDs using the 30 templates originally reported in Salvato et al. (2009) including templates with combinations of galaxy+AGN emission down to lower AGN fraction compared to the templates used in `LePhare`. For AGN templates we set the absolute r-band magnitude to vary between $r_{AB} = [-30 : -21]$.
The exctinction is taken into account considering the Calzetti et al. (2000) law, and allowing A_V to range [0,6] in step o $\Delta A_V = 0.02$.
No stellar templates are available in `Hyperz`.

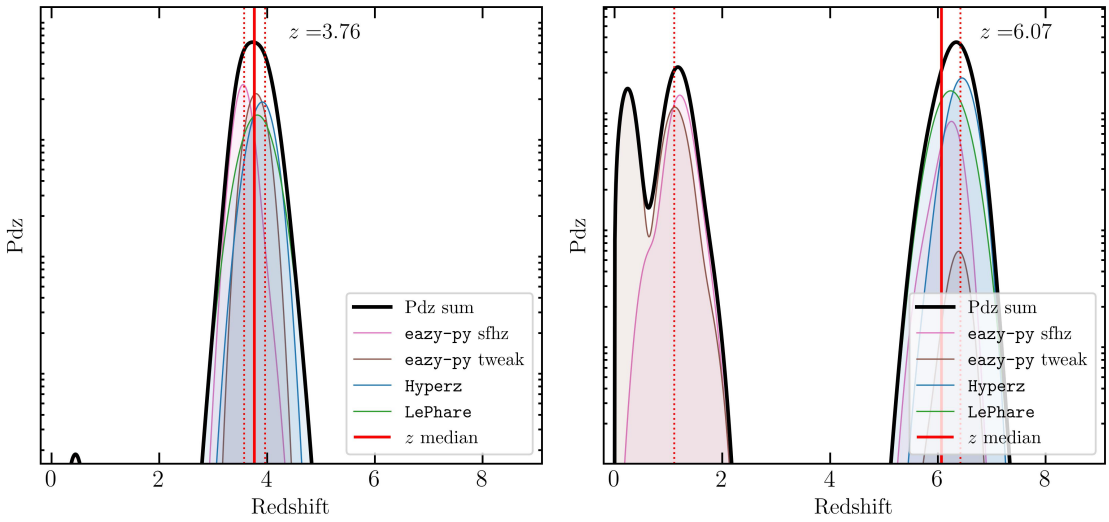

FIGURE 3.2: Examples of two Pdz obtained from the sum of the Pdz of each of the codes used to derive photometric redshift solutions, as described in Sect. 3.2.1. The red vertical line marks the position of the final photometric redshift solution (median of the Pdz), while the dashed vertical lines mark the upper and lower errors.

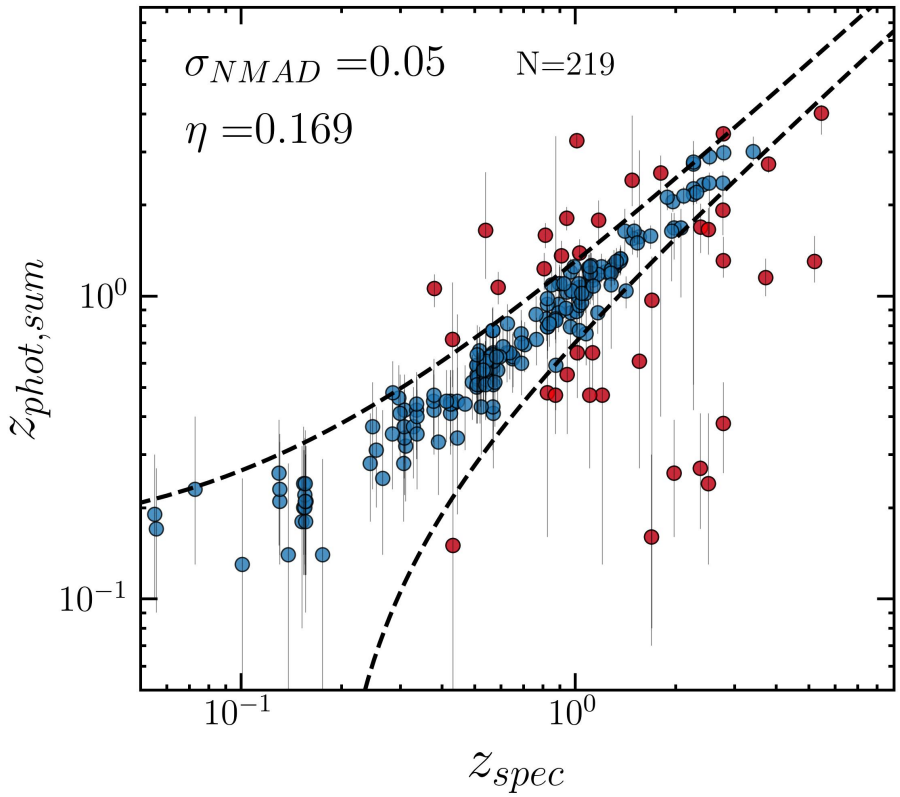

FIGURE 3.3: Photometric versus spectroscopic redshift. The photometric redshifts on the y-axis are those derived from the sum of the Pdz. The dashed line marks the curves $z_{phot} = z_{spec} \pm 0.15$. Red sources represent outliers with $|\Delta z| > 0.15(1 + z_{spec})$.

3.2.2 Validation and combination of photometric redshifts

In total, we derived six different photometric redshift solutions, two for each code, four derived only with galaxy templates, and two also including AGN emission. To validate these photometric redshifts, we used the sample of spectroscopic redshift available on the J1030 field. In total, we have 283 sources with a spectroscopic redshift, including BLAGN, NLAGN, passive and star-forming galaxies, and stars. A large fraction of the spectroscopic redshifts of the J1030 field were obtained with follow-up spectroscopic observations of the X-ray detected sources. Therefore, our spectroscopic sample is biased towards AGN, with 17% of the extragalactic sources having a spectroscopic redshift classified as BLAGN and 10% as NLAGN. However, these sources are not representative of the general population of extragalactic sources (they are expected to be $\sim 1\%$), and additionally, their photometry cannot be correctly reproduced by the galaxy templates. Consequently, we decided to exclude BLAGN from the reference spectroscopic sample used to validate the photometric redshifts of the J1030 catalog, given that they can bias the validation. The final spectroscopic catalog is made by 219 sources.
To quantify the precision of the photo-z, we used the normalized median absolute deviation (NMAD, Hoaglin et al., 1983), defined as:

$$\sigma_{NMAD} = 1.48 \times median\left(\frac{|\Delta z|}{1 + z_{spec}}\right) \tag{3.2}$$

where Δz is $z_{phot} - z_{spec}$. Then, to quantify the fraction of outliers η, following Hildebrandt et al. (2012), we considered those galaxies whose photo-z deviate from their spec-z by $|\Delta z| > 0.15(1 + z_{spec})$. We also considered the relative outlier fraction η_{rel} defined as the fraction of sources for which $|\Delta z| > 3\sigma_{NMAD}$.
The values of the different photo-z quality parameters for the different sets of code-template are reported in Table 3.1. All the photometric redshift code performs similarly in terms of σ_{NMAD} and η, except for the solutions obtained using only AGN templates that perform significantly worse, as expected.
Photo-z uncertainties are another important aspect of the assessment of the photo-z quality. If they are correctly estimated by the code (i.e., they correctly represent the 1σ uncertainty) the fraction of spec-z that falls into the interval $[z_{phot} - z_{err}, z_{phot} + z_{err}]$ should be 0.68. For all the solutions, this fraction was initially significantly smaller, because the codes underestimated photometric uncertainties (as it is often the case). Therefore, to enlarge the error bars, we applied a smoothing procedure to the Pdz, following the example reported in Dahlen et al. (2013). For each code, we iterated the smoothing procedure till the fraction of z_{spec} that belongs to the interval $[z_{phot} - z_{err}, z_{phot} + z_{err}]$ is 0.68.

Given that the different photometric redshift solutions perform similarly in terms of η and σ_{NMAD}, we decided to derive the final photometric redshift solution combining the results of the codes. As demonstrated in Dahlen et al. (2013), this approach, by combining outputs from multiple algorithms (that have different priors and/or templates), can reduce both the effect of systematic errors associated with single codes and also the scatter of the z_{phot} final solution around the true redshift. Therefore, we adopted the following strategy. For each source, we summed the full smoothed Pdz returned by the different codes and then

Code	σ_{NMAD}	η [%]	η_{rel} [%]
`eazy`-'tweak'	0.044	16.9	21.5
`eazy`-'sfhz'	0.051	18.3	18.3
`LePhare`-galaxy	0.085	23.7	13.7
`LePhare`-AGN	0.088	41.1	20.5
`Hyperz`-galaxy	0.05	17.8	17.8
`Hyperz`-AGN	0.112	32.9	22.4
`Pdz` sum	0.049	16.9	19.2

TABLE 3.1: Values of σ_{NMAD}, absolute outlier fraction η and relative outlier fraction η_{rel} for the different codes and set of templates. The last row reports the values for the final photometric redshift solution returned by the sum of the Pdz of the different codes.

took as best z_{phot} the value corresponding to the median of the summed Pdz. Contrary to using the strict median of the different redshifts, this approach allows the computation of the full summed probability distribution function, which can be used to compute the photometric redshift errors in a consistent way. Indeed, we set as upper and lower errors the values corresponding to the 16 and 84 percentile of the Pdz distribution.
Before summing the single Pdz of the codes, for the redshift solution obtained with `LePhare` and `Hyperz`, we considered for each source only the Pdz corresponding to the AGN or galaxy solution with the lowest χ^2 returned by the fit. This allows weighing more any possible AGN solution in the final Pdz. The examples of two sum Pdz are presented in Fig. 3.2.
We then checked the σ_{NMAD}, η, and η_{rel} values for the photometric redshift solutions obtained with this technique, which are reported in the last row of Table 3.1 and also shown in Fig. 3.3. The σ_{NMAD} and η values of the sum solution do not substantially improve the quality of the photo-z but still provide the overall best values of the quality assessment parameters. The fact that the adopted procedure does not significantly improve the results is probably due to the limit of the spectroscopic sample available, both in terms of number of sources and bias of the population dominating the spectroscopic catalog (X-ray sources).

3.2.3 Match with the radio catalog

The JVLA catalog of the J1030 field contains 1486 single sources, with 1102 falling into the multi-wavelength catalog footprint. Using the counterpart matching code `NWAY` (Salvato et al., 2018), we search the counterparts of the radio sources among the positions of the J1030 multiband catalog sources, using a matching radius of 1.5". `NWAY` is a likelihood matching

tool that allows setting a prior on the magnitude of the objects to look for the more probable counterpart in cases in which the counterpart association can be dubious. In particular, for each source the code returns two probabilities that can be used to determine the correct counterpart. One is p_{any}, which represents the probability that the source from the first catalog has a counterpart among the sources of the second catalog, while p_i is the probability of each possible match being the correct one. To further check all the associations, we visually inspect all those matches with $p_{any} < 0.5$ or $p_i < 0.5$. The matching procedure returned 1003 radio sources with a counterpart in the multiband catalog (91% of the radio sources in the multiband catalog footprint), while 99 sources are without a K-band counterpart. Of these 99 sources, 22 are without a counterpart in any of the optical or NIR bands, 45 are visible only starting from the IRAC bands (Sapori et al. in prep.), while the remaining are blended sources or are undetected in the K-band but have a counterpart only in bluer bands (14 sources).
Using the same procedure, we also cross-matched the position of sources in the radio catalog with the X-ray one. There are 763 radio sources with a multiband counterpart in the Chandra footprint; 101 of them are also associated with an X-ray-detected source. There are 2 additional radio sources with an X-ray counterpart that are not matched with any sources in the J1030 multiband catalog.
Among the 1003 radio sources, 95 have a spectroscopic redshift (of these, 62 are also X-ray sources).

3.3 Methods

In this section we describe the methods used to derive the main quantities needed for the identification of radio-excess AGN among the radio sources.

3.3.1 Radio and SED fitting star formation rates

Since all the sources are radio-detected, we can compute their 1.4GHz luminosities and then the SFR under the assumption that all of the radio luminosity is due to SF processes. Following Novak et al. (2017), we first computed the radio luminosity using:

$$L_{\nu,1.4GHz} = \frac{4\pi d_L^2 S_{1.4GHz}}{(1+z)^{1+\alpha}} \; WHz^{-1}, \tag{3.3}$$

where $L_{\nu,1.4GHz}$ is the radio luminosity density, d_L is the luminosity distance and α the radio spectral index. Then the corresponding radio SFR ($SFR_{1.4GHz}$) is given by:

$$\mathrm{SFR_{1.4GHz}} = \mathrm{f_{IMF}} \times 10^{-24} 10^{\mathrm{qTIR}} \frac{\mathrm{L_{\nu,1.4GHz}}}{\mathrm{WHz^{-1}}} \; \mathrm{M_\odot yr^{-1}}, \tag{3.4}$$

where f_{IMF} is a factor depending on the assumed IMF and corresponds to $f_{IMF} = 1$ for a Chabrier IMF and to $f_{IMF} = 1.7$ for a Salpeter one. For consistency with the results from the SED fitting, we used $f_{IMF} = 1$. In Eq. 3.4, q_{TIR} represents the ratio between the total infrared luminosity and the radio luminosity at 1.4GHz, as derived assuming the far infrared-radio correlation for SFG (presented in Sect. 1.3.1). This correlation was found to

depend both on the redshift and the stellar mass (M_*) of the sources (Delvecchio et al., 2021), as follows:

$$q_{TIR} = 2.646 \times (1+z)^{-0.023} - 0.148 \times \log\left(\frac{M_*}{10^{10} M_\odot}\right). \quad (3.5)$$

In particular, we used in Eq. 3.5 the value of M_* derived from the SED fitting with `CIGALE`. Indeed, to infer the physical properties of radio-detected galaxies, we performed a SED fitting analysis using `CIGALE` (Boquien et al., 2019; Yang et al., 2020), considering the eleven bands in the optical and NIR from the multiband photometric catalog. In particular, we used in `CIGALE` delayed star formation history (SFH) models that are able to reproduce both early-type and late-type galaxies. We adopted stellar templates from Bruzual & Charlot (2003), and a Chabrier initial mass function (Chabrier, 2003). We also include the nebular emission module, which is extremely important to account for the contribution of emission lines in the broad-band photometry (Schaerer & de Barros, 2012; Salvato et al., 2019). For the attenuation of the stellar continuum emission, we considered the `dustatt_modified_CF00` module (Charlot & Fall, 2000), which allows different attenuations for the young and old stellar populations, and exploring different levels of attenuation $0.01 < A_V < 3.5$. We also include dust emission in the IR following the empirical templates of Dale et al. (2014).
We did not include the radio data point in the fit because we wanted to derive the main physical parameters from the SED fitting considering only the optical-NIR emission of the sources. Indeed, excluding the radio information from the fit allows us not to contaminate the SFRs from a potential AGN-related radio excess emission, which would lead to a falsely larger SFR and limit the capabilities of the radio excess selection, as we will describe later. According to this, we also did not include in the SED fitting any AGN module, standardizing the fit of all the sources as if they were normal SFG, deferring the AGN identification to the radio-excess analysis shown in Sect. 3.3.2. Only for the X-ray detected sources, which we already know being AGN according to previous works, we performed the fit also including the AGN module in `CIGALE`. In particular, we employed the `skirtor2016` module introduced in Yang et al. (2020), which is widely used in the community and has demonstrated reliability in studying various aspects of AGN (e.g., Mountrichas et al., 2022; López et al., 2023; Yang et al., 2023a). The SED produced by this AGN module combines emissions from the accretion disk, torus, and polar dust.
We used the SED-fitting to derive two main parameters: the stellar mass M_* and the extinction corrected star formation rate SFR^{corr}_{SED}.

3.3.2 Radio Excess parameter

The radio-excess AGN selection techniques are based on the identification of those sources whose emission in the radio band can be justified only with an additional AGN component on top of the SF emission. We adopt as radio-excess parameter the ratio between the SFR derived directly from the radio luminosity and the SFR derived from the SED fitting to the optical-NIR photometry:

$$\mathrm{REX} = \frac{\mathrm{SFR}_{1.4\mathrm{GHz}}}{\mathrm{SFR}^{\mathrm{corr}}_{\mathrm{SED}}}. \quad (3.6)$$

Given that the stars emitting in the optical and NIR bands are the same producing the radio emission by means of supernovae explosions, we would expect normal SFG to be distributed

around $REX \simeq 1$, with a certain scatter. On the contrary, those sources hosting an AGN should show larger values of REX, because of the additional radio emission connected to the nuclear processes and not to SF. In the next section, we will describe how to define a threshold in REX to separate the AGN population from SFG.

3.4 Results

3.4.1 Radio excess AGN selection

In Fig. 3.4 we show the distribution of all the radio sources with a counterpart in the multi-band photometric catalog of the J1030 field in the redshift versus REX parameter plane (with REX values distributed in logarithmic space). As it is possible to see from the upper panel of Fig. 3.4, the peak of the marginal distribution in REX is centered almost exactly on $REX = 1$, as one would expect considering that for normal SFG the SFR computed from the optical-NIR photometry and the one from the radio luminosity corresponds. However, the REX distribution is highly asymmetric around $REX = 1$, showing a significant tail extending up to $REX \sim 10^4$, probably dominated by AGN. Therefore, if we consider the peak of this distribution as dominated by SFG, we can mirror the left-hand part of the log-histogram with respect to the peak and fit this mirrored distribution with a Gaussian function. The fit in the $\log REX$ space returned a $\sigma_{\log REX} = 0.31$ representing the intrinsic dispersion of the SFG population around the mean $\mu_{\log REX} = -0.03$ (equivalent to $\mu_{REX} = 0.92$). Given the distribution of SFG, we can identify as radio-excess sources all the sources that are 3σ outliers from the peak of the Gaussian, namely all those sources with:

$$\mathrm{REX} \geq 10^{\mu_{\log \mathrm{REX}} + 3\sigma_{\log \mathrm{REX}}} \simeq 8.5. \tag{3.7}$$

We can confidently assume that for radio-excess sources the radio emission that produces the excess in the observed $SFR_{1.4GHz}$ can instead be attributed to an AGN-related radio emission, as we are going to further show in the next Sections. We note that the uncertainties on the SFR^{corr}_{SED} can potentially scatter objects that are not AGN into the radio excess zone. However, these uncertainties have the primary effect of broadening the Gaussian around the peak value of the distribution of REX. Given that our radio-excess selection is not based on a fixed value but is defined in terms of σ of the SFG distribution, the impact of SED-fitting uncertainties on the radio-excess selection is largely mitigated. Among the 1003 radio sources with a multiwavelength counterpart, we identified 240 radio-excess sources, 188 of them falling inside the Chandra footprint and 145 without an X-ray counterpart.
In Fig. 3.4, we also show the positions of the 101 radio sources with an X-ray counterpart, which we already know to be AGN. As shown by the cumulative distribution of these sources in REX (upper panel), they are significantly shifted towards higher values of REX with respect to the cumulative distribution of all the radio sources. We performed a Kolmogorov-Smirnov (KS) test of the two distributions, finding a p-value~ 0.001, implying that the two distributions are different at the $> 99\%$ confidence level. This strongly supports the hypothesis that radio-excess sources are dominated by AGN.
In Sect. 3.7.1 of the supplementary material we further show that the large majority of the sources selected as radio-excess by the REX parameter are identified as radio-excess also

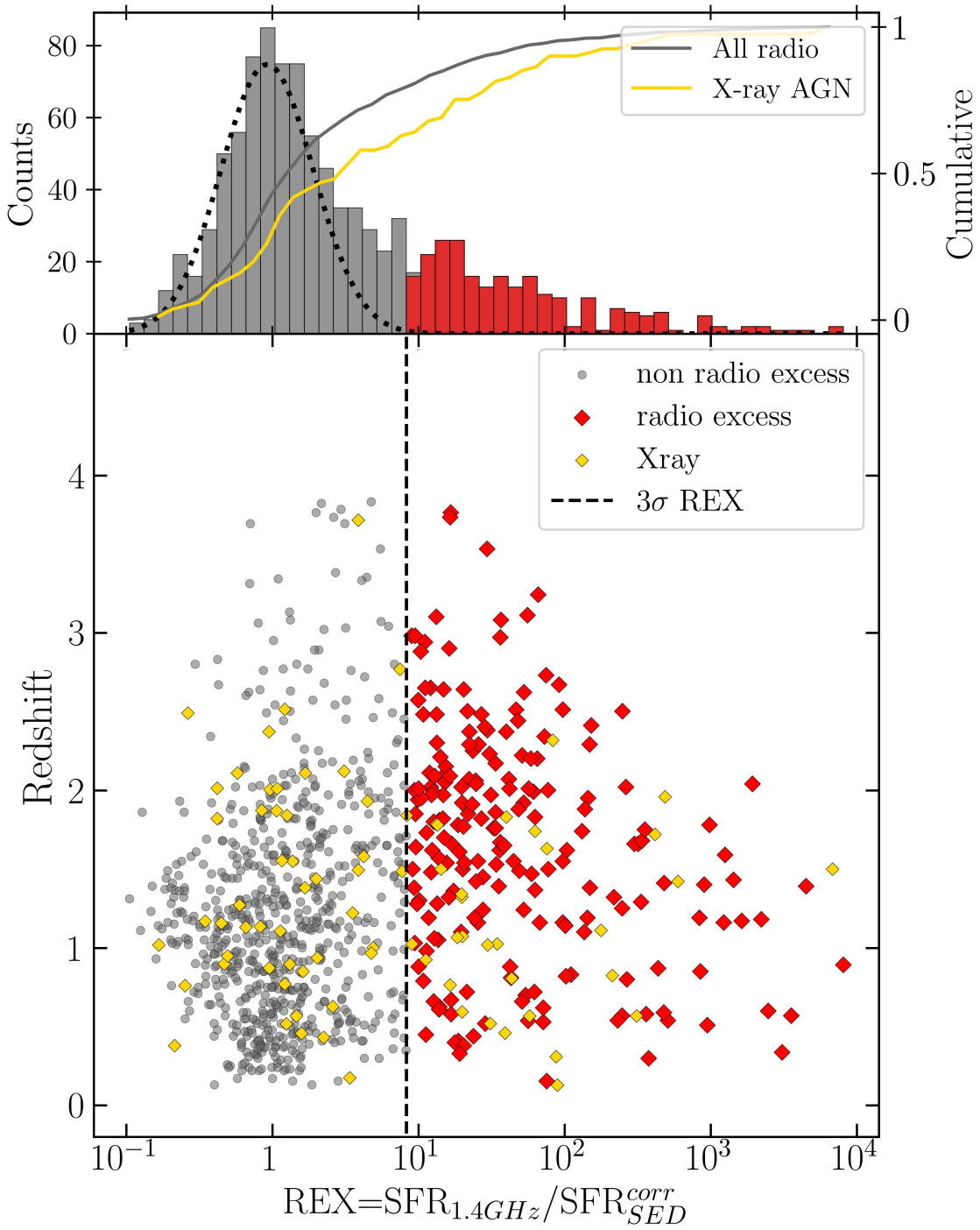

FIGURE 3.4: Redshift versus REX parameter of all the radio sources in the J1030 field with a counterpart in the K-band selected multiwavelength catalog. The red squares represent the radio-excess sources, while gold symbols are used for X-ray detected sources. The dashed line remarks the 3σ deviation from the peak of the Gaussian distribution of the REX parameter (upper panel), which is likely to be dominated by SFGs being $REX \simeq 1$. The histogram in the upper panel shows the distribution of the sources according to the REX parameter, while the gray and gold lines are the cumulative distribution of all the radio sources and of the X-ray detected sources, respectively.

using the q_{TIR} parameter derived combining the final `CIGALE` SED and the radio luminosity. Contrary to the q_{TIR}, the REX parameter adopted in this work allows the selection of radio-excess sources also when FIR photometry is not available, and without being significantly affected by contaminants (as we demonstrate in Sect. 3.5.1).

3.4.2 $SFR_{1.4GHz}$ and SFR_{SED}^{corr} distribution of radio sources

It is also interesting to investigate the distribution of the sources according to their SFR_{SED}^{corr} and $SFR_{1.4GHz}$, separately. In Fig. 3.5, we show the distribution of the radio-excess and non radio-excess sources according to their redshift and to the SFR_{SED}^{corr} derived from the SED fitting (upper panel) and from their radio luminosity (lower panel), normalized to their expected main-sequence (MS) SFR. In particular, to derive the expected MS SFR (SFR_{MS}), we used the result derived in Popesso et al. (2023), considering the values of redshift and M_* retrieved from the `CIGALE` SED-fitting.
We note that, in general, radio excess sources are characterized by lower SFR_{SED}^{corr} with respect to what is expected from their MS. This is particularly true for $z < 2$, while for sources at $z \sim 3$, the ratio between the two SFRs is ~ 1 (0 in Log). This distribution is expected for sources hosting AGNs at low redshifts. Indeed, different works (Bugiani et al., 2024; Piotrowska et al., 2022a; Bluck et al., 2023a) found that the integrated effect of the AGN feedback on the host galaxy tend to quench their SFR, This effect is observed in particular at low redshifts where many radio galaxies hosting AGN show a population of old stars associated with inefficient SF processes. Instead, at higher redshifts, this trend is less clear, and radio galaxies hosting AGN are observed also in star-forming systems.
On the contrary, when we consider the distribution of the $SFR_{1.4GHz}$ in the lower panel of Fig. 3.5, we note that the bulk of the radio-excess sources are distributed according to the general distribution of the radio sources, but that $\sim$ 30% of the radio-excess sources show $SFR_{1.4GHz}/SFR_{MS} > 10$, suggesting the presence of a consistent additional radio-emission.

3.4.3 Obscuration level

In this work, we do not want to investigate only the radio-excess AGN population, but we want to use the radio emission to uncover the population of heavily obscured AGN.
To do this we considered only those radio sources that fall in the Chandra X-ray image footprint of the J1030 field (763 sources) but that are not detected in the deep X-ray catalog (662 sources). The fact that these sources are not X-ray detected means that, if they are AGN, they should be heavily obscured given the depth of the 500ks Chandra observation. To quantify the level of obscuration, we need to find the minimum level of $\log N_H$ that, given the redshift and the intrinsic X-ray luminosity of the source, would return a non-detection in the X-ray image at the position of the source. Therefore, we first measure the X-ray flux limits in the soft band (SB, 0.5-2keV) at the position of the non-X-ray detected radio-excess sources (that are 145). Since we know their radio luminosity, we can compute their intrinsic X-ray luminosity using the $L_{1.4GHz} - L_{2-10keV}$ luminosity relation derived in D'Amato et al. (2022). This relation was derived considering a sample of 89 X-ray detected, spectroscopically confirmed AGN and Early Type Galaxies (ETG) on the J1030 field extending up

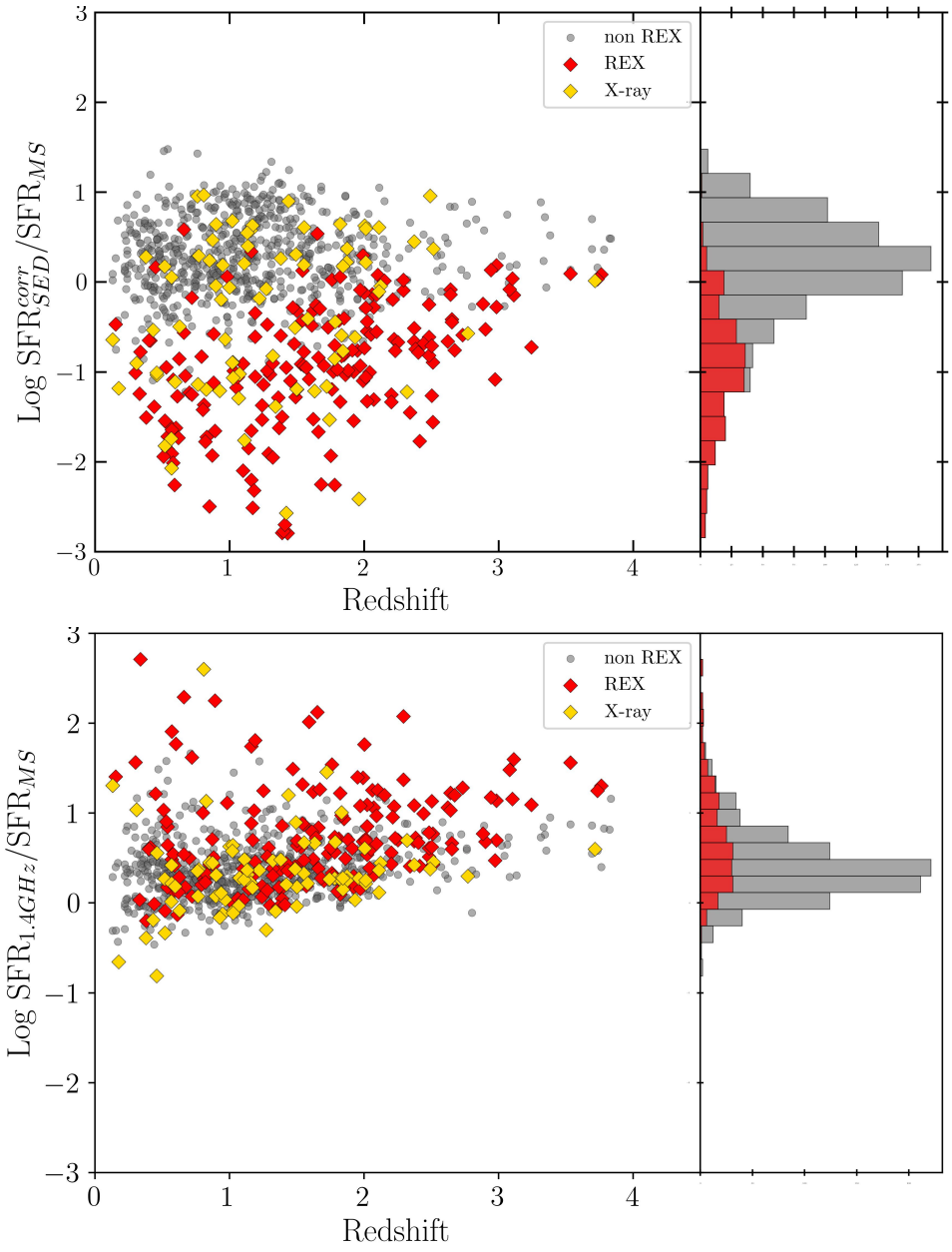

FIGURE 3.5: Distribution of the radio sources with a counterpart in the J1030 multiwavelength catalog according to their redshift and to the ratio between their radio (lower panel) and SED-fitting (upper panel) SFR, normalized to the expected MS SFR. Red data points represent radio-excess sources, while golden data points represent X-ray-detected sources.

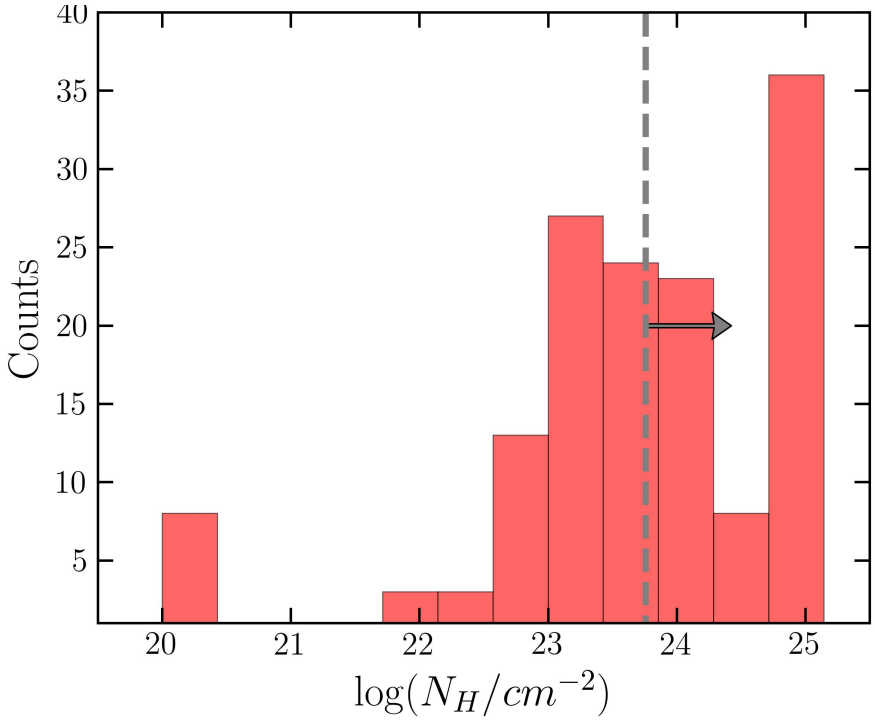

FIGURE 3.6: Histogram showing the distribution of the lower limit to the value of $\log N_H$ estimated as described in Sect. 3.4.3 for the population of radio-excess and X-ray undetected AGN candidates. The gray dashed line marks the median value of the distribution corresponding to $\log N_H = 23.8$.

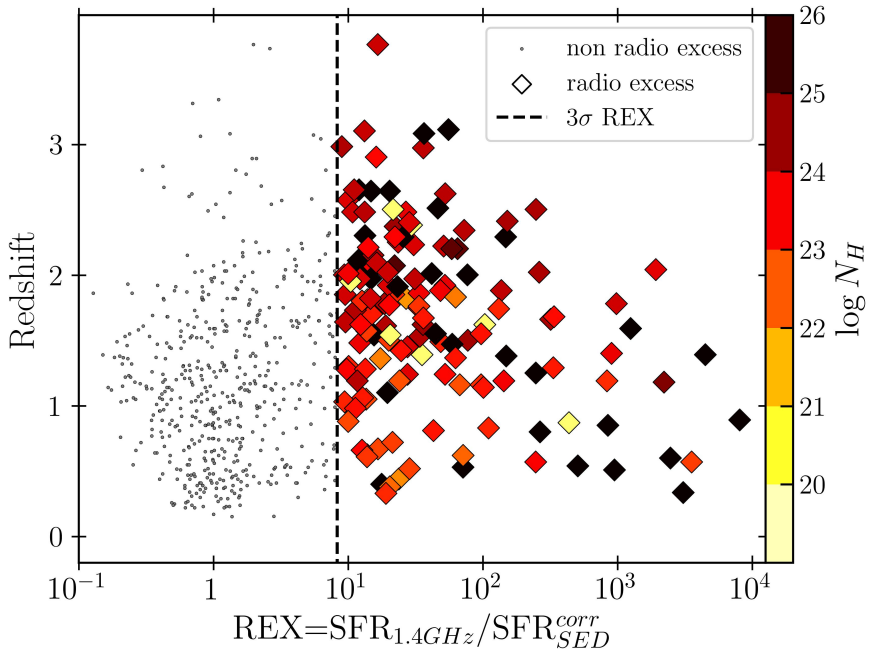

FIGURE 3.7: Distribution of the X-ray undetected radio sources of the J1030 field falling into the Chandra footprint in the z_{phot} vs REX-parameter space. The color code of the sources refers to the expected obscuration level as returned by the X-ray undetection.

to $z \sim 3$. Most of the sources from which the relation was derived are RQ AGN based on the ratio $L_{1.4GHz}/L_{2-10keV}$ (Terashima & Wilson, 2003). The assumption that also our sources are RQ is reasonable given that the RQ AGN population largely dominates over the RL one, the latter representing only $\sim$ 10% of the radio AGN population. However, since we are considering a sample of radio-excess sources, the fraction of RL AGN can be higher, as we will discuss later in this section. The intrinsic $L_{2-10keV}$ X-ray luminosities derived using the $L_{1.4GHz} - L_{2-10keV}$ relation were then converted into $L_{0.5-2keV}$ luminosities assuming an intrinsic X-ray spectrum with slope $\Gamma = 1.9$. To finally derive an estimate of the minimum column density obscuring the X-ray emission of the radio-excess AGN candidates, we used the mock catalogs presented in Marchesi et al. (2020). This catalog contains 5.4M sources simulated down to very faint X-ray fluxes, 10^{-20}erg cm^{-2} s^{-1}, over an area of 100deg^2 and with different obscuring hydrogen column densities, following the prescription of the Cosmic X-ray Background model (CXB) derived in Gilli et al. (2007). The catalog provides the SB flux and the corresponding intrinsic SB X-ray luminosity for each source at a given redshift and obscuration level. Therefore, interpolating the data of the mock at a given combination of z, $L_{0.5-2keV}$, $f_{0.5-2keV}$ (where the flux here corresponds to the flux limit at the position of an object) we were able to derive the correspondent lower limit to the value of $\log N_H$.
The histogram in Fig. 3.6 shows the distribution of the inferred values of $\log N_H$ for the 145 X-ray undetected radio-excess sources. The median value corresponds to $\log N_H \sim 23.8$. However, we recall that these values of $\log N_H$ have to be considered as lower limits to the actual obscuration. We note that there is a significant fraction of sources with $\log N_H \sim 25$. For these sources, the interpolation method allowed only solutions of the obscuration larger than the CTK one, and in a few cases this could be due to an overestimated intrinsic X-ray luminosity, as we will further discuss later.
Instead, in Fig. 3.7, we show the distribution of the same sources in the redshift versus REX plane and color-coded by their value of $\log N_H$. Assuming that all the radio-excess sources are AGN and that the procedure outlined above works for every source, there are in total 60 over 145 radio-excess X-ray undetected sources with an estimated $\log N_H > 24$, 44 over 89 at $z > 1.5$, and 12 over 16 sources at $z > 2.5$.
However, it is worth noting that not all the radio-excess sources are necessarily RQ AGN. If instead, a source is RL, for a fixed radio luminosity, the expected X-ray luminosity would be lower than the one estimated by the RQ AGN relation of D'Amato et al. (2022). This is because for typical RQ AGN SED $L_R/L_X \sim 10^{-4}$, while for RL AGN it can be $L_R/L_X \sim 10^{-1}$ (Fan & Bai, 2016; Shang et al., 2011). In turn a lower intrinsic $L_{0.5-2keV}$ would lead to a lower expected value of $\log N_H$.
Another class of sources that can overlap with RL AGN and show falsely larger obscuration values are low-excitation radio galaxies (LERGs). These are radio AGN (they might also be RL AGN) associated with passive galaxies dominated by an old stellar population and lacking large gas reservoirs that are needed to fuel the central SMBH. In these sources, the lack of X-ray emission is due to intrinsically faint X-ray luminosities connected to the highly inefficient accretion mechanism and not to the fact that they are obscured. This kind of sources are observed to constitute a not negligible fraction of the radio AGN population, in particular at low redshift ($z \lesssim 1$), while their fraction substantially decreases towards higher redshift Kondapally et al. (2022).

3.5 Discussion

3.5.1 Contamination of the radio-excess sample

In Sect. 3.4.1, we identified a threshold in REX above which the radio emission in excess of what is expected from SF can be reasonably explained assuming an AGN-related additional radio power. However, given the definition of the REX parameter in Eq. 3.6, there might be contaminants among the radio-excess sources that are not AGN.

The total SFR of a source is generally estimated taking into account both the direct emission from young stars as well as from stars obscured by interstellar dust. The total SFR can be computed in two ways. The first consists of adding to the unobscured SFR observed from the rest UV-optical emission (the SFR due to unobscured stars), the contribution absorbed by dust and reradiated in the IR, if photometry in the MIR and FIR is available. The second way uses the observed rest UV-optical emission, and applies the attenuation correction (always estimated from the rest-optical SED) to obtain the total SFR. This second possibility is generally applied when the MIR and FIR bands are not available.

The value of the SFR_{SED}^{corr} we derived from the SED fitting with CIGALE is computed in this second way, where the correction for the obscuration considers the value of A_V estimated from the attenuation of stellar emission in our photometric range. Indeed, our photometry does not cover the rest-frame MIR and FIR emission of the sources, where the dust reradiates the optical radiation absorbed from stars, which would allow the best estimate for the amount of obscured SF in a galaxy. Therefore, it is possible that for some sources the value of the obscuration A_V might be underestimated, consequently leading to an underestimation of $\mathrm{SFR}_{\mathrm{SED}}^{\mathrm{corr}}$ that moves these sources to larger values of REX. The average value of A_V derived from the SED-fitting for radio-excess sources is $A_V \sim 0.7$. Underestimating the obscuration A_V by $\delta A_V = 1$ mag would lead to an underestimation of the $\mathrm{SFR}_{\mathrm{SED}}^{\mathrm{corr}}$ by a factor of ~ 2.5, consequently overestimating the REX parameter by the same amount. This effect can be particularly important for the population of dust-obscured starburst galaxies, that are generally associated with a significant radio emission (related to the starburst) but whose rest UV and optical emission is almost completely absorbed by dust (Talia et al., 2021; Behiri et al., 2024; Gentile et al., 2024). However the number density of these galaxies is not high enough to explain the large majority of the radio-excess selected sources on the J1030 field (Sapori et al. in prep.).

We can further check the possible amount of dusty SFG contamination in our radio excess sample by taking advantage of the 1.1mm AzTEC observation of the J1030 field (Zeballos et al., 2018). This observation covers an area similar to the one covered by the Chandra image and has a median rms of 0.6 mJy. Assuming that all the radio excess sources in the Chandra footprint are dusty SFG, we can estimate their total infrared luminosities (L_{TIR}) as if all their radio emission comes from SF, and therefore using the radio luminosity of Eq. 3.3 and the q_{TIR} in Eq. 3.5. Taking a gray body model for the dust emission (with dust temperature $T = 40K$ and opacity coefficient $\beta = 1.6$ Venemans et al., 2018) normalized to the L_{TIR} values, we derived the expected 1.1mm flux densities of these sources. In total, we would expect $33 > 3\sigma$ detections in the AzTEC 1.1mm map of the field, while there is only one radio-excess source whose position and expected flux density are compatible with a 1.1mm-detected source (separation $d \sim 2$", and expected $f_{1.1mm} \sim 2$mJy compared to the measured

1.6±0.6mJy of the AzTEC source).
Another possible bias of our REX parameter can be the fact that the SFR traced by the radio luminosity ($SFR_{1.4GHz}$) and the one traced by the SED-fitting ($\mathrm{SFR}_{\mathrm{SED}}^{\mathrm{corr}}$) refer to different timescales of star formation. In particular Arango-Toro et al. (2023) showed that in some cases an excess in the radio SFR with respect to the one derived from the SED-fitting can be justified by the fact that the star formation time sensitivity of the radio frequency might be longer than 150 Myr, while the timescale of the SFR derived from the optical photometry can be assumed as instantaneous. Therefore sources with high values of REX can be sources that might have experienced a starburst in the last $\sim 200-300$Myr followed by a decreasing SFH. Even if this possibility does not exclude that the starburst or the following quenching might also be associated with AGN activity, we tested this eventuality by computing with CIGALE the average SFR over the last 100, 300 and 500 Myrs, and substituting these SFR to SFR_{SED}^{corr} in Eq. 3.6. The plots showing the new distribution of the sources in REX are analogous to the one in Fig. 3.4 and are shown in Sect. 3.7.2 of the supplementary material of this chapter. We did not find a significant shift in the position of the sources on the x-axis or in the REX threshold.
The SED-fitting with `CIGALE` was performed without including an AGN component. The AGN emission can potentially enhance the rest frame UV and optical emission, leading to an overestimation of the SFR_{SED}^{corr}, determining lower values of REX that can shift some true AGN out from the radio-excess selection. However this is not the case for the population of heavily obscured AGN we are interested in, whose emission in the rest frame UV and optical bands is almost completely negligible.

3.5.2 X-ray stacking analysis

To further check the goodness of our selection in terms of heavily obscured AGN, we performed an X-ray stacking analysis on the radio excess selected sources falling in the Chandra footprint and without an X-ray counterpart.
To perform the X-ray stack, we removed from the original X-ray image all the X-ray detected sources and also sub-threshold detected sources (for a total of 344 sources), and we filled the extraction circles (6" for all the sources) with uniform background noise. Then, we used the Chandra tool `CIAO`[1] to derive the X-ray stacking images corresponding to the sum, mean, and median stack. To derive the count rates and the signal to noise (S/N) of each stack, we used the following approach. For each of the sources involved in the stack, we define a circular source extraction region with the size determined by the 90% encircled counts fraction (ECF) radius (r_{90}) (fixing a minimum of 1"). We set the background region for each source to be an annulus with external radius 20 arcsec and internal radius $r_{90} + 1$". The background was computed from the cleaned X-ray image described above. Sources at a distance lower than $d < 6" + r_{90}$ from X-ray detected sources in the original image were excluded from the stack to avoid introducing artifacts in the stacking procedure.

We decided to perform the X-ray stacking analysis considering only sources for which $r_{90} < 6$", corresponding to a cut in off-axis angles at $\Theta \simeq 6.5$ arcmin. Given the central symmetry of the J1030 Chandra observations, this choice implies considering only sources in

[1] https://cxc.cfa.harvard.edu/ciao/

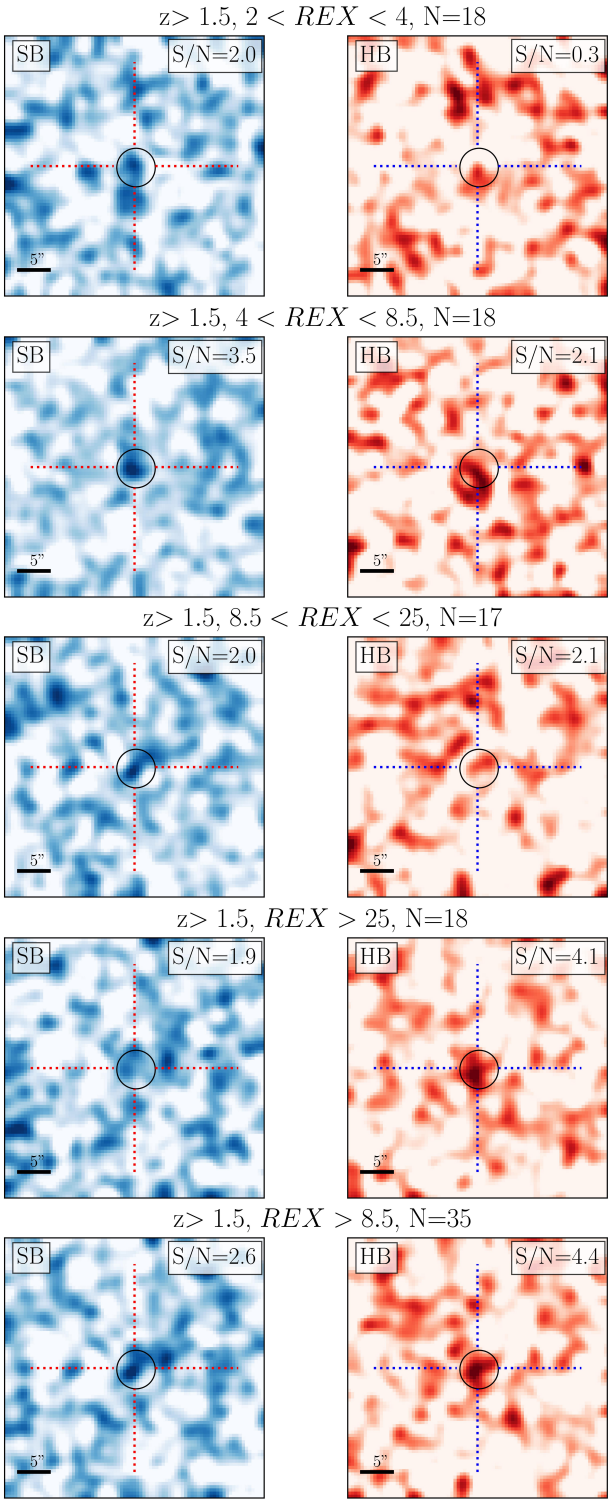

FIGURE 3.8: 40″×40″ X-ray stacking cutouts in the SB (left) and HB (right) of different samples of sources based on their REX, from the top to the bottom: $2 < REX < 4$, $4 < REX < 8.5$, $8.5 < REX < 25$, $REX > 25$, and $REX > 8.5$. The stack was performed considering only sources at $z > 1.5$, and $r_{90} < 6$”, this last corresponding to a cut in off-axis angle $\Theta \lesssim 6.5$ arcmin. The stack is presented using a Gaussian smoothing function. The black circle shows the 3″ radius extracting region for the computation of the S/N. In the upper right corner of each stack is shown the S/N of the detection.

the central region ($\sim$ 140arcmin^2) of the J1030 field, and this cut is motivated by two main reasons. First, for sources with larger r_{90}, any potential X-ray signal would be spread on such a large area that it would not significantly contribute to the stack but instead produce a noisier stacking image. Second, we are interested in the population of heavily obscured AGN, but, as reported in Fig. 3.6 there are $\sim$ 30 sources for which the level of obscuration estimated from the X-ray upper limit is low ($\log N_H < 23$). Almost all these sources are distributed along the edge of the X-ray image, where the X-ray flux limit is shallower, therefore excluding sources with $r_{90} > 6$" also removes the least obscured AGN candidates.
Furthermore, we considered only sources at $z > 1.5$ because at these redshifts the radio-excess selection is more reliable (largely avoiding the possible inclusion of LERGS, see Sect. 3.5.1, and Sect. 3.7.1), and also because we want to limit the analysis to the redshift range where the population of heavily obscured AGN is more unconstrained.

Following the approach described above we derived the count rates in the soft band (0.5-2keV, SB) and in the hard band (2-7keV, HB) in four equally populated REX bins: $2 < REX < 4$ (18 sources), $4 < REX < 8.5$ (18 sources), $8.5 < REX < 25$ (17 sources), $REX > 25$ (18 sources). From the net counts, we derived the S/N of the detection in each band considering the following equation:

$$S/N = \frac{S}{\sqrt{S + \tilde{B}}}, \tag{3.8}$$

where S are the net counts inside r_{90}, and $\tilde{B}$ are the background counts rescaled to the area of the source counts extraction.
In Fig. 3.8, we show the X-ray stacking cutouts in the SB and HB for the four different samples of sources. As it is possible to note from the first four rows of panels, there are only two samples showing a clear X-ray detection in at least one band. The stack of the sources with $4 < REX < 8$ shows a clear detection in the SB, while the sample of sources with $REX > 25$ is detected with enough significance only in the HB. The undetection in the stacks of sources with $2 < REX < 4$ can be easily explained, considering that these sources are mainly normal SFG (lying on the MS, see Fig. 3.5), and the limited statistic does not allow the detection of the X-emission coming from SF, that is probably low. On the contrary, sources with $4 < REX < 8$, that are clearly detected only in the SB, are probably galaxies that host dust-obscured SF, producing the soft X-ray emission, with a possible non-negligible AGN contamination. Indeed, we recall that our 3σ radio-excess selection has to be considered conservative in terms of AGN completeness. Instead, we would expect a detection in the X-ray stacking of the sources with $8 < REX < 25$, suggesting the presence of AGN. However, their X-ray emission is probably undetected because of the combined effect of obscuration and limited statistics, as we will see later. A clear HB-only detection is found in the sample of radio-excess sources with $REX > 25$, which strongly suggests that this sample is dominated by heavily obscured, possibly CTK, AGN.
To check that the radio-excess selection threshold $REX \geq 8.5$ does not fail in selecting heavily obscured AGN when X-ray undetected sources are considered, we performed a final X-ray stacking analysis merging the lst two samples, therefore considering all sources with $REX > 8.5$ (and also at $z > 1.5$ and with $r_{90} < 6$"). The X-ray stacks of these 35

sources are shown in the last two panels of Fig. 3.8 and confirm a clear detection in the HB, strongly supporting the presence of a heavily obscured AGN-dominated population among these sources, as we will further discuss in the next section.

3.5.3 X-ray obscuration

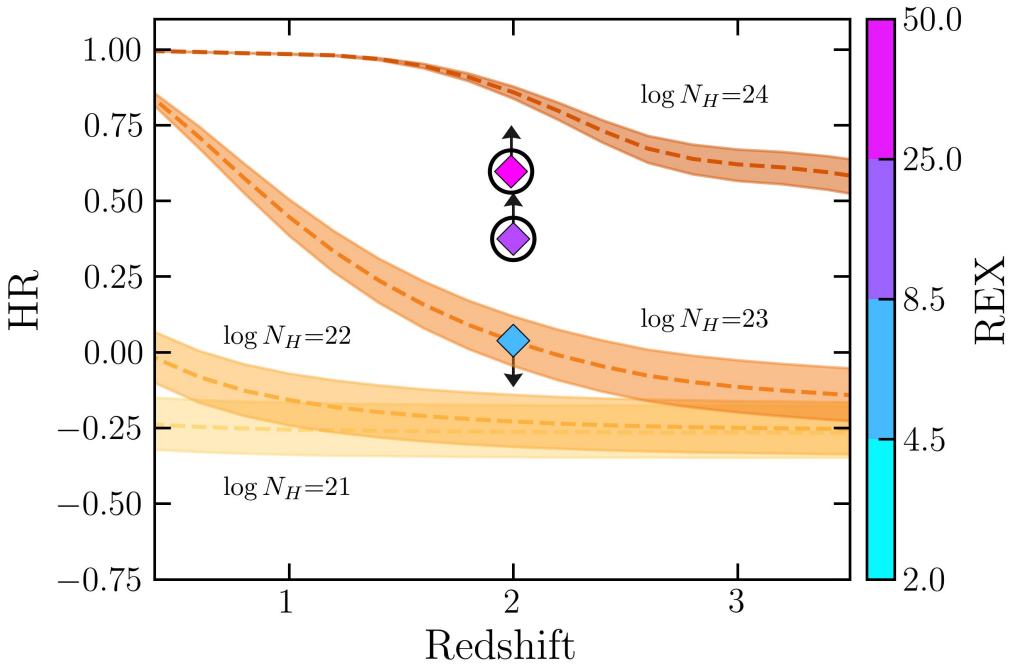

FIGURE 3.9: HR versus redshift distribution of the three samples $4 < REX < 8.5$, $REX > 8.5$ and $REX > 25$, color-coded based on their median REX value. We marked with a circle the two reference samples, $REX > 25$, $REX > 8.5$. The shaded areas indicate HR values derived with fixed $\Gamma = 1.7$ (top curves) and $\Gamma = 2.1$ (bottom curves), while the dotted lines represent models with $\Gamma = 1.9$, our baseline model. We used the response matrices at the aimpoint of the Chandra observations at Cycle 17 (Peca et al., 2021).

The X-ray detections shown in Sect. 3.5.2 suggest that the radio-excess selection is truly able to pick AGN from the general distribution of radio sources, supporting the reliability of our procedure and of the REX threshold. To further support the heavily obscured AGN nature of these radio-excess selected sources we computed the hardness ratio (HR) of the same samples described in Sect. 3.5.2 for which the stacking revealed a detection in at least one of the two bands. This parameter allows for an indication, given the median redshift of the sample, of its average obscuration level, and is defined as:

$$HR = \frac{H - S}{H + S}, \tag{3.9}$$

where S and H are the net counts (i.e. background subtracted) in the SB and HB, respectively.

Sample ($z > 1.5$& $r_{90} < 6$")	z_{med}	REX	$\log L_{1.4GHz}$	X-stack S/N	HR	$\log N_H$	$\log L_X$
$REX > 8.5$	2.0	25	40.1	2.6/4.4	>0.37	23.7	42.7
$REX > 25$	2.0	49	40.1	1.9/4.1	>0.59	23.9	42.9

TABLE 3.2: Summary of the median properties of the samples of sources at $z > 1.5$, having $r_{90} < 6$" and at $REX > 8.5$ and $REX > 25$. The two values in the X-ray S/N column refer to the SB and HB, respectively.

In Fig. 3.9 we show the distribution of the HR derived for three samples ($4 < REX < 8.5$, $REX > 8.5$ and $REX > 25$) presented in Sect. 3.5.2 as a function of the median redshift of each sample. The data points are also color-coded based on their median REX value. In Fig. 3.9 we also show HR values for typical AGN column densities ($21 < \log N_H < 24$) as a function of redshift and with a canonical photon index $\Gamma = 1.9 \pm 0.2$ (Peca et al., 2021). As it is possible to see, there is a clear increasing trend of the HR with increasing REX in our samples. In particular, the sample of sources with $REX > 25$ and with $REX > 8.5$ shows HRs that correspond to $\log N_H > 23.3$, with the first approaching $\log N_H \sim 24$ (considering the lower limit).

We further checked the average obscuration of the two samples with $REX > 8.5$ and $REX > 25$ using the Chandra tool PIMMS[2]. Considering the count rates in the SB and HB and the proper Cycle of the Chandra observations (Cycle 17 for the J1030 X-ray observations) we found average obscuration of $\log N_H \sim 23.7$ and 23.9 for sources with $REX > 8.5$ and $REX > 25$, respectively, in agreement with the results from the HR.

Then, using PIMMS with the average obscurations just computed, we converted the HB count rates of these two samples into median, unabsorbed X-ray fluxes in the 2-10 keV band, assuming an X-ray power spectrum with an intrinsic spectral index $\Gamma = 1.9$. We found 2-10keV intrinsic fluxes of 1.8×10^{-16}erg/s/cm^2 and 3^{-16}erg/s/cm^2 for the $REX > 8.5$ and $REX > 25$ sources, that correspond to X-ray luminosities of $L_{2-10keV} = 5 \times 10^{42}$erg/s and $L_{2-10keV} = 8 \times 10^{42}$erg/s, respectively. A summary of the physical properties of the $REX > 8.5$ and $REX > 25$ samples is reported in Table 3.2.

The derived X-ray luminosities are too high to be justified only by SF. The average SFR returned by `CIGALE` for sources at $REX > 8.5$ and $REX > 25$ are 11 $M_\odot yr^{-1}$ and 5 $M_\odot yr^{-1}$, respectively, while those returned by the radio luminosity are both $\sim 300 M_\odot yr^{-1}$. Considering the average stellar masses of the two samples, that is $\log M_* \sim 10.7$ $M_\odot$ for both, we can derive the correspondent X-ray luminosity using Eq.15 in Lehmer et al. (2016):

$$L_{2-10keV,gal} = a_0(1+z)^{\gamma_0} M_* + b_0(1+z)^{\delta_0} SFR \ erg/s, \tag{3.10}$$

where ($\log a_0$, $\log b_0$, γ_0, δ_0) = (29.30, 39.40, 2.19, 1.02). For both samples, the predicted X-ray luminosities considering SFR^{corr}_{SED} are $< 2 \times 10^{41}$erg/s, while taking $SFR_{1.4GHz}$ the expected X-ray luminosities are $< 2 \times 10^{42}$erg/s. In both cases, they are lower than the average intrinsic X-ray luminosities derived from the X-rays analysis.

[2] https://cxc.harvard.edu/toolkit/pimms.jsp

3.5.4 Comparison with radio predictions

In Chapter 2 (the work presented in Mazzolari et al., 2024c), we developed an analytical model to predict the number of radio-detectable AGN over a given radio image considering populations of AGN at different levels of obscuration. In particular, we extrapolated the predictions of the AGN population synthesis model of the CXB to the radio band by deriving the 1.4GHz luminosity functions of unobscured, obscured, and CTK AGN. Then, we used these functions to forecast the number of detectable AGN based on the area, flux limit, and completeness of a given radio survey and compare these results with the AGN number resulting from X-ray predictions. We found that, while X-ray selection is generally more effective in detecting unobscured AGN, the surface density of CTK AGN radio detected can be ~ 10 times larger than the X-ray one, and even greater at $z > 3$. Given that the intrinsic AGN population was assumed to be the same for the models in the two bands, the difference in detectability of heavily obscured AGN was a direct consequence of the fact that obscuration has no effect on radio emission (while it has on X-rays). In that work, we also considered the J1030 field, predicting over the whole radio image ($\sim 0.2\mathrm{deg}^2$) 533 radio detectable AGN, of which 222 CTK. Since in our analysis, we are only considering sources in the X-ray image footprint ($\sim 0.1\mathrm{deg}^2$), the expectations for the intrinsic radio AGN population of Mazzolari et al. (2024c) have to be halved. However, as mentioned above, the model in Mazzolari et al. (2024c) was based on the intrinsic population derived from the CXB model and, therefore, from X-ray observations, but it is possible that the intrinsic AGN population might be larger, in particular at $z > 3$, as suggested by recent JWST results (Yang et al., 2023a; Akins et al., 2024).
In the $0.09\mathrm{deg}^2$ of the J1030 field, our radio-excess selection returned 188 radio-excess sources, of which 145 are not X-ray detected, and 60 have an estimated $\log N_H > 24$. These numbers are lower (but not significantly) than the ~ 250 and ~ 100 radio AGN and CTK AGN predicted by the model in the same area. However, we have to recall that our selection is based only on radio-excess criteria, and as was shown in different works (Mazzolari et al., 2024c; Smolčić et al., 2017b; Bonzini et al., 2013), the radio-excess selection techniques are able to select only a fraction of the total radio AGN population, ranging between $30-60\%$ depending on the depth of the radio observations, on the availability of multiband information and on the purity of the selection. As we will further investigate in the next Section, our radio selection seems to indicate that the population of heavily obscured AGN is actually larger than what is predicted by X-ray observations and model, particularly at $z > 3$.

3.5.5 Number density

In this section, we want to compute the radio-excess selected CTK AGN number density and compare it with the predictions and the observational results derived from the X-rays. To compute the AGN number density we followed the same approach reported in Vito et al. (2014), passing through the binned luminosity function (Page & Carrera, 2000). The binned luminosity function for a given redshift and luminosity bins containing N sources is returned by:

$$\phi = \frac{N}{\int_{zmin}^{zmax} \int_{\log Lmin}^{\log Lmax} \Omega \, \frac{dV}{dz} \, dz \, dlogL}, \tag{3.11}$$

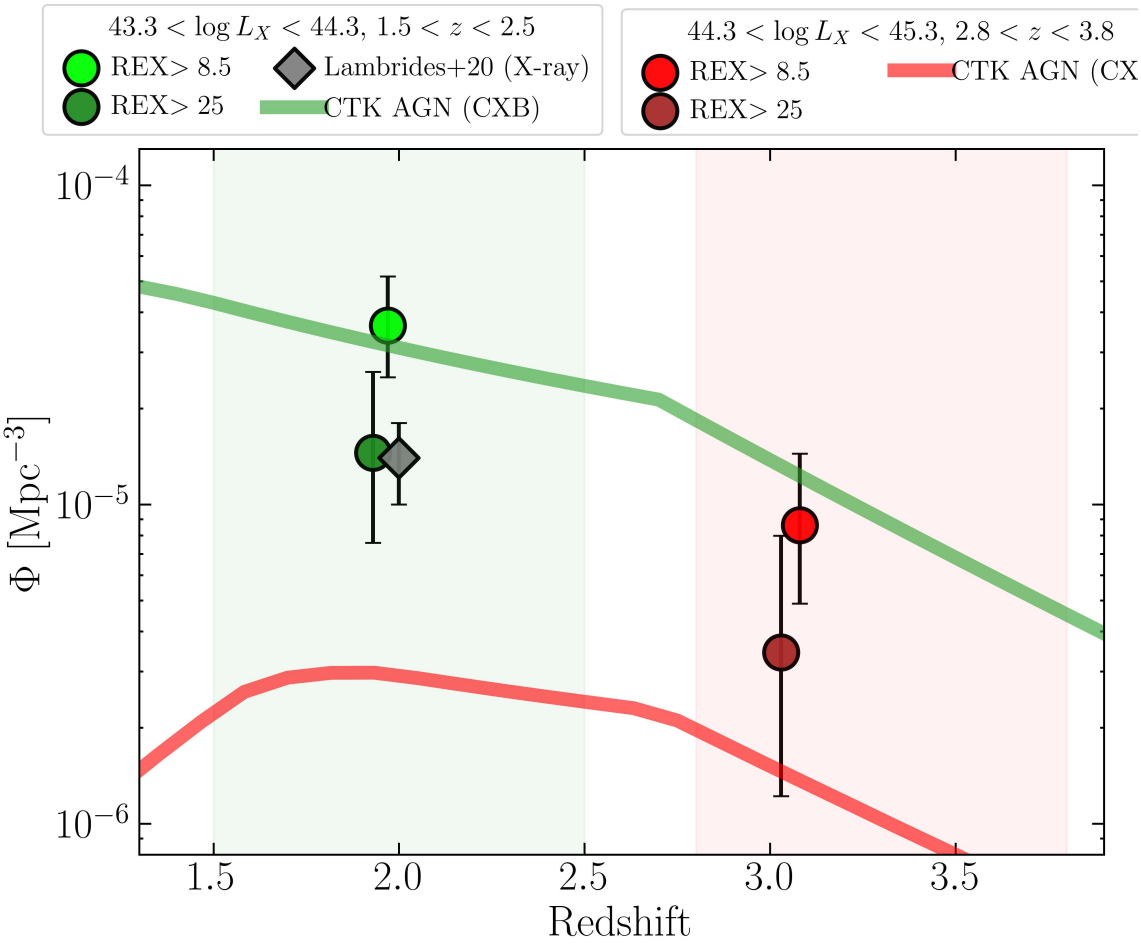

FIGURE 3.10: Comparison between the number densities of CTK AGN radio-selected derived from the two samples in Table 3.2, but selecting only sources with an estimated $\log N_H > 24$ (see Sect.3.4.3). Our results are compared with number densities of the same population derived from X-ray models and observations. Green colors refer to the $z \sim 2$ number densities, red colors to the $z \sim 3$ ones. $REX > 8.5$ and $REX > 25$ number densities are reported with lighter and darker colors, respectively, and they reasonably correspond to an upper and lower limit to the actual number density of the radio-excess selected CTK AGN. Instead, the green and red lines represent the trend of the number density derived by the CXB model of Gilli et al. (2007), while the gray data point corresponds to the observed X-ray CTK AGN number density computed in Lambrides et al. (2020). In the legend are shown the redshift and luminosity ranges used to derive the number densities. The green and red shaded areas mark the redshift range investigated by the two number densities. Data points are plotted in correspondence of the median value of the redshifts of the samples.

where $\Omega = \Omega(S) = \Omega(z, L)$ is the sky coverage, that is the fraction of the sky covered by a survey at a given observed flux S corresponding to a source luminosity $\log L$ at redshift z,

while dV/dz is the comoving volume element :

$$\frac{dV}{dz} = \frac{4\pi c}{H_0} \frac{D_{\rm M}^2}{\left(\Omega_{\rm m}(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda\right)^{1/2}}, \tag{3.12}$$

where $D_{\rm M}$ is the comoving distance. Then, the number density of sources is returned by multiplying Φ in Eq. 3.11 by $\Delta \log L$. In particular, we computed the heavily obscured AGN number densities considering two different redshift-luminosity bins: $1.5 < z < 2.5$ at $43.3 < \log L_{2-10keV} < 44.3$, and $2.8 < z < 3.8$ at $44.3 < \log L_{2-10keV} < 45.3$. The two redshift and luminosity intervals are almost contiguous, and the choice to use the X-ray luminosity as a reference allows us to compare our results with other measurements coming from X-ray observations. To compute the number densities of our sources in these X-ray luminosity bins, we translated these bins into 1.4GHz luminosity bins using the $L_{1.4GHz} - L_{2-10keV}$ relation of D'Amato et al. (2022). Since we are interested in the CTK AGN number density, we considered for both intervals only sources with an estimated $\log N_H > 24$ from those estimated in Sect.3.4.3. Given the results from the X-ray analysis reported in Sect. 3.5.2 and Sect. 3.5.3, we also conservatively considered two different radio-excess thresholds, sources at $REX > 8.5$ and $REX > 25$. In this way, we can provide a reliable upper and lower limit to the actual heavily obscured radio-selected AGN population. In the lower redshift-luminosity bin we counted 10 (4) sources considering $REX > 8.5$ ($REX > 25$), while in the other bin we have 5 (2) sources. We then derived the CTK AGN number densities using Eq. 3.11 and computing upper and lower errors at 68.3% confidence levels using the standard Gehrels approximation for low count statistics (Gehrels, 1986).

In Fig. 3.10 we show the number densities derived from our selection compared with other results computed in the literature considering X-ray observations or models in the same redshift and luminosity ranges. In particular, Lambrides et al. (2020), considered the same redshift and X-ray luminosity range as in our lower bin. They studied the multiwavelength properties of sources classified as low-luminosity AGN in the Chandra Deep Field South (CDFS, Luo et al., 2017), and found that many of these sources were heavily obscured AGN disguised as low-luminosity AGN, with column densities an order of magnitude higher than what previously derived. In particular, the X-ray number density reported in Fig. 3.10 was derived considering those sources with measured $\log N_H > 24$.
Then, we also report for both redshift-luminosity ranges the trend of the CTK AGN X-ray number density predicted by the population synthesis model of the CXB, selecting only sources with $24 < \log N_H < 26$.

In the lowest redshift bin ($z \sim 2$), the upper limit to the radio-excess CTK AGN number density (computed with the $REX > 8.5$ sample) is compatible with the intrinsic CTK AGN population returned by the CXB model, while the lower limit to the radio number density ($REX > 25$) is perfectly consistent with the X-ray number density derived in Lambrides et al. (2020). As expected for this redshift range, the CTK AGN population radio-selected is not in tension with the result from the X-rays. However, this also shows that radio observations are really able to unveil the population of heavily obscured AGN, probably even more efficiently than X-ray observations, being the datapoint from Lambrides et al. (2020) compatible with our lower limit.

At $z \sim 3$, the radio CTK AGN number density derived in this work is instead a factor of $3-5$ higher than what is predicted by the CXB model for the same redshift and luminosity range. Given the tension between X-ray observations and models in reproducing the correct AGN census at $z > 3$, our result suggests that the most obscured AGN population at high-z can actually be larger than what is expected from X-ray observations, as also supported by some recent works with JWST data in the redshift range $3 < z < 5$ (Yang et al., 2023a; Lyu et al., 2024).
This result further sustains the advantage of using radio emission to investigate the heavily obscured AGN population, which almost prevents any bias versus obscuration. We also stress that our radio-excess selected CTK AGN number densities have to be considered conservative, given that the selection was made to ensure the purity of the sample.

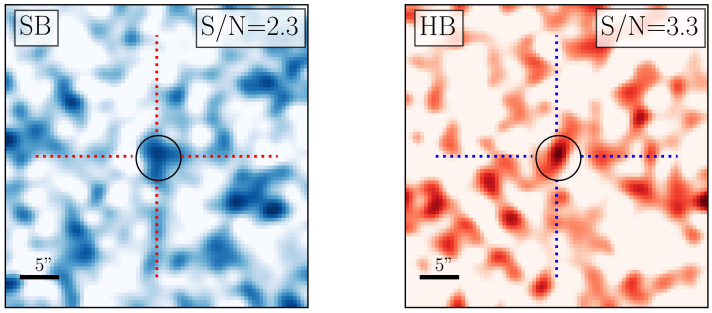

FIGURE 3.11: X-ray stacking in the SB and HB of the 40 radio detected sources with $r_{90} < 6$" and undetected in the J1030 multiband catalog.

3.5.6 Radio sources undetected in the multiband catalog

As reported in Sect. 3.2.3, there are 99 radio sources that are not detected in the K-band selected J1030 multiwavelength catalog, 22 of which are without a counterpart in any band, while most of the others are undetected (or extremely faint) in the K band but are visible at longer wavelengths. In any case, we could not attribute a redshift to any of these sources, but we still have the possibility to understand if this population can host a significant fraction of AGN using X-ray stacking. Following the same procedure described in Sect. 3.5.2, we considered only those in the Chandra footprint (76 sources) and we stacked only sources with $r_{90} < 6$" (40 sources), and the results are shown in Fig. 3.11. The stacking revealed a faint detection in the HB and not in the SB. While in the SB, the X-ray emission at the center of the stack is quite diffuse, the HB X-ray emission is highly focused. Using Eq. 3.9, we derived an HR$\sim 0.1 \pm 0.1$ that is indicative of a consistent obscuration, even if for a more robust estimate, redshifts are needed. We also performed an X-ray stacking analysis considering only the radio sources without a counterpart in any band (originally 22 sources 18 of which in the X-ray footprint). However, the X-ray stacking did not reveal any detection for these sources, probably due to the limited statistics or to the very heavy obscuration.
In any case, at least a fraction of these sources can contribute to the most obscured AGN population missed by the X-ray observation but revealed by deep continuum radio imaging, further contributing to the AGN number densities shown in Fig. 3.10.

3.6 Conclusions

In this work, we search for heavily obscured AGN among the population of radio sources of the J1030 field looking for heavily obscured AGN. We started computing the photometric redshift of the 1003 radio sources having a counterpart in the K-band selected multiwavelength catalog of the J1030 field, and we performed a radio-excess selection to identify AGN candidates. Then we focused on those that are undetected in the deep Chandra image to select only extremely obscured sources.
Part of the work is still ongoing, but the preliminary results can be summarized as follows:

- We defined a radio excess parameter REX that corresponds to the ratio between the $SFR_{1.4GHz}$ computed directly from the radio luminosity of the sources and the SFR^{corr}_{SED} derived from the SED fitting to the optical and NIR photometry (and corrected for dust attenuation). Large values of REX indicate that there is radio emission in excess with respect to what is expected from a normal SFG. Fitting the SFG population across the REX distribution, we found the (3σ) threshold to identify radio-excess sources in correspondence of $REX \sim 8.5$. Among the 1003 radio sources with a multiband counterpart, we identified 240 radio-excess AGN candidates.

- We then focus on the radio-excess sources falling into the X-ray image footprint but without an X-ray detection (145 sources). By performing a detailed X-ray stacking analysis, we found solid indications that this population is dominated by heavily obscured AGN. In particular, the stacks of samples with $REX > 8.5$ and $REX > 25$ show clear detections in the HB (S/N$>$ 4), and the obscurations, inferred both by modeling the X-ray undetection and using HRs, are compatible with CTK values. Furthermore, the median X-ray luminosities of the sources in the two samples are too high to be justified only by SF.

- We finally estimated the radio-excess selected CTK AGN number densities in two different redshift-luminosity bins one at $z \sim 2$ and the other at $z \sim 3$. By selecting only sources with estimated column densities larger than $\log N_H > 24$, we derived reliable lower and upper limits to the actual radio-selected CTK AGN number density considering sources at $REX > 25$ and at $REX > 8.5$. While at $z \sim 2$ the results are compatible with X-ray models and observations, at $z > 3$ our number densities are a factor of 3-5 larger than those predicted by the CXB model derived by Gilli et al. (2007) and more in line with recent JWST results.

There are also some aspects of the analysis on which I am still working on. In particular, I am performing a more detailed analysis of the contamination from LERGs and starburst galaxies among the radio-excess sources. I am also estimating the completeness corrections to apply to the number densities presented in Sect. 3.5.5 and related to the fact that the radio excess selection is incomplete, as discussed in Chapter 2 and Sect. 3.5.4.
In any case, this work shows the importance of finding new and complementary ways to select the population of the most obscured AGN, in particular at high redshift. The confirmation of the heavily obscured AGN nature of the selected sources will be further proved by an extensive spectroscopic follow-up that we obtained with LBT/MODS and ESO/VLT in the last cycles (PI: G. Mazzolari). The two programs, which consist of 25 hours of observation

each, will obtain rest-frame UV and optical spectra of more than 200 sources on the J1030 field, the large majority of which will be radio-excess selected galaxies. In this way, we will have the possibility not only to confirm their nature but also eventually trace any peculiar spectroscopic feature characterizing these radio-selected AGN.

Starting from the next chapter, we will explore the selection of obscured AGN using JWST spectroscopic data, moving towards the very high-z Universe.

3.7 Supplementary material

3.7.1 Comparison with other radio-excess parameters

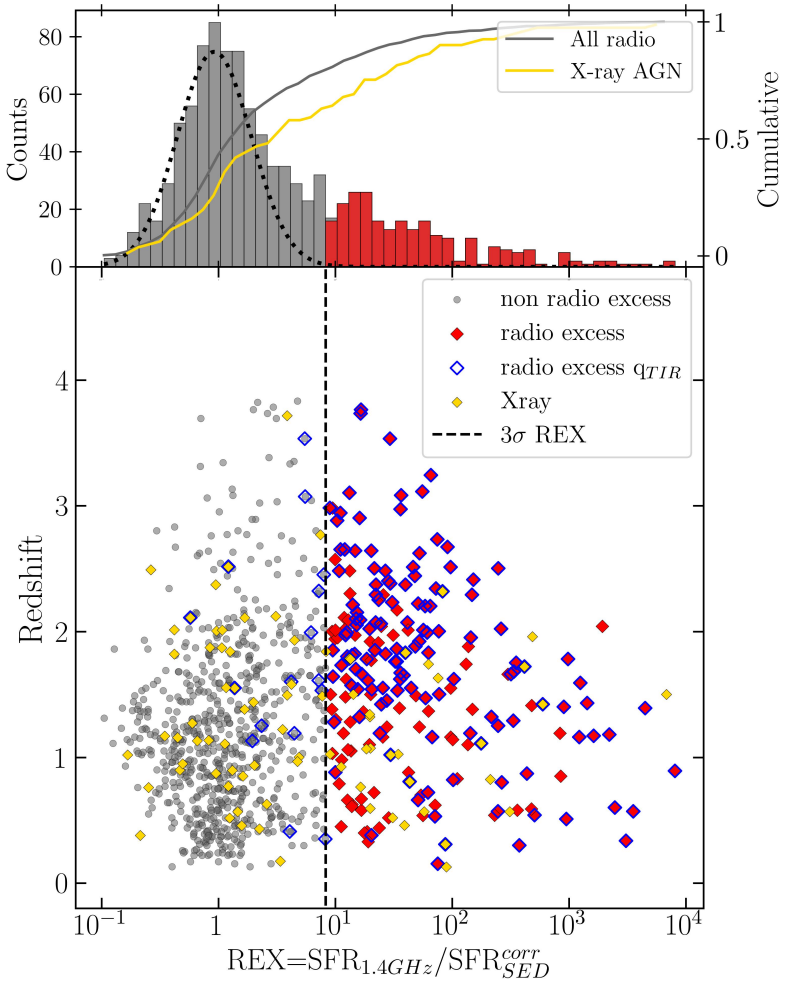

FIGURE 3.12: Same as in Fig. 3.4 but indicating with blue contours sources that are selected as radio-excess by the q_{TIR} parameter, as described in Sect. 3.7.1.

To test the validity of our radio-excess parameter and selection, we also considered a second well-known radio-excess indicator, the ratio between the total IR and the 1.4GHz

luminosity, $q_{\rm TIR} = \log(L_{TIR}) - \log(L_{1.4GHz})$. This radio-excess indicator is based on the well-established FIRC of SFG, which arises because the same population of massive stars that heats up dust, causing it to reradiate its energy in the far-infrared (FIR), produces supernovae that generate relativistic particles emitting synchrotron radiation at radio frequencies (Delvecchio et al., 2017, 2021; Delhaize et al., 2017).
From the available photometry, we do not have photometric constraints on the FIR part of the SED of these sources; however, since CIGALE assumes an energy balance between what is absorbed in the rest-frame UV-optical and what is re-radiated in the FIR, it also returns the best model SED extending up to the FIR wavelengths. Therefore we can compute the total infrared luminosity L_{TIR} integrating the SED between 8 and 1000μm and consequently compute also $q_{\rm TIR}$. Comparing our radio excess selection with the one returned by $q_{\rm TIR}$ allows us to prove the effectiveness of the REX parameter defined in Eq. 3.6 because the values of L_{TIR}, even if retrieved from the SED fitting, are not strictly related to the SFR. To identify those sources that host an AGN based on $q_{\rm TIR}$ we followed a procedure similar to the one applied for the REX parameter. Given the binned distribution of all the sources according to this parameter, we mirrored the bins on the right of the peak (that dominated by SFG) and then fitted it with a Gaussian, finally selecting only those sources with a $q_{\rm TIR} < q_{\rm TIR}^{peak} - 3\sigma$. This procedure was already adopted in Delvecchio et al. (2021). The sources selected with this radio-excess parameter are highlighted in blue in Fig. 3.12. As it is possible to note a large fraction of the radio-excess sources selected by our REX parameter are also selected by the $q_{\rm TIR}$ parameter ($> 75\%$), and only very few sources that are not selected as radio-excess in our selection result as radio-excess sources based on $q_{\rm TIR}$. It is worth noting that the fraction of sources that are selected as radio-excess by both the parameters rises to $> 85\%$ when only sources at $z > 1.5$ are taken into account.

3.7.2 REX with time-averaged SFR_{SED}^{corr}

Here we present the distribution of the radio-detected sources considering SFR_{SED}^{corr} derived from the SED-fitting with `CIGALE` and averaged over longer timescales: 100 Myrs, 300Myr, and 500 Myrs. Given that the distribution of the sources according to the different REX in Fig. 3.13 do not significantly change, we can rule out a dependence of the radio-excess selection on the different timescales characterizing the radio and the rest-optical emission, used to compute the two SFRs. In particular, the peak of the distribution over REX does not significantly vary in any of the three plots, as well as the radio-excess threshold is consistent for all the timescales with the one derived in Eq. 3.6. The number of radio-excess sources is 187, 193, 209 for REX_{100}, REX_{300}, and REX_{500}, respectively, where the differences have to be ascribed to the small shifts of the radio-excess threshold: REX_{100}=9.71, REX_{300}=8.44, and REX_{500}=7.03.

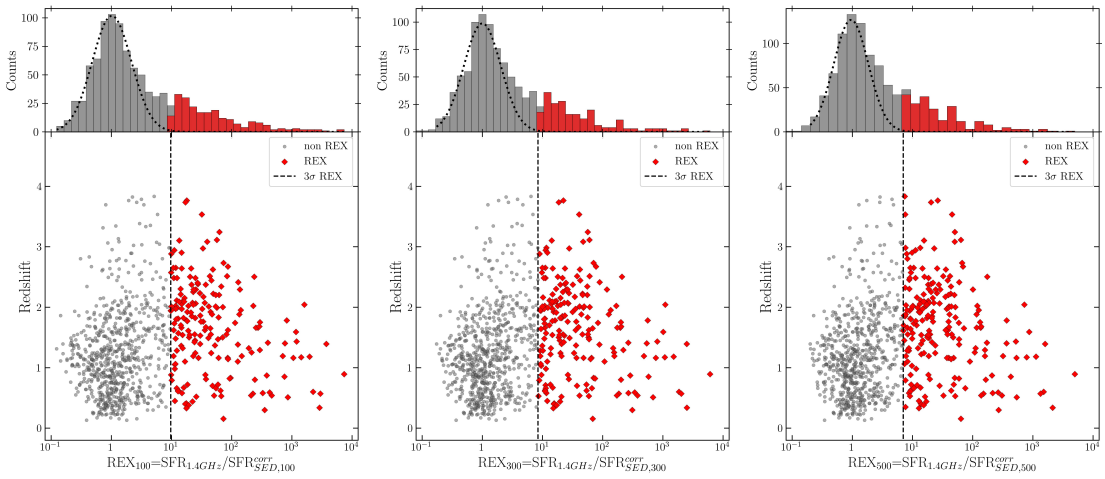

FIGURE 3.13: Same as in Fig. 3.4 but considering SFR^{corr}_{SED} averaged over different timescales, as described in Sect. 3.7.2.

Chapter 4

New diagnostic diagrams for high-redshift narrow line AGN

4.1 Introduction

As described in Sec. 1.4, thanks to the successful launch of the *James Webb Space Telescope* (JWST), we have the possibility to investigate with high resolution and unprecedented sensitivity both the photometric and spectroscopic properties of galaxies up to $z \sim 14$ (Carniani et al., 2024; Curtis-Lake et al., 2022; Robertson et al., 2022). Within this context, recent studies exploiting different kinds of JWST data have revealed a large population of AGN at high redshift. As presented in Sect. 1.4.1, studies on high-z AGN revealed unexpected properties such as a larger AGN number density than expected, ratios between the SMBH mass and the host galaxy stellar mass larger than what is found in the local Universe, and also an observed X-ray weakness.
However, most of these studies focused only on BLAGN, which probably constitutes only the tip of the iceberg of the AGN population, particularly at high-z (see Sect. 1.2.1). NLAGN, and therefore obscured AGN, are generally spectroscopically selected using BPTs diagrams (see Fig. 1.9). As we described in Sect. 1.4.2, different observational works have already demonstrated that these diagnostic diagrams are no longer effective in the high-z Universe, where environments are systematically more metal-poor and the stellar populations younger, ultimately determining the populations of AGN and SFG to overlap on these diagnostics.

To have a clear picture of the AGN properties and demographics at high-z, it is, therefore, crucial to identify new AGN selection techniques, also from the spectroscopic point of view, and this is exactly the context of this chapter.

Thanks to the wide wavelength coverage ($\sim 1 - 5\mu$m) and sensitivity of JWST, it is now possible to have access to a plethora of rest-UV and optical emission lines in high-z galaxies that were not accessible before. Among these lines, multiple JWST spectra have revealed the detection of the [O III] λ4363 line. In the metal-poor and highly ionized environments of many high-z galaxies, [O III] λ4363 is the strongest auroral line, a collisionally excited emission line generated from higher energy levels compared to the typical [O III] $\lambda\lambda$4959,5007 nebular lines. Due to its faintness and proximity to the (generally) stronger Hγ line, before

the launch of JWST the [O III] λ4363 had been detected and studied almost only in galaxies at $z < 2$. Given its sensitivity to the gas temperature (see Sect. 1.1.5), this line has been extensively used in the local Universe to determine the metallicity of the interstellar medium (ISM) via the direct temperature method, which exploits the possibility of computing the metallicity by fixing the gas electron temperature derived from the ratio between the [O III] λ4363 and the [O III] nebular lines (see e.g. the review by Maiolino & Mannucci, 2019). An enhancement of the [O III] λ4363 emission in the presence of an AGN is qualitatively expected. AGN ionizing photons are more energetic than ionizing photons produced in star-forming regions on average, and hence they can heat the gas more effectively by depositing a larger amount of energy per unit photon. In this work, we explore this scenario more quantitatively, with the ultimate goal of exploring the possibility of selecting AGN using the [O III] λ4363 line through both an observational-empirical approach and via photoionization models. As a result, we propose three new AGN diagnostic diagrams that can also be effective in selecting AGN at high-z.

In Sect. 4.2 we present the local and high-z observational samples and the set of photoionization models considered to test the validity of the new diagnostic diagrams. In Sect. 4.3 we describe the new diagnostic diagrams based on the line ratio [O III] λ4363 /Hγ and the demarcation lines that can be used to separate the AGN population from SFGs. Lastly, in Sect. 4.4 we provide interpretations of why these diagnostics work, and we show indications of the AGN nature of some newly identified AGN using these diagnostics. We also discuss the implications of this work on the metallicity estimates based on the [O III] λ4363 line and on the strong-line metallicity diagnostics. Finally, in Sect. 4.5, we draw the conclusions and outline a summary of the work presented in this chapter. In Sect. 4.6 we attach the supplementary material of this chapter.

The work presented in this chapter has been published in Mazzolari et al. (2024a).

4.2 Methods

We selected several AGN and SFG samples and two different sets of photoionization models to populate the diagnostic diagrams that we use to select AGN. All the observational samples described in the following (Sect. 4.2.1 and Sect. 4.2.2) are listed in Table 4.1.

4.2.1 Low-redshift samples

The sample of low-z SFGs is composed of both normal SFGs and also of the local analogues of high-z galaxies (i.e. local galaxies that show intrinsic properties similar to those observed in high-z galaxies), in terms of metallicities and emission lines ratios (e.g. Izotov et al., 2018, 2019, 2021). The Sloan Digital Sky Survey (SDSS) DR7 (Abazajian et al., 2009) gives the most numerous sample, mostly made of $z < 0.7$ galaxies. Starting from the whole catalogue provided by MPA/JHU, we considered only sources with a S/N > 5 in all the lines involved in the diagnostic diagrams we present: [O III] λ4363 , [O III] , [Ne III], [O II] , and Hγ . Then we distinguished between AGN and SFGs using the BPT diagram (Baldwin et al., 1981), taking the AGN demarcation line provided by Kewley et al. (2001) and after considering a

further cut in S/N> 5 also in the Hα , Hβ , and [N II] lines. Furthermore, we excluded from the selection AGN classified in the SDSS catalogue as BLAGN since the catalogue does not provide separately the narrow-line and broad-line components of the Balmer lines, and our diagnostics rely on the narrow-line emission only. We also corrected all the lines for dust attenuation using the Calzetti et al. (2000) attenuation law. The final SDSS SFG (AGN) sample considered in this work contains $\sim$ 2300 ($\sim$ 800) sources.

We identified several samples of local analogues of high-z galaxies with reported [O III] λ4363 line fluxes. From Yang et al. (2017b) we took the sample of 'blueberries', which are 40 dwarf starburst galaxies with small sizes ($<$ 1kpc), very high ionization values ([O III] /[O II] $\sim 10-60$), and low metallicities ($7.1 < 12+\log(\mathrm{O/H}) < 7.8$). They were selected from the SDSS DR12 at $z < 0.05$ and followed up spectroscopically using the *MMT* telescope. We also considered 43 'green pea' galaxies from Yang et al. (2017a) (i.e. nearby SFG with strong [O III] emission line), two-thirds of which also show a strong Lyα emission.

We included in the local analogue sample a compilation of $\sim$ 490 local SFGs with low metallicities from different works: Izotov et al. (2006), Berg et al. (2012), Izotov et al. (2018), Izotov et al. (2019), Pustilnik et al. (2020), Pustilnik et al. (2021), Nakajima et al. (2022). All these samples were selected considering local sources followed up with different telescopes to be sensitive to the faint [O III] λ4363 line, which allowed them to measure the galaxies' metallicity via the direct method. The metallicities of these sources span the range $12 + \log(O/H) \sim 6.9 - 8.9$.

We also included the sample of 165 extreme emission line galaxies (EELG) reported in Amorín et al. (2015), selected from the zCOSMOS survey (Lilly et al., 2007) in the redshift range $0.11 < z < 0.93$. These are very compact ($r_{50} \sim 1.3$ kpc), low-mass ($\mathrm{M_*} \sim 10^7 - 10^{10}\mathrm{M_\odot}$) galaxies characterized by specific star formation rates (sSFRs) above the main sequence for star-forming galaxies of the same stellar mass and redshift.

The local AGN sample comprises Type 2 AGN showing different features and selected for different aims. The most numerous AGN sample is represented by the SDSS sample mentioned above. We included the Dors et al. (2020) Type 2 AGN sample, selected from SDSS DR7 using multiple AGN diagnostic diagrams to guarantee a high purity and showing a range of metallicities between $8.0 < 12 + \log(\mathrm{O/H}) < 9.2$.

The compilation of Type 2 AGN includes three sources presented in Seyfert (1943): NGC1068, NGC1275, and the core of NGC4151. These spectra were obtained at the Mount Wilson Observatory and are among the first galaxies hosting active SMBH ever observed. These three sources clearly show the [O III] λ4363 emission line in their spectra.

We considered the sources described in Perna et al. (2017), selected from SDSS $z < 0.8$ spectra associated with X-ray emission and showing outflows, which can be spectrally decomposed from the narrow-line and broad-line region (BLR) emission. We also added Type 1 and Type 2 AGN from the S7 survey (Dopita et al., 2015; Thomas et al., 2017), where 131 nearby AGN were primarily selected because of their radio-detection, and then followed up using the WiFeS instrument at high resolution ($R = 7000$), which allows the faint [O III] λ4363 line to be resolved. In Thomas et al. (2017) they provide the narrow components of the emission lines, already subtracting the broad component, if present.

The 35 AGN selected in Armah et al. (2021) are Type 2 AGN in the local Universe ($z < 0.06$) characterized by quite high metallicities (average metallicity $12 + \log(\mathrm{O/H}) = 8.55$).

These sources were selected by the authors to study the neon-to-oxygen abundance and its evolution with metallicity.

The last sample of local AGN is from the MOSFIRE Deep Evolution Field (MOSDEF) survey (Kriek et al., 2015), which targeted the rest-frame optical spectra of ~ 1500 H-band-selected galaxies at $1.37 < z < 3.8$. We took the full MOSDEF catalogue and selected only those sources with a [O III] λ4363 line with a $S/N > 3$; then we distinguished between AGN and SFG using the BPT diagram. With these S/N cuts, we could identify only three sources, all classified as AGN. This indicates the difficulty in detecting such a faint line further than in the local Universe in the pre-JWST era.
For all the samples reported above we took dust-corrected fluxes from the works reported in Table 4.1.

4.2.2 High-redshift JWST samples

The high-z samples of AGN and SFGs only comprise sources observed spectroscopically using JWST, the only instrument with enough resolution, sensitivity, and IR-coverage to detect and disentangle the [O III] λ4363 line at $z > 3$. We considered the full publicly released sample of sources with a medium-resolution (MR, $R = 1000$) spectrum in the CEERS (PID: 1345; Finkelstein et al., 2022) and JADES (PID: 1210; Bunker et al., 2023a) surveys. For JADES, we considered the emission line fluxes and the AGN selection described in Scholtz et al. (2023b). Starting from a sample of 110 sources with MR spectroscopy, a reliable redshift, and sufficient wavelength coverage, the authors performed a NLAGN selection using multiple AGN rest-frame UV and optical diagnostic diagrams, finally selecting 28 reliable AGN candidates in the redshift range $1.8 < z < 9.4$. The diagnostic diagrams involved in the selection include the BPTs (Baldwin et al., 1981; Veilleux & Osterbrock, 1987), the CIV$\lambda\lambda$1549, 51/CIII]$\lambda\lambda$1907, 1909 versus CIII]$\lambda\lambda$1907, 1909 / HeIIλ1640, and the detection of high-ionization emission lines such as [NeV]λ3420 [NeIV]$\lambda\lambda$2422, 24, NV$\lambda\lambda$1239, 42. To select NLAGN the authors considered in these diagnostics both demarcation lines already defined in the literature, while in other cases (such as the BPTs) they defined new demarcation lines derived considering the distribution of photoionization models (Gutkin et al., 2016; Feltre et al., 2016; Nakajima & Maiolino, 2022). Of all the initial JADES sources, only 19 show a [O III] λ4363 detection, 7 of which are AGN. We followed the same approach for the CEERS program: starting from 313 MR spectra published by the CEERS collaboration, we selected 217 sources with a reliable spectroscopic redshift, and we selected 45 NLAGN candidates using the same diagnostics as in Scholtz et al. (2023b). The details on the selection and on the physical characterization of the NLAGN selected among the CEERS spectra will be presented in Chapter 5 (where we will select 52 NLAGN in total, also using the new diagnostics presented in this chapter). For this work, we considered 32 sources with the [O III] λ4363 line detected among the CEERS sample, 4 of which already classified as AGN. From the JADES and CEERS programmes, we also collected the BLAGN identified by Maiolino et al. (2023b) and Harikane et al. (2023) at $z > 4.5$, carefully checking for, and when necessary subtracting, the broad component of the Hγ line.

We further considered the three $z \sim 8$ galaxies in the galaxy cluster SMACS J0723.3-7327, observed in the Early Release Observations (ERO) JWST programme (Pontoppidan et al., 2022) and analysed in Curti et al. (2023b). Exploiting the [O III] λ4363 line detections

in the NIRSpec spectra, the authors measured metallicities ranging from extremely metal-poor ($12 + \log(\mathrm{O/H}) \sim 7$) to about one-third solar. Two of them were later identified as AGN in Brinchmann (2023), due to the presence of the high-ionization [Ne IV] emission line in one and based on the high ionization parameter ($\log \mathrm{U} \sim -1$) of the other.

The sample of high-z sources presented in Nakajima et al. (2023) includes galaxies from three different early JWST programmes (CEERS, ERO, GLASS Treu et al., 2022) at $6 < z < 9$. We selected only the ten sources with the [O III] λ4363 line detected, and we marked as BLAGN those that were selected based on their broad Hα line in Harikane et al. (2023). For the latter, we subtracted the broad Hγ component from the Hγ flux, if present.

We included in the high-z AGN sample the BLAGN reported in Übler et al. (2023) at $z \sim 5.55$, whose JWST/NIRSpec Integral Field Spectrograph (IFS) observation shows very high ionization lines and low metallicity (one-fourth solar). We also included both the components of the dual AGN candidate at $z \sim 7.15$ reported in Übler et al. (2024). For this source, thanks to IFS spectroscopy, the authors were able to study the displacement between the position of the Hβ broad-line region (BLR) and the strong [O III] emission line centroid, which were interpreted as the emissions coming from two distinct sources 620pc apart from each other. Interestingly, the authors also found an almost perfect alignment between the [O III] λ4363 emission line peak and the peak of the Hβ BLR, suggesting a correlation between the intensity of the [O III] λ4363 line and the presence of an AGN.

We considered the BLAGN reported in Kokorev et al. (2023) at $z = 8.50$ from the JWST UNCOVER Treasury survey (Labbe et al., 2021), showing a robust Hβ broad component and an unprecedented black hole to host a galaxy mass of at least $\sim 30\%$. An even larger black hole to host galaxy mass ratio was more recently found in the $z = 6.86$ BLAGN analysed in Juodžbalis et al. (2024a), representing an extreme example of dormant SMBH at high-z. We also considered the $z = 7.04$ BLAGN, triply imaged and lensed by the galaxy cluster Abell2744-QSO1, reported by Furtak et al. (2023). This source is probably undergoing a phase of rapid SMBH growth and is also heavily obscured since the authors measured an $A_V \sim 3$ from the Balmer decrement. It should be noted that in all these Type 1 AGN cases (most of which actually type 1.5–1.9) the broad component of the Balmer lines is not detected in Hγ, and hence the narrow component of Hγ can be easily measured.

We added to our high-z sample the $z = 6.1$ extreme SFG reported in Topping et al. (2024). Using distinct UV transitions, the authors found an electron density $\sim 10^4 - 10^5 \mathrm{cm}^{-3}$, a metal-poor ionized gas (i.e. $12 + \log(\mathrm{O/H}) = 7.43$), and a $\log \mathrm{U} \sim -1$ based on the ratio [O III] /[O II] . The authors ultimately classified this source as a SFG, based on the ratios of the rest-UV emission lines, but they also found some indications of possible AGN activity, for example the broad Hα component accounting for $\sim 20\%$ of the line flux.

We finally included in the sample GN-z11, an exceptionally luminous galaxy at z = 10.6. Its JWST spectrum reveals the presence of an AGN through the detection of semi-forbidden lines tracing very high densities (inconsistent with the ISM, but typical of the BLR), other transitions typical of AGN, fast outflows, ionization cones, and a larger Lyα halo, consistent with those seen in lower redshift quasars (Maiolino et al., 2024b, 2023a; Scholtz et al., 2023a).

Sample/References	Type	N included	Redshift
SDSS DR7 Abazajian et al. (2009)	SFG & AGN	3188 (2377/811)	$z_{med} \sim 0.1$
Yang et al. (2017b)	Blueberries Galaxies	40	$z < 0.05$
Yang et al. (2017a)	Green Pea Galaxies	43	$z < 0.05$
Low-OH compilation	Low-metallicities SFG	489	$z < 0.1$
Amorín et al. (2015)	Extreme Emission Line Galaxies	165	$z < 1$
Seyfert (1943)	AGN	3	$z \sim 0$
Dors et al. (2020)	AGN	26	$z < 0.05$
Perna et al. (2017)	AGN	44	$z < 0.8$
S7 Survey Dopita et al. (2015)	AGN	313	$z < 0.1$
Armah et al. (2021)	AGN	35	$z < 0.06$
Kriek et al. (2015)	AGN	3	$z \sim 2.5$
CEERS (R1000) Finkelstein et al. (2022)	SFG & AGN	32 (28/4)	$2.3 < z < 8.8$
JADES (R1000) Bunker et al. (2023a)	SFG & AGN	19 (12/7)	$0.66 < z < 9.43$
Curti et al. (2023a)	SFG & AGN	3 (1/2)	$z \sim 8$
Nakajima et al. (2023)	SFG & AGN	10 (6/4)	$6 < z < 9$
Topping et al. (2024)	SFG	1	6.1
Übler et al. (2023)	AGN	1	5.55
Übler et al. (2024)	AGN	2	7.15
Juodžbalis et al. (2024a)	AGN	1	6.86
Kokorev et al. (2023)	AGN	1	8.50
Furtak et al. (2023)	AGN	1	7.04
Maiolino et al. (2024b)	AGN	1	10.6

TABLE 4.1: Literature data of the low-z and high-z observational samples presented in Sects. 4.2.1 and 4.2.2, and used to test the effectiveness of the new diagnostic diagrams.

4.2.3 Photoionization models

The emission line fluxes coming from observational data were compared with the results coming from two different sets of photoionization models. The first set of models was initially described in Gutkin et al. (2016) (for star formation) and in Feltre et al. (2016) (for narrow-line AGN emission), and updated with more recent stellar spectra and with a better description of AGN cloud microturbulence (Mignoli et al., 2019; Hirschmann et al., 2019; Vidal-García et al., 2024). The photoionization models were built using the `Cloudy` code (version c13.03) (Ferland et al., 2013) for star formation and AGN narrow-line regions, assuming a wide range of parameters. In particular, they consider gas metallicities (Z, where $[O/H] = \log(Z/Z_{\odot})$) in the range $10^{-4} < Z < 0.07$, ionization parameters of the ionizing source $-4 < \log U < -1$, dust-to-metal ratios ($0.1 < \xi < 0.5$), hydrogen gas densities ($10^2 < (n_H/cm^{-3}) < 10^4$), different carbon-to-oxygen abundances (C/O), and two different initial mass functions (IMFs) for star formation models. For a complete list of the values of the different parameters, we refer to Table 1 in Feltre et al. (2016). These models include predictions for both the AGN and SFG populations; the main feature differentiating between the AGN and SFG models is the ionizing spectrum, which is the spectral energy distribution (SED) of the incident ionizing radiation, the first showing a harder radiation field.

As reported above, the photoionization models of Gutkin et al. (2016) and Feltre et al. (2016) cover a wide range of parameters and might actually include physical conditions rarely found in the general populations of galaxies and in the local analogues. For example, previous theoretical and observational studies have suggested a correlation between metallicity and the ionization parameter in SF regions (e.g. Dopita et al., 2006; Mingozzi et al., 2020; Ji & Yan, 2022), which implies certain combinations of metallicities and ionization parameters should rarely, if ever, be observed in SF regions. Since the demarcation lines in the new diagrams we propose rely on both the observational data and photoionization models, we performed the following check to make sure that the part of the models with highly unphysical combinations of parameters does not impact our results.

We performed a likelihood analysis comparing the observed line ratios in our samples of SFG and those predicted by the SFG of Gutkin et al. (2016) in order to limit the inclusion of highly improbable or unphysical combinations of the parameters. This method has been adopted in making model-based inferences for local galaxies (e.g. Blanc et al., 2015; Mingozzi et al., 2020; Ji & Yan, 2022).

We considered a total of three sets of line ratios, including [O III] λ4363 /Hγ , [O III] /Hβ , and [OII]$\lambda\lambda$3726, 3728/Hγ . The inclusion of collisionally excited lines from different ionization states of oxygen as well as recombination lines of hydrogen helps break the degeneracy between the metallicity and the ionization parameter in the models. Following the formalism of Blanc et al. (2015), we calculated the likelihood of each model given each data point, and combined the likelihood with a flat prior in the log space spanned by the metallicity and the ionization parameter to obtain the posterior distribution of these two parameters. From the posteriors, we obtained the weighted average metallicities and ionization parameters for all SFGs. We then selected a region in the metallicity-ionization parameter space not populated by the posteriors of our sample of SFGs. While the models of Gutkin et al. (2016) and Feltre et al. (2016) have a large set of parameters, we note that the primary drivers of the variations in the predicted line ratios are the metallicity and the ionization parameter. To

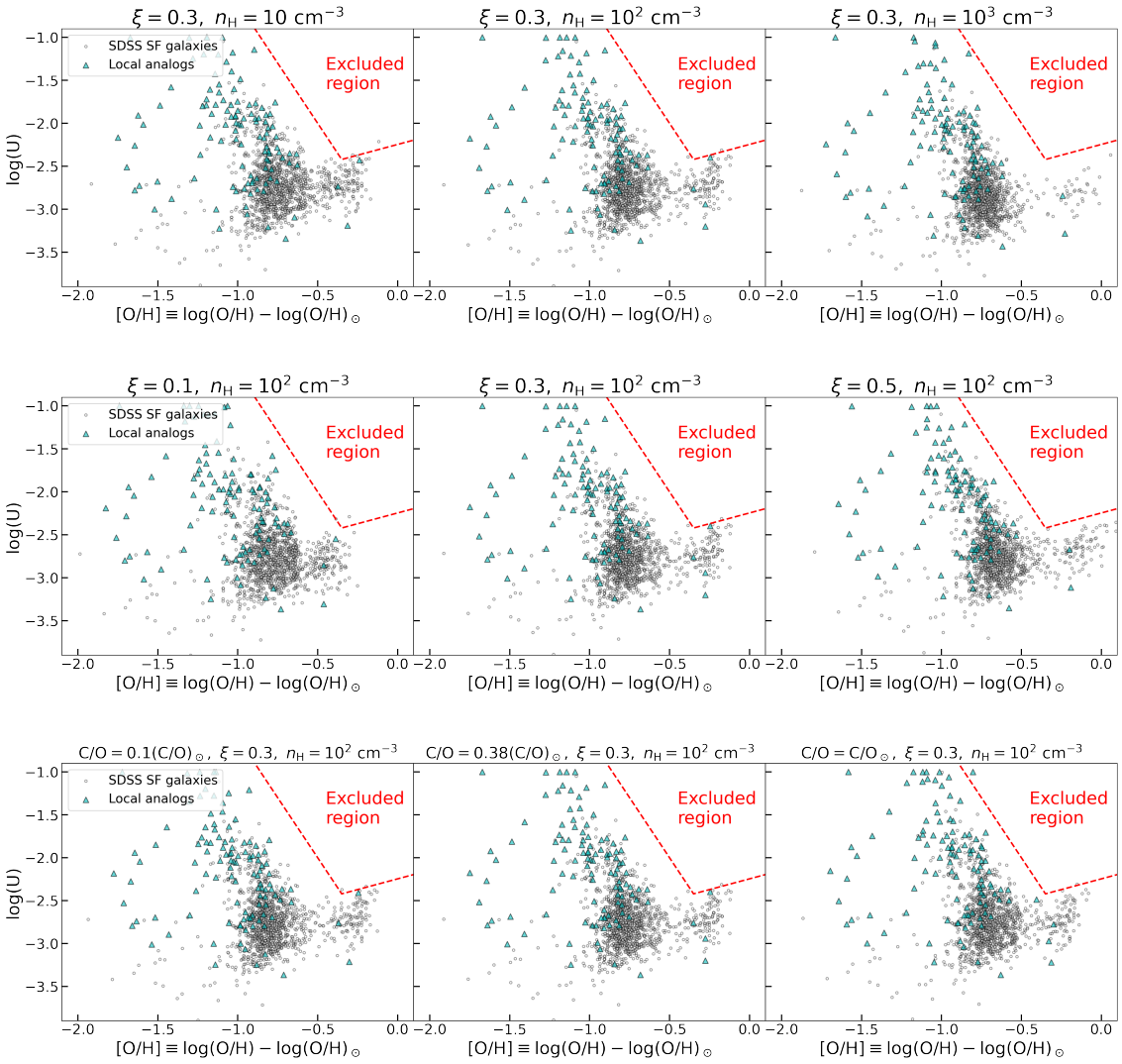

FIGURE 4.1: Distribution of the SDSS SFG sample (circles) and local analogues (triangles) in the log U vs log OH plane, according to the likelihood procedure with the Feltre et al. (2016) models described in Sect. 4.2.3. In each panel, we also show (red dashed line), the region of the parameter space that we decided to exclude, a posteriori, from the same photoionization models. The upper panels show the variation in the distribution of the sources by considering models with different electron densities n_H. The middle panels show different dust-to-metal ratios ξ, and the bottom panels show different carbon over oxygen (C/O) abundances. The excluded region is never significantly populated by any source in all the panels.

verify this point, we repeated the above calculation for models with different IMFs, densities, dust-to-gas ratios, and carbon-to-oxygen abundance ratios, which generally resulted in a similar distribution in the metallicity–ionization parameter space. With these results, we adopted a conservative cut to select realistic SFG models based on the 16th percentile of the inferred distribution in the metallicity–ionization parameter space that produces the highest boundary.[1]

Our final cut removes the region of the SFG parameter space spanned by $\log(Z/Z_\odot)$ and $\log U$ according to the following relations:

$$\log U > -2.8\ \log(Z/Z_\odot) - 3.4, \tag{4.1}$$

and

$$\log U > 0.5\ \log(Z/Z_\odot) - 2.245. \tag{4.2}$$

In Fig. 4.1 we plot our fiducial cut together with the inferred distribution of the metallicities and ionization parameters for our selected SFGs based on different sets of models. The cut basically removes regions having very high ionization parameters and high metallicities at the same time. From a physical point of view, due to the dependence of the ionizing spectra of young stellar populations and the mechanical feedback from young stellar populations on the metallicity, it is difficult to maintain a very high ionization parameter at a high enough metallicity (Dopita et al., 2006; Carton et al., 2017). While these physically motivated cuts do not impact the location of the AGN photoionization models in the diagnostic diagrams, the parameter space occupied by the SFG models is partially reduced. We note that 99.9% of the considered SFGs lie outside of the excluded parameter space, even assuming different values of ξ, n_H, and C/O.

We also compared the emission line ratios of the observational samples with the models described in Nakajima & Maiolino (2022). Using `Cloudy`, the authors investigated the emission line ratios of SFG, AGN, Population III stars, and direct collapse black holes (DCBHs) using the BPASS stellar population models (Eldridge et al., 2017) and a range of other parameters (similar to those considered in Feltre et al. 2016) listed in Table 1 of Nakajima & Maiolino (2022). For this work we only considered the models for SFG and AGN as the very low metallicity values of the other two models ($Z < 10^{-4}$) represent extreme scenarios that are not representative of the sources we want to identify with this analysis. In particular, for these models, the authors considered values of $10^{-5} < Z < 10^{-3}$ and values of $\log U$ from -0.5 to -3.5. Since the SFG models of Nakajima & Maiolino (2022) always cover regions of the diagnostic diagrams already covered by the physically motivated limited Gutkin et al. (2016) and Feltre et al. (2016) SFG models, thus confirming the goodness of our parameter cuts, we did not apply any cut in this case.

4.3 Results

In this section, we present the three new diagnostic diagrams proposed for selecting AGN using the [O III] λ4363 line. In these diagrams, we combine the auroral line, which is sensitive

[1]The boundary is determined by the common outer envelopes of SDSS SFGs and local analogues as these two samples differ significantly in number and occupy different regions in the parameter space.

to temperature, with various other lines that are sensitive to other additional properties of the ISM, such as the ionization parameter and the shape of the ionizing SED, and hence can further help to disentangle the source of excitation. Additionally, these lines are accessible with JWST at high redshift, and most of them are also close in wavelength, so little affected by dust reddening.

4.3.1 [O III] λ4363 /Hγ versus [O III] /[O II]

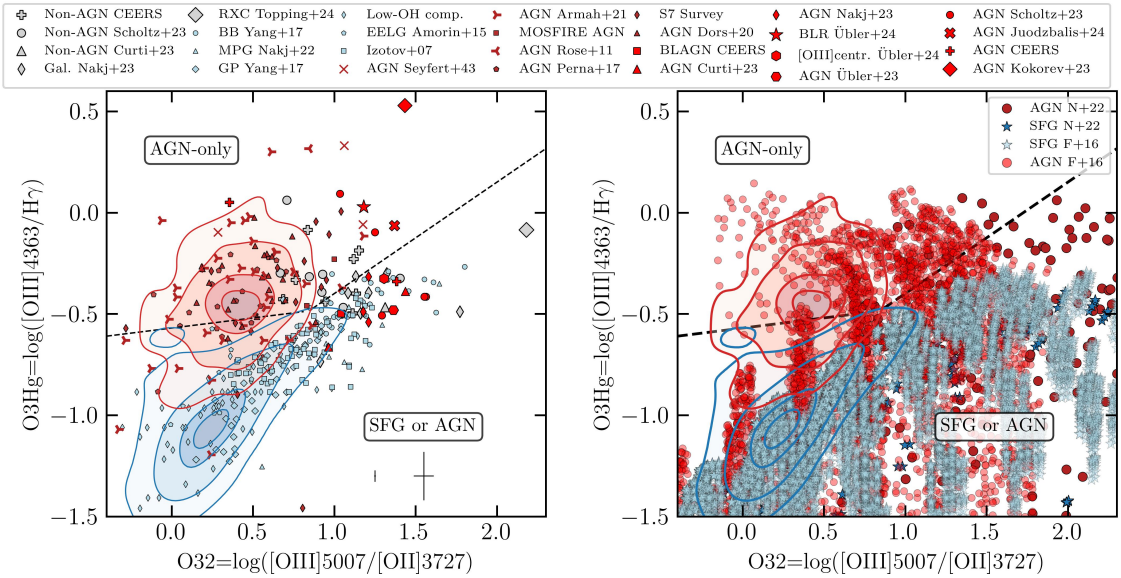

FIGURE 4.2: Diagnostic diagram showing the [O III] λ4363 /Hγ vs [O III] /[O II] line ratios. [O II] is the sum of the doublet [OII]$\lambda\lambda$3726, 3728. Left panel: Plot of all the observational samples described in Sects. 4.2.1 and 4.2.2. Red is for AGN, blue for low-z SFGs, and grey for high-z sources not classified as AGN (see legend for shapes). The red and blue contours show the distribution at SDSS AGN and SFGs, respectively, including the 90%, 70%, 30%, and 10% of the populations. Right panel: Same plot, but showing the AGN and SFG models computed by Gutkin et al. (2016), Feltre et al. (2016), and Nakajima & Maiolino (2022), as described in Sect. 4.2.3. The tracks of the AGN and SFG models according to $\log U$ and Z are shown in Fig. 4.5. The black dashed line indicates the demarcation defined in Sect. 4.3.4. The distribution of models and of observational samples both suggest that the dominant ionizing source in galaxies above the demarcation line is AGN. The error bars reported in the lower right corner represent the median errors of the low-redshift (left) and high-redshift (right) samples. The effect of dust attenuation on this diagnostic moves sources towards the right, without contaminating the AGN-only region with SFGs.

The diagnostic presented in Fig. 4.2 is based on the [O III] λ4363 /Hγ ratio compared to the [O III] /[O II] ratio. Hereafter [O II] refers to the sum of [OII]$\lambda\lambda$3726, 3728. In the left panel, where we plot the observational samples described in Sect. 4.2, the distribution of AGN and SFGs from SDSS are well separated, with AGN occupying a region with higher

[O III] λ4363 /Hγ ratios compared to normal SFGs. The distribution of the local analogues in this diagram seems to extend the distribution of the local SDSS galaxies towards higher [O III] λ4363 /Hγ and higher [O III] /[O II] , along the diagonal of the plot. On the contrary, the local AGN samples cover a wider area, mostly above the distribution of local analogues and SFGs.

High-z sources are generally distributed in the upper part of the diagnostic with respect to the local samples. We note that a non-negligible fraction of high-z sources not classified as AGN lie in the region of the diagnostic populated by local AGN, while there are also different high-z AGN falling in the same region covered by local analogues.

In the right panel, we show the distribution of the photoionization models presented in Sect. 4.2.3 according to the same line ratios. We see here that the SFG models are predicted to occupy a limited region of this diagram, whose upper boundary almost perfectly corresponds with the distribution of SDSS SFG and local analogues, while AGN are expected to cover both the region occupied by SFGs and also the part of the diagnostic covered by the local AGN samples. Therefore, SFG models are not able to reach the part of the diagram occupied by local AGN, as expected also from the distribution of the observational samples. The opposite is not true, since AGN can occupy the region populated by SFGs, as observed for some local and high-z AGN. These considerations remain true even considering the full grid of SFG models computed in Feltre et al. (2016), as shown in Sect. 4.6.1 of the supplementary material of this chapter.

As we discuss below, to a first approximation, the ratio [O III] /[O II] traces the ionization parameter. At a given [O III] /[O II] , AGN photoionization can reach much higher values of [O III] λ4363 /Hγ than photoionization from hot stars, which is likely a consequence, on average, of the much higher energy of ionizing photons produced by AGN; therefore, at a given ionization parameter, AGN can heat the gas much more effectively than hot stars.

It is worth noting the position of some high-z sources in this diagnostic. The AGN reported in Kokorev et al. (2023) shows an extremely strong [O III] λ4363 line and it is placed in a region that is not covered by any model, which probably suggests a very high electron temperature and log U, as also pointed out in Kokorev et al. (2023). Another significant high-z AGN falling in the AGN-dominated region of this diagnostic is the source corresponding to the BLR centroid in the dual source described in Übler et al. (2024), where, in the IFS map, the peak of the broad-line emission spatially corresponds to the peak of the [O III] λ4363 line emission. The other component, the [O III] centroid, showing a fainter [O III] λ4363 line emission, is close to the border between the local analogues and the local AGN distribution, together with most of the high-z BLAGN. The AGN region is populated by different AGN selected among the CEERS and JADES MR NIRSpec spectra, but we also have different sources not selected as AGN, but falling in the AGN-dominated region, as pointed out above. We demonstrate in Sect. 4.4 that, from the stack of their spectra, we have an indication that some of these sources may host a BLAGN as well.

An interesting source in this diagnostic is the SFG reported in Topping et al. (2024). The source shows both strong [O III] λ4363 /Hγ and [O III] /[O II] emission line ratios, following the extrapolation of the distribution of the local analogues, but at very high values of the ionization parameters and at low metallicities (see Fig. 4.5), as is indeed found by the analysis performed in Topping et al. (2024). Looking at the photoionization models, we see that the region is not covered by the models of Gutkin et al. (2016) and Feltre et al. (2016) (not even

taking the whole parameter grid), but is instead covered by some AGN models according to Nakajima & Maiolino (2022), suggesting that the main source of ionizing radiation in this extreme galaxy is ambiguous.

It is worth noting that the effect of dust reddening on the line ratio [O III] /[O II] , whose components can undergo a different amount of attenuation due to their distance in wavelengths, would move data points toward the right. This can potentially shift AGN into the SFG+AGN region under significant reddening, but not the opposite (i.e. galaxies dominated by star formation would not contaminate the AGN-only region).

4.3.2 [O III] λ4363 /Hγ versus [Ne III]/[O II]

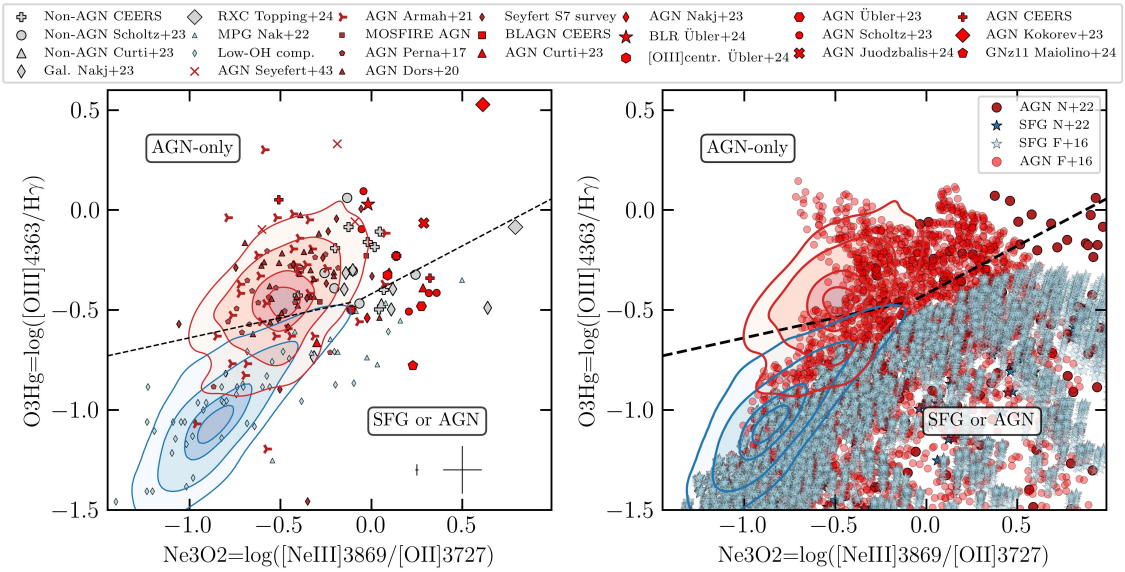

FIGURE 4.3: Same as Fig. 4.2, but for the line ratios [O III] λ4363 /Hγ vs [Ne III]/[O II] . Based on the distribution of observational samples and models, this diagnostic diagram also identifies a region that can be populated only by AGN.

The diagnostic presented in Fig. 4.3 is based on the ratio between the [O III] λ4363 /Hγ compared to [Ne III]/[O II] . Not surprisingly, the distribution of the local and high-z sources in this diagnostic is very similar to that described in the previous diagnostic. The [Ne III]/[O II] ratio correlates closely with [O III] /[O II] in AGN and in SFGs, as shown by Witstok et al. (2021), and as expected given that neon and oxygen are both α-elements and given that Ne^+ and O^+ have similar ionization potentials. The disadvantage of [Ne III]relative to [O III] is that it is typically much fainter. The advantage of [Ne III]is that it is at a shorter wavelength, and hence observable at higher redshift (out to z$\sim$12.3 with NIRSpec), and is closer in wavelength to [O II] , and hence the [Ne III]/[O II] ratio is much less affected by dust reddening. As in the previous diagram, normal SFGs and local analogues are distributed in the lower right of the diagram, in the same region covered by SFG models, while AGN can be distributed over a wider area, including the upper left part, which cannot be reached by any SFG models or

samples. As for the diagnostic described in Sect. 4.3.1, the separation between the two populations holds, even considering the whole grid of the parameter space of the Gutkin et al. (2016) SFG models, as presented in Sect. 4.6.1 of the supplementary material.

Even if the local analogues are less numerous than in Fig. 4.2 (simply because there are fewer reported [Ne III]fluxes in the literature), their distributions again seem to be an extrapolation of the distribution of local SFGs towards higher log U and lower metallicities (see Fig. 4.5). The sources that were clearly in the AGN region in the previous diagnostic are still above the SFG distribution in this diagram, and the sources not identified as AGN but lying in the AGN region of the diagnostic in Fig. 4.2 are in the same region in Fig. 4.3. In addition to this latter sample, in Fig. 4.3 there are a few other sources from the CEERS and JADES samples that were not in the diagnostic of Fig. 4.2 because the [O III] line was not available, due to the presence of a detector gap. The AGN region is also populated by three additional sources reported in Nakajima et al. (2023) and not classified as AGN in the literature. Their location, however, is still close to the border of the SFG distribution.

The diagnostic diagram presented in this section has some similarities with the OHNO diagram that involves the [O III] /Hβ versus [Ne III]/[O II] line ratios (Backhaus et al., 2022, 2023; Zeimann et al., 2015; Cleri et al., 2023; Trump et al., 2023; Feuillet et al., 2024; Killi et al., 2023). The latter is mainly an ionization-sensitive diagram (ionization parameter, shape of the ionizing continuum), but it is also sensitive to metallicity and at high redshift can be affected by low-metallicity galaxy contamination (Scholtz et al., 2023b). Instead, the O3Hg versus Ne3O2 diagnostic is more stable, also at high redshift, as the separation between the AGN and SFG populations is principally based on a different gas temperature, as we discuss further in Sect. 4.4.

4.3.3 [O III] λ4363 /Hγ versus [O III] λ5007/[O III] λ4363

The third diagnostic is presented in Fig. 4.4 and is based on the ratio of [O III] λ4363 /Hγ to [O III] λ5007/[O III] λ4363 . This diagnostic was also reported in Übler et al. (2024). Looking at the distribution of SDSS SFGs and AGN in the left panel, we note that there is more overlap between the two populations with respect to the previous diagnostics, but there is still a region, characterized by high values of [O III] λ4363 /Hγ and high [O III] λ5007/[O III] λ4363 , that is populated only by local AGN. The distribution of local analogues seems again to be an extrapolation of the trend of the SDSS SFG distribution. In this diagram there are very few high-z galaxies that are in the AGN-only region of the diagnostic, while most of them are shifted towards lower values of [O III] λ5007/[O III] λ4363 , even in a region that is not covered by the current photoionization models, as is evident in the right panel. This trend is emblematically represented by the two extreme AGN reported in Kokorev et al. (2023) and in Furtak et al. (2023). Given the strength of their [O III] λ4363 line emission, they are placed in the upper left part of the plot in a region that cannot be reproduced by any model. The same trend is shown by most of the galaxies that were not selected as AGN, but that lie in the AGN region of the two diagnostics presented above. The distribution of the SFG models matches the distribution of the local galaxies, in particular in its boundary with the AGN-only region, while AGN models cover both the region occupied by SFGs and the region populated only by AGN samples.

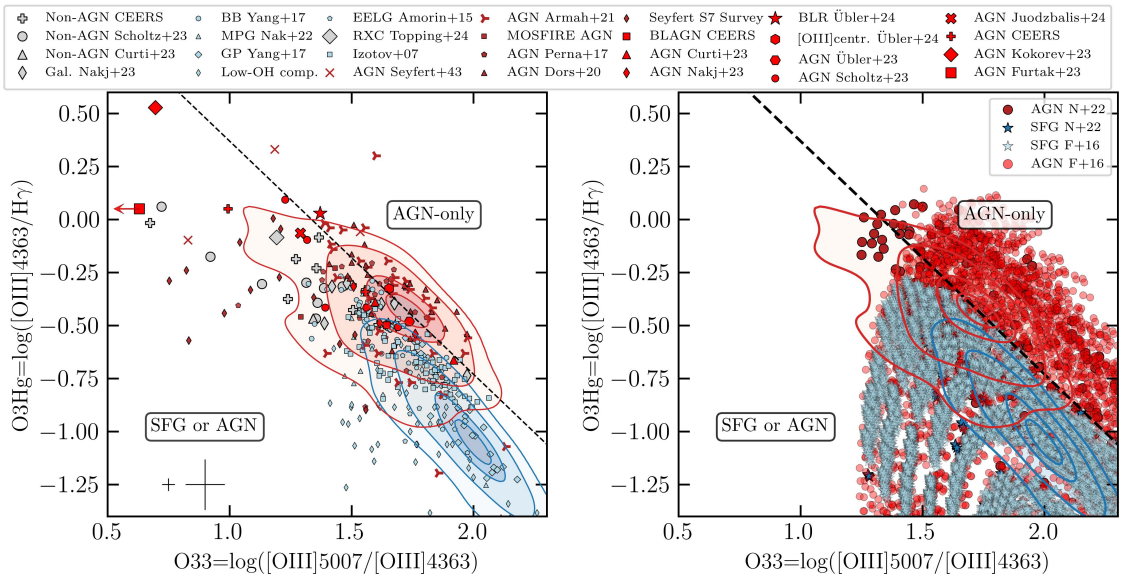

FIGURE 4.4: Same as Fig. 4.2, but for the line ratios [O III] λ4363 / Hγ vs [O III] λ5007/[O III] λ4363 . In the left panel the arrow on the AGN reported in Furtak et al. (2023) is for visualization purposes only since it would be located at O33$\sim$ 0. The cut in SFG models described in Sect. 4.2.3 allowed us to identify an AGN-only region of the diagnostic to the right of the black dashed line.

In this diagram, assuming the whole grid of parameters of the Feltre et al. (2016) models, we would have a complete overlap between the SFG and AGN models (see Sect. 4.6.1 of the supplementary material). However, the upper left part of the diagnostic would be covered by SFG models characterized by high values of log U and high metallicities, a condition not observed in any of our galaxy samples (see Fig. 4.1) nor in general in the literature (Kaasinen et al., 2018; Ji & Yan, 2022; Grasha et al., 2022).

4.3.4 Defining the locus of AGN-only region

For the three diagnostic diagrams described above, we were able to trace clear demarcation lines providing a separation between the AGN-dominated regions and the regions where SFGs and AGN can overlap. To do so, we considered both the photoionization models and the observational sample distributions. The demarcation lines for the O3Hg versus O32 diagnostic diagram are defined as follows:

$$Y = 0.55\,X - 0.95 \text{ for } X > 0.84 \tag{4.3}$$
$$Y = 0.1\,X - 0.57 \text{ for } X < 0.84. \tag{4.4}$$

Similarly, for the O3Hg versus Ne3O2 diagnostic diagram, the demarcation lines are

$$Y = 0.48\,X - 0.42 \text{ for } X > -0.07 \tag{4.5}$$
$$Y = 0.2\,X - 0.44 \text{ for } X < -0.07. \tag{4.6}$$

Finally, the demarcation line for the O3Hg versus O33 diagnostic diagram is the following:

$$Y = -1.1\,X + 1.47. \tag{4.7}$$

These demarcation lines ensure that the fraction of contaminating SDSS SFGs in the AGN part of the diagnostics is less than 1-2%, while none of the local analogues lie in the AGN-only part of the three diagnostics.
Such a tiny fraction of contaminants can probably be associated with sources that are not identified as AGN by the BPT diagram (used to distinguish between SDSS SFG and AGN in this work), but that host active SMBHs. The presence of AGN contaminants among BPT-selected SFG and local analogues has already been demonstrated by recent works (Svoboda et al., 2019; Harish et al., 2023).

Furthermore, it is important to note that the AGN selection based on these demarcation lines is not a necessary condition for a source to be an AGN, but rather a sufficient condition. Objects above the demarcation line are AGN-dominated with high confidence, while objects below the demarcation line can either be SFGs or AGN.

4.4 Discussion

By comparing the distribution of the observational samples with the distribution of the photoionization models, we can explain why these diagnostic diagrams can separate part of the AGN population from SFGs. In Fig. 4.5 we show the AGN and SFG models in the same diagnostic diagrams presented in Sect. 4.3, but highlighting the variation of log U and Z (i.e. the two main parameters affecting the distribution of the models in these plots). Since they are very local, we know that the SDSS AGN and SFG samples are populated mostly by solar metallicity and moderate- to low-ionization parameter sources, as shown also by the models occupying the same region of these samples in Fig. 4.5. On the contrary, local analogues are characterized, on average, by lower metallicities and higher ionization parameters (such as high-z galaxies). In Fig. 4.2 and Fig. 4.3 they fill the regions that extend from the SDSS SFG towards higher [O III] λ4363 /Hγ and higher [O III] /[O II] or [Ne III]/ [O II] , in close agreement with the trend of photoionization models for higher log U and lower Z in the same diagnostics (see Fig. 4.5).

The reason for the presence, in particular in Fig. 4.2 and Fig. 4.3, of a region that is populated only by AGN, at high [O III] λ4363 /Hγ values, can be explained in the following way. Since the [O III] λ4363 line is a collisionally excited line generated by high energy levels, its intensity relative to Hγ is directly related to the temperature and, secondly, to the metallicity of the ISM and the ionization parameter. The main difference between the AGN and SFGs is the SED of the incident ionizing radiation. For AGN the ionizing radiation is given by the emission of the accretion disc, while for SFGs it originates from young star clusters. The former are characterized by a harder spectrum, and therefore, on average, photons that are more energetic (see Fig. 1 in Feltre et al., 2016). Higher energy per ionizing photon, deposited into the ISM, results in higher effective heating, and in turn higher electron temperature (hence higher [O III] λ4363) at a given ionization parameter.

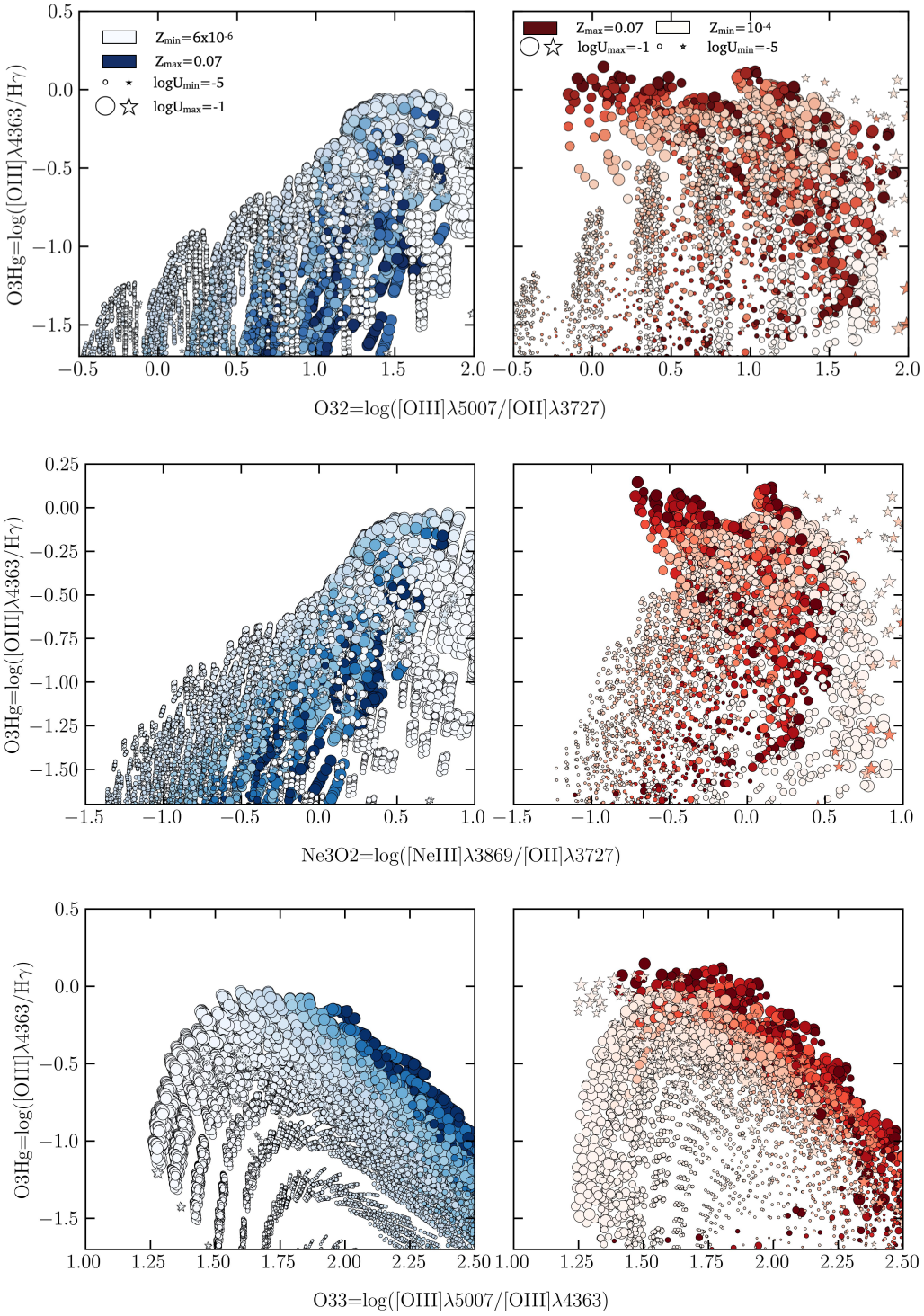

FIGURE 4.5: Distribution of the photoionization models computed in Feltre et al. (2016) (circles) and Nakajima & Maiolino (2022) (stars) with respect to the line ratios reported in Fig. 4.2 (top panel), Fig. 4.3 (central panel), and Fig. 4.4 (bottom panel). The photoionization models for SFG are reported on the left, while the AGN models are on the right. The points are colour-coded according to their metallicity (the stronger the colour, the higher the metallicity), and the marker size depends on the ionization parameter (the larger the marker, the higher the ionization parameter). The maximum and minimum values of Z and $\log \mathrm{U}$ for the two classes of models are reported in the top part of the top panels.

4.4.1 Stacking of the AGN candidates

To further demonstrate the effectiveness of the diagnostic diagrams proposed in this work, we explored potential tracers of AGN activity in those high-z galaxies lying in the AGN region of the diagnostics diagrams in Fig. 4.2 and Fig. 4.3, but not previously identified as AGN based on the selection performed in Scholtz et al. (2023b) (for JADES spectra) or using the standard NLAGN diagnostic diagrams for the CEERS spectra (see Chapter 5). Our stacking procedure was as follows. We first shifted the spectra to the rest frame and normalized them to the peak of the Hα line, which was available for all the sources involved. This ensured that the bright sources were not dominating the final stacks. We then resampled each of the spectra to a common wavelength grid. We verified that the wavelength grid did not impact our conclusions by repeating the analysis using wavelength bins ranging from the best resolution ($\sim$ 100km/s) to the worst ($\sim$ 150km/s). Finally, before stacking the spectra, we fitted and subtracted the continuum from the rest-frame rebinned spectra. We performed an inverse variance stacking of 15 sources (11 from CEERS and 4 from JADES).

We first fitted simultaneously the Hα and [N II] wavelength region using only a single Gaussian profile per emission line, with common full width at half maximum (FWHM) and redshift for all three emission lines. We fixed the [N II] doublet ratio to be 3. We show the narrow emission line fit in the top left panel of Fig. 4.6. We detected broad residuals around the Hα emission line. These broad components can arise from either a broad-line region or an outflow.

To model the broad wing in the Hα , we refitted the Hα and [N II] with two additional models, the BLR model (by fitting one additional broad Gaussian component to Hα only) and an outflow model (by adding broad Gaussian components to both Hα and [N II] with fixed FWHM and centroid). Using the Bayesian information criterion (BIC) parameter (Liddle, 2007), we found that the BLR model is strongly preferred to the narrow-only fit ($\Delta BIC = BIC_{H\alpha\ NL} - BIC_{H\alpha\ BL} = 90$). The broad component has a S/N of 10, making this a reliable detection. On the contrary, the fit with the outflow model did not return any broad component in the [N II] lines.

We then used the same approach with the Hβ and [O III] $\lambda\lambda$4959, 5007 complex, to test the possibility that this broad component could be associated with an outflow on larger scales, determining a broad [O III] . In the fit shown in the last panel of Fig. 4.6, we imposed the broad component of the [O III] $\lambda\lambda$4959, 5007 to have the same kinematics as the broad component of the Hα (i.e. the same FWHM and velocity offset). However, we did not find a significant broad line detection, and therefore outflow, in the [O III] $\lambda\lambda$4959, 5007 . In particular, by imposing a broad outflow component to the [O III] $\lambda\lambda$4959, 5007 and comparing it with the NL-only fit, it gives a $\Delta BIC = BIC_{[OIII]\ NL} - BIC_{[OIII]\ BL} = 4$, meaning that the two models are almost equally preferred, and hence a broad [O III] component is not required.

The absence of a counterpart in (brighter) [O III] strongly disfavors an outflow origin of the broad component of Hα . Therefore, the detection of a broad component in the Hα emission line suggests the presence of an AGN BLR in the spectra of at least some of these galaxies, which is too weak to be detected in the individual spectra. This indicates that galaxies located in the upper region of our diagnostic diagrams, even if originally not classified as AGN, actually may host faint Type 1 AGN. This result provides confidence that the

diagnostic diagrams presented in this work are effective in identifying AGN above our proposed demarcation lines, although they cannot distinguish between AGN and SFGs below the demarcation line.

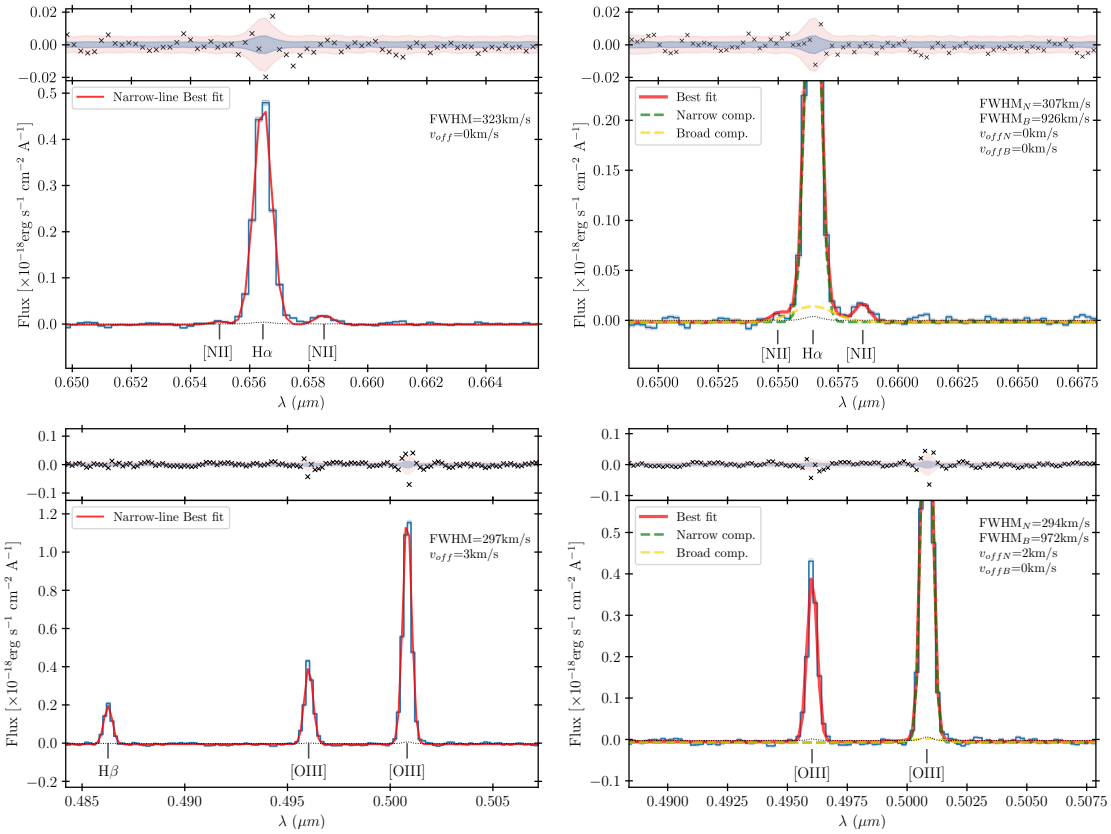

FIGURE 4.6: Different fits of the spectrum derived from the stack of the sources not identified as AGN in the literature, but lying in the AGN region of the diagnostic diagrams presented in Sect. 4.3. The stacked spectrum in these plots is resampled at the best resolution among those of the single spectra involved (i.e. $\sim$ 100km/s), but the results do not change considering the worst resolution (i.e. $\sim$ 150km/s). Top panels: Fit of the Hα and [N II] complex with only a narrow component (left) and adding a broad component to the Hα line (right). The global fit is presented in red, while the narrow and broad Hα components are in green and yellow, respectively. In the upper panel are shown the residuals of the fit compared with the distribution of the 1σ (blue) and 3σ (red) errors on the fluxes. In the upper right part of the plots are reported the FWHM and the velocity offset of the different components considered in the fits. Lower panels: Same as above, but for the Hβ and [O III] $\lambda\lambda$4959, 5007 doublet complex. In particular, on the right, we added to the fit of the [O III] a broad component with the same FWHM and velocity offset as the broad component of the Hα, but here it is not required by the fit.

4.4.2 Ruling out [FeII]λ4360 contamination

We investigate the possibility that the [O III] λ4363 emission line in the high-z sources located in the AGN-only region of the diagnostics could be contaminated by the [FeII]λ4360 emission line (Curti et al., 2017; Arellano-Córdova et al., 2022). Most of the high-z [O III] λ4363 detections come from JWST MR spectra, whose resolution is not high enough to disentangle the two lines. Curti et al. (2017) showed that this contamination is possible, but most likely in high-metallicity galaxies, while our high-z sample is generally metal poor.

To investigate the possible iron contamination of the [O III] λ4363 line, we consider the close [FeII]λ4288 emission line, which is isolated and originates from the same energy level as the [FeII]λ4360. We ran a grid of `Cloudy` photoionization models over a wide range of metallicities (with $-2 \leq \log(Z/Z_{\odot}) \leq 0.5$) and ionization parameters (with $-4 \leq \log(U) \leq -1$), and found that the flux ratio of [FeII]λ4288 / [FeII]λ4360[2] is roughly a constant of ~ 1.25. Therefore, we can use the [FeII]λ4288 emission line to reliably constrain the intensity of the [FeII]λ4360. We did not detect the [FeII]λ4288 line in any of the high-z sources in the AGN-only regions of the diagnostics. Even when we performed a weighted spectral stacking of all the spectra in the same region, the [FeII]λ4288 emission line remained undetected. Therefore, we conclude that any possible contribution to the [O III] λ4363 detections given by the [FeII]λ4360 emission line is negligible.

4.4.3 Impact on metallicity estimates and strong-line metallicity diagnostics

The [O III] λ4363 emission line has been widely used in the literature to derive metallicities via the T_e-method (see Maiolino & Mannucci, 2019, for a review), although Marconi et al. (2024) have shown that these metallicities may be biased low. The metallicities inferred from the T_e-method have then been used to calibrate the strong-line metallicity diagnostics (i.e. ratios of optical emission lines typically much stronger than the auroral lines, which can be used to estimate the metallicities on larger samples of galaxies). These calibrations have been inferred both by using local samples of galaxies (e.g. Maiolino et al., 2008; Bian et al., 2018; Curti et al., 2020) and, especially with JWST, at high redshift (Sanders et al., 2023; Laseter et al., 2024).

In most of these past studies care was taken to only use SFGs, or to assume that the gas ionization was dominated by young hot stars. The finding of consistent populations of AGN at high redshift by recent JWST studies, and especially our own finding in this work that a significant number of strong [O III] λ4363 detections at high redshift are due to AGN heating, may appear concerning when using this transition to infer the gas metallicity. However, the [O III] λ4363 line (together with [O III] , or [Ne III]as a proxy when the other is not available) provides a measurement of the temperature in the O^{+2} zone, regardless of the nature of the ionizing source. So its reliability remains unaffected in the case of photoionization by AGN (modulo the potential biases discussed in Marconi et al., 2024, which are an issue, also for SFGs). What is important is to properly use this information to infer the metallicity, and in particular by applying the proper ionization correction factors, which are different

[2]Our calculation takes into account six lines from Fe^{+} blended at 4288 Å and eight lines from Fe^{+} blended at 4360 Å, although most of the fluxes are contributed by [FeII]λ4287.39 and [FeII]λ4359.33.

in the case of AGN and SFGs. In particular, once the abundance of O^{+2}/H^{+} is derived, one has to estimate the contribution to the abundances from O^{+3}/H^{+} and O^{+}/H^{+}. The former is very difficult to assess, even in AGN, as there are no [OIV] strong transitions in the wavelength ranges typically accessible. Dors et al. (2020) estimate that, even in AGN, the contribution of O^{+3}/H^{+} is typically negligible. The contribution from O^{+}/H^{+} is estimated from the [O II] doublet; however, the key issue is that the temperature of the O^{+} region is different from the temperature of the O^{+2} region probed by the [O III] λ4363 line. If [OII] auroral lines are not available (as in the vast majority of the cases at high redshift), then one has to assume a relation between $T(O^{+})$ and $T(O^{+2})$. This relation has been extensively explored in local SFGs; however, it cannot be applied to AGN. Assuming in AGN the same temperature relation as for SFGs results in a large underestimation of the metallicities (Dors et al., 2015). This issue is greatly mitigated once AGN-bespoke relations between $T(O^{+})$ and $T(O^{+2})$ are adopted (Dors et al., 2020).

In summary, the fact that at high redshift [O III] λ4363 is boosted by AGN heating does not prevent it from being used to measure the metallicity in those galaxies, provided that the proper ionization correction factors are adopted and, in particular, the adequate $T(O^{+})$ and $T(O^{+2})$ is adopted (in absence of [OII] auroral lines).

However, the presence of AGN ionization, excitation, and heating certainly affects the empirical calibrations of the strong line diagnostics. At a given gas temperature the AGN radiation (with different ionizing shape and, typically, higher ionization parameter) results in a different ionization structure of the gas clouds, and also different collisional excitation rates of the various transitions typically used in the strong line diagnostics ([OIII], [OII], [NeIII], [NII], [SII]). As a consequence, separate and different empirical metallicity calibrations should be inferred for SFGs and AGN host galaxies. This has been attempted in the local universe (Dors, 2021). However, the same effort should be undertaken at high redshift, given the large abundance of AGN. Inferring empirical strong-line metallicity calibrations differentiating AGN and SFGs will be the focus of a separate dedicated work.

4.5 Conclusions

While JWST has revealed that some of the classical AGN diagnostics break down at high redshift, it has also opened the opportunity to explore new diagnostics. In this work we studied the possibility of selecting AGN using the [O III] λ4363 auroral line, whose detection in large numbers of high-z galaxies has become possible with JWST. In particular, we proposed three new diagnostic diagrams, and three corresponding demarcation lines, that allow a large population of AGN to be identified from SFGs by providing a sufficient (but not necessary) condition to claim the presence of an AGN.

To demonstrate the effectiveness of these diagnostics, we used multiple observational samples of local and high-z SFGs and AGN as well as photoionization models from Feltre et al. (2016) and Nakajima & Maiolino (2022). We specifically proposed the following three diagnostic diagrams:

- [O III] λ4363 /Hγ versus [OIII]5007/[OII]3727. This is the most thoroughly explored diagram from the empirical perspective (it has the largest observational test sample), and can be used with NIRSpec out to z<9.4;

- [O III] λ4363 /Hγ versus [NeIII]3869/[OII]3727. This diagram is the least sensitive to dust extinction given the wavelength proximity of both line ratios. Additionally, it can be used with NIRSpec out to z<10.9;

- [O III] λ4363 /Hγ versus [OIII]5007/[O III] λ4363 . Among the three, this is the one with the smallest region where AGN can be safely identified, but it is also relatively dust insensitive and requires the detection of only three lines instead of four, making it applicable to larger samples. Also this diagnostic diagram can be applied with NIRSpec data out to $z < 9.4$

In each of these cases we provided a demarcation line above which (i.e. with higher [O III] λ4363 /Hγ values) objects can be safely identified as hosting an AGN that dominates the nebular emission lines excitation.

We illustrated that at least some of the few objects falling into the AGN-only region and not previously identified as AGN actually show AGN signatures when stacked. This further supports the effectiveness of the diagnostics. At the same time, we did not find any indication of [FeII]λ4360 contamination of the [O III] λ4363 emission line in these sources, which could in principle artificially increase the [O III] λ4363 line due to blending effects.

The physics behind these diagnostics is tightly linked to the primary property of AGN (i.e. the hardness of the ionizing spectrum). The average energy of the AGN's ionizing photons is much higher than the energy of hot young stars in SFGs. Therefore, at a given ionizing radiation field intensity, AGN photons are more effective in heating the ionized gas than those in SF regions, hence yielding higher temperature, and hence higher [O III] λ4363 relative to Hγ (which accounts for the overall photoionizing radiation field).

We note that being above the demarcation line in the proposed diagrams is a sufficient but not necessary condition for an object to be identified as an AGN. Galaxies located below the demarcation lines can be powered by star formation and/or AGN. At the same time, the contamination from SFG is almost completely negligible in the upper part of the diagnostics, as demonstrated by the distribution of low-redshift observational samples and photoionization models. Moreover, the few high-redshift ambigous cases in the AGN-only region also turn out to be AGN based on the reliable (even if faint) detection of a broad Hα component in the stacking.

Finally, we note that the fact that strong auroral lines are often associated with AGN does not imply that they cannot be used for direct metallicity measurements (provided that proper ionization corrections are applied), but it does affect the calibration of strong line metallicity diagnostics, calling for new AGN-specific calibrations of these diagrams.

In the next chapter, we are going to apply these new diagnostic diagrams for the selection of NLAGN among the spectra of the CEERS survey on which we will also perform a detailed multi-wavelength analysis.

4.6 Supplementary material

4.6.1 Diagnostic diagrams including all photoionization models

The demarcation lines presented in Sec. 4.3.4 were defined considering both the distribution of the observational samples and of photoionization models with the cut described in Sec. 4.2.3 and shown in Fig. 4.1. If we consider the entire grid of parameters for the models of Gutkin et al. (2016) and Feltre et al. (2016), we obtain the distributions shown in Fig. 4.7. In this case, the demarcation lines for the O3Hg-O32 and O3Hg-Ne3O2 diagnostic plots still hold, except for a very small fraction of SFG models populating the AGN-only part of the two diagnostics. On the contrary, in the O3Hg-O33 diagnostic plot, there is an almost complete superposition between the SFG and AGN models. Even if this superposition is determined by highly unlikely combinations of parameters in the models, we cannot completely rule out the possibility that the AGN-only region of the O3Hg-O33 diagnostic diagram could actually be partially populated also by SFG.

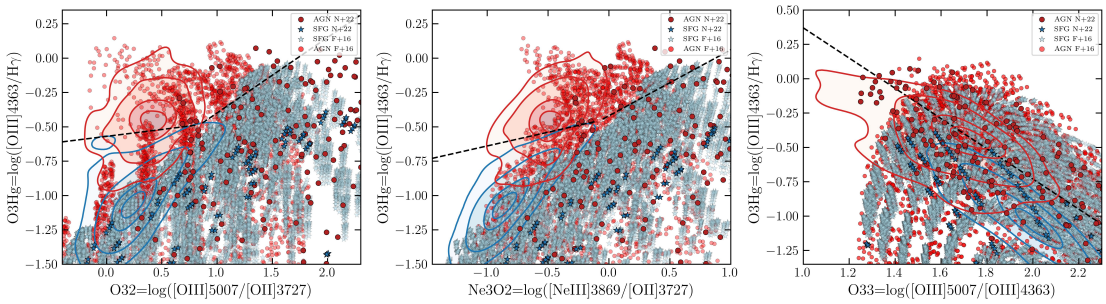

FIGURE 4.7: Same plots reported in the right panels of Fig. 4.2, Fig. 4.3 and Fig. 4.4 (from left to right), but showing all the models computed in Gutkin et al. (2016), Feltre et al. (2016), and Nakajima & Maiolino (2022), i.e. without the cut described in Sect. 4.2.3. The demarcation lines for the O3Hg-O32 and O3Hg-Ne3O2 still hold, while there is an almost complete superposition between SFG and AGN models in the O3Hg-O33 diagnostic diagram.

Chapter 5

Narrow line AGN selection in CEERS: physical properties of AGN at $2 \lesssim z \lesssim 9$

5.1 Introduction

In the previous chapter, we explored the possibility of identifying NLAGN at high redshift using new AGN diagnostic diagrams, not affected as the classical BPTs by the different physical properties of high-z galaxies compared to those in the local Universe.
As we discussed in Sect. 1.4.1 the identification of the elusive population of NLAGN is crucial to have a complete census of the AGN population, especially at early times, and to provide answers to the compelling questions about their unexpected physical properties observed at high-z.

So far, the systematic search for NLAGN at $z > 3$ has been attempted only by a few works. In Scholtz et al. (2023b) the authors selected 42 NLAGN among $\sim$ 200 medium resolution (MR) spectra of the JADES survey (Eisenstein et al., 2023; Bunker et al., 2023b). The selected sample allowed the authors to investigate the population of AGN up to $z \sim 9$, down to bolometric luminosities $L_{bol} \sim 10^{42}$ erg s^{-1} and host galaxy stellar masses of $10^7 M_\odot$, vastly expanding the region of the parameter space populated by these objects at early times.

In this chapter, we uncover the NLAGN population hidden among the sample of MR spectra observed as part of the CEERS survey. We apply, for the first time, the new NLAGN diagnostic diagrams presented in Chapter 4 (as well as many others), and we perform a comprehensive multiwavelength analysis to investigate in detail the physical properties of these sources across $\sim$ 3Gyrs of the cosmic history.

The Chapter is organized as follows. In Sect. 5.2, we present the data and the sample of spectra analyzed throughout this work and the techniques involved in the subsequent analysis: emission line fitting, spectral stacking, SED-fitting, and X-ray and radio stacking. In Sect. 5.3 we describe the different emission line diagnostic diagrams involved in the NLAGN selection and the results of the selection. In Sect. 5.4, we discuss the effectiveness of the different NLAGN diagnostic diagrams, and we compare our selection with other results presented in the literature on the same sample. Then in Sect. 5.5 we study the physical properties of the selected NLAGN: the distribution in velocity dispersion and obscuration (Sect. 5.5.1,

Sect. 5.5.2), the bolometric luminosities (Sect. 5.5.3), the host galaxies SFRs (Sect. 5.5.4). In Sect. 5.6 we extensively explore the average X-ray and radio properties of the AGN sample selected among the CEERS spectra. Finally in Sect. 5.7 we draw the conclusions and outline a summary of the NLAGN selection and characterization. In Sect. 5.8 we attach the supplementary material of this chapter.

The results discussed in this chapter have been presented in Mazzolari et al. (2024b) (accepted with minor revisions).

5.2 Data and Methods

5.2.1 CEERS Observational Data

We use publicly available medium resolution (MR) JWST NIRSpec (Jakobsen et al., 2022) Micro-Shutter Assembly (MSA; Ferruit et al., 2022a) data from the CEERS program (Program ID:1345, Finkelstein et al., 2022). The CEERS NIRSpec observations consist of six pointings in the 'Extended Groth Strip' field (EGS; Rhodes et al., 2000; Davis et al., 2007), all of which utilized the three grating/filter combination of G140M/F100LP, G235M/F170LP, and G395M/F290LP. This provided a spectral resolution of R $\sim$ 1000 over the wavelength range of approximately 1–5 μm (Jakobsen et al., 2022). Each grating/filter combination was observed for a total of 3107s for each pointing. The six NIRSpec pointing are shown (in red) in Fig. 5.1.

In particular, we considered the 313 spectra reduced and published by the CEERS collaboration in their latest data release DR-0.7[1]. The data reduction was performed by the CEERS collaboration using the JWST Calibration Pipeline version 1.8.5 (Bushouse et al., 2022), using the CRDS context "jwst_1029.pmap". The spectra were reduced using processing standard pipeline parameters, with some specific deviation in particular regarding the 'jump' parameters and the so-called "snowball" corrections. Nodded background subtraction was employed. The pipeline was instructed that all targets should be treated as point sources, determining each 1D spectrum to be extracted from the 2D spectral data file over a specified range of pixels in the cross-dispersion direction. Extraction apertures on the 2D spectra were determined for each object by interactive visual inspection and were specified explicitly for pipeline step 'extract_1d'. Most faint galaxies are compact, and the extraction apertures adopted for nearly all objects have heights ranging from 3 to 6 pixels (scale = 0.10 arcsec/pixel), with a median value of 4 pixels. The default pipeline path-loss correction was employed. This calibration is based primarily on pre-flight modeling, assuming the targets are point sources and determining these corrections to be incomplete for significantly extended sources, but this is not the case for the large majority of the targets. Flux calibration uses the default reference files for the adopted CRDS context.

For this work, we used the spectra with the three single grating spectra combined together. The data from the individual MR gratings are resampled to a common wavelength vector in the overlapping regions, adopting the wavelength grating which in the overlap region has the worst spectral resolution. Flux values in the overlap regions are the average of the individual

[1] https://ceers.github.io/dr07.html

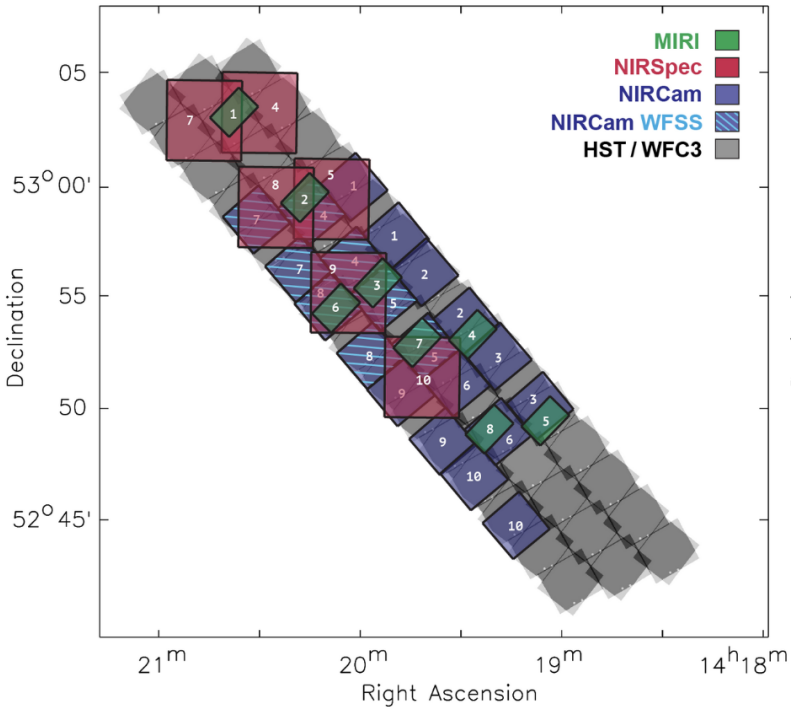

FIGURE 5.1: Footprint of the full JWST CEERS program. Different colors represent observations taken with different JWST instruments. The MR spectra analyzed in this work come from the six NIRSpec pointings (red). Image taken from CEERS collaboration website

grating values weighted by the flux errors, excluding pixels affected by masked artifacts.
After a careful visual inspection of the 313 single MR spectra available, we were able to attribute a secure redshift to 217 sources, constituting our parent sample throughout this work. Harikane et al. (2023) already investigated a limited sample of these spectra (i.e. only those at $z > 3.8$) to look for BLAGN, selecting 10 AGNs at z = 4.015 − 6.936 (2 of which are marked only as candidates) whose broad component is only seen in the permitted Hα or Hβ lines and not in the forbidden [O III] λ5007 line. In the rest of the chapter we refer to the 8 most reliable BLAGN selected in Harikane et al. (2023) as the sample of high-z BLAGN.

5.2.2 Emission line fitting

The NLAGN selection performed in this work is based on emission line diagnostics diagrams, where the AGN selection is provided by demarcation lines taken from the literature or by comparing the distribution of the sources with photoionization models, following the procedure done in Scholtz et al. (2023b).
We used a modified version of the publicly available code `QubeSpec` (Scholtz et al., 2023b; D'Eugenio et al., 2024) to fit the spectra of the sources. We fit the emission lines considering only small wavelength ranges around the single (or group of) emission lines ($\sim$ 500Å), with each emission line fitted using a single Gaussian component. The continuum was fit with a power law model, and in most of the spectra it was not detected. This approach for the fit of

the continuum is sufficient for describing a narrow continuum range around an emission line of interest, and we found no instances of a strong continuum that required more sophisticated (e.g., stellar emission) modelling. `QubeSpec` uses a Bayesian approach, which requires defining prior probability distributions for each model parameter. In our fit, we assume log-uniform priors for the peak of the Gaussian describing each line and also for the continuum normalization. The prior on the lines full-width half maximum (FWHMs) is set to a uniform distribution spanning from the minimum resolution of the NIRSpec/MSA ($\sim$ 200km/s) up to a maximum of 700 km/s. For the fit of the high-z BLAGN identified in Harikane et al. (2023) we also include in the fit a broad component in the Hα (with a FWHM uniform prior between 900 $-$ 5000km/s) to carefully disentangle the broad Hα emission from the narrow one, that is the only component used in the diagnostic diagrams. The prior on the redshift is a normal distribution centred on the redshift obtained from the visual inspection with a standard deviation of 100 km/s.
To estimate the posterior probability distribution of all these parameters, we use a Markov chain Monte Carlo integrator (`emcee`; Foreman-Mackey et al., 2013).
To address the NLAGN selection, we performed a fit of the following rest-frame optical and UV emission line blocks:

- Hα + [N II] $\lambda\lambda$6548,83 + [S II] $\lambda\lambda$6716,31
- [O I] λ6300
- Hβ + [O III] $\lambda\lambda$4959,5008 + He II λ4686
- Hγ + [O III] λ4363
- [O II] $\lambda\lambda$3726,29 + [Ne III] λ3869
- [Ne V] λ3426
- [Ne IV] $\lambda\lambda$2422,24
- C III] $\lambda\lambda$1907,10
- [O III] $\lambda\lambda$1660,66 + He II λ1640 +C IV $\lambda\lambda$ 1548,51
- N IV] $\lambda\lambda$1483,86
- Lyα + N V λ1240

For the components of the [O III] $\lambda\lambda$4959,5008 and [N II] $\lambda\lambda$6548,84 doublets, we fixed the ratios between the two components to be 1:3 given by the relative Einstein coefficients of the two transitions.
Given the resolution of the NIRSpec MR spectra, we fit the [O II] $\lambda\lambda$3726,29, [Ne IV] $\lambda\lambda$2422,24, C IV $\lambda\lambda$ 1548,51, [O III] $\lambda\lambda$1660,66, C III] $\lambda\lambda$1907,10, and N IV] $\lambda\lambda$1483,86 doublets as single Gaussians, centered at the mean wavelength between the two components.
The signal-to-noise ratio (S/N) of each line was evaluated considering the same equations reported in Mignoli et al. (2019) and Lenz & Ayres (1992), and requiring at least S/N$>$ 5 for the detection. Then, we visually inspected all the detected lines to exclude any possible false

detection. In case of undetection, we derived the upper limit to the flux of each line as 3σ the error on the flux returned by `QubeSpec`.
While fitting the Hα line, we found 4 sources at $z \leq 2$ for which the residuals from the narrow-line-only fit were too large, and a broad component was required. These sources were not selected as in Harikane et al. (2023) because in that work the authors investigated only sources at $z > 3.8$. We identify these sources as new BLAGN at low-z; their spectra are presented in Sect. 5.8.1 of the supplementary material of this chapter. In fitting all the BLAGN, we did not find any significant broad component either in the Hβ or in the [O III] λ5007 lines.

TABLE 5.1: Definitions of line ratios adopted throughout the chapter.

Diagnostics	Line Ratio
R3	$\log_{10}$([O III]λ5007 / Hβ)
N2	$\log_{10}$([N II]λ6583 / Hα)
S2	$\log_{10}$([S II]$\lambda\lambda$6716,31 / Hα)
O1	$\log_{10}$([O I]$\lambda\lambda$6300 / Hα)
He2	$\log_{10}$(He IIλ4686 / Hβ)
O3Hg	$\log_{10}$([O III]λ4363 / Hγ)
O32	$\log_{10}$([O III]λ5007 / [O II]$\lambda\lambda$3726,29)
Ne3O2	$\log_{10}$([Ne III] λ3869 / [O II] $\lambda\lambda$3726,29)
C43	$\log_{10}$(C IV$\lambda\lambda$1549,51/C III]$\lambda\lambda$1907,09)
C3He2	$\log_{10}$(C III] $\lambda\lambda$1907,09 / He IIλ1640)
Ne43	$\log_{10}$([Ne IV] $\lambda\lambda$2422,24/[Ne III] λ3869)
N5C3	$\log_{10}$(N V $\lambda\lambda$1239,42/C III]$\lambda\lambda$1907,09)
N4C3	$\log_{10}$(N IV] $\lambda\lambda$1483,86/C III]$\lambda\lambda$1907,09)

5.2.3 SED fitting

To investigate the physical properties of the host galaxies of the selected NLAGN and to compare them with the properties of the parent sample, we performed a SED fitting analysis using `CIGALE` (Boquien et al., 2019). Starting from the 217 spectra of the MR sample with a secure redshift, we first cross-match the target coordinates with the HST+JWST publicly available photometric catalog of the CEERS field reported in the DJA archive[2]. Considering a 0.5″ matching radius we found 117 counterparts. This catalog includes seven HST bands (F435W, F606W, F814W, F105W, F125W, F140W, F160W) and ten JWST/NIRCam and MIRI bands (F115W, F150W, 182M, F200W, 210M, F277W, F356W, F410M, F444W, F770W). Since not the whole CEERS MR sample is covered by JWST imaging (see Fig. 5.1), we also cross-match the remaining sources with the 3D-HST multiwavelength catalog (Momcheva et al., 2016; Skelton et al., 2014), covering the EGS field and using the same cross-matching radius to avoid false matches. We found 93 additional counterparts, bringing the total number of sources with an associated optical/NIR photometry to 210. For the remaining 7 spectra,

[2] https://s3.amazonaws.com/grizli-v2/JwstMosaics/v7/index.html

we didn't find a counterpart closer than 0.5", but their counterparts are returned when we considered a larger cross-matching radius ($0.5" < r < 1.2"$). For sources detected in the JWST+HST catalog, we didn't consider the additional photometry of the 3D-HST catalog since most of these sources are undetected in photometric bands bluer than the HST bands (being at high-z), while the photometry at longer wavelengths than HST bands is already covered by JWST (the deepest in the near and mid-infrared).
The SED fitting was performed considering two different parameter grids, reported in Sect. 5.8.2 of the supplementary material, one for sources at $z < 3$ and the other for sources at $z > 3$. We used delayed star formation history (SFH) models for both redshift groups because they are able to reproduce both early-type and late-type galaxies, with an additional term that allows for a recent (and constant) variation of the star-formation rate (SFR) that can be in the burst or in the quench phase. We adopted stellar templates from Bruzual & Charlot (2003), and a Chabrier initial mass function (Chabrier, 2003). We also include the nebular emission module that is extremely important to account for the contribution of emission lines in the broad-band photometry (Schaerer & de Barros, 2012; Salvato et al., 2019), in particular for the low-mass and high-z regimes, where emission lines can account for a consistent fraction of the wideband photometric flux. This module is computed by `CIGALE` based on a grid of user-provided parameters and in a self-consistent manner using a grid of `Cloudy` photoionization calculations Ferland et al. (2013). For the attenuation of the stellar continuum emission, we considered the `dustatt_modified_CF00` module (Charlot & Fall, 2000), which allows different attenuations for the young and old stellar populations. We also include dust emission in the IR following the empirical templates of Dale et al. (2014). The main differences in the $z > 3$ grid with respect to the low-z one consist in the larger parameter space explored for the possible final starburst state of high-z galaxies and in the lower metallicities and higher ionization parameters allowed in the stellar and nebular modules, according with recent JWST results (Endsley et al., 2023; Tang et al., 2023). Indeed, SFG at high-z are observed to be more frequent in a bursty SF regime (Faisst et al., 2020; Dressler et al., 2023; Looser et al., 2023) and also, the gas metallicities are generally 0.5-1 dex lower than in the local universe (Curti et al., 2023a; Nakajima et al., 2023).

For sources that are classified as AGN based on our NLAGN selection performed in Sect. 5.3, as well as for the BLAGN selected in Harikane et al. (2023), we included AGN modules in the SED fitting to account for the AGN emission. We employed the `skirtor2016` module introduced in Yang et al. (2020), which is widely used in the community and has demonstrated reliability in studying various aspects of AGN (e.g., Mountrichas et al., 2022; López et al., 2023; Yang et al., 2023b). The SED produced by the AGN combines emissions from the accretion disk, torus, and polar dust. The accretion disk, responsible for the UV-optical emission at the central region, is parameterized according to Schartmann et al. (2005), with a typical delta value of -0.36. Photons from the accretion disk can be obscured and scattered by dust in the vicinity, within the torus and/or in the polar direction. Specifically, for the torus, the `skirtor2016` module employs a clumpy two-phase model (Stalevski et al., 2016), based on the 3D radiative-transfer code SKIRT (Schaerer & de Barros, 2012; Salvato et al., 2019). As our photometry does not cover the restframe mid-IR, we fixed the optical depth at 9.7 μm at $\tau = 3$. For the polar dust component, we used the Small Magellanic Cloud (SMC) extinction curve, recommended due to its preference in AGN observations (e.g., Bongiorno et al., 2012). The extinction amplitude, parameterized as $E(B-V)$, is a user-defined

free parameter for which we chose a typical value of $E(B-V) = 0.04$. Emission from the polar dust maintains energy conservation, assuming isotropic emission and a 'grey body' model (Casey, 2012). Finally, we allowed inclination values from 50 to 80 degrees and varied the AGN fraction (defined as the ratio between the AGN luminosity and the host galaxy luminosity between 0.15 and 2 microns) from 0.1 (weak AGN contamination) to 0.7 (dominant AGN emission).

5.2.4 X-ray and radio counterparts

To investigate the multiwavelength properties of the CEERS sources, and in particular of those that are selected as NLAGN using the diagnostic diagrams reported in Sect. 5.3, we look for counterparts of the sources observed in CEERS in the X-rays and Radio images of the EGS field.
The Chandra AEGIS-XD field is the third deepest X-ray field ever performed, with an exposure that reaches ∼800 ks. The X-ray observations and the related X-ray catalogue are described in Nandra et al. (2015). The X-ray catalogue contains 937 X-ray sources, and 553 of these sources (those with enough photon counts) also have an X-ray spectral analysis performed by Buchner et al. (2015). By cross-matching the CEERS MR catalog with the X-ray catalogue, we found 7 matches, involving sources at $0.5 < z < 3$. The X-ray spectral analysis classified two of these sources as galaxies (CEERS-3060, CEERS-3051), the other 5 (CEERS-2919, CEERS-2808, CEERS-2904, CEERS-2989, CEERS-2900) as AGN. In particular, CEERS-2904, CEERS-2989, and CEERS-2919 are part of the new low-redshift BLAGN sample presented in Sect. 5.8.1 of the supplementary material.

The field targeted by the CEERS survey is also covered by the AEGIS20 1.4GHz observations described in Ivison et al. (2007). This radio survey was performed in 2006 with the Very Large Array (VLA) over an area of $\sim 0.75\mathrm{deg}^2$ and reaching, in the central region, a rms of $\sim 10\mu$Jy. The resulting radio catalog (Ivison et al., 2007) contains 1123 individual radio sources, of which only two are part of the CEERS MR sample, i.e., CEERS-2900 and CEERS-3129, the first (that is also X-ray detected) showing the typical morphology of a radio-loud AGN.

5.3 Results

Here we present the result of the NLAGN selection performed on the sample of the 217 MR CEERS spectra with a secure redshift identification. The NLAGN selection was performed using emission line diagnostic diagrams.
In all the diagrams, we plot both the observed data points and also the line ratios derived from the photoionization models described in Feltre et al. (2016) and Gutkin et al. (2016). These models were computed using the `Cloudy` code (Ferland et al., 2013) for SFG and AGN NL regions and for various gas metallicities, ionization parameters, dust content, and ISM densities, and considering a wide range of the parameters space (for the full grid of values covered by the parameter space we refer to Table 1 in Feltre et al., 2016). To allow a comprehensive enough coverage of the SFG models but to avoid physical conditions rarely (or never) found in the general populations of SFG in the local or high-z Universe, we applied to

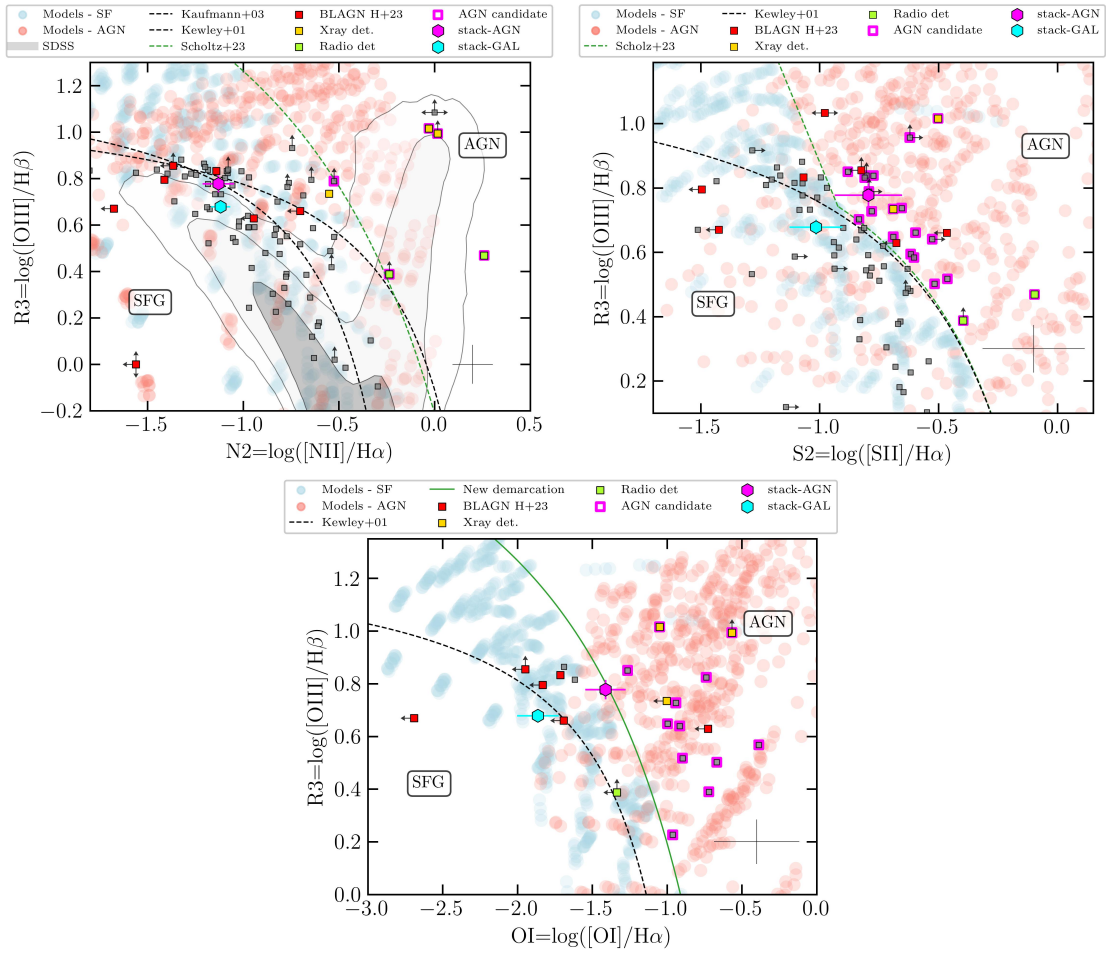

FIGURE 5.2: *Left:* R3N2 diagnostic diagram (also called BPT). The gray points represent the parent sample of galaxies analyzed in this work, gold, green and red colors are used for X-ray detected sources, radio detected sources and the high-z BLAGN reported in (Harikane et al., 2023). The NLAGN selected using this diagram are marked with magenta edges. In magenta and cyan, we show the line ratios derived from the stacked spectra of all the selected NLAGN (52 sources) and non-AGN sources. The dashed lines refer to three different demarcation lines, as labeled, the one in green is the more conservative demarcation line derived in Scholtz et al. (2023b). The blue and red shaded area represents the regions covered by the SFG and AGN photoionization models computed in Gutkin et al. (2016) and Feltre et al. (2016). The gray contours mark the distribution of SDSS sources (taken from SDSS DR7 Abazajian et al., 2009). In the lower-right corner are reported the median errors of the sample. *Right:* R3S2 diagnostic diagram (also called VO87) with the demarcation line originally presented in Kewley et al. (2001) (in black) and the new (more conservative) demarcation derived in Scholtz et al. (2023b). *Bottom:* R3O1 diagnostic diagram with the demarcation line presented in Kewley et al. (2001) (black dashed line) and the new one derived in this work (green solid line).

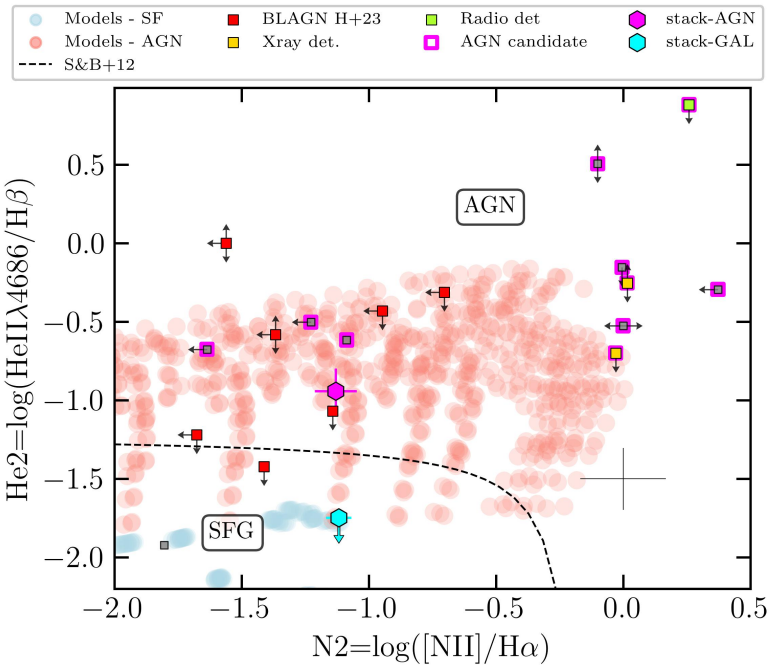

FIGURE 5.3: He2N2 diagnostic diagram. The colors follow the same scheme as in Fig. 5.2

these models the same cut in the parameter space as presented and discussed in Scholtz et al. (2023b). Through the work, photoionization models for AGN and SFG will be shown in red and in blue, respectively. We also mark with distinctive colours the seven sources that are X-ray detected (gold), the two radio-detected sources (green), and the eight sources that were selected as BLAGN at $z > 4.5$ by Harikane et al. (2023) (red). In each diagnostic diagram, the sources with magenta edges are those selected as NLAGN in that specific diagnostic.
It is worth noting that the line fluxes involved in the diagnostic diagrams described in the next Sections are not corrected for the effect of dust. This is because we are always considering ratios between lines close to each other, hence subjected to almost the same reddening, making the ratio insensible to the presence of dust. The few exceptions are discussed and justified. The acronyms used for the line ratios are reported in Tab. 5.1.

5.3.1 Optical diagnostics

In Fig. 5.2, we show the traditional BPT (Baldwin et al., 1981), the S2-VO87, and the R3O1 (Veilleux & Osterbrock, 1987) diagnostics diagrams based on the R3-N2, R3-S2, and R3-O1 line ratios, respectively. The traditional demarcation lines between SFG and AGN provided by Kauffmann et al. (2003) and Kewley et al. (2001) already proved to be way less effective in selecting high-z AGN compared to the local universe (Scholtz et al., 2023b; Übler et al., 2023; Maiolino et al., 2023b). Indeed, the lower metallicities of high-z sources make [N II] (but also [S II]) emission lines fainter, shifting the objects (AGN included) towards the left part of the BPT diagram. At the same time, the higher ionization parameter of the high-z SFG (Cameron et al., 2023; Sanders et al., 2023; Topping et al., 2024), due to, on average, the presence of younger stars and lower metallicities compared to the local universe, moves the

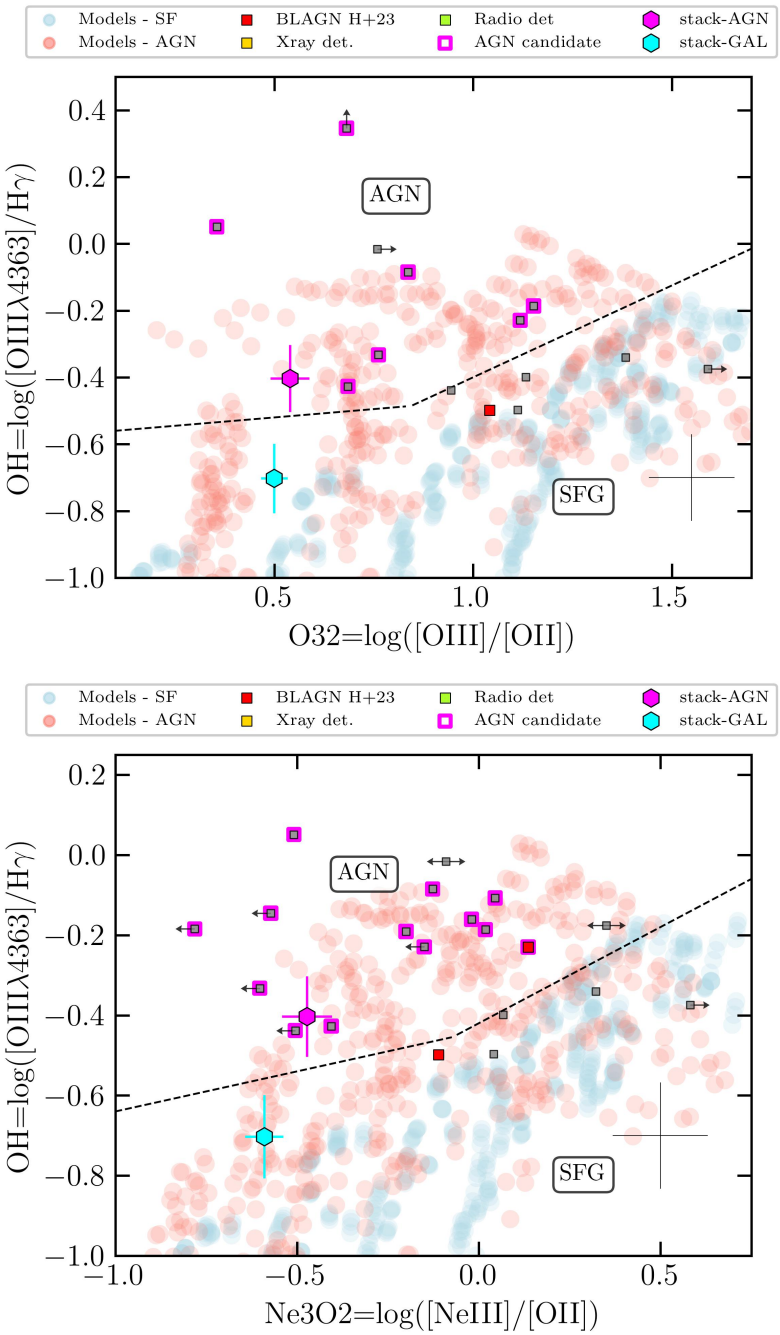

FIGURE 5.4: M1 and M2 diagnostic diagrams, firstly presented and discussed in Mazzolari et al. (2024a). The colors follow the same scheme as in Fig. 5.2

SFG towards higher R3 ratios, generating a large overlap between AGN and SFGs. Therefore, in the R3N2 and R3S2 diagnostics, we used the *conservative* demarcation lines provided in Scholtz et al. (2023b), derived considering the distribution of the photoionization models of Gutkin et al. (2016); Feltre et al. (2016) and Nakajima & Maiolino (2022).
Since the R3O1 diagnostic diagram was not included in the analysis of Scholtz et al. (2023b), we followed the same fitting method to the photoionization models to find more conservative demarcation lines with respect to those presented in Kewley et al. (2001). We derived the following dividing line between the AGN and SFG population:

$$\mathrm{R3} = 2.5 + \frac{2.65}{\mathrm{O1} - 0.15}. \tag{5.1}$$

As it is possible to see from the three plots, none of the BLAGN at high-z are in the AGN region of the three diagnostics, based on the new demarcation lines, and only one would be selected considering the traditional demarcation lines in the R3S2 and R3O1 diagrams (is the same source, CEERS-1665).
The five AGNs selected with the R3N2 diagram are all at $z < 2.9$. Three of them are X-ray-detected AGN (one is also radio-detected, i.e. CEERS-2900), and one is another radio-detected source (CEERS-3129). It is worth noting that even the X-ray source at the highest redshift (i.e. CEERS-2808 at $z = 3.38$), does not fall in the pure-AGN region of this diagnostic diagram.
The R3S2 diagnostic diagram appears effective in selecting AGN, also at a higher redshift with respect to the BPT. The new demarcation line provided in Scholtz et al. (2023b) partially overlaps with the traditional one proposed in Kewley et al. (2001). In this case, we select 21 AGN, six of which at $z > 3$, one at $z = 5.27$. None of the BLAGN is selected as AGN based on this diagnostic, while three X-ray sources and the two radio sources are confirmed as AGN also in this diagnostic.
Also the R3O1 diagnostic diagram is more effective in identifying NLAGN compared to the R3N2. In this diagram, we select 12 AGN, 7 at $z < 3$ (including 2 X-ray sources), and 5 at $z > 3$, one of these at $z \sim 6$. We will further discuss the effectiveness of this diagnostic and of the R3S2 in Sect. 5.4.
In Fig. 5.3 we show the diagnostic based on the He2 vs N2 lines ratio. In this case, the demarcation line is the original one provided by Shirazi & Brinchmann (2012), which still holds even at high-z, as also reported in Dors et al. (2024) and Tozzi et al. (2023). Overall, we detect He II λ4686 in seven sources: all except one of these are classified as AGN. None of the CEERS BLAGN show detection of the He II λ4686 line. In this diagnostic, there are also five sources that are selected as AGN only based on the N2 ratio.
The last diagnostics exploiting rest-frame optical lines are the two diagnostics based on the O3Hg ratio and described in Chapter 4 (Mazzolari et al., 2024a, and published in). Hereafter, we will refer with M1 to the O3Hg versus O32 diagnostic diagram and with M2 to the O3Hg versus Ne3O2 diagram. As explained in the previous chapter, these diagnostics proved to be effective in selecting those AGN characterized by high O3Hg ratios at a given ratio of O32 or Ne3O2 and were able to select NLAGN also at $z > 6$. The [O III] λ4363 line is sensitive to the electron temperature of the ISM. The effectiveness of these diagnostics is related to the fact that the average energy of AGN's ionizing photons is higher than that of young stars in SFGs, because the AGN ionizing source (the accretion disk) produces a harder

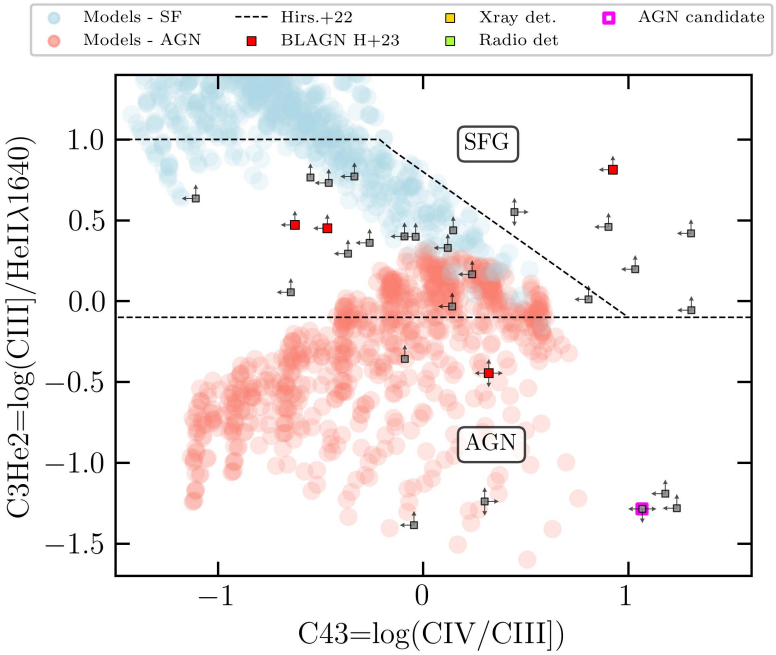

FIGURE 5.5: C3He2 vs C43 diagnostic diagrams, selecting only one NLAGN due to the difficulties in detecting rest-UV lines in high-z galaxies with the short exposure times of the CEERS survey. The stacking line ratios and the average errors are not reported because of the poor statistics. The colors follow the same scheme as in Fig. 5.2. The demarcation lines are those defined in Hirschmann et al. (2022).

SED with respect to stars. Therefore, AGN can more efficiently heat the gas, boosting the [O III] λ4363 line. Using the two diagnostic diagrams reported in Fig. 5.4, we were able to select 15 distinct NLAGN, up to $z \sim 6$. In the case of the left panel of Fig. 5.4, the O32 line ratio can suffer from the effect of dust reddening due to the wavelength distance of the two lines involved. However, the effect of dust attenuation on this diagnostic moves sources toward the right, without contaminating the AGN-only region with SFGs, allowing us to select only pure AGN.

We note that in the diagnostics presented in Fig. 5.2, 5.3, 5.4 there are some sources lying close enough to the demarcation lines that their line ratios' error can potentially place them outside from the AGN-only region of the diagrams. We considered all the selected NLAGN whose 1σ uncertainties are compatible with SF-driven photoionization, and we double-check if they were safely identified as NLAGN in other diagnostics (i.e. without 1σ errors crossing the demarcation line). We identified 4 NLAGN with errors compatible with SFG ionization: CEERS-2668, CEERS-3535, CEERS-1836, CEERS-3585 and we conservatively decided to mark them only as candidates NLAGN. Also CEERS-2900 should be included in this sample, but given its detection in both the X-ray and radio image of the field it can be safely considered a NLAGN.

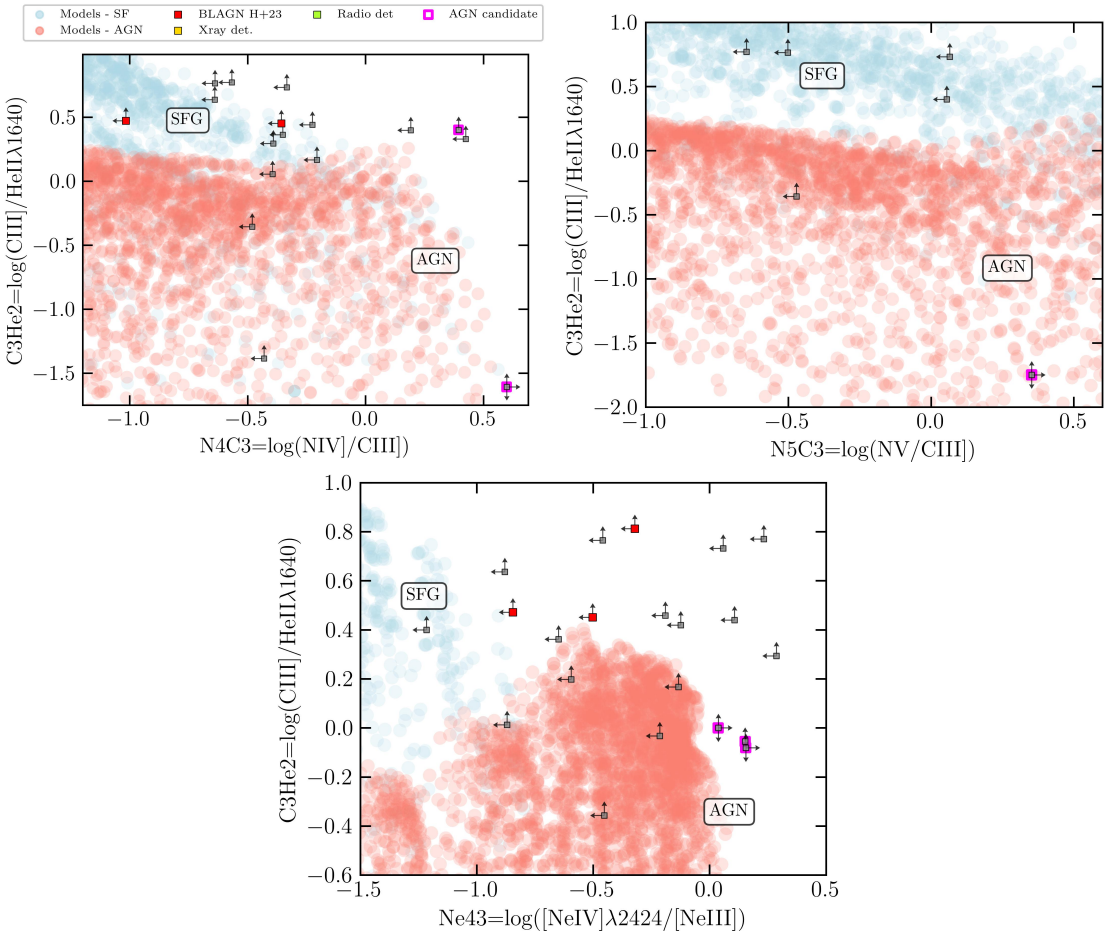

FIGURE 5.6: From top left to the bottom (clockwise): C3He2 vs N4C3, C3He2 vs N5C3, C3He2 vs Ne4C3 diagnostic diagrams. The stacking line rations and the average errors are not reported because of the poor statistics. The colors follow the same scheme as in Fig. 5.2

5.3.2 UV diagnostic

In this section, we will focus on the diagnostic diagrams involving rest-frame UV lines (Mingozzi et al., 2024; Mascia et al., 2024; Scholtz et al., 2023b; Feltre et al., 2016). In particular, we considered the UV diagnostic diagram C3He2-C43 (where He2 is He II λ1640) and the diagnostic diagrams involving the so-called high-ionization emission lines, i.e. N IV] λ1486, N V λ1243, [Ne IV] λ2425.
The C3He2-C43 diagnostic diagram reported in Fig. 5.5 allows us to select one single NLAGN (CEERS-613), whose selection is based on a clear detection of the He II λ1640 line, that places the source well into the AGN-dominated region according to the demarcation lines presented in Hirschmann et al. (2022). The detection of the UV lines involved in this diagram is challenging with only $\sim$ 50min of JWST exposure (the average on-source exposure time of the CEERS survey). Also in Scholtz et al. (2023b) the number of detections of these lines is low, even if the exposure time per target of the JADES survey was $\sim$ 8 times longer than the CEERS observations.
The high-ionization emission lines are characterized by high-photoionization energies that can be produced more probably in AGN, due to their harder ionizing radiation, and therefore can be used as indicators of the presence of an AGN, even if hidden. In the diagnostics diagrams involving these lines, we base the identification of NLAGN on the position of the sources compared to the distribution of AGN and SFG photoionization models of Gutkin et al. (2016) and Feltre et al. (2016). In the top panel of Fig. 5.6 we show the diagnostic diagram involving the N IV] λ1486 emission line, which has an ionization potential of 47eV. This diagnostic allows us to select two NLAGN (CEERS-496, CEERS-1019), based on their extreme position with respect to the distribution of both SFG and AGN photoionization models. However, given their unclear detection of the He II λ1640 line in both these sources, and given the uncertainties in the extension of the SFG region in the diagnostic, we decided to mark these NLAGNs only as candidates. In particular, the source in the upper part of the diagram is CEERS-1019, whose AGN nature has already been discussed in multiple works in the literature (Mascia et al., 2024; Larson et al., 2023), some supporting the presence of an AGN and some discarding it (Tang et al., 2023). In the middle panel of the same figure is represented the diagnostic diagram involving the N V λ1240 emission line (ionization potential of 77eV). In this case, we could select only one source with a N V λ1240 detection, CEERS-23, that again we marked as a candidate due to the undetection of the He II λ1640 line. In the last panel of Fig. 5.6, we show the [Ne IV] 2424 emission line diagnostic, whose ionization potential is 63eV. We detect the [Ne IV] 2424 line in three different sources (CEERS-1029, CEERS-13061, CEERS-4210), and all of them occupy a region of the diagram covered only by the AGN models, even considering the upper limits in He II λ1640. In this diagnostic diagram, where the [Ne IV] and [Ne III]lines are $\sim$ 1000 apart one from the other, the dust correction would move the points towards the right, to even more AGN-extreme values.
Furthermore, we detect the [Ne V] 3426 emission line (97eV of ionization potential) in one source (CEERS-8299, at $z = 2.15$) that we marked as AGN given the extremely high-ionization energy required to produce this line. Even if at high-z it has been proposed that the [Ne V] 3426 could also be associated with SF processes Cleri et al. (2023), it was always related to AGN activity at $z < 3$ Mignoli et al. (2013); Barchiesi et al. (2024).

Overall we selected 52 NLAGN among the initial 217 sources.

5.3.3 Spectral Stacking

Once the final sample of 52 NLAGN have been selected, to get the average emission line properties of the SFG and NLAGN samples, we stacked the MR spectra for each of the emission line complexes investigated in this work. The spectral stacking was performed firstly, normalizing all the NLAGN and SFG spectra by the peak flux of the [O III] λ5007 line, in order not to be biased towards the brightest objects. The [O III] λ5007 line is the most common and well-detected line among the CEERS MR spectra, but is not available in 8 AGN and in 14 SFG spectra because for these sources the line falls into a detector gap. Therefore we excluded them from the stacking. We shifted all the spectra to the rest frame, and then we resampled each of the spectra to a uniform and common wavelength grid, given that the three components of each spectrum (corresponding to the three different gratings) have different resolutions and wavelength bins. The new, and uniform, wavelength grid was defined, for each line complex, by the wavelength bin allowing for the best spectral resolution among the different resolutions of the spectra involved in the stacking of the line(s). We checked that the results do not significantly change taking the wavelength bin corresponding to the worst spectral resolution among the spectra. Before stacking the spectra, we also fit and subtract the continuum from the rest-frame rebinned spectra.
We finally opted for spectral stacking involving a bootstrap procedure, given the number of sources involved. The stacked spectrum and uncertainties of each spectral bin were derived taking the median and standard deviation of 250 bootstrap realizations of the input spectra. The spectra resulting from the stacking of all NLAGN and SFG are presented in Fig. 5.7.

In each of the diagnostic diagrams reported in Sect. 5.3, we also report the line ratios of the stacked spectra of all the selected NLAGN and of the non-AGN selected sources. The two stacked spectra, shown in Fig. 5.7, were fitted to measure emission line fluxes using `Qubspec`, following the same procedure as for all the other single spectra and described in Sect. 5.2.2. As it is possible to see from the BPT (R3N2) diagram, the position of the stacked NLAGN and the non-AGN sources are very close together, with AGN having a slightly larger R3 line ratio. This clearly shows how the BPT diagram and its traditional demarcation lines are no longer useful in effectively separating the AGN population from SFG, probably even for relatively low redshifts ($z > 2$). We will further discuss this point in Sect. 5.4. On the contrary, in the R3S2 diagnostic diagram (VO87), the line ratios of the stacked NLAGN spectrum clearly lie in the AGN region of the diagnostic, while the non-AGN one is in the SFG-dominated region. The R3S2 line ratio seems to be still informative for the AGN selection also at (relatively) high-z, as we will discuss in Sect. 5.4. The same applies to the R3O1 diagnostic diagram. As for the He2N2 diagram, we note that while the NLAGN stacked spectrum shows a clear detection of the He II 4686 line, which places the final NLAGN stack clearly in the AGN region of the diagnostic, for the non-AGN sample the He II λ4686 line is not detected and the line ratios fall in the SFG region, as expected. This is also a strong point in favor of the goodness of our NLAGN selection since the He II λ4686 line was detected with high-enough S/N only in 5/52 of the NLAGN, 4 or them at $z < 3$. The fact that this line is clearly detected in the NLAGN spectral stack and that the stack line ratios fall in the AGN region of the diagnostic means that signatures of AGN emission are widely present among our NLAGN sample.
In both the diagnostic diagrams involving the [O III] λ4363 auroral line, NLAGN are characterized by a larger ($\sim$ 0.3 dex) [O III] λ4363/Hγ line ratio compared to the non-AGN, while the two have a similar O3O2 or Ne3O2 line ratios, meaning that the NLAGN selection

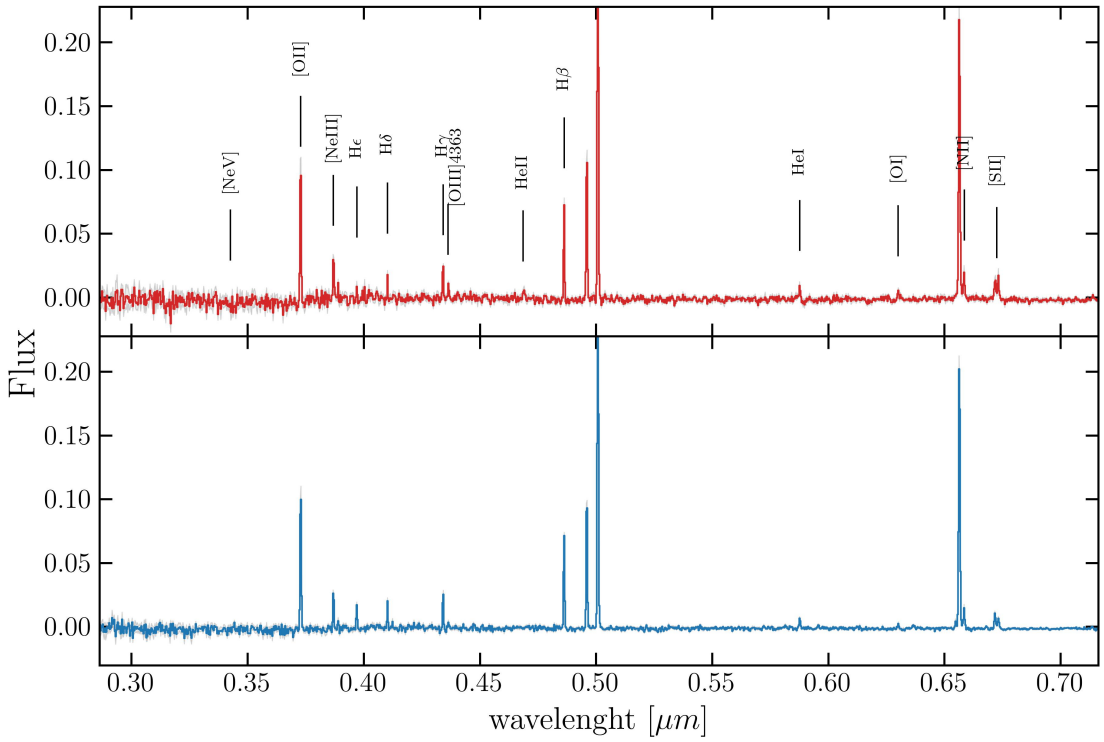

FIGURE 5.7: JWST NIRSpec stacked spectra of the final sample of NLAGN (red) and non-AGN (blue) derived as described in Sect. 5.3.3. We also marked the positions of some relevant lines.

is mainly driven by a stronger [O III] λ4363 auroral line in AGN, as discussed in Mazzolari et al. (2024a).
We did not put the results of the stack in the rest-UV diagnostic diagrams. This is because in the case of the C3He2-C43 diagnostics, the He II λ1640 and C IV lines are available only for 33/155 non-AGN and 7/52 NLAGN, and both these lines are not clearly detected in none of the two stacked spectra. This is expected given the faintness of these lines and the observing time of the CEERS survey, and indeed this diagnostic was able to identify only 1 NLAGN. The same goes for the high-ionization emission line diagnostics: none of the high-ionization emission lines involved in the diagnostics in Fig. 5.6 is clearly detected in our stack.

5.4 Comparison of NLAGN selection methods

The IDs and the main physical properties (that will be presented and discussed in the next Sections) of all the selected 52 NLAGN are reported in Table 5.4 of the supplementary material. In this section, we discuss the effectiveness of the different diagnostic diagrams reported in Sect. 5.3. For the discussion below, the term effectiveness refers to the ability of a diagnostic diagram to efficiently select (i.e. isolate in the diagram) NLAGN among all the considered sources.
As presented in the top panel of Fig. 5.8, most of the NLAGN are selected based on only one single diagnostic diagram (35/52, the 65%), 13 are selected in two different AGN diagnostics, while 4 are selected based on more than two diagnostics diagrams. The approach chosen for the NLAGN selection in this work clearly favors the high-completeness instead of the high-purity criterium, selecting as NLAGN all the sources showing the dominance of the AGN photoionization in at least one diagnostic diagram. However, the fact that most of the NLAGN are selected based on one single NLAGN diagnostic is not surprising. The NLAGN diagnostic diagrams we used for our selection involve emission lines that cover a wide wavelength range (from rest-UV to rest-optical), over which the spectral features of sources at $1 \lesssim z \lesssim 9$ can vary a lot depending on many parameters (metallicities, stellar population in the host galaxy, radio or X-ray emission, etc.). Given this, different NLAGN diagnostic diagrams are intrinsically sensitive to different AGN properties and, in turn, can become ineffective (and therefore less useful in separating AGN from SFG) under different conditions. Additionally, given the wide redshift range of the sources analyzed in this work, not all of them have in their observed spectra all the lines used for the diagnostic reported in Sect.5.3.1,5.3.2, mostly because lines might be redshift out from the observed wavelength range or fall into a spectral gap.
Looking at the lower panel of Fig. 5.8, we see the distribution of the sources selected by the different diagnostic diagrams with respect to their redshift and their selection diagram. The diagnostic diagram that selected the larger number of the AGN is the R3S2, in particular at $z < 3$, but with a non-negligible number of NLAGN also selected at $z \sim 5$. The effectiveness of the R3S2 (but also of the R3O1) diagnostic diagram in the low-metallicity (and therefore high-z) regime was already pointed out in some recent works investigating the effectiveness of traditional NLAGN diagnostic diagrams to select AGN among low-metallicity dwarf galaxies (Polimera et al., 2022). In particular, they show that the S2 and the O1 line ratios are less metallicity sensitive and more successful in identifying AGN in these dwarf

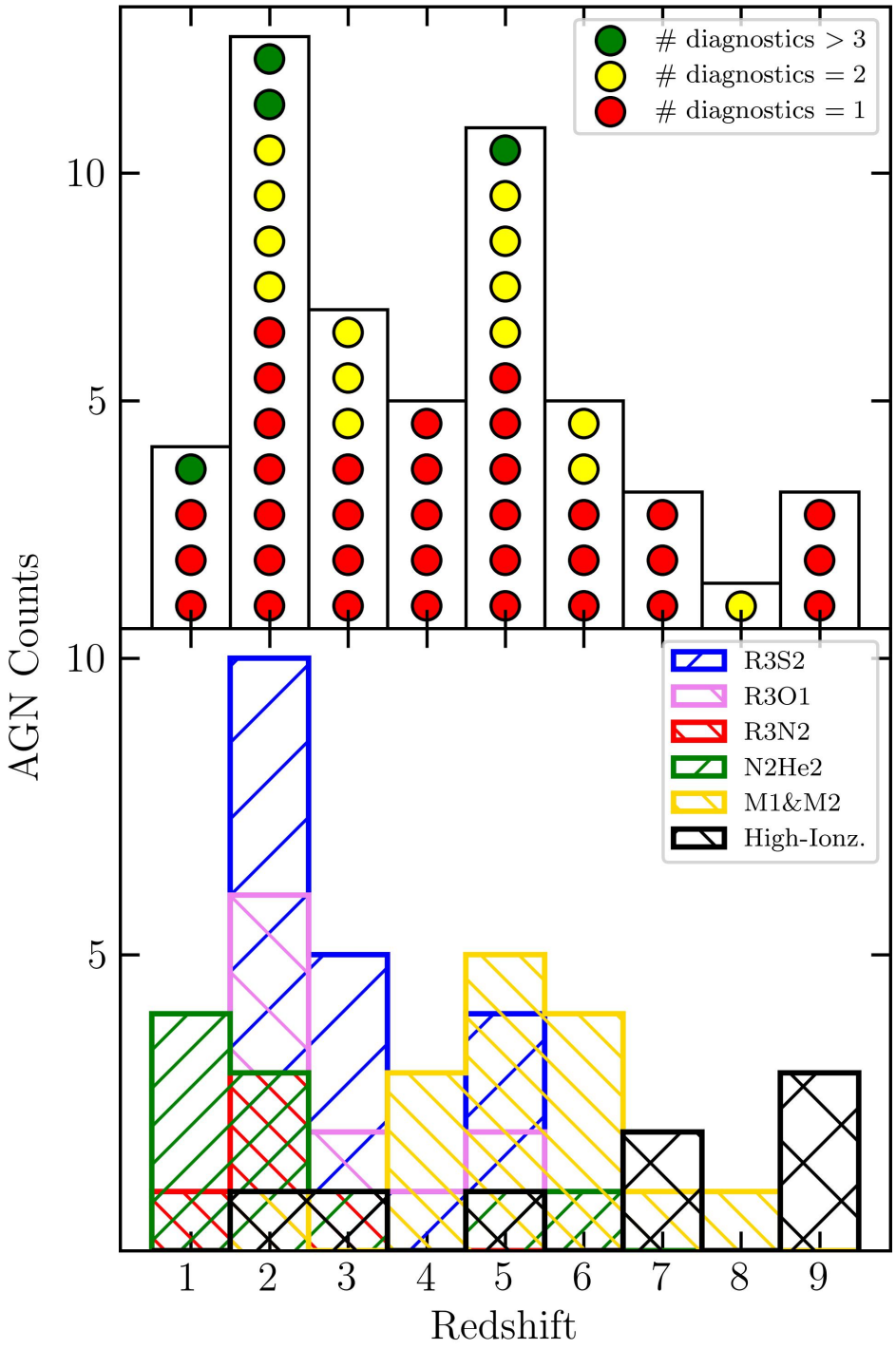

FIGURE 5.8: Summurizing plots of the NLAGN selection. *Upper*: NLAGN counts in the different redshift bins, colorcoded by the number of diagnostics in which each NLAGN has been selected. *Lower*: Distribution of the NLAGN according to the different diagnostic diagrams presented in Sect. 5.3, as labeled.

environments (more similar to those at high-z) than the R3N2 diagnostic. The lower effectiveness of the R3N2 at low metallicities (hence at high redshift) can be a consequence of the nitrogen production channel (Henry et al., 2000) that can determine nitrogen abundance to scale about quadratically with metallicity (at $Z > 0.1\ Z_{\odot}$), hence the [NII]/Hα ratio drops strongly. On the contrary, Sulfur and oxygen aboundances, being directly produced by massive stars through α-processes, can be less dependent on metallicity (Dopita et al., 2013). We also notice that [N II] and [S II] get redshifted out from the reddest grating of the spectra at $z \sim 6.7$, while the [O I] gets redshifted out at $z \sim 7$.
In general, at $z < 3$ the most effective AGN selection techniques are the R3S2, N2He2, and R3N2, while the large majority of the AGN at $z > 3$ are selected based on the auroral line diagnostics (both M1 and M2), or based on the detection of high-ionization emission lines. This supports the picture in which the traditional AGN diagnostic diagrams (in particular the R3N2) become less effective at high-z, mainly due to metallicity-related effects, while the detection of high-ionization emission lines and the new [O III] λ4363-based AGN diagnostics prove to be effective in selecting AGN also in the early Universe, reinforcing the results pointed out by Maiolino et al. (2023b), Übler et al. (2023) and Scholtz et al. (2023b).
The spectra of the NLAGN selected at $z > 6$ (together with the main spectral features that lead to the NLAGN classification) are presented in Sect. 5.8.4 of the supplementary material of this chapter.

5.4.1 Selection of X-ray and radio sources in the diagnostics

In this section, we want to summarize the results related to the X-rays and radio sources involved in our analysis. Among the seven X-ray sources with a spectrum in the CEERS MR sample, two sources (CEERS-3050 and CEERS-3061) are at $z < 0.5$, and therefore, the rest-frame optical and UV lines are not available in JWST spectra. These sources were classified as SFG by the X-ray spectral analysis performed in Buchner et al. (2015). Among the other five X-ray sources, four are selected as NLAGN in at least one of the diagnostics. In particular, the one at highest redshift (CEERS-2808, $z = 3.384$) is selected as NLAGN only in the R3S2 diagnostic diagram and not in the R3N2 (considering the more conservative demarcation line reported in Scholtz et al., 2023b). The only three X-ray sources selected as NLAGN in the BPT diagram are CEERS-2919, CEERS-2904, and CEERS-2900, all at $z \lesssim 2$, the first two also showing a broad Hα emission line (see Sect. 5.8.1 of supplementary material). The only X-ray source not selected in any of the diagnostics is CEERS-2989, at $z = 1.433$, whose spectrum shows only clear Hα and [N II] detections, and therefore, we could not define the R3 line ratio to place it in the above-mentioned diagnostics. It is worth noting that the lack of [O III] λ5007 and Hβ detections in this source is probably due to high obscuration levels. Considering the analysis in Sect. 5.5.2, we found a lower limit to the value of $A_V \sim 6$ mag. The large AGN obscuration of this source is further supported by the Compton-thick obscuration level derived by the X-ray analysis performed in Buchner et al. (2015). However, this source shows a faint broad Hα component, as reported in Sect. 5.8.1 of the supplementary material of this chapter.
The two radio-detected sources are CEERS-2900 (that is also X-ray detected) and CEERS-3129, both selected in the R3N2 diagram as well as in the R3S2 one. They also both show strong [N II] emission, in particular CEERS-3129, probably indicative of a shock (Allen et al.,

2008; Nesvadba et al., 2017). Furthermore, CEERS-3129 shows a significant broad Hα component, but no indication of X-ray emission, as we will further explore in Sect. 5.6.1.

5.4.2 Comparison with previous CEERS AGN selections

Calabrò et al. (2023) attempted to select AGN among the CEERS sources with a MR spectrum using near-infrared emission line diagnostics. To do so, they restricted the redshift range of the sources at $1 < z < 3$, therefore considering only 65 single sources. Using the R3N2 and R3S2 diagnostic diagrams, the authors selected 8 NLAGN that were classified as NLAGN also based on at least one NIR diagnostic, i.e. CEERS-2919, CEERS-3129, CEERS-2904, CEERS-2754, CEERS-5106, CEERS-12286, CEERS-16406, and CEERS-17496. The first three were marked as BLAGN, as we also found in Sect. 5.8.1. Five of the sources selected as NLAGN by Calabrò et al. (2023) are also selected in our work, while the other three sources were not selected because of the more conservative demarcation lines used in the R3S2 and R3N2 diagnostics or because the [S II] line was not considered detected with enough significance by our fit. The authors also selected 5 NLAGN using near-infrared diagnostic diagrams that were classified as SFG based on the optical diagnostic diagrams, i.e. CEERS-2900, CEERS-8515, CEERS-8588, CEERS-8710, CEERS-9413. Among these, the only one that is classified as NLAGN also in our selection is CEERS-2900; the other three were selected as SFG also by our optical diagnostics (even if CEERS-8588 shows a tentative [Ne V] line emission, but we conservatively decided to mark it as a non-detection).
In Davis et al. (2023) the authors investigate the presence of extreme emission-line galaxies (EELG) at $4 < z < 9$ using JWST NIRCam photometry in the CEERS program. They used a method to photometrically identify EELGs with Hβ + [O III] or Hα emission of observed-frame equivalent width > 5000. Among the photometrically selected EELGs there are 39 sources with a NIRSpec PRISM or MR spectrum, 25 with a match in our MR parent sample. Among these, 6 are selected as NLAGN in our work, i.e., $\sim 25\%$, supporting the non-negligible AGN contamination in the photometric selection of this kind of sources, as already shown also by other works (Amorín et al., 2015).

5.4.3 AGN prevalence

In Fig. 5.9, we show the fraction of the selected NLAGN among the CEERS MR sample. In particular, we consider as parent sample all the galaxies with a MR spectrum and with a secure redshift (217 sources). Then we divided the distribution into four redshift bins of $\Delta z = 2$ between z=0 and z=9, and for each bin, we computed the fraction of AGN compared to the parent sample. The number of sources in the different redshift bins is 122, 85, 59 and 49, going from the lower to the higher redshift bin. We derived an almost constant fraction of AGN among the CEERS MR sources $\sim 20\%$. This fraction do not represent intrinsic AGN fractions at these redshift, given the fact the the original selection function of the spectroscopic survey is extremely hard to derive. However, this result agrees well with what was found for the NLAGN selection in the JADES survey by Scholtz et al. (2023b). Given the different sensitivities reached by the CEERS and JADES surveys, one would probably expect a lower fraction of NLAGN selected in the CEERS sample since, in most cases, the AGN selection is based on the detection of faint emission lines. At the same time, however,

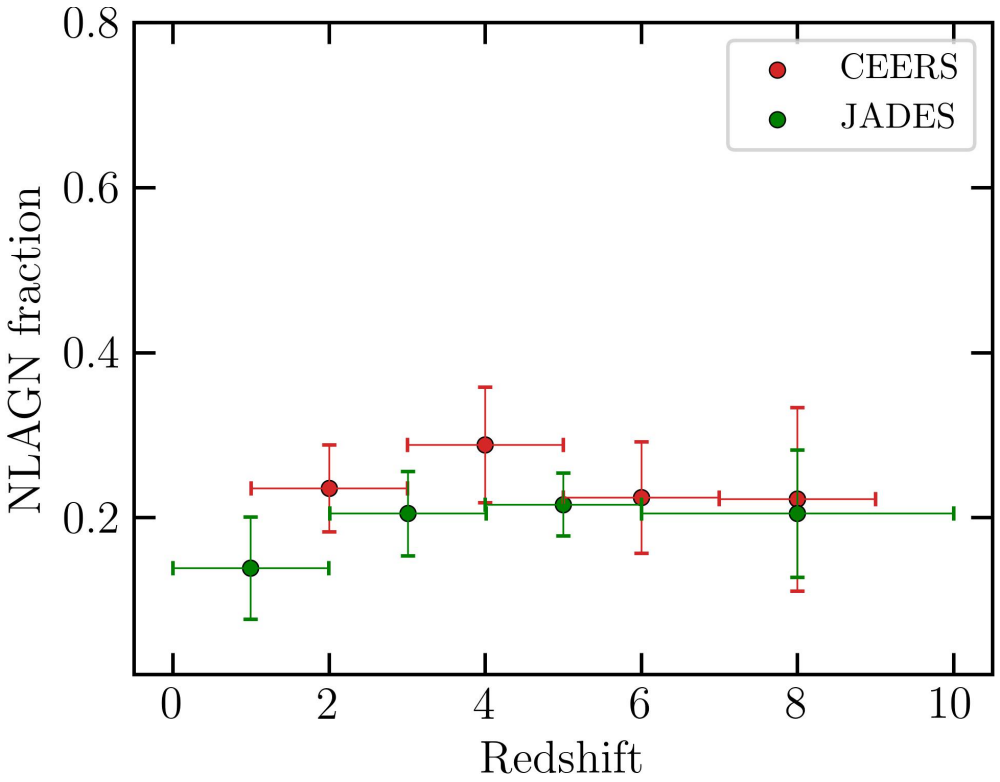

FIGURE 5.9: Fraction of the spectroscopically selected NLAGN with respect to the parent sample in this work (CEERS, in red) and in the JADES survey (green) (program ID 1210&3215 Scholtz et al., 2023b) in different redshift bins. The errors account for the statistical uncertainties and are computed considering a Poissonian noise.

as we are going to show in Sect. 5.5.3, the median bolometric luminosity of the NLAGN in the CEERS sample is $\sim$ 1 dex higher compared to those selected in JADES, and so the fact that the targets are brighter partially compensate for the shorter exposure times.

5.5 AGN and host galaxies properties from spectral and SED fitting

5.5.1 Velocity dispersion

In Fig. 5.10 we show the distribution of the intrinsic narrow line FWHM of the [O III] λ5007 emission line compared to the redshift of the sources. In particular, we considered only sources with the [O III] λ5007 line detected at S/N$>$ 5. To estimate the intrinsic FWHM, the observed FWHM retrieved by the fit was deconvolved by subtracting in quadrature the instrumental resolution of the grating at the observed wavelength of the [O III] λ5007 line. The instrumental velocity resolution was derived from the [point_source_lsf_f290lp_g395m_QD1_i185_j85.csv] line spread function file, which was calculated from the instrument model (Ferruit et al., 2022b), assuming a point-source geometry and a target located in the first MSA quadrant, at the center of shutter (i,j)=(185,85); this procedure is described in de Graaff et al. (2024). The

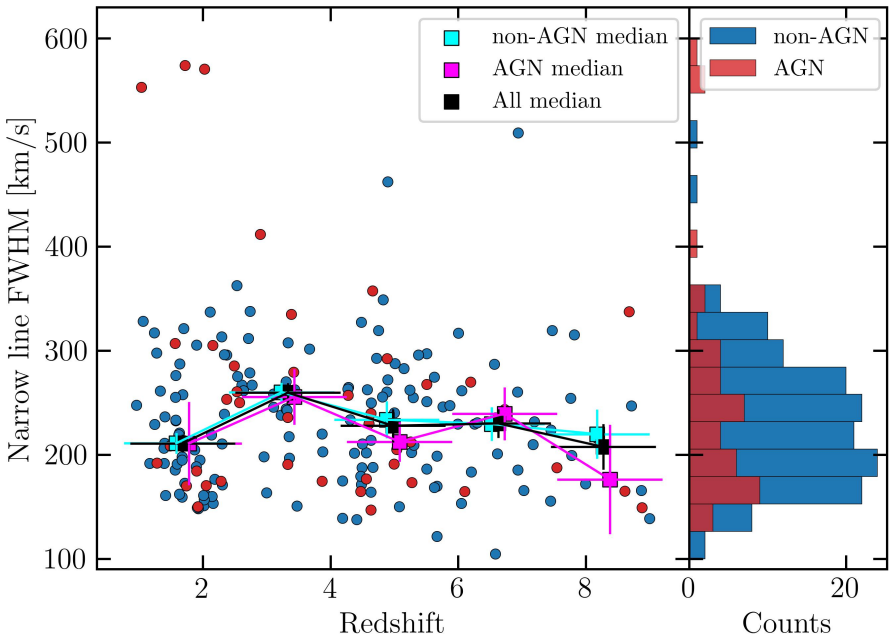

FIGURE 5.10: Redshift distribution of the intrinsic [O III] λ5007 line FWHM, resulting from the line-fitting procedure and subtracting the instrumental FWHM. On the left, we also report the median values of the FWHM of NLAGN (magenta), non-AGN (cyan) and of the global population (black) in five different redshift bins. Errors are derived using a bootstrap procedure. There is no significant trend of the FWHM of the sources with redshift. On the right, we report the histogram of the NL-FWHM distribution for AGN (red) and non-AGN (blue).

median value of the instrumental resolution is $\sim$ 160km/s.
The lower and upper boundaries of the intrinsic FWHM distribution are limited by the prior on the width of the line given to `Qubespec`, i.e. 200km/s $<$ FWHM $<$ 700km/s. None of the sources, after the subtraction of the instrumental resolution, shows a negative FWHM, but there is a non-negligible number of sources with an FWHM$<$ 200km/s, meaning that in these cases we are really detecting emission lines at the limits of JWST resolution. Indeed, the cut in S/N and a careful visual inspection confirm beyond any doubt that these lines are real, and, actually, such narrow FWHM are expected given the trend of decreasing host galaxy stellar mass with redshift and given the fact that with JWST high-resolution spectrosocpy were recovered FWHM up to $\sim$ 100km/s (Maiolino et al., 2023b). On the other hand, there are six sources with an intrinsic FWHM$>$ 400 km/s. Four out of these six were classified as AGN using the diagnostic diagrams discussed above; in particular, three of these NLAGN are also low-redshift BLAGN presented in Sect. 5.8.1. Among the analyzed spectra, we did not find any significant residual from the fit of the narrow [O III] λ5007 line that could be indicative of the presence of an outflow component.
In Fig. 5.10, we also report the median values of the FWHM of the NLAGN population and of the non-AGN in five equally spaced redshift bins. The distribution of AGN and non-AGN, considering the errors, do not differ significantly from the other in any of the redshift bins. We also perform the Kolmogorov-Smirnov (KS) test on the global marginal distributions in FWHM of the two populations (histogram on the right panel), finding a *p-value*= 0.62, which does not point towards different parent samples of the two distributions. This indicates that the NLAGN we spectroscopically selected among JWST spectra do not significantly impact their host-galaxy ISM, contrary to what was found in other studies on NLAGN samples, where, however, the NLAGN were selected based on other AGN activity tracers. For example, in the X-ray selected AGN sample of the KASHz survey, investigating sources at Cosmic Noon, (Harrison et al., 2016) found that $\sim$ 50% of the targets have ionized gas velocities indicative of gas dominated by outflows and/or highly turbulent material, with [O III] λ5007 FWHM$\geq$ 600km/s. We will further discuss the AGN impact on the sources studied in this work in the next sections.

5.5.2 Obscuration

In Fig. 5.11, we show the distribution of the A_V values obtained from the Balmer line decrement. In particular, to derive the A_V we considered the Small Magellanic Cloud (SMC) attenuation law (Gordon et al., 2003)[3] ($R_V = 2.74$) for sources at $z > 3$ and the Calzetti et al. (2000) attenuation law for sources at $z < 3$ (with $R_V = 3.1$). While the Calzetti et al. (2000) attenuation law has been extensively used in the low-z Universe to account for the effects of dust, the choice of the SMC attenuation law is more appropriate for the high-z Universe (Reddy et al., 2015; Shapley et al., 2023), where galaxies are smaller and more compact than in the local Universe (Ono et al., 2023). In particular, we select only those sources that have both Hα and Hβ detected. We assumed CASE B recombination, i.e. an intrinsic

[3]We used for the attenuation curve a fit to the average SMC-bar A_λ/A_V data points reported in Table 4 of Gordon et al. (2003). We found that the A_λ/A_V analytic expression reported in their Eq. 4-5 does not correctly fit their observed data points at $\lambda > 3030$ and has $A(\lambda = 5500)/A_V \neq 1$.

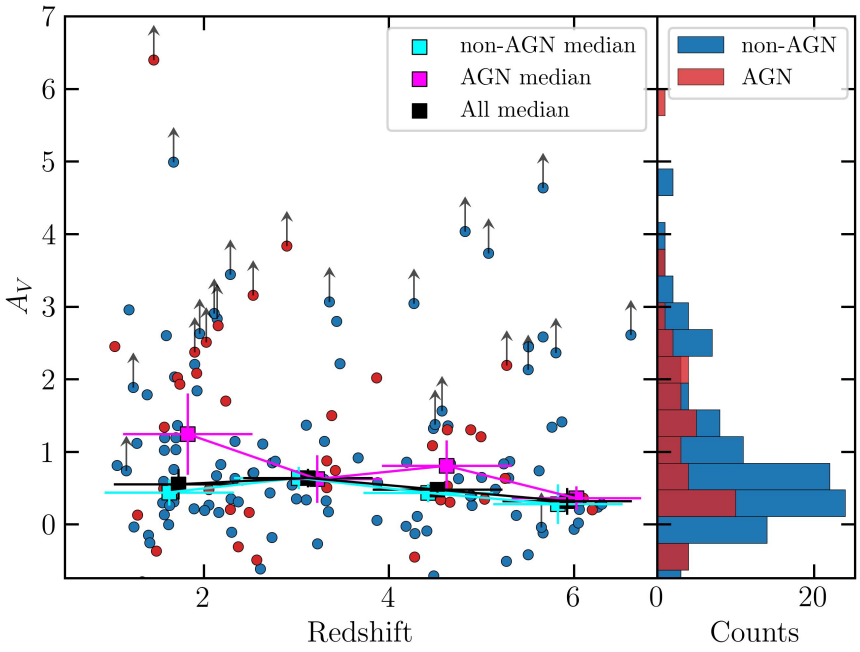

FIGURE 5.11: Redshift distribution of the values of A_V inferred from the Balmer decrement in sources with detected Hα and Hβ lines. For sources with upper limits on Hβ we derived lower limits on A_V. On the left, we also report the median values of A_V for NLAGN (magenta), non-AGN (cyan) and for the global population (black) in four different redshift bins. Errors are derived using the bootstrap procedure. On the right, we report the histogram of the A_V distribution for AGN (red) and non-AGN (blue).

Hα /Hβ ratio of 2.86[4]. We also considered sources with upper limits in Hβ to derive lower limits in A_V.
We note that a small number of galaxies in Fig. 5.11 scatter to either surprisingly high values of A_V or else negative values, meaning an Hα /Hβ line ratio lower than the dust-free minimum value of 2.86. Sources with a dust-free Hα /Hβ line ratio lower than CASE-B recombination have also been reported recently in the literature (Scarlata et al., 2024; Yanagisawa et al., 2024; McClymont et al., 2024), but in our case these sources all have Hα /Hβ line ratios compatible at 1-2σ with 2.86. However, there is also the possibility that a small number of these outliers is due to small systematics in the NIRSpec grating-to-grating flux calibration, i.e. when Hβ and Hα are measured in different gratings. A similar result was indeed also observed in Shapley et al. (2023) on the same sample analyzed here. Even if present, these calibration issues do not affect our NLAGN selection, because the line ratios of the diagnostics presented above involve lines close to each other, and so generally in the same grating.
In Fig. 5.11, we also plot the median value of the obscuration considering four different redshift bins and taking separately the two populations of NLAGN and sources not identified as AGN. We did not consider the upper limits in A_V for the median values. We did not find a significant evolution of A_V with redshift. We also note that the AGN population is characterized, in the lower redshift bin ($1 < z < 2.5$), by a $\sim$ 0.7mag higher obscuration. At $z < 2.5$ the average NLAGN obscuration is 1.15 mag compared to the 0.5 mag of non-AGN. On the contrary, when the whole NLAGN and non-AGN populations are considered, the average value of the obscuration across all redshifts is very similar, around 0.5 mag.

5.5.3 Bolometric luminosities

In this section, we investigate the bolometric luminosities of the NLAGN selected in this work. The way to compute the AGN bolometric luminosity is not unique, and it can be done starting from the X-ray luminosity, the luminosity of the BL emission or the UV continuum emission determined by the accretion disk. However, in our case, the majority of the sources are not X-ray detected, they do not have a BL emission, and their continuum is generally undetected. Therefore to estimate the bolometric luminosities, we rely on the dust-corrected narrow line fluxes of the Hβ line, using the calibrations reported in Netzer (2009). These fiducial L_{bol} for the sample of NLAGN (with available Hβ line) are reported with red squares in Fig. 5.12. In the same figure, we also report, with a fainter marker, the value of L_{bol} estimated from the same line fluxes but with the calibration used in Scholtz et al. (2023b) and taken from Hirschmann et al. (in prep). In this second case, the bolometric luminosity depends quadratically on the luminosity of the Hβ line, while the Netzer (2009) calibration is linear, determining a difference in the estimated L_{bol} that can go up to $\sim$ 1 dex (on average the bolometric luminosities computed using Hirschmann et al. calibrations are 0.8 dex lower). To verify the reliability of our fiducial values of L_{bol} we also considered the bolometric luminosity calibrations derived in Lamastra et al. (2009) where the NLAGN bolometric luminosities are derived from the dust corrected [O III] λ5007 line. The L_{bol} computed using the calibration in Netzer (2009) and those computed using the calibrations in Lamastra et al. (2009) agree well, with a median discrepancy of only $\sim$ 0.25 dex.
However, It is worth noting that all these calibrations assume that the lines used to derive

[4] $T_e = 10^4$, $n_e = 100\mathrm{cm}^{-3}$ (Osterbrock & Ferland, 2006)

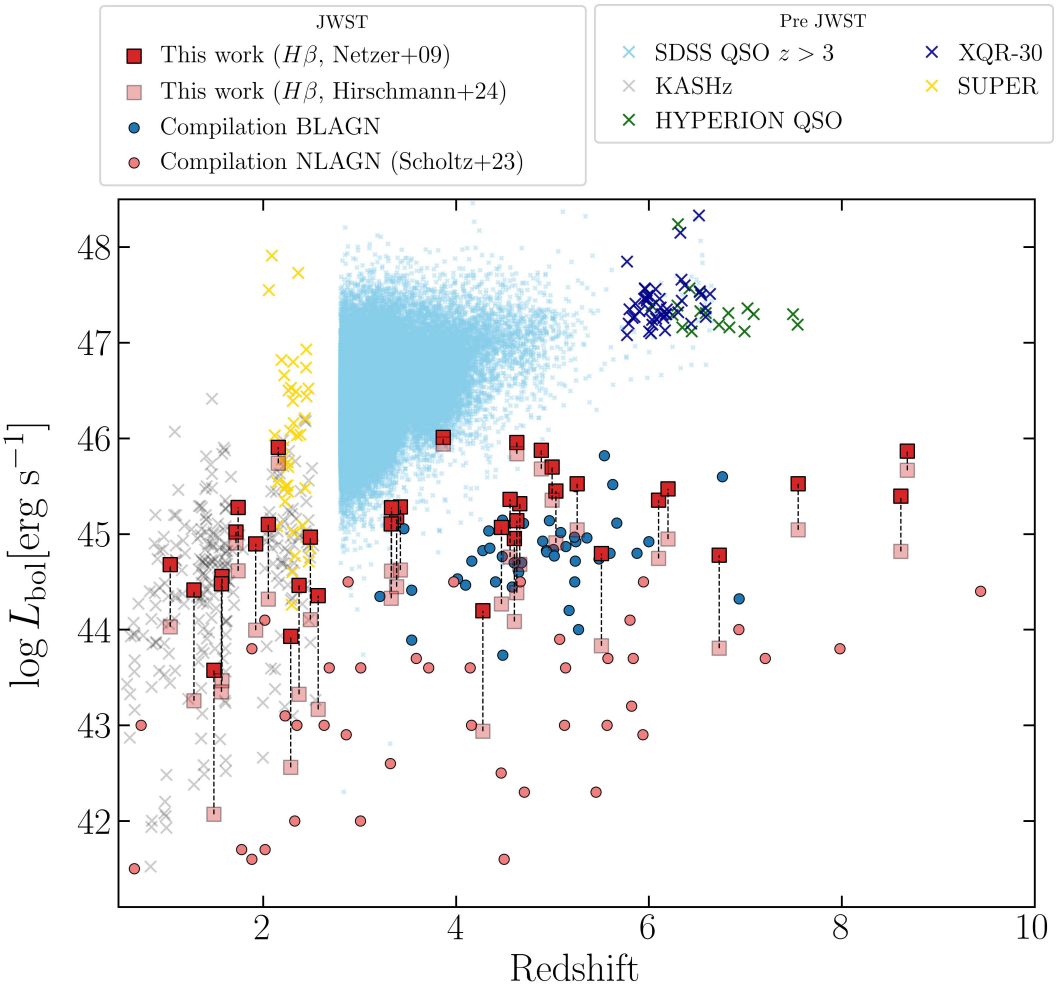

FIGURE 5.12: Bolometric luminosities versus redshift of the sample of NLAGN selected in this work (red squares) compared to the bolometric luminosities of other AGN samples selected using JWST spectroscopic observations or pre-JWST, as labeled. The darker red squares represent the baseline bolometric luminosities, derived using Netzer (2009) calibration, while the fainter red squares show the values of the bolometric luminosities obtained using the same calibration adopted in Scholtz et al. (2023b). The compilation of JWST selected BLAGN includes the sources taken from Maiolino et al. (2023b); Harikane et al. (2023); Matthee et al. (2023); Übler et al. (2023); Kocevski et al. (2023). SDSS BLAGN at $z > 3$ (light-blue crosses) are taken from Wu & Shen (2022). AGN from the KASHz and SUPER surveys at Cosmic Noon (grey and red gold crosses) are taken from Harrison et al. (2016) and Kakkad et al. (2020), respectively. QSOs samples at Epoch of Reionization are taken from Zappacosta et al. (2023) (green crosses, HYPERION sample), and Mazzucchelli et al. (2023) (dark-blue crosses, XQR-30).

L_{bol} are dominated by the AGN emission. This assumption is not necessarily true for the whole sample of NLAGN, and therefore, our L_{bol} should be generally taken as upper limits. For the BLAGN at $z \leq 2$ presented in Sect. 5.8.1, we computed the bolometric luminosities considering the same relation used in Harikane et al. (2023) to determine the bolometric luminosities of the BLAGN at $z > 4.5$ of this sample, i.e the bolometric luminosity calibration derived in Greene & Ho (2005) from the broad Hα emission.
In Fig. 5.12, we compare the distribution of L_{bol} of our sample with the AGN bolometric luminosities taken from the literature. In particular, we report AGN bolometric luminosities derived both from JWST spectroscopic studies and from pre-JWST surveys. Given the bolometric luminosities of $z > 3$ AGN detected before the advent of JWST, it is clear that we are now able to sample a completely new regime in the luminosity-redshift space. Comparing the bolometric luminosities of our sample with those of the NLAGN selected sample from Scholtz et al. (2023b), we note that our targets are on average ~ 1.5 dex more luminous considering our fiducial calibration, but still more luminous even considering the luminosities derived with the calibration of Hirshman et al. (in prep). On the opposite, the AGN luminosities of our sample are comparable, in the same redshift range, with the distribution of L_{bol} derived from multiple samples of BLAGN detected with JWST.

5.5.4 Host Galaxies Properties

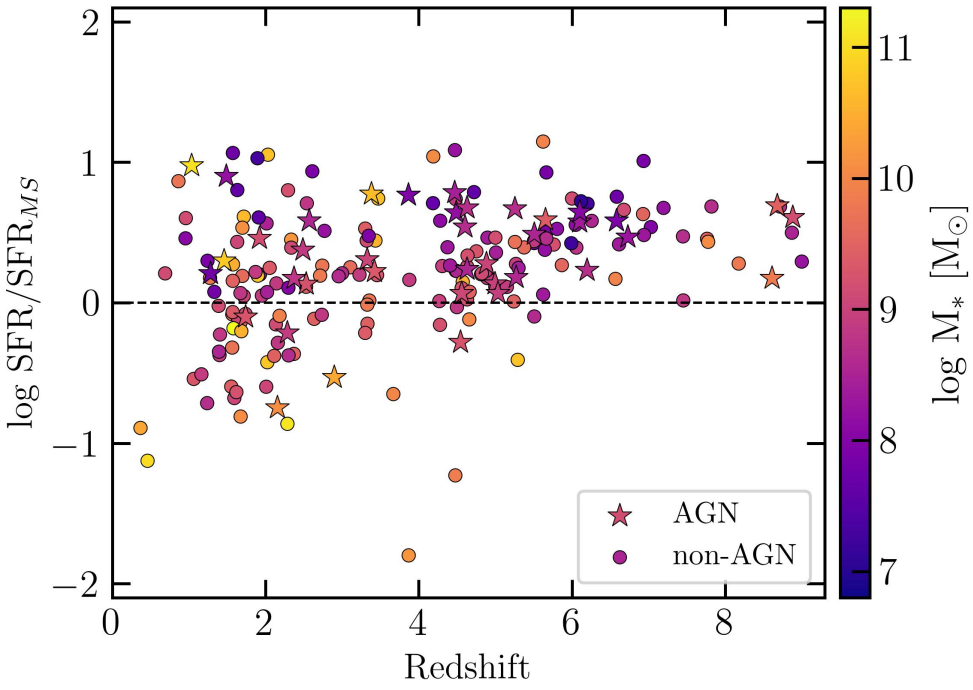

FIGURE 5.13: Redshift distribution of the ratio between the SFR derived from the SED-fitting and the SFR computed from the MS relation derived by Popesso et al. (2023) at the redshift and M_* of each source. We plot AGN and non-AGN sources with stars and circles, respectively. The sources are color-coded based on the stellar mass, as derived from the SED-fitting.

In Fig. 5.13, we show the results from the SED-fitting performed on the whole sample of 217 sources as described in Sect. 5.2.3. In particular, we compare the trend with redshift of the ratio between the SFR derived from the `CIGALE` SED fitting and the one obtained from the main sequence (MS) reported in Popesso et al. (2023) at the stellar mass (M_*, obtained from the fit) and redshift of each source. We note that sources at $z < 4$ are generally distributed almost symmetrically with respect to the MS, with the median value of $\log \mathrm{SFR}/\mathrm{SFR}_{\mathrm{MS}} \sim 0.1$. On the contrary, at $z > 4$, sources appear to be systematically above the corresponding MS, with a median value of $\log \mathrm{SFR}/\mathrm{SFR}_{\mathrm{MS}} \sim 0.4$, i.e. SFRs 2-3 times larger than the respective MS SFR. Sources with larger SFRs are also the sources with the lower M_*, some of them reaching $\log M_* \lesssim 8$. This means that the high-z galaxies analyzed in this sample seem to have, on average, a higher MS normalization, being less massive and more star-forming than expected from the MS. Different works, taking advantage of JWST spectroscopy, already showed that high-z galaxies are frequently in a state of intense or bursty SF (Dressler et al., 2023; Looser et al., 2023; Endsley et al., 2023), probably driven by a higher SF efficiency related to the lower metallicities of high-z galaxies. However, another possibility that can determine this apparent bursty behavior in our sample is a selection effect. Indeed, given that we are investigating a spectroscopic sample, we are probably more biased towards high SFR, in particular at high-z.

For sources not selected as AGN, we also compared the SFR inferred from `CIGALE` to the values obtained from the Hα line (using the measured dust-corrected Hα luminosity and the relations reported in Shapley et al., 2023). On average, we find a good agreement, with $\log \mathrm{SFR}_{\mathrm{H}\alpha} = 0.7 \log \mathrm{SFR}_{\mathrm{SED}} - 0.1$ with 0.25dex of scatter.

The AGN distribution in Fig. 5.13 is in agreement with the distribution of the other sources not selected as AGN: considering the median distribution of the ratios of the two SFRs, even in different redshift bin, AGN and SFG are indistinguishable, as already found by other works (Ramasawmy et al., 2019). The AGN quenching effect on the host galaxy SF is still a largely debated topic (Bugiani et al., 2024; Man & Belli, 2018; Beckmann et al., 2017; Scholtz et al., 2018), and in this case, we do not note a negative impact of the AGN feedback on SF in the host galaxy. This finding is also in line with recent studies based on various cosmological simulations, according to which star formation quenching is not primarily driven by the instantaneous AGN activity but the integrated black hole accretion (as traced by black hole mass Piotrowska et al., 2022b; Bluck et al., 2023b; Scholtz et al., 2024). The AGN contribution to the final SED is measured in `CIGALE` by the parameter f_{AGN}, which corresponds to the fraction of the AGN luminosity with respect to the galaxy one in the rest-frame 0.1-2 μm. As expected, given the NLAGN nature of our sources, for most of the NLAGN, the fit returned a $f_{AGN} < 0.5$ according to the scenario, already discussed in Sect. 5.5.1 and Sect. 5.5.2, that the AGN activity of the selected NLAGN is significantly buried by the host galaxy, in particular in the rest-frame optical part of the SED traced by the available photometry. This scenario, and the absence of a significant impact of the AGN activity over the host galaxy properties, is even more supported by the fact that we did not find relevant signatures of outflows in our spectroscopic analysis, in particular in the high-z galaxies.

From the NLAGN SED-fitting, we also derived the AGN bolometric luminosities (sum of the disk emission plus the one reprocessed by the tours). However, due to the limited photometric range (the reddest filter is F770W for sources with JWST photometry and IRAC4 for sources with 3D-HST photometry), in most of the cases, we were not able to explore the

rest-frame MIR NLAGN emission and the AGN bolometric luminosity is usually largely unconstrained. By selecting only those sources with a reliable bolometric luminosity from the SED-fitting, and comparing these values with the L_{bol} estimated in Sect. 5.5.3, we found that the median distributions agree well: the median value of the bolometric luminosity from the Hβ line is $\sim$0.23dex larger than the one derived from the SED-fitting, but the scatter goes up to $\sim$ 1 dex.

5.6 Average X-ray and radio properties

5.6.1 X-ray weakness

Four out of five of the X-ray sources classified as AGN by the X-ray analysis in Buchner et al. (2015) are also classified as AGN by our NLAGN diagnostics (see Sect. 5.4.1). However, all the other 48 selected NLAGN do not show any indication of X-ray emission. Therefore, we decided to perform an X-ray stacking analysis of the NLAGN not X-ray detected using the publicly available code CSTACK v4.5 Miyaji et al. (2008). For each target, CSTACK provides the net (background-subtracted) count rate in the soft (0.5–2 keV) and hard (2–8 keV) bands using all the observations from the AEGIS-XD survey and the associated exposure maps. Since the survey is a mosaic, multiple observations at different off-axis angles can cover the same position. Due to the variation of the Chandra PSF with the off-axis angle, for each observation of an object, CSTACK defines a circular source extraction region with the size determined by the 90% encircled counts fraction (ECF) radius (r_{90}) (with a minimum of 1″), to optimize the signal-to-noise ratio of the stacked signals. We set the background region for each source to be a 30×30 arcsec^2 area centered on the object, excluding the region around the object used to compute the net counts. Moreover, the code excludes the photon counts of the background circular regions around all the detected X-ray sources, with radii that depend on the net counts of the X-ray source. We derived the count rates in the soft band (0.5-2keV, SB), the band for which the detector is most sensitive and can, therefore, provide tighter constraints on the X-ray analysis. Then we converted the count rates into fluxes assuming an X-ray power spectrum with an intrinsic spectral index $\Gamma = 1.9$ and using the Chandra tool PIMMS[5].

In particular, we decided to divide the sample of NLAGN in two redshift bins, equally populated, using $z = 4.5$ as the dividing threshold. We also include in our analysis the BLAGN selected at $z > 4.5$ in CEERS by Harikane et al. (2023), since none of them show any indication of X-ray emission and to perform a more complete analysis on the AGN sample among the CEERS MR spectra. We did not stack the sample of low-z BLAGN presented in Sect. 5.8.1 because 3 out of 4 of these sources are X-ray detected (except CEERS-3129, which will be discussed later). The median redshifts of the three AGN samples are 2.37 for the low-redshift NLAGN, 5.50 for the high-z NLAGN, and 5.43 for the high-z BLAGN sample. In Fig. 5.14, we show the stacking maps in the SB and HB for the two NLAGN redshift bins and for the BLAGN, respectively. Given the median redshift of the samples, and considering the SB as the reference observing band since it provides the deepest constraint, we are investigating, on average, the following rest-frame X-ray energy bands: $\sim$1.7-6.5 keV for the low-z

[5] https://cxc.harvard.edu/toolkit/pimms.jsp

NLAGN sample and $\sim 3.5 - 13$keV for the other two samples. The stacking analysis did not reveal any X-ray detection in any of the sub-samples. This is quite surprising given the rest-frame energy range tested and the depth of the X-ray observations. To check whether the non-detections in the stacking of the NLAGN were due to possible galaxy contamination in the NLAGN sample, we also performed the same stacking but considering only NLAGN selected in more than two diagnostics diagrams, but again, the X-ray analysis did not reveal any detection.

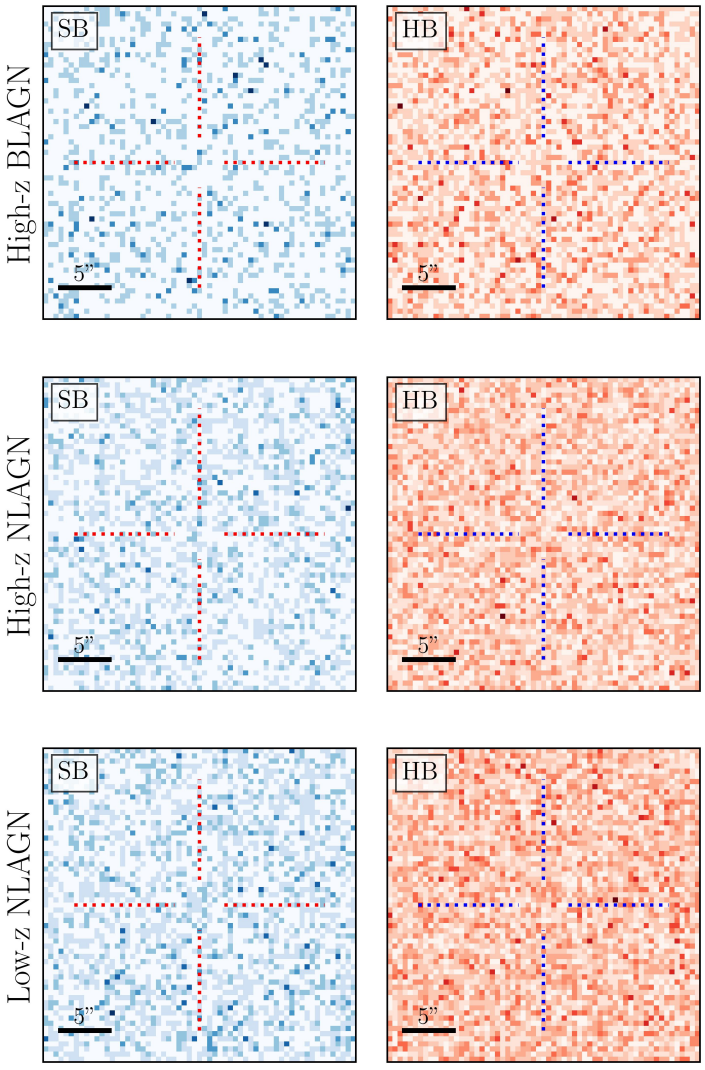

FIGURE 5.14: 30″×30″ cutouts of the X-ray stacking image in the observed SB (0.5-2keV) and HB (2-8keV) of the three samples considered in this work. From the top to the bottom: BLAGN, NLAGN at $z > 4.5$, NLAGN at $z < 4.5$. None of the samples show a detection in either bands.

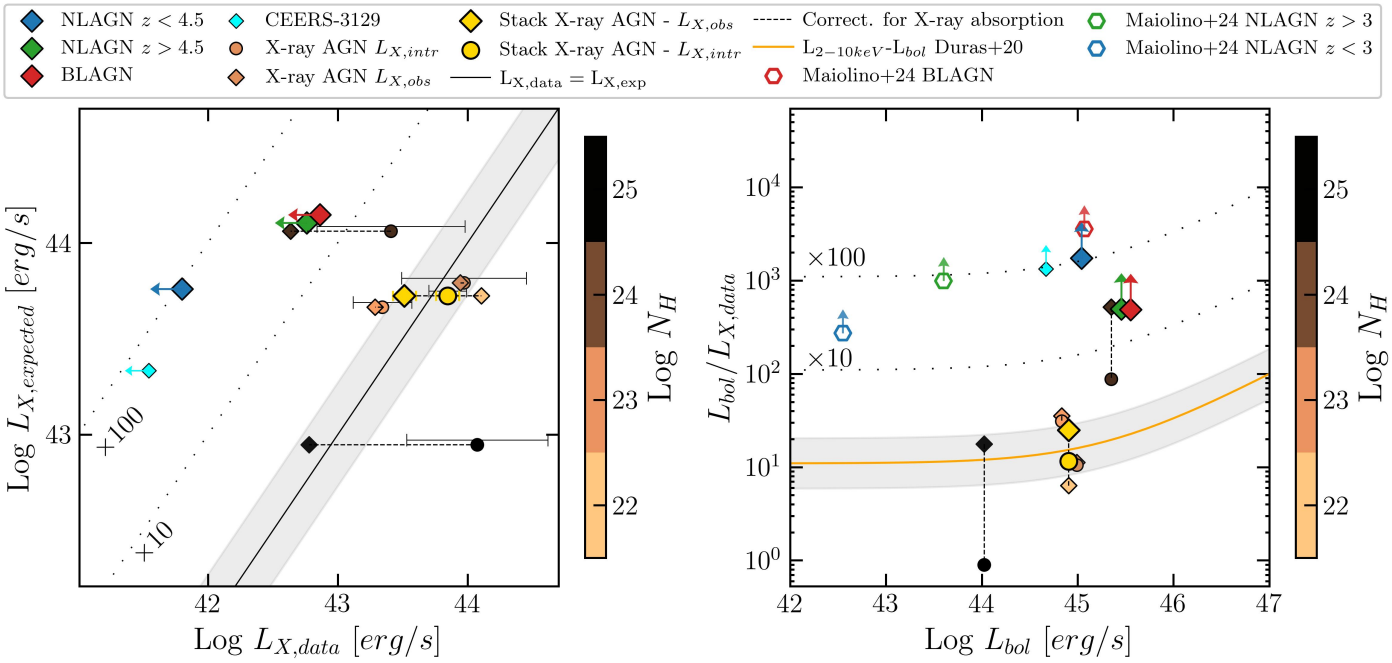

FIGURE 5.15: *Left:* expected X-ray luminosity versus observed X-ray luminosity (in the 2-10keV band) of the three AGN samples we used for the stacking, as labeled. We also plot the sources in the CEERS MR sample with an X-ray detection color-coded by their obscuration level, as derived from the X-ray spectral analysis (Buchner et al., 2015). Diamonds refer to observed X-ray luminosities, while circles refer to the intrinsic X-ray luminosities derived from the X-ray analysis. Errors bars mark the error on the intrinsic X-ray luminosities of the sources reported in Buchner et al. (2015). The result from the X-ray stacking of these X-ray detected sources is reported as a gold diamond, while the gold circle considers the median of the intrinsic X-ray luminosities of the X-ray sources. The dashed lines correspond to the shift from the observed to intrinsic X-ray luminosities. In cyan, we report the $z = 1.037$ BLAGN CEERS-3129, presented in Sect. 5.8.1. *Right:* ratio between the mean bolometric luminosity of the three AGN samples (derived using the values computed in Sect. 5.5.3) and the observed X-ray luminosity upper limits versus the mean bolometric luminosities. The orange line shows the reference $L_{2-10keV} - L_{bol}$ relation derived by Duras et al. (2020) and used to derive the expected X-ray luminosity of the sources. The empty hexagons represent the lower limits derived from the X-ray stack of the JWST selected BLAGN (red) and NLAGN (green for $z > 3$ and blue for $z < 3$) on the CDFS and CDFN (Maiolino et al., 2024a).

Therefore, we wanted to test if such X-ray non-detections would be compatible or not with the expected X-ray luminosities of these objects.
We used the $L_{bol} - L_{2-10keV}$ luminosity relations derived in Duras et al. (2020) to compute the expected X-ray luminosities given the baseline dust-corrected bolometric luminosities computed in Sect. 5.5.3 for our sample of NLAGN. For the BLAGN, instead, we used the bolometric luminosities computed in Harikane et al. (2023) using the Hα broad emission. The comparison between the expected X-ray luminosities and the upper limits derived from the stacking is shown in Fig. 5.15. For all the AGN, the upper limits in their observed X-ray luminosities are largely below the expected X-ray luminosities, posing severe constraints on their X-ray production and emission. The X-ray deficit is $\sim$ 1.3 dex for BLAGN and for the sample of NLAGN at $z > 4.5$, and is $\sim$ 2 dex for the NLAGN at $z < 4.5$. The X-ray deficit of this last sample is particularly surprising, given the fact that most of these low-z NLAGN were selected using the traditional BPT and VOT diagnostic diagrams, which are not expected to fail the NLAGN selection at low-z.
We further check that even taking the bolometric luminosities derived using the calibration of Hirschmann et al. in prep. (shown as faint red squares in Fig. 5.12) the results remain almost unchanged, with the median expected X-ray luminosities of the two NLAGN samples decreasing only of $\sim$ 0.2 dex.
In Fig. 5.15 we also show the distribution of the X-ray detected sources in our MR spectroscopic sample together with the position of their X-ray stack. In particular, for each X-ray source, we plot with a diamond the observed X-ray luminosity derived from their X-ray flux assuming $\Gamma = 1.9$, namely applying the same procedure as was done for the stack of the other sources. As expected, X-ray sources whose spectral analysis revealed larger column densities are on the left of the 1:1 relation, with a deviation up to $\sim$ 2 dex for CEER-2989, whose spectral analysis revealed a column density of $\log(N_H/\mathrm{cm}^{-2}) \sim 25$. This is because the spectral index $\Gamma = 1.9$, does not properly account for the harder spectrum of these heavily obscured sources. For each X-ray source, we also plot with circles the intrinsic, absorption corrected, X-ray luminosity derived in Buchner et al. (2015), which generally reduces the deviation from the 1:1 relation, especially if we take into account the large uncertainties on the intrinsic X-ray luminosities of the most obscured X-ray AGN. The same applies to the position of the stack of these X-ray sources: when we considered the observed X-ray luminosity (derived from the X-ray stacking) the point slightly deviates from the median expected X-ray luminosity, but if we correct for the absorption the X-ray luminosities (considering the median of the intrinsic X-ray luminosities on the x-axis), then the golden circle perfectly aligns with the median expected X-ray luminosity of these sources.
The same conclusions about the different samples can also be drawn by looking at the right panel in Fig. 5.15 where we compare the position of the sources with respect to the X-ray-bolometric luminosity relation of Duras et al. (2020). In the plot on the right we also show the lower limits derived in Maiolino et al. (2024a) from the X-ray stacking analysis performed on JWST selected BLAGN and NLAGN on the CDFS and Chandra Deep Field North (CDFN, Xue et al., 2016).
In both figures, we also show in cyan the position of the peculiar source CEERS-3129. This is a BLAGN at $z = 1.037$ (whose broad Hα component is shown in Sect. 5.8.1), it is not detected in the X-ray image and shows an X-ray deficit of $\sim$ 2 dex. This source is also characterized by a consistent dust obscuration: from the Balmer decrement derived in Sect. 5.5.2 we found

$A_V = 2.27$. Since the broad Hβ component is not detected and its upper limit would return an extremely large A_V, we assumed for the dust correction of the broad Hα emission (and therefore for the bolometric luminosity) the value of A_V derived from the NLR. However, this implies that the bolometric luminosity of CEERS-3129 might be underestimated and, consequently, also its expected X-ray luminosity and the amount of its X-ray weakness. Considering the dust-corrected bolometric luminosity, the lower limit to the X-ray weakness of CEERS-3129 is $\sim$2 dex.
Assuming that these sources are really AGN, the two possible scenarios that can justify such a lack of X-ray emission are:

- Obscuration coming from large column densities, larger than the CTK limit ($\log(N_H/\mathrm{cm}^{-2}) > 24$).
- An intrinsic X-ray weakness due to a possible different accretion disk/corona structure. This can determine, for example, a larger ratio between optical and X-ray fluxes due to a lower efficiency of the corona in producing X-ray photons.
- A combination of the two possibilities above

Yue et al. (2024) analyzing a sample of so-called 'Little Red Dots' (LRD) showing clear AGN signatures (as the detection of a broad Hα component) and selected on various fields covered by JWST imaging and spectroscopy, found a similar result. The amount of the X-ray weakness of the LRD is comparable to what we found in this work for the sample of BLAGN at high-z, while they did not investigate the population of NLAGN. In Yue et al. (2024), the authors suggest an intrinsic X-ray weakness as a possible origin of this discrepancy or a wrong AGN classification of these sources; in the latter case, the broad Hα component might be due to fast galactic outflows (with velocities $\sim$ 1000km/s). In our case, however, we did not detect any signature of any outflow component in the forbidden lines of the BLAGN. More evidence and strong arguments against the outflow scenario have been presented by Juodžbalis et al. (2024b).
Also Maiolino et al. (2024a), considering a large sample of BLAGN and NLAGN spectroscopically selected in the JADES survey over the Chandra Deep Field South (CDFS, with 7Ms of Chandra observations Liu et al., 2017) show that almost all the AGN discovered by JWST lack any X-ray detection, even in the stack. In this work, the authors investigated different possible origins of this X-ray weakness, focusing in particular on the three possibilities reported above and finding indications in support of one or the other. In particular, they found an enhanced EW(H$\alpha_{\,broad}$) in these sources that can suggest the Broad-line region to be distributed with a large covering factor around the central source acting as a dense dust-free absorbing medium, that can potentially reach Compton-thick column densities. On the other hand, the authors also find that many of the AGN newly discovered by JWST have features in common with the population of Narrow-line Seyfert 1 (NLS1), which are known to have a steep X-ray spectrum (and therefore more shifted towards a softer X-ray emission) and/or are characterized by high accretion rates, that are expected to result into a weaker X-ray emission, especially at high redshift, where higher rest frame energies are probed with Chandra.

5.6.2 Radio stacking to discriminate the nature of X-ray weakness

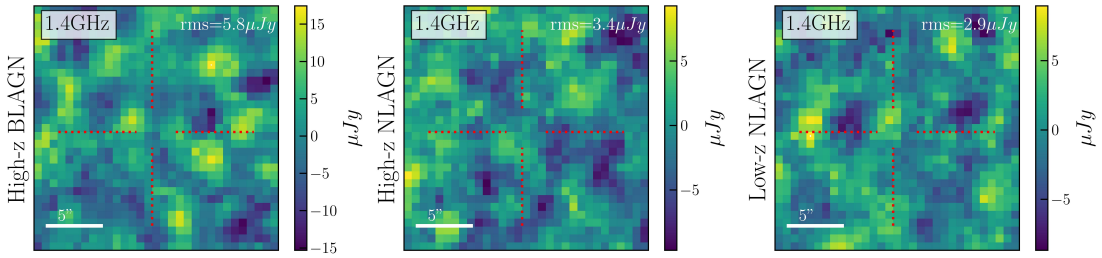

FIGURE 5.16: Median radio stacking of the three AGN samples considered in this work. From the top to the bottom we report 25″×25″ cutouts of BLAGN, NLAGN at $z > 4.5$, NLAGN at $z < 4.5$. None of the samples show radio detection. For each stack, we report the noise (rms) value in the top right corner

To attempt to discriminate between the heavily obscured versus intrinsic X-ray weakness scenario to explain the observed X-ray weakness of these AGNs, we took into account the 1.4GHz radio image of the EGS field (Ivison et al., 2007). Different works showed that the physical processes and the magnetic fields at the origin of the hot corona, where the inverse Compton scattering produces the AGN X-ray emission, are also responsible for (at least part of) the synchrotron radiation at the origin of the AGN-related radio emission (Laor & Behar, 2008; Panessa et al., 2019). The connection between these two emissions is also at the basis of the so-called fundamental plane relations (Merloni et al., 2003) or of X-ray - radio luminosity relations reported in different works in the literature (D'Amato et al., 2022; Panessa et al., 2019; Fan & Bai, 2016). Radio emission is generally largely unaffected by obscuration and can pass freely even from the most obscured environments, contrary to the X-rays that are almost completely absorbed at CTK hydrogen column densities.
However, some works show that even the radio emission can be partially or completely absorbed by the BLR clouds if their uniform distribution is also dense enough to determine free-free absorption (Juodžbalis et al., 2024b, Mazzolari et al. in prep.). However, it is still not clear whether the X-ray and radio emissions originate and are distributed on the same scale, and, therefore, if the effect of obscuration from BLR clouds can suppress both or not (Paul et al., 2024). For example, in Laor & Behar (2008) the authors found that the smallest possible scale of radio emission ($\nu \sim 1 - 5\ GHz$) in radio-quiet (RQ) AGN, e.g., coming from a compact radio corona or jet base, should originate from a region $\geq$ 100 times the extent of the X-ray emitting core. In this case, the radio emission would probably originate (or be being mainly distributed) outside of the BLR without being absorbed. Therefore in the following computations we assume that there is not a free-free absorption effect on the radio emission of the NLAGN and BLAGN investigated in this work, and we leave a more detailed discussion to Mazzolari et al. in prep.
Consequently, if these AGN are normal AGN (in terms of bolometric to intrinsic X-ray luminosity), and heavily obscured (but without absorbing also the radio emission), we might expect not to detect their X-ray emission, but we would expect to detect their radio emission at the level returned by the intrinsic X-ray to radio luminosity relations derived in the literature. On the contrary, if the radio image is deep enough to detect the expected AGN

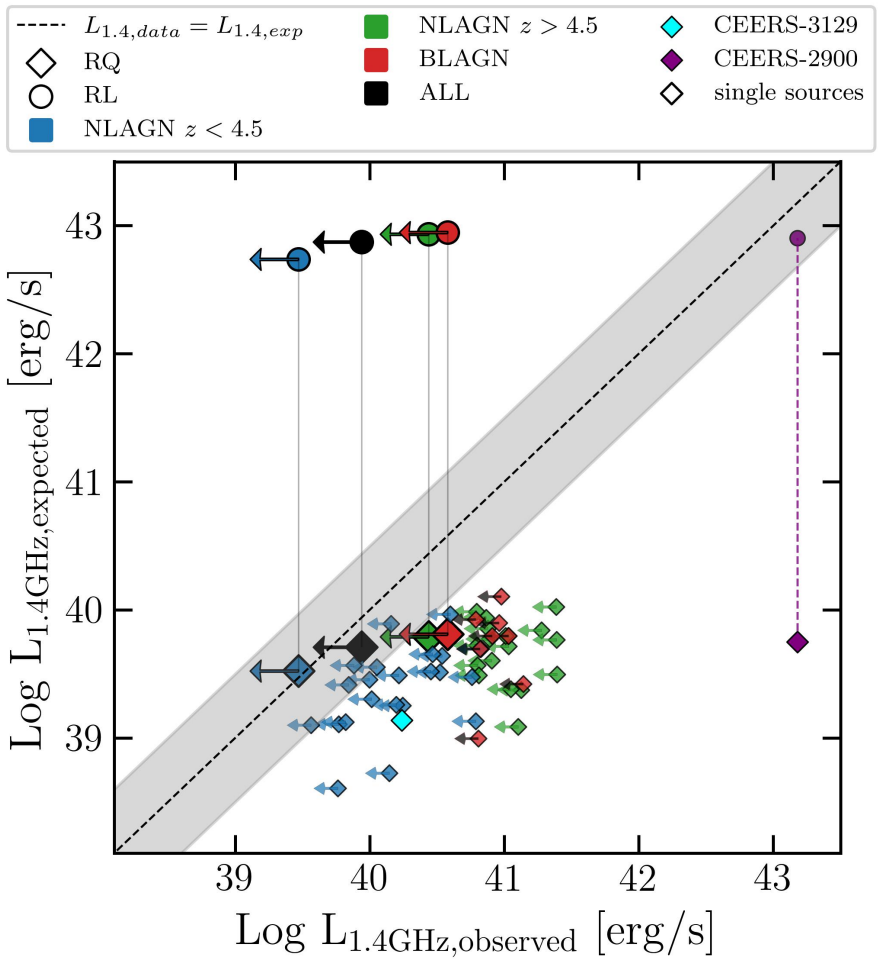

FIGURE 5.17: Expected 1.4GHz radio luminosity versus observed radio luminosity from the stack of the three AGN samples, as labeled. The diamonds refer to the assumption that AGN are RQ, while circles to the hypothesis that sources are RL. We also show the positions of the single sources according to the color of the respective sample. With a cyan diamond, we report the $z = 1.037$ BLAGN CEERS-3129 (presented in Sect. 5.8.1). The purple diamond represents the position of CEERS-2900 with the expected $L_{1.4GHz}$ derived using the RQ AGN relation, while the purple circle marks the expected $L_{1.4GHz}$ derived using the RL AGN relation. The gray shaded region represents the dispersion of the $L_{1.4GHz} - L_{2-10keV}$ luminosity relation derived for radio-quiet AGN by D'Amato et al. (2022).

radio emission, but none is detected, we would prefer a scenario in which both the radio and X-ray emissions are intrinsically lower with respect to what is observed in normal AGN at lower redshifts.
The only two sources with a radio detection among the CEERS MR spectra are CEERS-2900 (that is also X-ray detected) and CEERS-3129, both identified as NLAGN in our selection (see Sect. 5.4.1). Therefore, we perform a detailed stacking analysis at the position of the other NLAGN and BLAGN, considering the same three samples as for the X-ray stacking.
We applied a pixel-by-pixel median stacking, taking 30×30 arcsec^2 cutouts around the sources we wanted to stack. This method was proved to be more robust (compared to the mean stacking, for example) towards systematic effects (i.e. confusion) and also less affected by possible high-flux density contaminants (White et al., 2007). Since the CEERS survey covers only a very small region of the radio image on which the rms is almost uniform, we do not weigh each radio cutout for the corresponding noise value in the stacking procedure. We also estimated the final rms of the stacking images using a median absolute deviation procedure, following Keller et al. (2024).

The radio stacking analysis didn't recover any 1.4GHz detection, but allowed us to derive upper limits on the median radio luminosities of these sources in the three samples. We compared these values with the expected AGN radio luminosities. Given the expected X-ray luminosities of the sources derived in Sect. 5.6.1, we can estimate the expected 1.4GHz luminosities using the $L_{1.4GHz} - L_{2-10keV}$ luminosity relation derived for radio-quiet (RQ) AGN by D'Amato et al. (2022). The assumption that the sources are RQ is reasonable given that the fraction of radio-loud (RL) AGN has been found to be $\sim 10\%$ of the radio AGN population by different works, and it is not observed to vary with redshift (Liu et al., 2021; Williams & Röttgering, 2015).
It is worth noting that the effectiveness of this test also depends on the fact that the AGN-related radio emission has to be dominant over the radio emission coming from star formation in the host galaxy. To check this, we considered the SFR derived for these AGNs from the SED fitting in Sect. 5.5.4, and we converted them into radio luminosities using the relations reported in Novak et al. (2017); Delvecchio et al. (2021). For all the three AGN samples the median 1.4GHz luminosities due to star formation are $\sim$1.5 dex lower than the AGN radio luminosity derived from the $L_{1.4GHz} - L_{2-10keV}$ luminosity relation of D'Amato et al. (2022).
The comparison between the observed and the expected 1.4GHz luminosities of the three AGN samples is presented in Fig. 5.17, together with the upper limits of the single sources in the three samples. Unfortunately, the AEGIS20 radio image is too shallow to give conclusive constraints on the nature of the X-ray weakness of these samples of AGN, being the upper limits of the stack still compatible with the expected radio luminosities. We also explore the possibility that the bulk of the sources are instead RL, and, therefore, we computed the expected radio luminosities of the three stacks considering the $L_{1.4GHz} - L_{2-10keV}$ luminosity relation of Fan & Bai (2016), derived for RL AGN. In this case, the upper limits on the 1.4GHz luminosities are different orders of magnitudes lower than expected, excluding the RL scenario, which was, however, already considered improbable implying all (or most of) the sources to be RL.
In Fig. 5.17, we also report the position of the two radio-detected sources, i.e., CEERS-2900

and CEERS-3129, at $z = 1.9$ and $z = 1.03$, respectively. The first has a 1.4GHz radio flux of 58mJy, which corresponds to an observed radio luminosity of $L_{1.4GHz} \sim 10^{43}$erg/s. If we assume that CEERS-2900 is a RQ AGN, the expected radio luminosity given its intrinsic 2-10keV X-ray luminosity reported in Buchner et al. (2015) (where the source is classified as a CTK AGN) is only $L_{1.4GHz,exp} = 6 \times 10^{39}$erg/s. This strongly points toward the RL nature of this source. Indeed, considering the RL $L_{1.4GHz}/\log L_{2-10keV}$ luminosity relation reported in Fan & Bai (2016), the observed and expected radio luminosities correspond almost perfectly.
CEERS-3129 has instead a 1.4GHz radio flux of 286μJy that translates into a radio luminosity $L_{1.4GHz} \sim 10^{40}$erg/s. The value of the expected $L_{1.4GHz}$, computed starting from the expected $L_{2-10keV}$ luminosity, is almost consistent with the hypothesis of CEERS-3129 being a RQ AGN, since the observed and expected radio luminosities are quite consistent (close to the scatter of the relation). This is even more true if we recall that the expected $L_{2-10kev}$, derived from the dust-corrected L_{bol} computed using the broad Hα emission, is probably slightly underestimated because the obscuration of the broad Hα component is probably larger than the one derived from the narrow-line Balmer decrement (see Sect. 5.6.1).
The fact that the expected and observed radio luminosities of CEERS-3129 almost correspond supports the extreme CTK scenario to explain the X-ray weakness of this source. However, this scenario is not free from problematics. For example, it requires the CTK envelope around the central SMBH absorbing the X-ray emission to be almost completely dust-free to allow the broad Hα emission to escape. A natural solution to this problem is that the BLR clouds themselves, with large column densities distributed with approximately a 4π covering factor, absorb the X-ray photons, as discussed and analyzed in Maiolino et al. (2024a).

5.7 Conclusions

In this work, we analyzed the public *JWST*/NIRSpec medium resolution (MR) spectra of the CEERS survey (PID.1340) to identify NLAGN using multiple emission-line diagnostic diagrams. In particular, we identify a secured spectroscopic redshift of 217/313 targets observed with MR setup. We performed spectral line fitting using `QubeSpec`, and fitting separately different complexes of emission lines, ranging from rest-UV to rest-optical. We considered 10 different emission line diagnostic diagrams to perform the NLAGN selection, using both demarcation lines taken from the literature (including the new one presented in Chapter 4) and also exploiting the distribution of SFG and AGN photoionization models taken from Gutkin et al. (2016) and Feltre et al. (2016). This process led to the identification of 52 NLAGN, whose IDs and main physical parameters are reported in Table 5.4 to allow the scientific community to perform further studies. The main results are the following:

- We confirm that the traditional AGN diagnostic diagrams, in particular the BPT (R3N2), are less effective in selecting AGN at $z > 3$, because high-z galaxies and AGN generally overlap in these diagnostics due to the different physical conditions of high-z sources. For the high-z regime, the most effective AGN diagnostic diagrams are those based on the [O III] λ4363 line (Mazzolari et al., 2024a) or on the high-ionization emission lines.
- We select 10 NLAGN at $z > 6$ (4 of them marked as candidates NLAGN), whose spectra are reported in Sect. 5.8.4.

- The NLAGN selection includes only one out of the eight BLAGN at $z > 4.5$ selected in Harikane et al. (2023), i.e. CEERS-1244, four out of five of the X-ray sources with a CEERS MR spectrum and classified as AGN by the X-ray spectral analysis in Buchner et al. (2015), and both the radio detected sources reported in the radio catalog of Ivison et al. (2007).

- The fraction of NLAGN selected among the CEERS spectra analyzed in this work is $\sim 20\%$, a fraction that is almost constant across different redshifts and agrees well with what was found by Scholtz et al. (2023b) for the JADES survey.

- We analyzed the distribution of [O III] λ5007 NL FWHM of the AGN and non-AGN samples without finding any significantly different distribution for the two populations, suggesting that the NLAGN identified in this work do not significantly impact the ISM of their host galaxies. This probably happens because, thanks to the sensitivity of JWST, we are sampling a different population of NLAGN without outflows compared to previous NLAGN samples. Furthermore, in many cases, and in particular for low-mass high-redshift sources, the FWHM are very close to the resolution limit of JWST.

- We also investigated the average obscuration level of the sources by computing the A_V values for those sources having both the Hα and Hβ lines available in their spectra. We did not find a significant evolution with a redshift of the value of A_V. We also note that the AGN population is characterized, in the lower redshift bin $1 < z < 2.5$, by a ~ 0.7 mag larger obscuration value, while in the other redshift bins the median value of obscuration of NLAGN and SFG is almost the same.

- We computed the bolometric luminosities of the sources using the dust-corrected Hβ fluxes and the scaling relation reported in Netzer (2009). The range of L_{bol} spanned by our NLAGN sample at $z > 3$ is much lower than the AGN bolometric luminosities reported in any pre-JWST samples, but $\sim 1-2$ dex larger than bolometric luminosities of the sample of JWST detected NLAGN reported in Scholtz et al. (2023b) selected in the JADES survey (where the observing times are significantly longer). We tested the robustness of the derived bolometric luminosities by comparing our baseline results with those derived using the calibrations reported in Lamastra et al. (2009) for the [O III] λ5007 line, finding a good agreement between the two.

- We performed a SED-fitting analysis using `CIGALE` to derive the main physical properties of the host galaxies of the sample NLAGN and non-AGN. We found that, on average, sources at $z > 4$ deviate from the main sequence of Popesso et al. (2023), moving toward higher values of SFR than what is predicted by the MS. The range of stellar masses extends down to $\sim 10^7 M_*$. We didn't find any significant difference in the SFR or M_* distribution of the NLAGN compared to those of the parent sample, suggesting that any potential AGN positive or negative effect on the host-galaxy SFR is not significantly taking place (see Fig. 5.13). This is further supported by the lack of significant outflow detections in their spectra.

- We investigate the X-ray properties of the selected NLAGN and of the BLAGN selected in Harikane et al. (2023). We found that none of the BLAGN is detected in the

X-ray image of the field and that only 4 of the 52 NLAGN are X-ray detected. We performed an X-ray stacking analysis on the AGN that are not X-ray detected, finding an X-ray deficit of ~ 1.3 dex for the high-z BLAGN and for the sample of NLAGN at $z > 4.5$, and ~ 2 dex for the NLAGN at $z < 4.5$ between their expected X-ray luminosities (given their bolometric luminosity) and what comes from the X-ray stacking (see Fig. 5.15).

- We investigated the origin of this X-ray weakness by studying the radio emission of these sources, which can help to disentangle the heavily obscured AGN scenario from the intrinsic X-ray weak one. Unfortunately, the radio observation covering the region of the CEERS survey is not deep enough to provide a conclusive answer (see Fig.5.17). However, this shows the key role that present and future deep radio continuum observations can have in understanding the problem of the X-ray weakness of high-z AGN.

- Interestingly, one of the low-redshift BLAGN, which has also been selected by the NLAGN selection, CEERS-3129, is undetected in the deep X-ray image (showing an X-ray deficit of ~ 2 dex), but it is detected in the radio image of the field, with a radio luminosity compatible with the RQ AGN nature. The radio detection of CEERS-3129 supports the CTK AGN scenario at the origin of the observed X-ray weakness.

JWST is allowing us to reveal the population of heavily obscured AGN up to very high redshift ($z > 8$) and, at the same time, to investigate the NLAGN population at low redshift with a spectroscopic sensitivity not explored by pre-JWST surveys. This work highlights the necessity of defining new NLAGN diagnostic diagrams and, in general, different and complementary AGN selection techniques suited for the identification at high redshift AGN. Future works will have to explore in more detail the actual number density and the demography of these objects in the early Universe, carefully accounting for the selection function of each spectroscopic campaign. We also stress that a multiwavelength approach, such as the one outlined in this work, is necessary to explore and potentially provide answers to the peculiar features characterizing high redshift AGN.

5.8 Supplementary material

5.8.1 Broad-line AGN at $z \lesssim 2$

Here we present the spectra and line fitting of the 4 broad line AGN at $z \leq 2$ selected in our work. The 4 sources are CEERS-2904 (z=2.024), CEERS-2919 (z=1.717), CEERS-3129 (z=1.037), and CEERS-2898 (z=1.433). Three of these sources are X-ray detected (CEERS-2904, CEERS-2919, CEERS-2989), while one is radio detected (CEERS-3129). These sources were first fitted in the range around the Hα line using a narrow-line only fit (as described in Sect. 5.2.2) and then adding a single broad component accounting for the emission of the BLR around the AGN, as presented in Fig. 5.18. In order to test the significance of the broad Hα component we used the Bayesian Information Criterion (BIC) parameter (Liddle, 2007), we found that the BLR model is strongly preferred to the narrow-only fit ($\Delta BIC = BIC_{H\alpha\ NL} - BIC_{H\alpha\ BL} > 40$ for all the sources). The ΔBIC values for the single fit are shown in the right part of each panel in Fig. 5.18.

In fitting the Hβ and [O III] λ5007 lines complex of these sources we didn't find any significant broad component in none of the two lines.

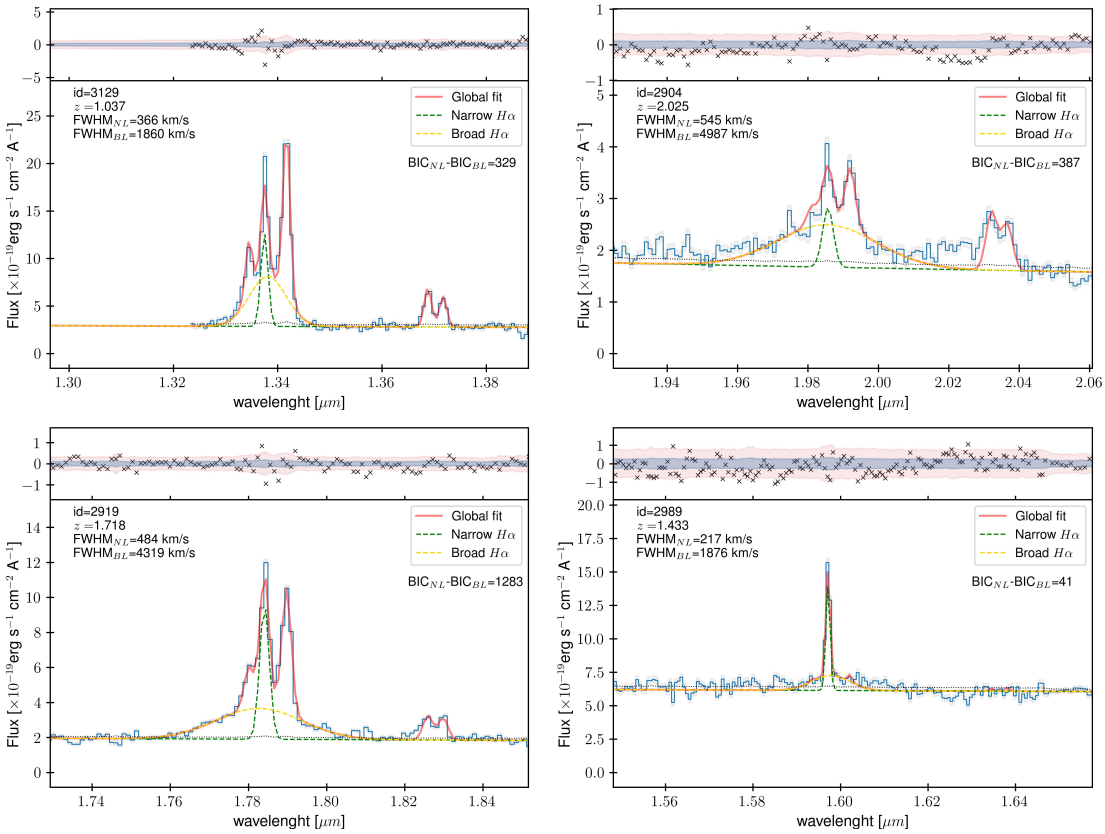

FIGURE 5.18: Fit of the Hα , [N II] , and [S II] complex including both a broad and a narrow components of the Hα emission line, as described in Sect. 5.2.2. The global fit is presented in red, while the narrow and broad Hα components are in green and yellow, respectively. In the upper panel are shown the residuals of the fit compared with the distribution of the 1σ (blue) and 3σ (red) errors on the fluxes. On the upper left part of the plots are reported the id, redshift and FWHMs derived from the fits. On the upper right we show the difference between the BIC value computed from the narrow-only line fit and the one from the fit with the additional broad Hα component

5.8.2 CIGALE SED fitting parameters grid

In Table 5.3 we report the grid of parameters used for the SED-fitting described in Sect. 5.2.3. Values of the parameters reported in italics refer to the SED-fitting of sources at $z > 3$.

5.8.3 Table of the selected Narrow-line AGN

TABLE 5.2: **CIGALE** parameters used for the SED fitting.

Parameter	Values	Description
SFH (sfhdelayedbq)		
τ_{main}	0.2, 0.3, 0.5, 0.7, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0 [Gyr]	e-folding time of the main stellar population
	0.15, 0.2, 0.3, 0.5, 0.7, 1.0, 1.5 [Gyr]	
age_{main}	0.5, 0.7, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0 [Gyr]	Age of the main stellar population in the galaxy
	0.2, 0.3, 0.5, 0.7, 1.0, 1.5 [Gyr]	
age_{bq}	0.01, 0.03, 0.05, 0.1 [Gyr]	Age of the burst/quench episode
r_{sfr}	0.1, 0.2, 0.5, 1, 2, 5	Ratio of the SFR after/before age_bq
	0.1, 0.5, 1, 2, 5, 10, 30, 50	
Stellar component (bc03)		
IMF	1	Initial mass function: 1 (Chabrier 2003)
Z	0.008, 0.02	Metallicity
	0.0004, 0.008, 0.02	
Separation age	10 [Myr]	Age of the separation between the young and the old star populations
Nebular emission (nebular)		
logU	-3.0, -2.0, -1.0	Ionisation parameter
	-2.0, -1.0	
zgas	0.008, 0.02	Gas metallicity
	0.0004, 0.008, 0.02	
ne	100	Electron density
	100, 1000	
Dust attenuation (dustatt_modified_cf00)		
A_V^{ISM}	0.01, 0.1, 0.3, 0.5, 0.7, 1.0, 1.5, 2.0, 3.0, 4.0	V-band attenuation in the interstellar medium
μ	0.44	Ratio of the BC-to-ISM attenuation
slope ISM	-0.7	Power law slope of the attenuation in the ISM
slope BC	-1.3	Power law slope of the attenuation in the BC
Dust emission (dale2014)		
fracAGN	0.0	AGN fraction
α	2.0	Slope of the dust emission
AGN component (skirtor2016)		
r_{ratio}	20.0	Ratio of the outer to inner radius of the dusty torus
τ	3.0	Equatorial optical depth at 9.7 μm
i	50, 60, 70, 80	Angle between equatorial axis and line of sight
f_{AGN}	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7	AGN fraction at 0.1-1.0 μm
EBV	0.04	E(B-V) for the extinction in the polar direction in magnitudes

TABLE 5.3: The first column reports the name of the templates as well as each individual parameter. In the second column, the parameters are listed, and in the third column, the descriptions of the parameters are provided. In cases where the grid is different for galaxies at $z > 3$, the values are reported in italics.

TABLE 5.4: List of NLAGN from CEERS MR spectra selected in this work. Columns: ID, redshift, diagnostic diagram(s) in which the source is selected, AGN bolometric luminosity (and the emission line used to derive it), ISM obscuration A_V, stellar mass, SFR, and any other additional notes. For bolometric luminosities derived from the Hβ line we used the Netzer (2009) calibration, while for those derived from the [O III] λ5007 we used the Lamastra et al. (2009) calibration. Low-z and high-z BLAGN have instead L_{bol} computed using the broad Hα lines and the calibrations of Greene & Ho (2005). When is reported a 'x' it means that it was not possible to compute that quantity for that source.

ID	Redshift	selection method	Log L_{bol} ergs s^{-1}	A_V mag	Log M_* $M_\odot$	SFR $M_\odot$ yr^{-1}	Notes
CEERS-23*	8.8802	N5	44.92 ([OIII])	x	9.17	36	
CEERS-355	6.0995	R3O1 He2N2	45.31 (Hβ)	0.32	8.37	5	
CEERS-428	6.1022	M2	x	x	8.06	2	
CEERS-496*	6.5809	N4	43.93 ([OIII])	x	7.93	1	
CEERS-613*	6.7288	C3He2 Ne5 N5	44.78 (Hβ)	x	8.28	3	
CEERS-669	5.2714	R3S2	45.22 ([OIII])	>1.92	8.59	3	
CEERS-689	7.5457	M1 M2	45.53 (Hβ)	x	x	x	
CEERS-1019*	8.6785	N4	45.87 (Hβ)	x	9.49	88	
CEERS-1029	8.61	Ne4	45.4 (Hβ)	x	9.86	58	
CEERS-1064	6.7905	M2	x	x	x	x	
CEERS-1173	4.9955	R3S2	45.55 (Hβ)	1.06	8.97	6	
CEERS-1212	4.2772	R3O1	44.44 (Hβ)	0	9.22	5	
CEERS-1244	4.4778	M2	45.28 (Hα)	x	9.86	1	BLAGN
CEERS-1400	4.4893	M2	x	0	8.19	3	
CEERS-1477	4.6307	M1 M2	45.8 (Hβ)	1.14	8.54	3	
CEERS-1536	5.0338	M1 M2	45.41 (Hβ)	0.3	8.93	5	

* Tentative selection

Table 5.5: Continuation of Table. 5.4

CEERS-1539	4.8838	R3S2 R3O1	45.71 (Hβ)	1.14	9.03	10	
CEERS-1561	6.1965	M1 M2	45.45 (Hβ)	0.18	8.68	4	
CEERS-1605	4.6305	R3S2 M1 M2	45.07 (Hβ)	0.46	8.69	11	
CEERS-1658	4.6035	He2N2	44.95 (Hβ)	x	8.65	7	
CEERS-1732	4.6598	R3O1	45.28 (Hβ)	0.27	10.02	33	
CEERS-1746	4.5602	M1 M2	45.32 (Hβ)	0.3	9.08	6	
CEERS-1767	4.5455	M2	x	x	9.28	4	
CEERS-1836*	4.4694	R3S2	44.93 (Hβ)	0.95	8.44	8	
CEERS-2168	5.6541	M2	x	x	9.4	51	
CEERS-2668*	2.8984	R3N2 R3S2	45.5 ([OIII])	>3.32	10.42	19	
CEERS-2754	2.238	He2N2	x	1.47	x	x	
CEERS-2808	3.3834	R3S2	44.99 (Hβ)	1.31	10.68	619	X-ray
CEERS-2900	1.9004	R3N2 R3S2	43.62 ([OIII])	>2.06	11.17	0	X-ray & Radio
CEERS-2904	2.0249	R3N2 R3S2 R3O1 He2N2 M1 M2	44.9 (Hα)	>2.17	10.78	26	X-ray & BLAGN
CEERS-2919	1.7177	R3N2 R3S2 R3O1 He2N2	44.84 (Hα)	1.75	10.62	203	X-ray & BLAGN
CEERS-3129	1.0373	R3N2 R3S2 He2N2	44.46 (Hα)	2.12	11.04	309	Radio & BLAGN
CEERS-3187	1.4896	He2N2	43.76 (Hβ)	0	8.23	3	
CEERS-3223	2.2862	R3S2	43.91 (Hβ)	0.18	8.9	1	
CEERS-3535*	2.4886	R3S2	44.95 (Hβ)	0.14	8.87	6	

* Tentative selection

Table 5.6: Continuation of Table. 5.4

CEERS-3585*	3.8662	M2	45.76 (Hβ)	1.77	7.97	2
CEERS-3838	1.9221	R3S2 R3O1	44.68 (Hβ)	1.8	9.16	11
CEERS-4210	5.2549	Ne4	45.45 (Hβ)	0.56	8.69	12
CEERS-4385	3.4219	He2N2	45.19 (Hβ)	0.65	9.33	15
CEERS-4402	3.3293	R3O1	45.21 (Hβ)	0.44	9.82	23
CEERS-5117	1.5728	R3S2 R3O1	44.41 (Hβ)	1.16	9.26	2
CEERS-6689	1.4608	He2N2	x	>5.53	10.82	95
CEERS-8041	2.5319	R3S2	44.96 ([OIII])	>2.73	9.24	8
CEERS-8110	1.5662	R3O1	44.43 (Hβ)	0.43	9.72	5
CEERS-8299	2.1567	Ne5	45.62 (Hβ)	2.37	10.1	5
CEERS-9156	2.0534	R3S2	45.05 (Hβ)	0.42	9.54	6
CEERS-9238	1.7415	R3S2	45.07 (Hβ)	1.67	9.29	3
CEERS-13661	3.3279	R3S2 Ne4	45.0 (Hβ)	0.76	9.03	9
CEERS-16551	2.5714	R3S2 R3O1	44.61 (Hβ)	0	8.57	5
CEERS-17496	2.3732	R3S2 R3O1	44.62 (Hβ)	0	8.79	3
CEERS-31075	1.2831	He2N2	44.4 (Hβ)	0.11	7.8	0
CEERS-44804	5.508	M1	44.79 (Hβ)	x	8.54	5

* Tentative selection

5.8.4 Spectra of the NLAGN at $z > 6$

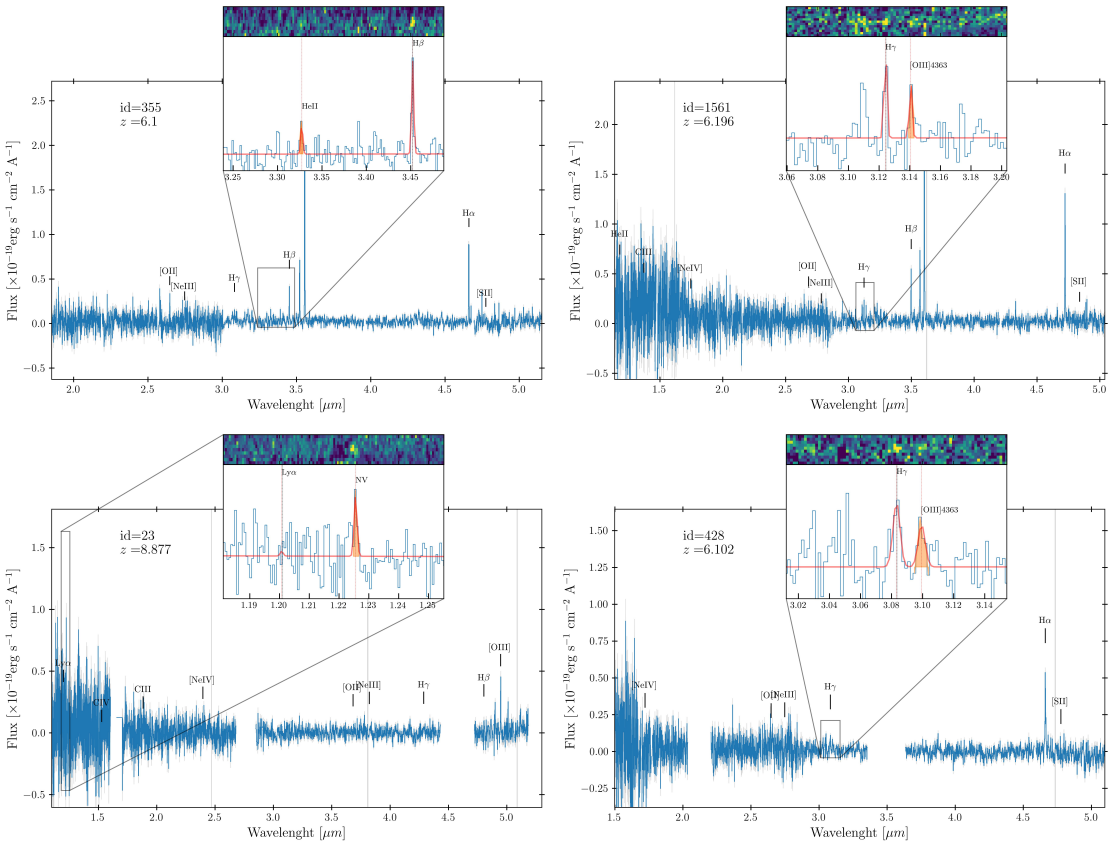

FIGURE 5.19: Spectra of the sources selected as NLAGN at $z > 6$. In each plot, we highlight in the box the main spectral features that led to the identification as a NLAGN. CEERS-1019, CEERS-23 and CEERS-496 are marked as tentative selections due to their dubious position in the diagnostic diagrams reported in Sect. 5.3

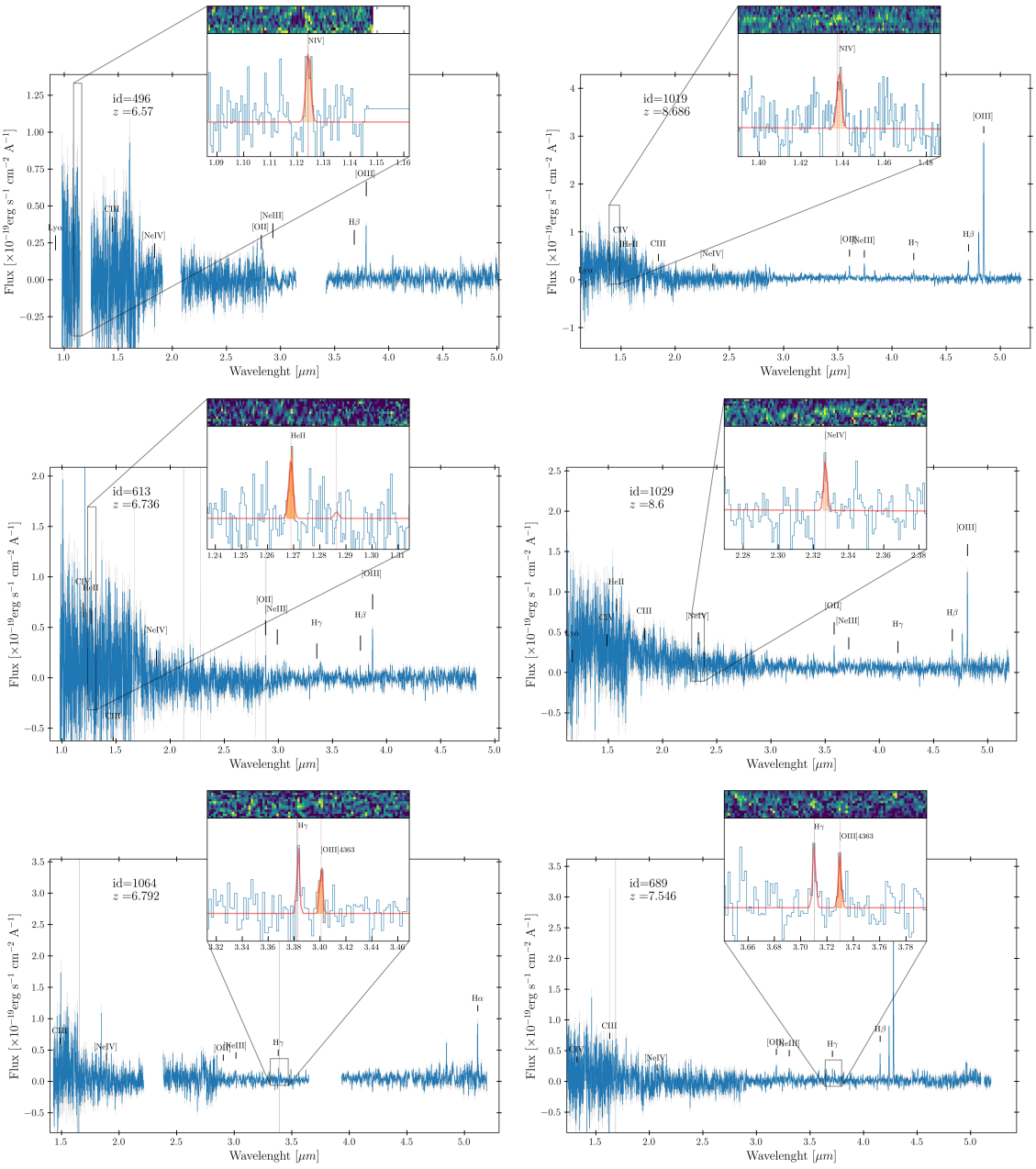

FIGURE 5.20: Continuation of Figure 5.19

Chapter 6

Concluding remarks

The existence of a large population of hidden AGN, in particular at high-redshift, poses several questions about their origin, their co-evolution with their host galaxy, and their impact on the evolution of the SMBH accretion across Cosmic Time. Having a proper census of the AGN population at all redshifts is crucial to test theories and models regarding SMBH seeding, to properly understand the origin of the SMBH-galaxy scaling relations observed in the local Universe, and to have a clearer picture of the properties of the accretion physics at different epochs.
In this PhD thesis, I presented a comprehensive exploration of many aspects characterizing the population of obscured AGN, focusing in particular on three: selection methods, demographic characterization, and their physical properties.
In these three years, I tried to pursue these lines of research, keeping a multiwavelength and comprehensive approach that I think is necessary to investigate these aspects and provide answers to the most compelling questions mentioned above. The advent of JWST marked a key step for the investigation of all these aspects up to the very high redshift $z > 6$, and the synergies between this and other present and future facilities will be crucial to have a better understanding of the demography and physical properties of this hidden population of SMBH.

In particular, in the chapters of this Thesis I presented and discussed the following main results:

- In Chapter 2 I demonstrated that a promising way to select distant obscured AGN is through radio emission, which is largely unaffected by obscuration and can be used as a proxy for nuclear activity. Translating into the radio band the population synthesis model of the Cosmic X-ray background, I developed an analytical model to predict the number of radio-detectable AGN considering different levels of obscuration, redshifts, and bolometric luminosities. By comparing these predictions with the results coming from X-ray forecasts, I showed that the surface density of CTK AGN radio detected is, on average, ~ 10 times larger than the X-ray one and even larger at $z > 3$. I also showed forecast for the number of AGN and CTK AGN that will be detected by future SKAO continuum observations, predicting thousands of AGN and CTK AGN that will be detected at $z > 6$. This work has been published in Mazzolari et al. (2024c).

- In Chapter 3, I tested the results presented in Chapter 2 from an observational point of view on the J1030 field. I performed a radio-excess selection to identify AGN candidates, focusing in particular on the radio-excess sources falling into the X-ray image footprint and without an X-ray detection. By performing a detailed X-ray stacking analysis we found solid indications that this population is dominated by heavily obscured (probably CTK) AGN. Finally, estimating the radio-excess selected CTK AGN number density at $z \sim 3$ we found that the radio number density is a factor $\sim 3-5$ larger than the one predicted by the CXB models, supporting the hypothesis that X-ray observation are probably missing part of the AGN population (the most obscured one) at $z > 3$. The results presented in this chapter will be part of two papers in preparation: Mazzolari et al. in prep and Mignoli, Mazzolari et al. in prep.

- In Chapter 4 I proposed three new NLAGN optical diagnostic diagrams based on the auroral emission line [OIII]λ4363, whose detection at $z > 3$ has become possible only thanks to JWST. The need for new NLAGN diagnostic diagrams is always becoming more compelling, given the limitations of the traditional diagnostic diagrams in the high-z Universe, where the physical conditions of the host galaxy are very different compared to the local universe. These diagrams exploit the sensitivity of the [OIII]λ4363 auroral line to the gas temperature and the fact that AGN, thanks to their harder ionizing radiation, can increase the electron temperature of their surrounding gas to higher values than what SFG can do, ultimately determining a stronger [OIII]λ4363 line emission, also at high-z. These new diagnostics were published in Mazzolari et al. (2024a).

- In Chapter 5 I present the selection, the properties, and the detailed multiwavelength analysis of the 52 NLAGN selected at $2 \lesssim z \lesssim 9$ among the spectra of the JWST CEERS survey. We confirmed that the traditional BPT diagnostic diagrams are not suited to identify high-z AGN, while most of the high-z NLAGN are selected using the recently proposed AGN diagnostic diagrams based on the [O III] λ4363 auroral line or high-ionization emission lines. We computed the emission line velocity dispersion and the ISM extinction levels of the sample of NLAGN without finding significant differences with those of the parent sample, suggesting a population of AGN heavily buried and not significantly impacting the host galaxies' physical properties, as further confirmed by SED-fitting. The bolometric luminosities of the high-z NLAGNs selected in this work are well below those sampled by surveys before JWST. Finally, we investigate the X-ray properties of CEERS NLAGN and of the sample of high-z BLAGN selected among the CEERS spectra by Harikane et al. (2023). We find that all but 4 NLAGN are undetected in the deep X-ray image of the field, as well as all the high-z BLAGN. We do not obtain a detection even by stacking the undetected sources, resulting in an X-ray weakness of $\sim 1-2$ dex from what is expected based on their bolometric luminosities. To discriminate between a heavily obscured AGN scenario or an intrinsic X-ray weakness of these sources, we performed a radio (1.4GHz) stacking analysis, which did not reveal any detection and left open questions about the origin of the X-ray weakness. These results were presented in Mazzolari et al. (2024b).

6.1 Next steps and Future perspectives

The work done so far allowed me to investigate many different aspects concerning the obscured AGN selection and their characterization across Cosmic Time. Each of the projects presented and discussed in this PhD Thesis, showing new and solid findings, suggests promising directions for further study and exploration that could significantly advance our understanding of SMBH evolution.

Spectroscopic follow-up of radio-excess sources on the J1030 field

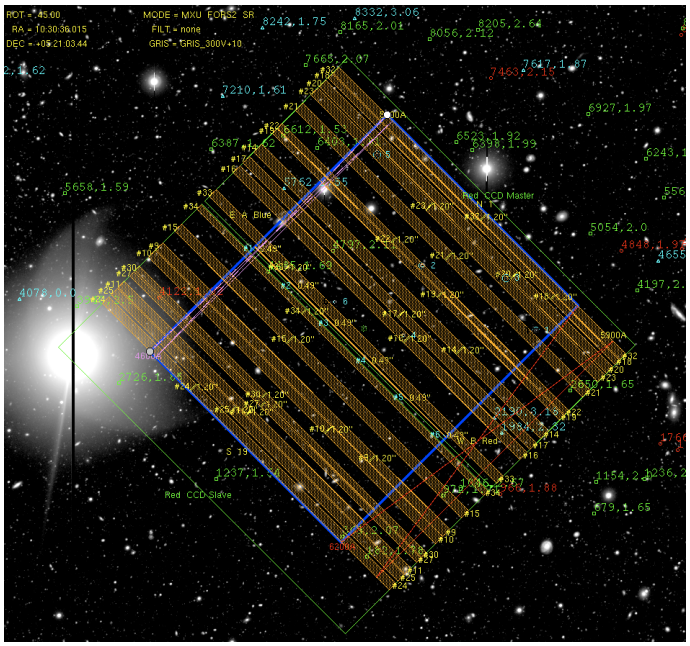

FIGURE 6.1: Example of one mask setup ($\sim 6 \times 6$ arcmin) for the observations that will be performed by VLT/FORS2 on the J1030 field.

Regarding the work done on the J1030 field, the next step will be the analysis of the spectroscopic follow-up observations I obtained as PI at GEMINI/GMOS, VLT/FORS2 and LBT/MODS facilities. As mentioned in Chapter 3, the last two programs consist of 25 hours of observations each and will target, with 10 different masks (in total) and in multi-object spectroscopy mode, more than 200 sources, most of them candidate high-z radio-excess AGN. One example of mask configuration for VLT/FORS2 is presented in Fig. 6.1. The GEMINI/GMOS observations have been recently performed, but only one out of the four requested masks was observed. The aim of these observations is twofold: first, constrain their redshift spectroscopically, and second, prove their AGN nature by detecting high-ionization emission lines typical of AGN, or using the AGN emission-line diagnostic diagrams presented in Chapter 5.
These observations will start this Winter, so I expect to perform the data reduction and analysis by next Spring and to be ready to discuss and present the results by the end of 2025.
Furthermore, JWST slitless spectroscopy observations that have been performed on the central $\sim$ 25 arcmin^2 of the J1030 field will significantly contribute to expanding the number

of sources with available spectra.
By proving spectroscopically that most of the targets host heavily obscured AGN, we will validate the effectiveness of the radio selection technique in identifying heavily obscured AGN, strengthening the analysis reported in Chapter 3. A successful validation would have far-reaching implications, making next-generation wide-and-deep radio surveys (e.g. with the SKA) a revolutionary tool to probe the bulk of the obscured AGN population up to very early times ($z > 6$).

Extending the radio-excess selection to the GOODS fields

Given the excellent results obtained so far on the J1030 field, our plan is also to extend the radio-excess selection of CTK AGN to other well-known extragalactic fields with very deep X-ray and radio coverage, such as GOODS-S and GOODS-N fields, where there are already many spectroscopic informations (including JWST observations). By uniformly defining a radio-excess parameter and by selecting sources undetected in the deep X-ray images, we can evaluate the efficiency of radio selection for revealing heavily obscured AGN up to redshifts $z \sim 5 - 6$, potentially bridging the gap between X-ray and JWST findings.

Constraints to the BHARD at $3 < z < 9$ with JWST spectroscopic samples

As reported in Sect. 1.4.1, recent studies, such as those by Yang et al. (2023a) and Akins et al. (2024), leveraging on JWST NIRCam and MIRI photometry to derive BHARD functions at redshifts $3 < z < 5$, revealed values that exceed those from X-ray observations and better align with simulation predictions. While these findings are promising, they may be affected by significant contamination since the AGN identification relied solely on photometric colors or SED-fitting techniques. However, with the release of data from numerous large JWST spectroscopic programs, a more robust approach is now possible. Newly published high-redshift AGN catalogs, based on spectroscopic selection, provide highly reliable samples. Spectroscopy also enables direct and precise measurements of AGN bolometric luminosities, which are essential for calculating the BHARD without the extensive multiwavelength data typically required for reliable measurements with SED-based methods.
My plan involves compiling a high-redshift spectroscopically selected AGN sample using publicly available data from JWST surveys such as JADES, CEERS, RUBIES, FRESCO, and GLASS. Some of these surveys already have published AGN selections; for those that do not, it is feasible to quickly apply a selection using a method similar to Mazzolari et al. (2024b). This approach should yield a sample of approximately 100 BLAGN and 300 NLAGN in the range $3 < z < 9$, constituting the largest spectroscopically selected AGN dataset for these redshifts to date. Although spectroscopic selection can be influenced by the selection function of targets, the use of slitless spectroscopy in some observations helps mitigate this bias. For other datasets, completeness corrections can be applied by performing completeness simulations of each field (Taylor et al., 2024). This robust sample, with thorough completeness corrections, will allow for an accurate and contamination-free BHARD estimation, for the first time across $3 < z < 9$. Comparing these results with X-ray findings and results from simulations will definitively provide an answer to the compelling questions about the missing AGN population at early times.

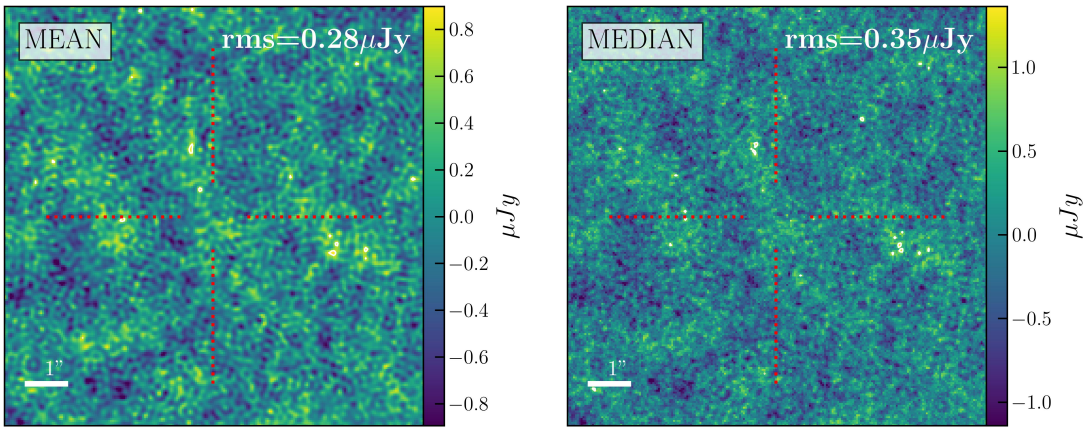

FIGURE 6.2: 10" × 10" cutouts of the mean (left) and median (right) radio stacking of the 22 JWST selected BLAGN on the GOODS-N field performed on the 1.5GHz radio image of the e-MERGE survey (Muxlow et al., 2020). The values of the rms of the two stacked images are reported in the upper right corners. No detections are returned.

Investigate the origin of X-ray weakness of BLAGN discovered by JWST

Another project that I am leading and am currently working on consists of analyzing the radio properties of the BLAGN selected by JWST, similar to what I have presented in Chapter 5 for the CEERS AGN (see Sect. 5.6.2). As described in Sect. 1.4.1, most of these JWST-selected AGN show a significant X-ray weakness, and investigating their radio properties could shed light on the origin of the lack of X-ray emission. If the lack of X-rays is due to absorption by a dense medium, "standard" AGN would still be detectable in the radio band according to known radio-to-X-ray or radio-to-bolometric luminosity relation. Instead, if both radio and X-ray emissions are much fainter than expected, this could indicate an intrinsic origin of the weak X-ray and radio emissions, for example driven by the absence of an X-ray corona or by a weaker magnetic field.

Based on publicly available radio images and the location of JWST-selected BLAGN discovered so far, the GOODS-N field provides the best combination of radio sensitivity and BLAGN statistics. There are 22 BLAGN, spanning a redshift range $2.26 < z < 10.6$ (with a median $z_{med} = 5.2$), a SMBH mass range $6.2 < \log M_{BH} < 8.5$ (with a median $\log M_{BH,med} = 7.2$), and have bolometric luminosities $43.6 < \log L_{bol} < 45.8$ (median $\log L_{bol} = 44.6$). None of them is detected in X-rays, despite being located in one of the deepest X-ray fields to date (CDFN, with a flux limit of $\sim 10^{-17}$erg cm^{-2} s^{-1} Xue et al., 2016). From X-ray stacking Maiolino et al. (2024a) infer an upper limit to the X-ray luminosity $\sim$ 2 dex lower than expected. To investigate the radio properties of these sources we used the 1.5GHz radio images of the *e*-MERLIN Galaxy Evolution (*e*-MERGE) Survey (Muxlow et al., 2020), with a resolution (0.28×0.26 arcsec^2), and a sensitivity reaching a rms of 1.17μJybeam^{-1}.

All 22 BLAGN fall inside the aforementioned radio images, but none of them has a detected counterpart in the associated radio catalogs. Hence, we performed a radio stacking analysis. In Fig. 6.2 we show both the median and mean stack performed on the 1.5GHz radio

image, in which we reached an rms of $\sim 0.3\mu$Jy/beam. The non-detection in the stack corresponds to a 3σ upper limit of the median rest-frame radio luminosity of these sources of $\log L_{1.5GHz} = 39.1$ erg/s.
To verify whether the non-detection of the radio emission of these BLAGN (both as single sources and in the stack) is indicative of a radio weakness, we have to estimate their expected radio luminosities.
Essentially, we are assuming the scenario whereby the intrinsic X-ray and radio emission follow the standard SED of AGN, and we want to test if this hypothesis is compatible with the radio non-detection. The fact that these sources are characterized by an observed X-ray weakness does not necessarily imply that their SED is different from that of 'standard' AGN because this X-ray weakness can also be explained by obscuration. Therefore, we first computed their expected intrinsic X-ray luminosities considering their bolometric luminosities and using the standard $L_{bol} - L_{2-10keV}$ relation derived in Duras et al. (2020). Then, to estimate the intrinsic radio luminosity, we are using both direct L_X to L_{rad} relations (D'Amato et al., 2022; Panessa et al., 2015; Fan & Bai, 2016) as well as different fundamental plane relations (Wang et al., 2024; Li et al., 2008; Bariuan et al., 2022; Dong et al., 2014; Merloni et al., 2003; Xie & Yuan, 2017). We are also computing the expected radio luminosity of these sources considering relations between the AGN optical luminosity and their radio luminosity, leveraging on the intrinsic distribution of the radio loudness parameters found by Cirasuolo et al. (2003) and Baloković et al. (2012).
Preliminary results show that expected radio luminosities of many of these BLAGN are below their 3σ radio upper limit (meaning that the radio image is still too shallow to detect their radio luminosity, if they are RQ), but the radio stack results to be only marginally compatible with the expected median luminosity of these objects. Moreover, there are a couple of sources for which the observed radio upper limit is significantly below its expected radio luminosity, showing a deviation of ~ 1 dex for all the relations taken into account, much larger than the scatter of the single relations. Even if preliminary, these results may suggest that these sources are probably characterized not only by an X-ray weakness but also by a radio weakness.

***

Overall, the work presented in this thesis tried to advance our understanding of obscured AGN, especially at high redshifts. It highlights the importance of multiwavelength approaches and the role of new facilities like JWST in uncovering the hidden AGN population, offering fresh perspectives on their selection, demography, and influence on cosmic evolution. As the next generation of radio (SKAO, ngVLA), X-rays (Athena, AXIS), and optical/infrared (ELT, EUCLID) observatories come online, the groundwork laid in this thesis will provide useful instruments in guiding future studies of AGN demographics and their role in shaping the universe.

Bibliography

Abazajian K. N., et al., 2009, "The Seventh Data Release of the Sloan Digital Sky Survey" ApJS, 182, 543 (arXiv:0812.0649)

Abuter R., et al., 2024, "A dynamical measure of the black hole mass in a quasar 11 billion years ago" Nature, 627, 281 (arXiv:2401.14567)

Aird J., Coil A. L., Georgakakis A., Nandra K., Barro G., Pérez-González P. G., 2015, "The evolution of the X-ray luminosity functions of unabsorbed and absorbed AGNs out to z 5" MNRAS, 451, 1892 (arXiv:1503.01120)

Akins H. B., et al., 2024, "COSMOS-Web: The over-abundance and physical nature of "little red dots"–Implications for early galaxy and SMBH assembly" arXiv e-prints, p. arXiv:2406.10341 (arXiv:2406.10341)

Alberts S., Rujopakarn W., Rieke G. H., Jagannathan P., Nyland K., 2020, "Completing the Census of AGN in GOODS-S/HUDF: New Ultradeep Radio Imaging and Predictions for JWST" ApJ, 901, 168 (arXiv:2008.11208)

Algera H. S. B., et al., 2020, "A Multiwavelength Analysis of the Faint Radio Sky (COSMOS-XS): the Nature of the Ultra-faint Radio Population" ApJ, 903, 139 (arXiv:2009.13531)

Allen M. G., Groves B. A., Dopita M. A., Sutherland R. S., Kewley L. J., 2008, "The MAPPINGS III Library of Fast Radiative Shock Models" ApJS, 178, 20 (arXiv:0805.0204)

Alonso-Tetilla A. V., et al., 2024, "Probing the roles of orientation and multiscale gas distributions in shaping the obscuration of active galactic nuclei through cosmic time" MNRAS, 527, 10878 (arXiv:2310.16854)

Amorín R., et al., 2015, "Extreme emission-line galaxies out to z ~1 in zCOSMOS. I. Sample and characterization of global properties" A&A, 578, A105 (arXiv:1403.3441)

Ananna T. T., et al., 2019, "The Accretion History of AGNs. I. Supermassive Black Hole Population Synthesis Model" ApJ, 871, 240 (arXiv:1810.02298)

Ananna T. T., Bogdán Á., Kovács O. E., Natarajan P., Hickox R. C., 2024, "X-Ray View of Little Red Dots: Do They Host Supermassive Black Holes?" ApJ, 969, L18 (arXiv:2404.19010)

Andonie C., et al., 2022, "A panchromatic view of infrared quasars: excess star formation and radio emission in the most heavily obscured systems" MNRAS, 517, 2577 (arXiv:2209.13321)

Annunziatella M., et al., 2018, "Complete IRAC Mapping of the CFHTLS-DEEP, MUSYC, and NMBS-II Fields" PASP, 130, 124501 (arXiv:1810.11508)

Antonucci R., 1993, "Unified models for active galactic nuclei and quasars." ARA&A, 31, 473

Arango-Toro R. C., Ciesla L., Ilbert O., Magnelli B., Jiménez-Andrade E. F., Buat V., 2023, "Probing the timescale of the 1.4 GHz radio emissions as a star formation tracer" A&A, 675, A126 (arXiv:2304.13758)

Aravena M., et al., 2020, "The ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: The Nature of the Faintest Dusty Star-forming Galaxies" ApJ, 901, 79 (arXiv:2006.04284)

Arellano-Córdova K. Z., et al., 2022, "A First Look at the Abundance Pattern-O/H, C/O, and Ne/O-in z > 7 Galaxies with JWST/NIRSpec" ApJ, 940, L23 (arXiv:2208.02562)

Armah M., et al., 2021, "Chemical abundances in Seyfert galaxies - VII. Direct abundance determination of neon based on optical and infrared emission lines" MNRAS, 508, 371 (arXiv:2109.04596)

Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, "Measuring and modelling the redshift evolution of clustering: the Hubble Deep Field North" MNRAS, 310, 540 (arXiv:astro-ph/9902290)

Arnouts S., et al., 2002, "Measuring the redshift evolution of clustering: the Hubble Deep Field South" MNRAS, 329, 355 (arXiv:astro-ph/0109453)

Assef R. J., et al., 2013, "Mid-infrared Selection of Active Galactic Nuclei with the Wide-field Infrared Survey Explorer. II. Properties of WISE-selected Active Galactic Nuclei in the NDWFS Boötes Field" ApJ, 772, 26 (arXiv:1209.6055)

Bañados E., et al., 2016, "The Pan-STARRS1 Distant z > 5.6 Quasar Survey: More than 100 Quasars within the First Gyr of the Universe" ApJS, 227, 11 (arXiv:1608.03279)

Backhaus B. E., et al., 2022, "CLEAR: Emission-line Ratios at Cosmic High Noon" ApJ, 926, 161 (arXiv:2109.08147)

Backhaus B. E., et al., 2023, "CLEAR: Spatially Resolved Emission Lines and Active Galactic Nuclei at 0.6 < z < 1.3" ApJ, 943, 37 (arXiv:2207.11265)

Bahcall J. N., Kozlovsky B.-Z., Salpeter E. E., 1972, "On the Time Dependence of Emission-Line Strengths from a Photoionized Nebula" ApJ, 171, 467

Baldwin J. A., Phillips M. M., Terlevich R., 1981, "Classification parameters for the emission-line spectra of extragalactic objects." PASP, 93, 5

Ballantyne D. R., 2009, "The Contribution of Active Galactic Nuclei to the Microjansky Radio Population" ApJ, 698, 1033 (arXiv:0904.0996)

Balmaverde B., Gilli R., Mignoli M., et al., 2017, "The primordial environment of supermassive black holes (II): deep Y and J band images around the z 6.3 quasar SDSS J1030+0524" Astronomy & Astrophysics, 606, A23
Baloković M., Smolčić V., Ivezić Ž., Zamorani G., Schinnerer E., Kelly B. C., 2012, "Disclosing the Radio Loudness Distribution Dichotomy in Quasars: An Unbiased Monte Carlo Approach Applied to the SDSS-FIRST Quasar Sample" ApJ, 759, 30 (arXiv:1209.1099)
Barchiesi L., et al., 2021, "The role of SPICA-like missions and the Origins Space Telescope in the quest for heavily obscured AGN and synergies with Athena" PASA, 38, e033 (arXiv:2106.08345)
Barchiesi L., et al., 2024, "COSMOS2020: Investigating the AGN-obscured accretion phase at z ∼ 1 via [Ne V] selection" A&A, 685, A141 (arXiv:2403.03251)
Bariuan L. G. C., Snios B., Sobolewska M., Siemiginowska A., Schwartz D. A., 2022, "The Fundamental Planes of black hole activity for radio-loud and radio-quiet quasars" MNRAS, 513, 4673 (arXiv:2201.04666)
Barvainis R., 1987, "Hot Dust and the Near-Infrared Bump in the Continuum Spectra of Quasars and Active Galactic Nuclei" ApJ, 320, 537
Becker R. H., White R. L., Helfand D. J., 1995, "The FIRST Survey: Faint Images of the Radio Sky at Twenty Centimeters" ApJ, 450, 559
Becker R. H., et al., 2001, "The FIRST Bright Quasar Survey. III. The South Galactic Cap" ApJS, 135, 227 (arXiv:astro-ph/0104279)
Beckmann R. S., et al., 2017, "Cosmic evolution of stellar quenching by AGN feedback: clues from the Horizon-AGN simulation" MNRAS, 472, 949 (arXiv:1701.07838)
Behiri M., Giulietti M., Galluzzi V., Lapi A., Liuzzo E., Massardi M., 2024, "Teaming up Radio and Sub-mm/FIR Observations to Probe Dusty Star-Forming Galaxies" Galaxies, 12, 14 (arXiv:2403.17276)
Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, "What Shapes the Luminosity Function of Galaxies?" ApJ, 599, 38 (arXiv:astro-ph/0302450)
Bentz M. C., Peterson B. M., Netzer H., Pogge R. W., Vestergaard M., 2009, "The Radius-Luminosity Relationship for Active Galactic Nuclei: The Effect of Host-Galaxy Starlight on Luminosity Measurements. II. The Full Sample of Reverberation-Mapped AGNs" ApJ, 697, 160 (arXiv:0812.2283)
Berg D. A., et al., 2012, "Direct Oxygen Abundances for Low-luminosity LVL Galaxies" ApJ, 754, 98 (arXiv:1205.6782)
Bian F., Kewley L. J., Dopita M. A., 2018, ""Direct" Gas-phase Metallicity in Local Analogs of High-redshift Galaxies: Empirical Metallicity Calibrations for High-redshift Star-forming Galaxies" ApJ, 859, 175 (arXiv:1805.08224)
Biggs A. D., Ivison R. J., 2006, "A catalogue of μJy radio sources in northern legacy fields" MNRAS, 371, 963 (arXiv:astro-ph/0606595)
Blanc G. A., Kewley L., Vogt F. P. A., Dopita M. A., 2015, "IZI: Inferring the Gas Phase Metallicity (Z) and Ionization Parameter (q) of Ionized Nebulae Using Bayesian Statistics" ApJ, 798, 99 (arXiv:1410.8146)
Blandford R. D., Begelman M. C., 1999, "On the fate of gas accreting at a low rate on to a black hole" MNRAS, 303, L1 (arXiv:astro-ph/9809083)
Blandford R. D., McKee C. F., 1982, "Reverberation mapping of the emission line regions of Seyfert galaxies and quasars." ApJ, 255, 419
Blandford R. D., Znajek R. L., 1977, "Electromagnetic extraction of energy from Kerr black holes." MNRAS, 179, 433
Blecha L., et al., 2016, "Recoiling black holes: prospects for detection and implications of spin alignment" MNRAS, 456, 961 (arXiv:1508.01524)
Blecha L., Snyder G. F., Satyapal S., Ellison S. L., 2018, "The power of infrared AGN selection in mergers: a theoretical study" MNRAS, 478, 3056 (arXiv:1711.02094)
Bluck A. F. L., Piotrowska J. M., Maiolino R., 2023a, "The Fundamental Signature of Star Formation Quenching from AGN Feedback: A Critical Dependence of Quiescence on Supermassive Black Hole Mass, Not Accretion Rate" ApJ, 944, 108 (arXiv:2301.03677)
Bluck A. F. L., Piotrowska J. M., Maiolino R., 2023b, "The Fundamental Signature of Star Formation Quenching from AGN Feedback: A Critical Dependence of Quiescence on Supermassive Black Hole Mass, Not Accretion Rate" ApJ, 944, 108 (arXiv:2301.03677)
Bogdán Á., et al., 2023, "Evidence for heavy-seed origin of early supermassive black holes from a z ≈ 10 X-ray quasar" Nature Astronomy, (arXiv:2305.15458)
Bolzonella M., Miralles J. M., Pelló R., 2000, "Photometric redshifts based on standard SED fitting procedures" A&A, 363, 476 (arXiv:astro-ph/0003380)
Bonaldi A., Bonato M., Galluzzi V., Harrison I., Massardi M., Kay S., De Zotti G., Brown M. L., 2019, "The Tiered Radio Extragalactic Continuum Simulation (T-RECS)" MNRAS, 482, 2 (arXiv:1805.05222)
Bonato M., Prandoni I., De Zotti G., Brienza M., Morganti R., Vaccari M., 2021, "New constraints on the 1.4 GHz source number counts and luminosity functions in the Lockman Hole field" MNRAS, 500, 22 (arXiv:2010.08748)
Bongiorno A., et al., 2010, "The [O iii] emission line luminosity function of optically selected type-2 AGN from zCOSMOS" A&A, 510, A56 (arXiv:0911.3914)
Bongiorno A., et al., 2012, "Accreting supermassive black holes in the COSMOS field and the connection to their host galaxies" MNRAS, 427, 3103 (arXiv:1209.1640)
Bonning E. W., Cheng L., Shields G. A., Salviander S., Gebhardt K., 2007, "Accretion Disk Temperatures and Continuum Colors in QSOs" ApJ, 659, 211 (arXiv:astro-ph/0611263)
Bonzini M., Padovani P., Mainieri V., Kellermann K. I., Miller N., Rosati P., Tozzi P., Vattakunnel S., 2013,

"The sub-mJy radio sky in the Extended Chandra Deep Field-South: source population" MNRAS, 436, 3759 (arXiv:1310.1248)
Bonzini M., et al., 2015, "Star formation properties of sub-mJy radio sources" MNRAS, 453, 1079 (arXiv:1508.01905)
Boquien M., Burgarella D., Roehlly Y., Buat V., Ciesla L., Corre D., Inoue A. K., Salas H., 2019, "CIGALE: a python Code Investigating GALaxy Emission" A&A, 622, A103 (arXiv:1811.03094)
Brammer G. B., van Dokkum P. G., Coppi P., 2008, "EAZY: A Fast, Public Photometric Redshift Code" ApJ, 686, 1503 (arXiv:0807.1533)
Brenneman L., 2013, "Measuring Supermassive Black Hole Spins in AGN" Acta Polytechnica, 53, 652
Brinchmann J., 2023, "High-z galaxies with JWST and local analogues - it is not only star formation" MNRAS, 525, 2087 (arXiv:2208.07467)
Brunner H., Cappelluti N., Hasinger G., Barcons X., Fabian A. C., Mainieri V., Szokoly G., 2008, "XMM-Newton observations of the Lockman Hole: X-ray source catalogue and number counts" A&A, 479, 283 (arXiv:0711.4822)
Brusa M., et al., 2009, "Black hole growth and starburst activity at z = 0.6-4 in the Chandra Deep Field South. Host galaxies properties of obscured AGN" A&A, 507, 1277 (arXiv:0910.1007)
Bruzual G., Charlot S., 2003, "Stellar population synthesis at the resolution of 2003" MNRAS, 344, 1000 (arXiv:astro-ph/0309134)
Buchner J., et al., 2015, "Obscuration-dependent Evolution of Active Galactic Nuclei" ApJ, 802, 89 (arXiv:1501.02805)
Bugiani L., et al., 2024, "AGN Feedback in Quiescent Galaxies at Cosmic Noon Traced by Ionized Gas Emission" arXiv e-prints, p. arXiv:2406.08547 (arXiv:2406.08547)
Bunker A. J., et al., 2023a, "JADES NIRSpec Initial Data Release for the Hubble Ultra Deep Field: Redshifts and Line Fluxes of Distant Galaxies from the Deepest JWST Cycle 1 NIRSpec Multi-Object Spectroscopy" arXiv e-prints, p. arXiv:2306.02467 (arXiv:2306.02467)
Bunker A. J., et al., 2023b, "JADES NIRSpec Spectroscopy of GN-z11: Lyman-α emission and possible enhanced nitrogen abundance in a z = 10.60 luminous galaxy" A&A, 677, A88 (arXiv:2302.07256)
Bushouse H., et al., 2022, JWST Calibration Pipeline, doi:10.5281/zenodo.7429939
Calabrò A., et al., 2023, "Near-infrared emission line diagnostics for AGN from the local Universe to z $\sim$ 3" A&A, 679, A80 (arXiv:2306.08605)
Calzetti D., Kinney A. L., Storchi-Bergmann T., 1994, "Dust Extinction of the Stellar Continua in Starburst Galaxies: The Ultraviolet and Optical Extinction Law" ApJ, 429, 582
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, "The Dust Content and Opacity of Actively Star-forming Galaxies" ApJ, 533, 682 (arXiv:astro-ph/9911459)
Cameron A. J., et al., 2023, "JADES: Probing interstellar medium conditions at $z \sim 5.5 - 9.5$ with ultra-deep JWST/NIRSpec spectroscopy" arXiv e-prints, p. arXiv:2302.04298 (arXiv:2302.04298)
Carniani S., et al., 2024, "A shining cosmic dawn: spectroscopic confirmation of two luminous galaxies at $z \sim 14$" arXiv e-prints, p. arXiv:2405.18485 (arXiv:2405.18485)
Carton D., et al., 2017, "Inferring gas-phase metallicity gradients of galaxies at the seeing limit: a forward modelling approach" MNRAS, 468, 2140 (arXiv:1703.01090)
Casey C. M., 2012, "Far-infrared spectral energy distribution fitting for galaxies near and far" MNRAS, 425, 3094 (arXiv:1206.1595)
Castellano M., et al., 2022, "Early Results from GLASS-JWST. III. Galaxy Candidates at z 9-15" ApJ, 938, L15 (arXiv:2207.09436)
Chabrier G., 2003, "Galactic Stellar and Substellar Initial Mass Function" Publications of the Astronomical Society of the Pacific, 115, 763 (arXiv:astro-ph/0304382)
Charlot S., Fall S. M., 2000, "A Simple Model for the Absorption of Starlight by Dust in Galaxies" ApJ, 539, 718 (arXiv:astro-ph/0003128)
Chartas G., et al., 2016, "Gravitational lensing size scales for quasars" Astronomische Nachrichten, 337, 356 (arXiv:1509.05375)
Chen C. T. J., et al., 2018, "The XMM-SERVS survey: new XMM-Newton point-source catalogue for the XMM-LSS field" MNRAS, 478, 2132 (arXiv:1804.07763)
Chen N., et al., 2023, "Properties and evolution of dual and offset AGN in the ASTRID simulation at z 2" MNRAS, 522, 1895 (arXiv:2208.04970)
Chen S., et al., 2024, "Windy or not: Radio pc-scale evidence for a broad-line region wind in radio-quiet quasars" arXiv e-prints, p. arXiv:2408.15934 (arXiv:2408.15934)
Chisholm J., et al., 2024, "[Ne v] emission from a faint epoch of reionization-era galaxy: evidence for a narrow-line intermediate mass black hole" arXiv e-prints, p. arXiv:2402.18643 (arXiv:2402.18643)
Cirasuolo M., Celotti A., Magliocchetti M., Danese L., 2003, "Is there a dichotomy in the radio loudness distribution of quasars?" MNRAS, 346, 447 (arXiv:astro-ph/0306415)
Civano F., et al., 2016, "The Chandra Cosmos Legacy Survey: Overview and Point Source Catalog" ApJ, 819, 62 (arXiv:1601.00941)
Cleri N. J., et al., 2023, "Using [Ne V]/[Ne III] to Understand the Nature of Extreme-ionization Galaxies" ApJ, 953, 10 (arXiv:2301.07745)
Collaboration E. H. T., 2019, "First M87 Event Horizon Telescope Results. IV. Imaging the Central Supermassive Black Hole" The Astrophysical Journal, 875, L1

Comastri A., Setti G., Zamorani G., Hasinger G., 1995, "The contribution of AGNs to the X-ray background." A&A, 296, 1 (arXiv:astro-ph/9409067)
Comastri A., Gilli R., Marconi A., Risaliti G., Salvati M., 2015, "Mass without radiation: Heavily obscured AGNs, the X-ray background, and the black hole mass density" A&A, 574, L10 (arXiv:1501.03620)
Conroy C., Gunn J. E., 2010, FSPS: Flexible Stellar Population Synthesis, Astrophysics Source Code Library, record ascl:1010.043
Curti M., Cresci G., Mannucci F., Marconi A., Maiolino R., Esposito S., 2017, "New fully empirical calibrations of strong-line metallicity indicators in star-forming galaxies" MNRAS, 465, 1384 (arXiv:1610.06939)
Curti M., et al., 2020, "The KLEVER Survey: spatially resolved metallicity maps and gradients in a sample of 1.2 < z < 2.5 lensed galaxies" MNRAS, 492, 821 (arXiv:1910.13451)
Curti M., et al., 2023a, "JADES: Insights on the low-mass end of the mass–metallicity–star-formation rate relation at $3 < z < 10$ from deep JWST/NIRSpec spectroscopy" arXiv e-prints, p. arXiv:2304.08516 (arXiv:2304.08516)
Curti M., et al., 2023b, "The chemical enrichment in the early Universe as probed by JWST via direct metallicity measurements at z 8" MNRAS, 518, 425 (arXiv:2207.12375)
Curtis-Lake E., et al., 2022, "Spectroscopy of four metal-poor galaxies beyond redshift ten" arXiv e-prints, p. arXiv:2212.04568 (arXiv:2212.04568)
D'Amato Q., et al., 2020, "Dust and gas content of high-redshift galaxies hosting obscured AGN in the Chandra Deep Field-South" A&A, 636, A37 (arXiv:2003.08631)
D'Amato Q., et al., 2022, "A deep 1.4 GHz survey of the J1030 equatorial field: A new window on radio source populations across cosmic time" A&A, 668, A133 (arXiv:2210.15595)
D'Eugenio F., et al., 2024, "JADES Data Release 3 – NIRSpec/MSA spectroscopy for 4,000 galaxies in the GOODS fields" arXiv e-prints, p. arXiv:2404.06531 (arXiv:2404.06531)
Daddi E., et al., 2007, "Multiwavelength Study of Massive Galaxies at z~2. I. Star Formation and Galaxy Growth" ApJ, 670, 156 (arXiv:0705.2831)
Dahlen T., et al., 2013, "A Critical Assessment of Photometric Redshift Methods: A CANDELS Investigation" ApJ, 775, 93 (arXiv:1308.5353)
Dale D. A., Helou G., Magdis G. E., Armus L., Díaz-Santos T., Shi Y., 2014, "A Two-parameter Model for the Infrared/Submillimeter/Radio Spectral Energy Distributions of Galaxies and Active Galactic Nuclei" ApJ, 784, 83 (arXiv:1402.1495)
Davidzon I., et al., 2019, "HORIZON-AGN virtual observatory - 2. Template-free estimates of galaxy properties from colours" MNRAS, 489, 4817 (arXiv:1905.13233)
Davis M., et al., 2007, "The All-Wavelength Extended Groth Strip International Survey (AEGIS) Data Sets" ApJ, 660, L1 (arXiv:astro-ph/0607355)
Davis K., et al., 2023, "A Census from JWST of Extreme Emission Line Galaxies Spanning the Epoch of Reionization in CEERS" arXiv e-prints, p. arXiv:2312.07799 (arXiv:2312.07799)
Dayal P., et al., 2020, "Reionization with galaxies and active galactic nuclei" MNRAS, 495, 3065 (arXiv:2001.06021)
De Marco B., Ponti G., Miniutti G., Belloni T., Cappi M., Dadina M., Muñoz-Darias T., 2013, "Time lags in the ultraluminous X-ray source NGC 5408 X-1: implications for the black hole mass" MNRAS, 436, 3782 (arXiv:1310.0042)
Decarli R., et al., 2018, "An ALMA [C II] Survey of 27 Quasars at z > 5.94" ApJ, 854, 97 (arXiv:1801.02641)
Decarli R., et al., 2019, "Testing the paradigm: First spectroscopic evidence of a quasar-galaxy Mpc-scale association at cosmic dawn" A&A, 631, L10 (arXiv:1910.12887)
Del Moro A., et al., 2013, "GOODS-Herschel: radio-excess signature of hidden AGN activity in distant star-forming galaxies" A&A, 549, A59 (arXiv:1210.2521)
Delhaize J., et al., 2017, "The VLA-COSMOS 3 GHz Large Project: The infrared-radio correlation of star-forming galaxies and AGN to $z \lesssim 6$" A&A, 602, A4 (arXiv:1703.09723)
Delvecchio I., et al., 2017, "The VLA-COSMOS 3 GHz Large Project: AGN and host-galaxy properties out to $z \lesssim 6$" A&A, 602, A3 (arXiv:1703.09720)
Delvecchio I., et al., 2021, "The infrared-radio correlation of star-forming galaxies is strongly $M_\star$-dependent but nearly redshift-invariant since $z \sim 4$" A&A, 647, A123 (arXiv:2010.05510)
Di Matteo T., 1998, "Magnetic reconnection: flares and coronal heating in active galactic nuclei" MNRAS, 299, L15 (arXiv:astro-ph/9805347)
DiPompeo M. A., Myers A. D., Hickox R. C., Geach J. E., Hainline K. N., 2014, "The angular clustering of infrared-selected obscured and unobscured quasars" MNRAS, 442, 3443 (arXiv:1406.0778)
DiPompeo M. A., Myers A. D., Hickox R. C., Geach J. E., Holder G., Hainline K. N., Hall S. W., 2015, "Weighing obscured and unobscured quasar hosts with the cosmic microwave background" MNRAS, 446, 3492 (arXiv:1411.0527)
DiPompeo M. A., Hickox R. C., Eftekharzadeh S., Myers A. D., 2017, "The characteristic halo masses of half-a-million WISE-selected quasars" MNRAS, 469, 4630 (arXiv:1705.05306)
DiPompeo M. A., Hickox R. C., Carroll C. M., Runnoe J. C., Mullaney J. R., Fischer T. C., 2018, "The [O III] Profiles of Infrared-selected Active Galactic Nuclei: More Powerful Outflows in the Obscured Population" ApJ, 856, 76 (arXiv:1803.00083)
Doi A., Oyama T., Kono Y., Yamauchi A., Suzuki S., Matsumoto N., Tazaki F., 2016, "A radio detection survey of narrow-line Seyfert 1 galaxies using very long baseline interferometry at 22 GHz" PASJ, 68, 73 (arXiv:1603.04593)

Dong A.-J., Wu Q., Cao X.-F., 2014, "A New Fundamental Plane for Radiatively Efficient Black-hole Sources" ApJ, 787, L20 (arXiv:1404.5317)
Dong A.-J., Ge K., Liu X., Zhi Q.-J., 2021, "The fundamental plane of black hole activity for quiescent sources" Astronomische Nachrichten, 342, 191
Donley J. L., et al., 2012, "Identifying Luminous Active Galactic Nuclei in Deep Surveys: Revised IRAC Selection Criteria" ApJ, 748, 142 (arXiv:1201.3899)
Donoso E., et al., 2012, "Origin of 12 μm Emission across Galaxy Populations from WISE and SDSS Surveys" ApJ, 748, 80 (arXiv:1201.2943)
Dopita M. A., et al., 2006, "Modeling the Pan-Spectral Energy Distribution of Starburst Galaxies. III. Emission Line Diagnostics of Ensembles of Evolving H II Regions" ApJS, 167, 177 (arXiv:astro-ph/0608062)
Dopita M. A., Sutherland R. S., Nicholls D. C., Kewley L. J., Vogt F. P. A., 2013, "New Strong-line Abundance Diagnostics for H II Regions: Effects of κ-distributed Electron Energies and New Atomic Data" ApJS, 208, 10 (arXiv:1307.5950)
Dopita M. A., et al., 2015, "Probing the Physics of Narrow Line Regions in Active Galaxies. II. The Siding Spring Southern Seyfert Spectroscopic Snapshot Survey (S7)" ApJS, 217, 12 (arXiv:1501.02022)
Dors O. L., 2021, "Chemical abundances in Seyfert galaxies - VI. Empirical abundance calibration" MNRAS, 507, 466
Dors O. L., Cardaci M. V., Hägele G. F., Rodrigues I., Grebel E. K., Pilyugin L. S., Freitas-Lemes P., Krabbe A. C., 2015, "On the central abundances of active galactic nuclei and star-forming galaxies" MNRAS, 453, 4102 (arXiv:1508.07802)
Dors O. L., Maiolino R., Cardaci M. V., Hägele G. F., Krabbe A. C., Pérez-Montero E., Armah M., 2020, "Chemical abundances of Seyfert 2 AGNs - III. Reducing the oxygen abundance discrepancy" MNRAS, 496, 3209 (arXiv:2006.09152)
Dors O. L., Cardaci M. V., Hägele G. F., Ilha G. S., Oliveira C. B., Riffel R. A., Riffel R., Krabbe A. C., 2024, "Cosmic metallicity evolution of Active Galactic Nuclei: implications for optical diagnostic diagrams" MNRAS, 527, 8193 (arXiv:2311.14026)
Dressler A., et al., 2023, "Early Results from GLASS-JWST. XVII. Building the First Galaxies-Chapter 1. Star Formation Histories for 5 < z < 7 Galaxies" ApJ, 947, L27 (arXiv:2208.04292)
Duras F., et al., 2020, "Universal bolometric corrections for active galactic nuclei over seven luminosity decades" A&A, 636, A73 (arXiv:2001.09984)
Eddington A. S., 1913, "On a formula for correcting statistics for the effects of a known error of observation" MNRAS, 73, 359
Edelson R. A., Malkan M. A., 1986, "Spectral Energy Distributions of Active Galactic Nuclei between 0.1 and 100 Microns" ApJ, 308, 59
Eisenstein D. J., et al., 2023, "Overview of the JWST Advanced Deep Extragalactic Survey (JADES)" arXiv e-prints, p. arXiv:2306.02465 (arXiv:2306.02465)
Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, "Binary Population and Spectral Synthesis Version 2.1: Construction, Observational Verification, and New Results" PASA, 34, e058 (arXiv:1710.02154)
Endsley R., et al., 2023, "The Star-forming and Ionizing Properties of Dwarf z_6-9 Galaxies in JADES: Insights on Bursty Star Formation and Ionized Bubble Growth" arXiv e-prints, p. arXiv:2306.05295 (arXiv:2306.05295)
Event Horizon Telescope Collaboration et al., 2021, "First M87 Event Horizon Telescope Results. VIII. Magnetic Field Structure near The Event Horizon" ApJ, 910, L13 (arXiv:2105.01173)
Event Horizon Telescope Collaboration et al., 2024, "First Sagittarius A* Event Horizon Telescope Results. VII. Polarization of the Ring" ApJ, 964, L25
Fabian A. C., 2012, "Observational Evidence of Active Galactic Nuclei Feedback" ARA&A, 50, 455 (arXiv:1204.4114)
Fabian A. C., Ballantyne D. R., Merloni A., Vaughan S., Iwasawa K., Boller T., 2002, "How the X-ray spectrum of a narrow-line Seyfert 1 galaxy may be reflection-dominated" MNRAS, 331, L35 (arXiv:astro-ph/0202297)
Fabian A. C., Lohfink A., Kara E., Parker M. L., Vasudevan R., Reynolds C. S., 2015, "Properties of AGN coronae in the NuSTAR era" MNRAS, 451, 4375 (arXiv:1505.07603)
Faisst A. L., et al., 2020, "The ALPINE-ALMA [C II] Survey: Multiwavelength Ancillary Data and Basic Physical Measurements" ApJS, 247, 61 (arXiv:1912.01621)
Falcke H., Biermann P. L., 1995, "The jet-disk symbiosis. I. Radio to X-ray emission models for quasars." A&A, 293, 665 (arXiv:astro-ph/9411096)
Fan X.-L., Bai J.-M., 2016, "The Radio/X-Ray Correlation and Black Hole Fundamental Plane for Young Radio Sources: Implications for X-Ray Origin and Accretion Mode" ApJ, 818, 185
Feltre A., Charlot S., Gutkin J., 2016, "Nuclear activity versus star formation: emission-line diagnostics at ultraviolet and optical wavelengths" MNRAS, 456, 3354 (arXiv:1511.08217)
Ferland G. J., 1993, Hazy, A Brief Introduction to Cloudy 84
Ferland G. J., et al., 2013, "The 2013 Release of Cloudy" Rev. Mexicana Astron. Astrofis., 49, 137 (arXiv:1302.4485)
Ferruit P., et al., 2022a, "The Near-Infrared Spectrograph (NIRSpec) on the James Webb Space Telescope. II. Multi-object spectroscopy (MOS)" A&A, 661, A81 (arXiv:2202.03306)
Ferruit P., et al., 2022b, "The Near-Infrared Spectrograph (NIRSpec) on the James Webb Space Telescope. II. Multi-object spectroscopy (MOS)" A&A, 661, A81 (arXiv:2202.03306)

Feuillet L. M., Meléndez M., Kraemer S., Schmitt H. R., Fischer T. C., Reeves J. N., 2024, "Classifying Intermediate-redshift Galaxies in SDSS: Alternative Diagnostic Diagrams" ApJ, 962, 104 (arXiv:2312.17381)
Finkelstein S. L., et al., 2022, "A Long Time Ago in a Galaxy Far, Far Away: A Candidate z ∼ 12 Galaxy in Early JWST CEERS Imaging" ApJ, 940, L55 (arXiv:2207.12474)
Finkelstein S. L., et al., 2023, "CEERS Key Paper. I. An Early Look into the First 500 Myr of Galaxy Formation with JWST" ApJ, 946, L13 (arXiv:2211.05792)
Fiore F., et al., 2009, "Chasing Highly Obscured QSOs in the COSMOS Field" ApJ, 693, 447 (arXiv:0810.0720)
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, "emcee: The MCMC Hammer" PASP, 125, 306 (arXiv:1202.3665)
Franzen T. M. O., et al., 2015, "ATLAS - I. Third release of 1.4 GHz mosaics and component catalogues" MNRAS, 453, 4020 (arXiv:1508.03150)
Fukushima K., Nagamine K., Shimizu I., 2023, "Star formation and chemical enrichment in protoclusters" MNRAS, 525, 3760 (arXiv:2212.12281)
Furtak L. J., et al., 2023, "A supermassive black hole in the early universe growing in the shadows" arXiv e-prints, p. arXiv:2308.05735 (arXiv:2308.05735)
GRAVITY Collaboration et al., 2024, "The size-luminosity relation of local active galactic nuclei from interferometric observations of the broad-line region" A&A, 684, A167 (arXiv:2401.07676)
Gardner J. P., et al., 2023, "The James Webb Space Telescope Mission" PASP, 135, 068001 (arXiv:2304.04869)
Gaskell C. M., 2009, "What broad emission lines tell us about how active galactic nuclei work" New A Rev., 53, 140 (arXiv:0908.0386)
Gebhardt K., et al., 2000, "A Relationship between Nuclear Black Hole Mass and Galaxy Velocity Dispersion" ApJ, 539, L13 (arXiv:astro-ph/0006289)
Gehrels N., 1986, "Confidence Limits for Small Numbers of Events in Astrophysical Data" ApJ, 303, 336
Gentile F., et al., 2024, "Dark progenitors and massive descendants: A first ALMA perspective of radio-selected near-IR-dark galaxies in the COSMOS field" A&A, 687, A288 (arXiv:2402.05994)
Giallongo E., et al., 2019, "Space Densities and Emissivities of Active Galactic Nuclei at z > 4" ApJ, 884, 19
Gilli R., Comastri A., Hasinger G., 2007, "The synthesis of the cosmic X-ray background in the Chandra and XMM-Newton era" A&A, 463, 79 (arXiv:astro-ph/0610939)
Gilli R., et al., 2019, "Discovery of a galaxy overdensity around a powerful, heavily obscured FRII radio galaxy at z = 1.7: star formation promoted by large-scale AGN feedback?" A&A, 632, A26 (arXiv:1909.00814)
Gilli R., et al., 2022, "Supermassive black holes at high redshift are expected to be obscured by their massive host galaxies' interstellar medium" A&A, 666, A17 (arXiv:2206.03508)
Gordon K. D., Clayton G. C., Misselt K. A., Landolt A. U., Wolff M. J., 2003, "A Quantitative Comparison of the Small Magellanic Cloud, Large Magellanic Cloud, and Milky Way Ultraviolet to Near-Infrared Extinction Curves" ApJ, 594, 279 (arXiv:astro-ph/0305257)
Goulding A. D., et al., 2023, "UNCOVER: The Growth of the First Massive Black Holes from JWST/NIRSpec-Spectroscopic Redshift Confirmation of an X-Ray Luminous AGN at z = 10.1" ApJ, 955, L24 (arXiv:2308.02750)
Grasha K., et al., 2022, "Metallicity, Ionization Parameter, and Pressure Variations of H II Regions in the TYPHOON Spiral Galaxies: NGC 1566, NGC 2835, NGC 3521, NGC 5068, NGC 5236, and NGC 7793" ApJ, 929, 118 (arXiv:2203.02522)
Greene J. E., Ho L. C., 2005, "Estimating Black Hole Masses in Active Galaxies Using the Hα Emission Line" ApJ, 630, 122 (arXiv:astro-ph/0508335)
Greene J. E., et al., 2023, "UNCOVER spectroscopy confirms a surprising ubiquity of AGN in red galaxies at $z > 5$" arXiv e-prints, p. arXiv:2309.05714 (arXiv:2309.05714)
Gültekin K., et al., 2009, "The M-σ and M-L Relations in Galactic Bulges, and Determinations of Their Intrinsic Scatter" ApJ, 698, 198 (arXiv:0903.4897)
Gutkin J., Charlot S., Bruzual G., 2016, "Modelling the nebular emission from primeval to present-day star-forming galaxies" MNRAS, 462, 1757 (arXiv:1607.06086)
Haardt F., Maraschi L., Ghisellini G., 1994, "A Model for the X-Ray and Ultraviolet Emission from Seyfert Galaxies and Galactic Black Holes" ApJ, 432, L95 (arXiv:astro-ph/9405059)
Habouzit M., et al., 2022, "Co-evolution of massive black holes and their host galaxies at high redshift: discrepancies from six cosmological simulations and the key role of JWST" MNRAS, 511, 3751 (arXiv:2201.09892)
Hale C. L., et al., 2023, "MIGHTEE: deep 1.4 GHz source counts and the sky temperature contribution of star-forming galaxies and active galactic nuclei" MNRAS, 520, 2668 (arXiv:2211.05741)
Hamann F., Korista K. T., Morris S. L., 1993, "On the Geometry, Covering Factor, and Scattering-Emission Properties of QSO Broad Absorption-Line Regions" ApJ, 415, 541
Harikane Y., et al., 2023, "JWST/NIRSpec First Census of Broad-Line AGNs at z=4-7: Detection of 10 Faint AGNs with M_BH_10^{6}-10^{7} M_sun and Their Host Galaxy Properties" arXiv e-prints, p. arXiv:2303.11946 (arXiv:2303.11946)
Harish S., Malhotra S., Rhoads J. E., Jiang T., Yang H., Knorr K., 2023, "Evidence for Black Holes in Green Peas from WISE Colors and Variability" ApJ, 945, 157
Harrison C. M., 2017, "Impact of supermassive black hole growth on star formation" Nature Astronomy, 1, 0165 (arXiv:1703.06889)
Harrison C. M., Ramos Almeida C., 2024, "Observational Tests of Active Galactic Nuclei Feedback: An Overview of Approaches and Interpretation" Galaxies, 12, 17 (arXiv:2404.08050)

Harrison C. M., et al., 2016, "The KMOS AGN Survey at High redshift (KASHz): the prevalence and drivers of ionized outflows in the host galaxies of X-ray AGN" MNRAS, 456, 1195 (arXiv:1511.00008)
Hasinger G., 2008, "Absorption properties and evolution of active galactic nuclei" A&A, 490, 905 (arXiv:0808.0260)
Hasinger G., Miyaji T., Schmidt M., 2005, "Luminosity-dependent evolution of soft X-ray selected AGN. New Chandra and XMM-Newton surveys" A&A, 441, 417 (arXiv:astro-ph/0506118)
Heckman T. M., Best P. N., 2014, "The Coevolution of Galaxies and Supermassive Black Holes: Insights from Surveys of the Contemporary Universe" ARA&A, 52, 589 (arXiv:1403.4620)
Heckman T. M., Ptak A., Hornschemeier A., Kauffmann G., 2005, "The Relationship of Hard X-Ray and Optical Line Emission in Low-Redshift Active Galactic Nuclei" ApJ, 634, 161 (arXiv:astro-ph/0507674)
Henry R. B. C., Edmunds M. G., Köppen J., 2000, "On the Cosmic Origins of Carbon and Nitrogen" ApJ, 541, 660 (arXiv:astro-ph/0004299)
Heywood I., Hale C. L., Jarvis M. J., Makhathini S., Peters J. A., Sebokolodi M. L. L., Smirnov O. M., 2020, "VLA imaging of the XMM-LSS/VIDEO deep field at 1-2 GHz" MNRAS, 496, 3469 (arXiv:2006.08551)
Hickox R. C., Alexander D. M., 2018, "Obscured Active Galactic Nuclei" ARA&A, 56, 625 (arXiv:1806.04680)
Hildebrand R. H., 1983, "The determination of cloud masses and dust characteristics from submillimetre thermal emission." QJRAS, 24, 267
Hildebrandt H., et al., 2012, "CFHTLenS: improving the quality of photometric redshifts with precision photometry" MNRAS, 421, 2355 (arXiv:1111.4434)
Hirschmann M., Charlot S., Feltre A., Naab T., Somerville R. S., Choi E., 2019, "Synthetic nebular emission from massive galaxies - II. Ultraviolet-line diagnostics of dominant ionizing sources" MNRAS, 487, 333 (arXiv:1811.07909)
Hirschmann M., et al., 2022, "Emission-line properties of IllustrisTNG galaxies: from local diagnostic diagrams to high-redshift predictions for JWST" arXiv e-prints, p. arXiv:2212.02522 (arXiv:2212.02522)
Hoaglin D. C., Mosteller F., Tukey J. W., 1983, Understanding robust and exploratory data analysis. Wiley
Hönig S. F., Kishimoto M., Antonucci R., Marconi A., Prieto M. A., Tristram K., Weigelt G., 2012, "Parsec-scale Dust Emission from the Polar Region in the Type 2 Nucleus of NGC 424" ApJ, 755, 149 (arXiv:1206.4307)
Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, "A Cosmological Framework for the Co-Evolution of Quasars, Supermassive Black Holes, and Elliptical Galaxies. I. Galaxy Mergers and Quasar Activity" ApJS, 175, 356 (arXiv:0706.1243)
Hopkins P. F., Torrey P., Faucher-Giguère C.-A., Quataert E., Murray N., 2016, "Stellar and quasar feedback in concert: effects on AGN accretion, obscuration, and outflows" MNRAS, 458, 816 (arXiv:1504.05209)
Hopkins P. F., et al., 2023, "What causes the formation of discs and end of bursty star formation?" MNRAS, 525, 2241 (arXiv:2301.08263)
Huertas-Company M., et al., 2024, "Galaxy morphology from z $\sim$ 6 through the lens of JWST" A&A, 685, A48 (arXiv:2305.02478)
Ichimaru S., 1977, "Bimodal behavior of accretion disks: theory and application to Cygnus X-1 transitions." ApJ, 214, 840
Ilbert O., et al., 2006, "Accurate photometric redshifts for the CFHT legacy survey calibrated using the VIMOS VLT deep survey" A&A, 457, 841 (arXiv:astro-ph/0603217)
Ilbert O., et al., 2009, "Cosmos Photometric Redshifts with 30-Bands for 2-deg^2" ApJ, 690, 1236 (arXiv:0809.2101)
Ilbert O., et al., 2013, "Mass assembly in quiescent and star-forming galaxies since z = 4 from UltraVISTA" A&A, 556, A55 (arXiv:1301.3157)
Inayoshi K., Nakatani R., Toyouchi D., Hosokawa T., Kuiper R., Onoue M., 2022, "Rapid Growth of Seed Black Holes during Early Bulge Formation" ApJ, 927, 237 (arXiv:2110.10693)
Inoue Y., Doi A., 2014, "Unveiling the nature of coronae in active galactic nuclei through submillimeter observations" PASJ, 66, L8 (arXiv:1411.2334)
Inoue Y., Doi A., 2018, "Detection of Coronal Magnetic Activity in nearby Active Supermassive Black Holes" ApJ, 869, 114 (arXiv:1810.10732)
Ivison R. J., et al., 2007, "AEGIS20: A Radio Survey of the Extended Groth Strip" ApJ, 660, L77 (arXiv:astro-ph/0607271)
Iwasawa K., et al., 2012, "The XMM deep survey in the CDF-S. II. A 9-20 keV selection of heavily obscured active galaxies at z > 1.7" A&A, 546, A84 (arXiv:1209.0916)
Izotov Y. I., Stasińska G., Meynet G., Guseva N. G., Thuan T. X., 2006, "The chemical composition of metal-poor emission-line galaxies in the Data Release 3 of the Sloan Digital Sky Survey" A&A, 448, 955 (arXiv:astro-ph/0511644)
Izotov Y. I., Thuan T. X., Guseva N. G., Liss S. E., 2018, "J0811+4730: the most metal-poor star-forming dwarf galaxy known" MNRAS, 473, 1956 (arXiv:1709.00202)
Izotov Y. I., Guseva N. G., Fricke K. J., Henkel C., 2019, "Low-redshift lowest-metallicity star-forming galaxies in the SDSS DR14" A&A, 623, A40 (arXiv:1902.01775)
Izotov Y. I., Guseva N. G., Fricke K. J., Henkel C., Schaerer D., Thuan T. X., 2021, "Low-redshift compact star-forming galaxies as analogues of high-redshift star-forming galaxies" A&A, 646, A138 (arXiv:2103.01505)
Jakobsen P., et al., 2022, "The Near-Infrared Spectrograph (NIRSpec) on the James Webb Space Telescope. I. Overview of the instrument and its capabilities" A&A, 661, A80 (arXiv:2202.03305)

Ji X., Yan R., 2022, "Correlation between the gas-phase metallicity and ionization parameter in extragalactic H II regions" A&A, 659, A112 (arXiv:2110.00612)
Jiang L., et al., 2016, "The Final SDSS High-redshift Quasar Sample of 52 Quasars at z>5.7" ApJ, 833, 222 (arXiv:1610.05369)
Jin C., Ward M., Done C., 2012, "A combined optical and X-ray study of unobscured type 1 active galactic nuclei - II. Relation between X-ray emission and optical spectra" MNRAS, 422, 3268 (arXiv:1203.0239)
Juodžbalis I., et al., 2024a, "A dormant, overmassive black hole in the early Universe" arXiv e-prints, p. arXiv:2403.03872 (arXiv:2403.03872)
Juodžbalis I., et al., 2024b, "JADES – The Rosetta Stone of JWST-discovered AGN: deciphering the intriguing nature of early AGN" arXiv e-prints, p. arXiv:2407.08643 (arXiv:2407.08643)
Kaasinen M., Kewley L., Bian F., Groves B., Kashino D., Silverman J., Kartaltepe J., 2018, "The ionization parameter of star-forming galaxies evolves with the specific star formation rate" MNRAS, 477, 5568 (arXiv:1804.10621)
Kakkad D., et al., 2020, "SUPER. II. Spatially resolved ionised gas kinematics and scaling relations in z ∼ 2 AGN host galaxies" A&A, 642, A147 (arXiv:2008.01728)
Kashino D., Lilly S. J., Matthee J., Eilers A.-C., Mackenzie R., Bordoloi R., Simcoe R. A., 2023, "EIGER. I. A Large Sample of [O III]-emitting Galaxies at 5.3 < z < 6.9 and Direct Evidence for Local Reionization by Galaxies" ApJ, 950, 66 (arXiv:2211.08254)
Kaspi S., Smith P. S., Netzer H., Maoz D., Jannuzi B. T., Giveon U., 2000, "Reverberation Measurements for 17 Quasars and the Size-Mass-Luminosity Relations in Active Galactic Nuclei" ApJ, 533, 631 (arXiv:astro-ph/9911476)
Kaspi S., Maoz D., Netzer H., Peterson B. M., Vestergaard M., Jannuzi B. T., 2005, "The Relationship between Luminosity and Broad-Line Region Size in Active Galactic Nuclei" ApJ, 629, 61 (arXiv:astro-ph/0504484)
Kauffmann G., et al., 2003, "The host galaxies of active galactic nuclei" MNRAS, 346, 1055 (arXiv:astro-ph/0304239)
Keller P. M., Thyagarajan N., Kumar A., Kanekar N., Bernardi G., 2024, "The radio-loud fraction of quasars at z > 6" MNRAS, 528, 5692 (arXiv:2402.08732)
Kellermann K. I., Sramek R., Schmidt M., Shaffer D. B., Green R., 1989, "VLA observations of objects in the Palomar Bright Quasar Survey" AJ, 98, 1195
Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, "Theoretical Modeling of Starburst Galaxies" ApJ, 556, 121 (arXiv:astro-ph/0106324)
Kewley L. J., Nicholls D. C., Sutherland R., Rigby J. R., Acharya A., Dopita M. A., Bayliss M. B., 2019, "Theoretical ISM Pressure and Electron Density Diagnostics for Local and High-redshift Galaxies" ApJ, 880, 16 (arXiv:1908.05504)
Killi M., et al., 2023, "Deciphering the JWST spectrum of a 'little red dot' at $z \sim 4.53$: An obscured AGN and its star-forming host" arXiv e-prints, p. arXiv:2312.03065 (arXiv:2312.03065)
Kocevski D. D., et al., 2023, "Hidden Little Monsters: Spectroscopic Identification of Low-Mass, Broad-Line AGN at $z > 5$ with CEERS" arXiv e-prints, p. arXiv:2302.00012 (arXiv:2302.00012)
Kocevski D. D., et al., 2024, "The Rise of Faint, Red AGN at $z > 4$: A Sample of Little Red Dots in the JWST Extragalactic Legacy Fields" arXiv e-prints, p. arXiv:2404.03576 (arXiv:2404.03576)
Kokorev V., et al., 2023, "UNCOVER: A NIRSpec Identification of a Broad-line AGN at z = 8.50" ApJ, 957, L7 (arXiv:2308.11610)
Kondapally R., et al., 2022, "Cosmic evolution of low-excitation radio galaxies in the LOFAR two-metre sky survey deep fields" MNRAS, 513, 3742 (arXiv:2204.07588)
Kormendy J., Ho L. C., 2013, "Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies" ARA&A, 51, 511 (arXiv:1304.7762)
Kovacs O. E., et al., 2024, "A candidate supermassive black hole in a gravitationally-lensed galaxy at $z \approx 10$" arXiv e-prints, p. arXiv:2403.14745 (arXiv:2403.14745)
Kriek M., et al., 2015, "The MOSFIRE Deep Evolution Field (MOSDEF) Survey: Rest-frame Optical Spectroscopy for ~1500 H-selected Galaxies at 1.37 < z < 3.8" ApJS, 218, 15 (arXiv:1412.1835)
La Franca F., et al., 2005, "The HELLAS2XMM Survey. VII. The Hard X-Ray Luminosity Function of AGNs up to z = 4: More Absorbed AGNs at Low Luminosities and High Redshifts" ApJ, 635, 864 (arXiv:astro-ph/0509081)
La Franca F., Melini G., Fiore F., 2010, "Tools for Computing the AGN Feedback: Radio-loudness Distribution and the Kinetic Luminosity Function" ApJ, 718, 368 (arXiv:1006.1247)
Labbe I., et al., 2021, UNCOVER: Ultra-deep NIRCam and NIRSpec Observations Before the Epoch of Reionization, JWST Proposal. Cycle 1, ID. #2561
Laigle C., et al., 2016, "The COSMOS2015 Catalog: Exploring the 1 < z < 6 Universe with Half a Million Galaxies" ApJS, 224, 24 (arXiv:1604.02350)
Lamastra A., Bianchi S., Matt G., Perola G. C., Barcons X., Carrera F. J., 2009, "The bolometric luminosity of type 2 AGN from extinction-corrected [OIII]. No evidence of Eddington-limited sources" A&A, 504, 73 (arXiv:0905.4439)
Lambrides E. L., Chiaberge M., Heckman T., Gilli R., Vito F., Norman C., 2020, "A Large Population of Obscured AGN in Disguise as Low-luminosity AGN in Chandra Deep Field South" ApJ, 897, 160 (arXiv:2002.00955)
Lambrides E., et al., 2024, "The Case for Super-Eddington Accretion: Connecting Weak X-ray and UV Line Emission in JWST Broad-Line AGN During the First Gyr of Cosmic Time" arXiv e-prints, p. arXiv:2409.13047 (arXiv:2409.13047)
Langeroodi D., Hjorth J., 2023, "Ultraviolet Compactness of High-Redshift Galaxies as a Tracer of Early-Stage

Gas Infall, Stochastic Star Formation, and Offset from the Fundamental Metallicity Relation" arXiv e-prints, p. arXiv:2307.06336 (arXiv:2307.06336)
Laor A., Behar E., 2008, "On the origin of radio emission in radio-quiet quasars" MNRAS, 390, 847 (arXiv:0808.0637)
Lapi A., Pantoni L., Boco L., Danese L., 2020, "New Analytic Solutions for Galaxy Evolution. II. Wind Recycling, Galactic Fountains, and Late-type Galaxies" ApJ, 897, 81 (arXiv:2006.01643)
Larson R. L., et al., 2023, "A CEERS Discovery of an Accreting Supermassive Black Hole 570 Myr after the Big Bang: Identifying a Progenitor of Massive z > 6 Quasars" ApJ, 953, L29 (arXiv:2303.08918)
Laseter I. H., et al., 2024, "JADES: Detecting [OIII]λ4363 emitters and testing strong line calibrations in the high-z Universe with ultra-deep JWST/NIRSpec spectroscopy up to z ∼ 9.5" A&A, 681, A70 (arXiv:2306.03120)
Lawrence A., 1991, "The relative frequency of broad-lined and narrow-lined active galactic nuclei : implications for unified schemes." MNRAS, 252, 586
Lawrence A., Elvis M., 1982, "Obscuration and the various kinds of Seyfert galaxies." ApJ, 256, 410
Lee J. C., et al., 2023, "The PHANGS-JWST Treasury Survey: Star Formation, Feedback, and Dust Physics at High Angular Resolution in Nearby GalaxieS" ApJ, 944, L17 (arXiv:2212.02667)
Lehmer B. D., et al., 2016, "The Evolution of Normal Galaxy X-Ray Emission through Cosmic History: Constraints from the 6 MS Chandra Deep Field-South" ApJ, 825, 7 (arXiv:1604.06461)
Leja J., Johnson B. D., Conroy C., van Dokkum P. G., Byler N., 2017, "Deriving Physical Properties from Broadband Photometry with Prospector: Description of the Model and a Demonstration of its Accuracy Using 129 Galaxies in the Local Universe" ApJ, 837, 170 (arXiv:1609.09073)
Lenz D. D., Ayres T. R., 1992, "Errors Associated with Fitting Gaussian Profiles to Noisy Emission-Line Spectra" PASP, 104, 1104
Li Z.-Y., Wu X.-B., Wang R., 2008, "The Black Hole Fundamental Plane: Revisited with a Larger Sample of Radio and X-Ray-Emitting Broad-Line AGNs" ApJ, 688, 826 (arXiv:0808.0541)
Li K., Bogdanović T., Ballantyne D. R., Bonetti M., 2023, "Massive Black Hole Binaries from the TNG50-3 Simulation. II. Using Dual AGNs to Predict the Rate of Black Hole Mergers" ApJ, 959, 3 (arXiv:2207.14231)
Liddle A. R., 2007, "Information criteria for astrophysical model selection" MNRAS, 377, L74 (arXiv:astro-ph/0701113)
Lilly S. J., et al., 2007, "zCOSMOS: A Large VLT/VIMOS Redshift Survey Covering 0 < z < 3 in the COSMOS Field" ApJS, 172, 70 (arXiv:astro-ph/0612291)
Liu T., et al., 2017, "X-Ray Spectral Analyses of AGNs from the 7Ms Chandra Deep Field-South Survey: The Distribution, Variability, and Evolutions of AGN Obscuration" ApJS, 232, 8 (arXiv:1703.00657)
Liu Y., et al., 2021, "Constraining the Quasar Radio-loud Fraction at z ∼ 6 with Deep Radio Observations" ApJ, 908, 124 (arXiv:2012.07301)
Longair M. S., 2011, High Energy Astrophysics
Looser T. J., et al., 2023, "JADES: Differing assembly histories of galaxies – Observational evidence for bursty SFHs and (mini-)quenching in the first billion years of the Universe" arXiv e-prints, p. arXiv:2306.02470 (arXiv:2306.02470)
López I. E., et al., 2023, "The miniJPAS survey: AGN and host galaxy coevolution of X-ray-selected sources" A&A, 672, A137 (arXiv:2302.01358)
Luo B., et al., 2017, "The Chandra Deep Field-South Survey: 7 Ms Source Catalogs" ApJS, 228, 2 (arXiv:1611.03501)
Lupi A., Quadri G., Volonteri M., Colpi M., Regan J. A., 2024, "Sustained super-Eddington accretion in high-redshift quasars" A&A, 686, A256 (arXiv:2312.08422)
Lusso E., et al., 2013, "The Obscured Fraction of Active Galactic Nuclei in the XMM-COSMOS Survey: A Spectral Energy Distribution Perspective" ApJ, 777, 86 (arXiv:1309.0814)
Lusso E., et al., 2020, "Quasars as standard candles. III. Validation of a new sample for cosmological studies" A&A, 642, A150 (arXiv:2008.08586)
Lusso E., Valiante R., Vito F., 2023, in , Handbook of X-ray and Gamma-ray Astrophysics. Edited by Cosimo Bambi and Andrea Santangelo. p. 122, doi:10.1007/978-981-16-4544-0_116-1
Lyu J., et al., 2024, "Active Galactic Nuclei Selection and Demographics: A New Age with JWST/MIRI" ApJ, 966, 229 (arXiv:2310.12330)
Madau P., Dickinson M., 2014, "Cosmic Star-Formation History" ARA&A, 52, 415 (arXiv:1403.0007)
Madau P., Haardt F., 2024, "X-Ray Weak AGNs from Super-Eddington Accretion onto Infant Black Holes" arXiv e-prints, p. arXiv:2410.00417 (arXiv:2410.00417)
Magorrian J., et al., 1998, "The Demography of Massive Dark Objects in Galaxy Centers" AJ, 115, 2285 (arXiv:astro-ph/9708072)
Maiolino R., Mannucci F., 2019, "De re metallica: the cosmic chemical evolution of galaxies" A&A Rev., 27, 3 (arXiv:1811.09642)
Maiolino R., Salvati M., Bassani L., Dadina M., della Ceca R., Matt G., Risaliti G., Zamorani G., 1998, "Heavy obscuration in X-ray weak AGNs" A&A, 338, 781 (arXiv:astro-ph/9806055)
Maiolino R., Shemmer O., Imanishi M., Netzer H., Oliva E., Lutz D., Sturm E., 2007, "Dust covering factor, silicate emission, and star formation in luminous QSOs" A&A, 468, 979 (arXiv:0704.1559)
Maiolino R., et al., 2008, "AMAZE. I. The evolution of the mass-metallicity relation at z > 3" A&A, 488, 463 (arXiv:0806.2410)

Maiolino R., et al., 2023a, "JWST-JADES. Possible Population III signatures at z=10.6 in the halo of GN-z11" arXiv e-prints, p. arXiv:2306.00953 (arXiv:2306.00953)
Maiolino R., et al., 2023b, "JADES. The diverse population of infant Black Holes at 4<z<11: merging, tiny, poor, but mighty" arXiv e-prints, p. arXiv:2308.01230 (arXiv:2308.01230)
Maiolino R., et al., 2024a, "JWST meets Chandra: a large population of Compton thick, feedback-free, and X-ray weak AGN, with a sprinkle of SNe" arXiv e-prints, p. arXiv:2405.00504 (arXiv:2405.00504)
Maiolino R., et al., 2024b, "A small and vigorous black hole in the early Universe" Nature, 627, 59 (arXiv:2305.12492)
Man A., Belli S., 2018, "Star formation quenching in massive galaxies" Nature Astronomy, 2, 695 (arXiv:1809.00722)
Mancuso C., et al., 2017, "Galaxy Evolution in the Radio Band: The Role of Star-forming Galaxies and Active Galactic Nuclei" ApJ, 842, 95 (arXiv:1705.06539)
Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, "A fundamental relation between mass, star formation rate and metallicity in local and high-redshift galaxies" MNRAS, 408, 2115 (arXiv:1005.0006)
Marchesi S., et al., 2020, "Mock catalogs for the extragalactic X-ray sky: Simulating AGN surveys with ATHENA and with the AXIS probe" A&A, 642, A184 (arXiv:2008.09133)
Marchesi S., et al., 2021, "Redshift identification of X-ray-selected active galactic nuclei in the J1030 field: searching for large-scale structures and high-redshift sources" A&A, 656, A117 (arXiv:2109.08162)
Marchesi S., et al., 2023, "LBT-MODS spectroscopy of high-redshift candidates in the Chandra J1030 field. A newly discovered z ∼ 2.8 large-scale structure" A&A, 673, A97 (arXiv:2303.13575)
Marconi A., et al., 2024, "HOMERUN a new approach to photoionization modelling. I – reproducing observed emission lines with percent accuracy and obtaining accurate physical properties of the ionized gas" arXiv e-prints, p. arXiv:2401.13028 (arXiv:2401.13028)
Mascia S., et al., 2024, "New insight on the nature of cosmic reionizers from the CEERS survey" A&A, 685, A3 (arXiv:2309.02219)
Masini A., et al., 2020, "The Chandra Deep Wide-field Survey: A New Chandra Legacy Survey in the Boötes Field. I. X-Ray Point Source Catalog, Number Counts, and Multiwavelength Counterparts" ApJS, 251, 2 (arXiv:2009.03317)
Massardi M., Bonaldi A., Negrello M., Ricciardi S., Raccanelli A., de Zotti G., 2010, "A model for the cosmological evolution of low-frequency radio sources" MNRAS, 404, 532 (arXiv:1001.1069)
Matsuoka Y., et al., 2019, "Subaru High-z Exploration of Low-luminosity Quasars (SHELLQs). X. Discovery of 35 Quasars and Luminous Galaxies at $5.7 \leq z \leq 7.0$" ApJ, 883, 183 (arXiv:1908.07910)
Matthee J., et al., 2023, "Little Red Dots: an abundant population of faint AGN at $z \sim 5$ revealed by the EIGER and FRESCO JWST surveys" arXiv e-prints, p. arXiv:2306.05448 (arXiv:2306.05448)
Mayer L., Kazantzidis S., Escala A., Callegari S., 2010, "Direct formation of supermassive black holes via multi-scale gas inflows in galaxy mergers" Nature, 466, 1082 (arXiv:0912.4262)
Mazzolari G., et al., 2024a, "New AGN diagnostic diagrams based on the [OIII]λ4363 auroral line" arXiv e-prints, p. arXiv:2404.10811 (arXiv:2404.10811)
Mazzolari G., et al., 2024b, "Narrow line AGN selection in CEERS: spectroscopic selection, physical properties, X-ray and radio analysis" arXiv e-prints, p. arXiv:2408.15615 (arXiv:2408.15615)
Mazzolari G., et al., 2024c, "Heavily obscured AGN detection: A radio versus X-ray challenge" A&A, 687, A120 (arXiv:2402.00109)
Mazzucchelli C., et al., 2017, "Physical Properties of 15 Quasars at $z \gtrsim 6.5$" ApJ, 849, 91 (arXiv:1710.01251)
Mazzucchelli C., et al., 2023, "XQR-30: Black hole masses and accretion rates of 42 $z \gtrsim 6$ quasars" A&A, 676, A71 (arXiv:2306.16474)
McAlpine K., et al., 2015, in Advancing Astrophysics with the Square Kilometre Array (AASKA14). p. 83 (arXiv:1412.5771), doi:10.22323/1.215.0083
McClymont W., et al., 2024, "The density-bounded twilight of starbursts in the early Universe" arXiv e-prints, p. arXiv:2405.15859 (arXiv:2405.15859)
Merloni A., Fabian A. C., 2002, "Coronal outflow dominated accretion discs: a new possibility for low-luminosity black holes?" MNRAS, 332, 165 (arXiv:astro-ph/0112451)
Merloni A., Heinz S., 2008, "A synthesis model for AGN evolution: supermassive black holes growth and feedback modes" MNRAS, 388, 1011 (arXiv:0805.2499)
Merloni A., Heinz S., 2013, in Oswalt T. D., Keel W. C., eds, , Vol. 6, Planets, Stars and Stellar Systems. Volume 6: Extragalactic Astronomy and Cosmology. p. 503, doi:10.1007/978-94-007-5609-0_11
Merloni A., Heinz S., di Matteo T., 2003, "A Fundamental Plane of black hole activity" MNRAS, 345, 1057 (arXiv:astro-ph/0305261)
Merloni A., et al., 2014, "The incidence of obscuration in active galactic nuclei" MNRAS, 437, 3550 (arXiv:1311.1305)
Messias H., Afonso J. M., Salvato M., Mobasher B., Hopkins A. M., 2014, "The dependency of AGN infrared colour-selection on source luminosity and obscuration. An observational perspective in CDFS and COSMOS" A&A, 562, A144 (arXiv:1312.3336)
Meyer R. A., et al., 2020, "The role of galaxies and AGN in reionizing the IGM - III. IGM-galaxy cross-correlations at z ∼ 6 from eight quasar fields with DEIMOS and MUSE" MNRAS, 494, 1560 (arXiv:1912.04314)

Mignoli M., et al., 2013, "Obscured AGN at z ∼1 from the zCOSMOS-Bright Survey. I. Selection and optical properties of a [Ne v]-selected sample" A&A, 556, A29 (arXiv:1305.6167)
Mignoli M., et al., 2019, "Obscured AGN at 1.5 < z < 3.0 from the zCOSMOS-deep Survey . I. Properties of the emitting gas in the narrow-line region" A&A, 626, A9 (arXiv:1903.11085)
Mignoli M., et al., 2020, "Web of the giant: Spectroscopic confirmation of a large-scale structure around the z = 6.31 quasar SDSS J1030+0524" A&A, 642, L1 (arXiv:2009.00024)
Miller S. H., Sullivan M., Ellis R. S., 2013, "Uncovering Drivers of Disk Assembly: Bulgeless Galaxies and the Stellar Mass Tully-Fisher Relation" ApJ, 762, L11 (arXiv:1211.3415)
Mingozzi M., et al., 2020, "SDSS IV MaNGA: Metallicity and ionisation parameter in local star-forming galaxies from Bayesian fitting to photoionisation models" A&A, 636, A42 (arXiv:2002.05744)
Mingozzi M., et al., 2024, "CLASSY. VIII. Exploring the Source of Ionization with UV Interstellar Medium Diagnostics in Local High-z Analogs" ApJ, 962, 95 (arXiv:2306.15062)
Miyaji T., Griffiths R. E., C-COSMOS Team 2008, in AAS/High Energy Astrophysics Division #10. p. 4.01
Momcheva I. G., et al., 2016, "The 3D-HST Survey: Hubble Space Telescope WFC3/G141 Grism Spectra, Redshifts, and Emission Line Measurements for ˜ 100,000 Galaxies" ApJS, 225, 27 (arXiv:1510.02106)
Morabito L. K., et al., 2022, "Identifying active galactic nuclei via brightness temperature with sub-arcsecond international LOFAR telescope observations" MNRAS, 515, 5758 (arXiv:2207.13096)
Morselli L., et al., 2014, "Primordial environment of super massive black holes: large-scale galaxy overdensities around z ∼6 quasars with LBT" A&A, 568, A1 (arXiv:1406.3961)
Mountrichas G., Buat V., Yang G., Boquien M., Burgarella D., Ciesla L., Malek K., Shirley R., 2022, "Comparison of the star formation in X-ray-selected AGN in eFEDS with that of star-forming galaxies" A&A, 663, A130 (arXiv:2205.01451)
Mushotzky R. F., et al., 2019, "The Advanced X-ray Imaging Satellite" arXiv e-prints, p. arXiv:1903.04083 (arXiv:1903.04083)
Muxlow T. W. B., et al., 2020, "The e-MERGE Survey (e-MERLIN Galaxy Evolution Survey): overview and survey description" MNRAS, 495, 1188 (arXiv:2005.02407)
Nakajima K., Maiolino R., 2022, "Diagnostics for PopIII galaxies and direct collapse black holes in the early universe" MNRAS, 513, 5134 (arXiv:2204.11870)
Nakajima K., et al., 2022, "EMPRESS. V. Metallicity Diagnostics of Galaxies over 12 + log(O/H) = 6.9-8.9 Established by a Local Galaxy Census: Preparing for JWST Spectroscopy" ApJS, 262, 3 (arXiv:2206.02824)
Nakajima K., Ouchi M., Isobe Y., Harikane Y., Zhang Y., Ono Y., Umeda H., Oguri M., 2023, "JWST Census for the Mass-Metallicity Star-Formation Relations at z=4-10 with Self-Consistent Flux Calibration and the Proper Metallicity Calibrators" arXiv e-prints, p. arXiv:2301.12825 (arXiv:2301.12825)
Nandra K., et al., 2013, "The Hot and Energetic Universe: A White Paper presenting the science theme motivating the Athena+ mission" arXiv e-prints, p. arXiv:1306.2307 (arXiv:1306.2307)
Nandra K., et al., 2015, "AEGIS-X: Deep Chandra Imaging of the Central Groth Strip" ApJS, 220, 10 (arXiv:1503.09078)
Nanni R., et al., 2020, "The deep Chandra survey in the SDSS J1030+0524 field" A&A, 637, A52 (arXiv:2003.13710)
Nesvadba N. P. H., Drouart G., De Breuck C., Best P., Seymour N., Vernet J., 2017, "Gas kinematics in powerful radio galaxies at z 2: Energy supply from star formation, AGN, and radio jets" A&A, 600, A121 (arXiv:1610.01627)
Netzer H., 2009, "Accretion and star formation rates in low-redshift type II active galactic nuclei" MNRAS, 399, 1907 (arXiv:0907.3575)
Netzer H., 2019, "Bolometric correction factors for active galactic nuclei" MNRAS, 488, 5185 (arXiv:1907.09534)
Newman J. A., Gruen D., 2022, "Photometric Redshifts for Next-Generation Surveys" ARA&A, 60, 363 (arXiv:2206.13633)
Ni Y., Di Matteo T., Gilli R., Croft R. A. C., Feng Y., Norman C., 2020, "QSO obscuration at high redshift (z ≳ 7): predictions from the BLUETIDES simulation" MNRAS, 495, 2135 (arXiv:1912.03780)
Ni Q., et al., 2021, "The XMM-SERVS Survey: XMM-Newton Point-source Catalogs for the W-CDF-S and ELAIS-S1 Fields" ApJS, 256, 21 (arXiv:2106.10572)
Nims J., Quataert E., Faucher-Giguère C.-A., 2015, "Observational signatures of galactic winds powered by active galactic nuclei" MNRAS, 447, 3612 (arXiv:1408.5141)
Novak M., et al., 2017, "The VLA-COSMOS 3 GHz Large Project: Cosmic star formation history since z 5" A&A, 602, A5 (arXiv:1703.09724)
O'Dea C. P., Saikia D. J., 2021, "Compact steep-spectrum and peaked-spectrum radio sources" A&A Rev., 29, 3 (arXiv:2009.02750)
Ono Y., et al., 2018, "Great Optically Luminous Dropout Research Using Subaru HSC (GOLDRUSH). I. UV luminosity functions at z ∼ 4-7 derived with the half-million dropouts on the 100 deg^2 sky" PASJ, 70, S10 (arXiv:1704.06004)
Ono Y., et al., 2023, "Morphologies of Galaxies at z ≳ 9 Uncovered by JWST/NIRCam Imaging: Cosmic Size Evolution and an Identification of an Extremely Compact Bright Galaxy at z 12" ApJ, 951, 72 (arXiv:2208.13582)
Osterbrock D. E., 1989, Astrophysics of gaseous nebulae and active galactic nuclei
Osterbrock D. E., Ferland G. J., 2006, Astrophysics of gaseous nebulae and active galactic nuclei
Owen F. N., 2018, "Deep JVLA Imaging of GOODS-N at 20 cm" ApJS, 235, 34 (arXiv:1803.05455)
Owen F. N., Morrison G. E., 2008, "The Deep Swire Field. I. 20 cm Continuum Radio Observations: A Crowded Sky" AJ, 136, 1889 (arXiv:0809.0314)

Pacifici C., et al., 2023, "The Art of Measuring Physical Parameters in Galaxies: A Critical Assessment of Spectral Energy Distribution Fitting Techniques" ApJ, 944, 141 (arXiv:2212.01915)
Pacucci F., Loeb A., 2022, "The search for the farthest quasar: consequences for black hole growth and seed models" MNRAS, 509, 1885 (arXiv:2110.10176)
Pacucci F., Narayan R., 2024, "Mildly Super-Eddington Accretion Onto Slowly-Spinning Black Holes Explains the X-Ray Weakness of the Little Red Dots" arXiv e-prints, p. arXiv:2407.15915 (arXiv:2407.15915)
Pacucci F., Nguyen B., Carniani S., Maiolino R., Fan X., 2023, "JWST CEERS and JADES Active Galaxies at z = 4-7 Violate the Local M -M $_\star$ Relation at >3σ: Implications for Low-mass Black Holes and Seeding Models" ApJ, 957, L3 (arXiv:2308.12331)
Padovani P., 2016, "The faint radio sky: radio astronomy becomes mainstream" A&A Rev., 24, 13 (arXiv:1609.00499)
Padovani P., Miller N., Kellermann K. I., Mainieri V., Rosati P., Tozzi P., 2011, "The VLA Survey of Chandra Deep Field South. V. Evolution and Luminosity Functions of Sub-millijansky Radio Sources and the Issue of Radio Emission in Radio-quiet Active Galactic Nuclei" ApJ, 740, 20 (arXiv:1107.2759)
Padovani P., Bonzini M., Kellermann K. I., Miller N., Mainieri V., Tozzi P., 2015, "Radio-faint AGN: a tale of two populations" MNRAS, 452, 1263 (arXiv:1506.06554)
Padovani P., et al., 2017, "Active galactic nuclei: what's in a name?" A&A Rev., 25, 2 (arXiv:1707.07134)
Page M. J., Carrera F. J., 2000, "An improved method of constructing binned luminosity functions" MNRAS, 311, 433 (arXiv:astro-ph/9909434)
Panessa F., et al., 2015, "The 1.4-GHz radio properties of hard X-ray-selected AGN" MNRAS, 447, 1289 (arXiv:1411.7829)
Panessa F., Baldi R. D., Laor A., Padovani P., Behar E., McHardy I., 2019, "The origin of radio emission from radio-quiet active galactic nuclei" Nature Astronomy, 3, 387 (arXiv:1902.05917)
Paul J. D., et al., 2024, "Radio Scrutiny of the X-ray-Weak Tail of Low-Mass Active Galactic Nuclei: A Novel Signature of High-Eddington Accretion?" arXiv e-prints, p. arXiv:2404.02423 (arXiv:2404.02423)
Peca A., et al., 2021, "X-Ray Redshifts for Obscured AGN: A Case Study in the J1030 Deep Field" ApJ, 906, 90 (arXiv:2011.05983)
Peca A., et al., 2022, "On the cosmic evolution of AGN obscuration and the X-ray luminosity function: XMM-Newton and Chandra spectral analysis of the 31.3 deg^2 Stripe 82X" arXiv e-prints, p. arXiv:2210.08030 (arXiv:2210.08030)
Peca A., et al., 2023, "On the Cosmic Evolution of AGN Obscuration and the X-Ray Luminosity Function: XMM-Newton and Chandra Spectral Analysis of the 31.3 deg^2 Stripe 82X" ApJ, 943, 162 (arXiv:2210.08030)
Pérez-González P. G., et al., 2023, "CEERS Key Paper. IV. A Triality in the Nature of HST-dark Galaxies" ApJ, 946, L16 (arXiv:2211.00045)
Perna M., Lanzuisi G., Brusa M., Cresci G., Mignoli M., 2017, "An X-ray/SDSS sample. II. AGN-driven outflowing gas plasma properties" A&A, 606, A96 (arXiv:1705.08388)
Peterson B. M., 1997, An Introduction to Active Galactic Nuclei
Petric A. O., Carilli C. L., Bertoldi F., Fan X., Cox P., Strauss M. A., Omont A., Schneider D. P., 2003, "Sensitive Observations at 1.4 and 250 GHz of z>5 QSOs" AJ, 126, 15 (arXiv:astro-ph/0304097)
Piconcelli E., Jimenez-Bailón E., Guainazzi M., Schartel N., Rodríguez-Pascual P. M., Santos-Lleó M., 2005, "The XMM-Newton view of PG quasars. I. X-ray continuum and absorption" A&A, 432, 15 (arXiv:astro-ph/0411051)
Piotrowska J. M., Bluck A. F. L., Maiolino R., Peng Y., 2022a, "On the quenching of star formation in observed and simulated central galaxies: evidence for the role of integrated AGN feedback" MNRAS, 512, 1052 (arXiv:2112.07672)
Piotrowska J. M., Bluck A. F. L., Maiolino R., Peng Y., 2022b, "On the quenching of star formation in observed and simulated central galaxies: evidence for the role of integrated AGN feedback" MNRAS, 512, 1052 (arXiv:2112.07672)
Polimera M., et al., 2022, in American Astronomical Society Meeting #240. p. 241.25
Polletta M., et al., 2007, "Spectral Energy Distributions of Hard X-Ray Selected Active Galactic Nuclei in the XMM-Newton Medium Deep Survey" ApJ, 663, 81 (arXiv:astro-ph/0703255)
Pontoppidan K. M., et al., 2022, "The JWST Early Release Observations" ApJ, 936, L14 (arXiv:2207.13067)
Popesso P., et al., 2023, "The main sequence of star-forming galaxies across cosmic times" MNRAS, 519, 1526 (arXiv:2203.10487)
Prandoni I., Seymour N., 2015, in Advancing Astrophysics with the Square Kilometre Array (AASKA14). p. 67 (arXiv:1412.6512), doi:10.22323/1.215.0067
Prandoni I., Gregorini L., Parma P., de Ruiter H. R., Vettolani G., Wieringa M. H., Ekers R. D., 2001, "The ATESP radio survey. III. Source counts" A&A, 365, 392 (arXiv:astro-ph/0010444)
Prandoni I., Guglielmino G., Morganti R., Vaccari M., Maini A., Röttgering H. J. A., Jarvis M. J., Garrett M. A., 2018, "The Lockman Hole Project: new constraints on the sub-mJy source counts from a wide-area 1.4 GHz mosaic" MNRAS, 481, 4548 (arXiv:1810.03738)
Prevot M. L., Lequeux J., Maurice E., Prevot L., Rocca-Volmerange B., 1984, "The typical interstellar extinction in the Small Magellanic Cloud." A&A, 132, 389
Proga D., 2005, "On how much X-ray and UV radiation processes are coupled in accretion disks: AGN case" ApJL, 630, L9 (Arxiv:astro-ph/0507152v1)

Pustilnik S. A., Kniazev A. Y., Perepelitsyna Y. A., Egorova E. S., 2020, "XMP gas-rich dwarfs in nearby voids: results of SALT spectroscopy" MNRAS, 493, 830 (arXiv:1912.06857)
Pustilnik S. A., Egorova E. S., Kniazev A. Y., Perepelitsyna Y. A., Tepliakova A. L., Burenkov A. N., Oparin D. V., 2021, "XMP gas-rich dwarfs in nearby voids: results of BTA spectroscopy" MNRAS, 507, 944 (arXiv:2108.12065)
Radcliffe J. F., Beswick R. J., Thomson A. P., Garrett M. A., Barthel P. D., Muxlow T. W. B., 2019, "An insight into the extragalactic transient and variable microJy radio sky across multiple decades" MNRAS, 490, 4024 (arXiv:1909.12588)
Raginski I., Laor A., 2016, "AGN coronal emission models - I. The predicted radio emission" MNRAS, 459, 2082 (arXiv:1604.00646)
Ramasawmy J., Stevens J., Martin G., Geach J. E., 2019, "A flat trend of star formation rate with X-ray luminosity of galaxies hosting AGN in the SCUBA-2 Cosmology Legacy Survey" MNRAS, 486, 4320 (arXiv:1904.07880)
Ranalli P., Comastri A., Setti G., 2003, "The 2-10 keV luminosity as a Star Formation Rate indicator" A&A, 399, 39 (arXiv:astro-ph/0211304)
Rawlings S., Saunders R., 1991, "Evidence for a common central-engine mechanism in all extragalactic radio sources" Nature, 349, 138
Reddy N. A., et al., 2015, "The MOSDEF Survey: Measurements of Balmer Decrements and the Dust Attenuation Curve at Redshifts z ~1.4-2.6" ApJ, 806, 259 (arXiv:1504.02782)
Reines A. E., Volonteri M., 2015, "Relations between Central Black Hole Mass and Total Galaxy Stellar Mass in the Local Universe" ApJ, 813, 82 (arXiv:1508.06274)
Reines A. E., Greene J. E., Geha M., 2013, "Dwarf Galaxies with Optical Signatures of Active Massive Black Holes" ApJ, 775, 116 (arXiv:1308.0328)
Rémy-Ruyer A., et al., 2014, "Gas-to-dust mass ratios in local galaxies over a 2 dex metallicity range" A&A, 563, A31 (arXiv:1312.3442)
Rhodes J., Refregier A., Groth E. J., 2000, "Weak Lensing Measurements: A Revisited Method and Application to-Hubble Space Telescope Images" ApJ, 536, 79 (arXiv:astro-ph/9905090)
Richards G. T., et al., 2002, "Spectroscopic Target Selection in the Sloan Digital Sky Survey: The Quasar Sample" AJ, 123, 2945 (arXiv:astro-ph/0202251)
Richards G. T., et al., 2004, "Efficient Photometric Selection of Quasars from the Sloan Digital Sky Survey: 100,000 z < 3 Quasars from Data Release One" ApJS, 155, 257 (arXiv:astro-ph/0408505)
Richards G. T., et al., 2006, "Spectral Energy Distributions and Multiwavelength Selection of Type 1 Quasars" ApJS, 166, 470 (arXiv:astro-ph/0601558)
Rinaldi P., et al., 2023, "MIDIS: Strong (Hβ+[O III]) and Hα Emitters at Redshift z = 7-8 Unveiled with JWST NIRCam and MIRI Imaging in the Hubble eXtreme Deep Field" ApJ, 952, 143 (arXiv:2301.10717)
Risaliti G., Maiolino R., Salvati M., 1999, "The Distribution of Absorbing Column Densities among Seyfert 2 Galaxies" ApJ, 522, 157 (arXiv:astro-ph/9902377)
Robertson B. E., et al., 2022, "Discovery and properties of the earliest galaxies with confirmed distances" arXiv e-prints, p. arXiv:2212.04480 (arXiv:2212.04480)
Salvato M., et al., 2009, "Photometric Redshift and Classification for the XMM-COSMOS Sources" ApJ, 690, 1250 (arXiv:0809.2098)
Salvato M., et al., 2011, "Dissecting Photometric Redshift for Active Galactic Nucleus Using XMM- and Chandra-COSMOS Samples" ApJ, 742, 61 (arXiv:1108.6061)
Salvato M., et al., 2018, "Finding counterparts for all-sky X-ray surveys with NWAY: a Bayesian algorithm for cross-matching multiple catalogues" MNRAS, 473, 4937 (arXiv:1705.10711)
Salvato M., Ilbert O., Hoyle B., 2019, "The many flavours of photometric redshifts" Nature Astronomy, 3, 212 (arXiv:1805.12574)
Sandage A., Véron P., Wyndham J. D., 1965, "Optical Identification of New Quasi-Stellar Radio Sources." ApJ, 142, 1307
Sanders D. B., Phinney E. S., Neugebauer G., Soifer B. T., Matthews K., 1989, "Continuum Energy Distributions of Quasars: Shapes and Origins" ApJ, 347, 29
Sanders R. L., Shapley A. E., Topping M. W., Reddy N. A., Brammer G. B., 2023, "Excitation and Ionization Properties of Star-forming Galaxies at z=2.0-9.3 with JWST/NIRSpec" arXiv e-prints, p. arXiv:2301.06696 (arXiv:2301.06696)
Scarlata C., Hayes M., Panagia N., Mehta V., Haardt F., Bagley M., 2024, "On the universal validity of Case B recombination theory" arXiv e-prints, p. arXiv:2404.09015 (arXiv:2404.09015)
Schaerer D., de Barros S., 2012, in Tuffs R. J., Popescu C. C., eds, IAU Symposium Vol. 284, The Spectral Energy Distribution of Galaxies - SED 2011. pp 20–25 (arXiv:1111.6373), doi:10.1017/S1743921312008630
Schartmann M., Meisenheimer K., Camenzind M., Wolf S., Henning T., 2005, "Towards a physical model of dust tori in Active Galactic Nuclei. Radiative transfer calculations for a hydrostatic torus model" A&A, 437, 861 (arXiv:astro-ph/0504105)
Schaye J., et al., 2015, "The EAGLE project: simulating the evolution and assembly of galaxies and their environments" MNRAS, 446, 521 (arXiv:1407.7040)
Scholtz J., et al., 2018, "Identifying the subtle signatures of feedback from distant AGN using ALMA observations and the EAGLE hydrodynamical simulations" MNRAS, 475, 1288 (arXiv:1712.02708)
Scholtz J., et al., 2023a, "GN-z11: The environment of an AGN at z =10.603" arXiv e-prints, p. arXiv:2306.09142 (arXiv:2306.09142)

Scholtz J., et al., 2023b, "JADES: A large population of obscured, narrow line AGN at high redshift" arXiv e-prints, p. arXiv:2311.18731 (arXiv:2311.18731)
Scholtz J., et al., 2024, "Net-zero gas inflow: deconstructing the gas consumption history of a massive quiescent galaxy with JWST and ALMA" arXiv e-prints, p. arXiv:2405.19401 (arXiv:2405.19401)
Seyfert C. K., 1943, "Nuclear Emission in Spiral Nebulae." ApJ, 97, 28
Shakura N. I., Sunyaev R. A., 1973, "Black holes in binary systems. Observational appearance." A&A, 24, 337
Shang Z., et al., 2011, "The Next Generation Atlas of Quasar Spectral Energy Distributions from Radio to X-Rays" ApJS, 196, 2 (arXiv:1107.1855)
Shankar F., et al., 2014, "On the Intermediate-redshift Central Stellar Mass-Halo Mass Relation, and Implications for the Evolution of the Most Massive Galaxies Since z ˷1" ApJ, 797, L27 (arXiv:1411.2597)
Shapley A. E., Sanders R. L., Reddy N. A., Topping M. W., Brammer G. B., 2023, "JWST/NIRSpec Balmer-line Measurements of Star Formation and Dust Attenuation at z 3-6" ApJ, 954, 157 (arXiv:2301.03241)
Shirazi M., Brinchmann J., 2012, "Strongly star forming galaxies in the local Universe with nebular He IIλ4686 emission" MNRAS, 421, 1043 (arXiv:1201.1290)
Signorini M., et al., 2023, "X-ray properties and obscured fraction of AGN in the J1030 Chandra field" A&A, 676, A49 (arXiv:2305.13368)
Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, "The Illustris simulation: the evolving population of black holes across cosmic time" MNRAS, 452, 575 (arXiv:1408.6842)
Silverman J. D., et al., 2008, "The Luminosity Function of X-Ray-selected Active Galactic Nuclei: Evolution of Supermassive Black Holes at High Redshift" ApJ, 679, 118 (arXiv:0710.2461)
Silverman J. D., et al., 2009, "The Environments of Active Galactic Nuclei within the zCOSMOS Density Field" ApJ, 695, 171 (arXiv:0812.3402)
Silverman J. D., et al., 2023, "Resolving Galactic-scale Obscuration of X-Ray AGNs at z $\gtrsim$ 1 with COSMOS-Web" ApJ, 951, L41 (arXiv:2306.03205)
Simet M., Chartab N., Lu Y., Mobasher B., 2021, "Comparison of Observed Galaxy Properties with Semianalytic Model Predictions Using Machine Learning" ApJ, 908, 47 (arXiv:1905.08996)
Simmonds C., Buchner J., Salvato M., Hsu L. T., Bauer F. E., 2018, "XZ: Deriving redshifts from X-ray spectra of obscured AGN" A&A, 618, A66 (arXiv:1807.01782)
Skelton R. E., et al., 2014, "3D-HST WFC3-selected Photometric Catalogs in the Five CANDELS/3D-HST Fields: Photometry, Photometric Redshifts, and Stellar Masses" ApJS, 214, 24 (arXiv:1403.3689)
Smolčić V., et al., 2017a, "The VLA-COSMOS 3 GHz Large Project: Continuum data and source catalog release" A&A, 602, A1 (arXiv:1703.09713)
Smolčić V., et al., 2017b, "The VLA-COSMOS 3 GHz Large Project: Multiwavelength counterparts and the composition of the faint radio population" A&A, 602, A2 (arXiv:1703.09719)
Smolčić V., et al., 2017c, "The VLA-COSMOS 3 GHz Large Project: Cosmic evolution of radio AGN and implications for radio-mode feedback since z 5" A&A, 602, A6 (arXiv:1705.07090)
Soltan A., 1982, "Masses of quasars." MNRAS, 200, 115
Stalevski M., Ricci C., Ueda Y., Lira P., Fritz J., Baes M., 2016, "The dust covering factor in active galactic nuclei" MNRAS, 458, 2288 (arXiv:1602.06954)
Stern J., Laor A., 2012, "Type 1 AGN at low z- I. Emission properties" MNRAS, 423, 600 (arXiv:1203.3158)
Stern D., et al., 2005, "Mid-Infrared Selection of Active Galaxies" ApJ, 631, 163 (arXiv:astro-ph/0410523)
Stiavelli M., et al., 2005, "Evidence of Primordial Clustering around the QSO SDSS J1030+0524 at z=6.28" ApJ, 622, L1 (arXiv:astro-ph/0502223)
Sutherland R. S., Dopita M. A., 1993, "Cooling Functions for Low-Density Astrophysical Plasmas" ApJS, 88, 253
Svoboda J., Douna V., Orlitová I., Ehle M., 2019, "Green Peas in X-Rays" ApJ, 880, 144 (arXiv:1810.09318)
Tacconi L. J., et al., 2018, "PHIBSS: Unified Scaling Relations of Gas Depletion Time and Molecular Gas Fractions" ApJ, 853, 179 (arXiv:1702.01140)
Talia M., Cimatti A., Giulietti M., Zamorani G., Bethermin M., Faisst A., Le Fèvre O., Smolçić V., 2021, "Illuminating the Dark Side of Cosmic Star Formation Two Billion Years after the Big Bang" ApJ, 909, 23 (arXiv:2011.03051)
Tang M., et al., 2023, "JWST/NIRSpec spectroscopy of z = 7-9 star-forming galaxies with CEERS: new insight into bright Lyα emitters in ionized bubbles" MNRAS, 526, 1657 (arXiv:2301.07072)
Taylor A. J., et al., 2024, "Broad-Line AGN at $3.5 < z < 6$: The Black Hole Mass Function and a Connection with Little Red Dots" arXiv e-prints, p. arXiv:2409.06772 (arXiv:2409.06772)
Terashima Y., Wilson A. S., 2003, "Chandra Snapshot Observations of Low-Luminosity Active Galactic Nuclei with a Compact Radio Source" ApJ, 583, 145 (arXiv:astro-ph/0209607)
Thomas A. D., et al., 2017, "Probing the Physics of Narrow-line Regions in Active Galaxies. IV. Full Data Release of the Siding Spring Southern Seyfert Spectroscopic Snapshot Survey (S7)" ApJS, 232, 11 (arXiv:1708.02683)
Thorne J. E., et al., 2021, "Deep Extragalactic VIsible Legacy Survey (DEVILS): SED fitting in the D10-COSMOS field and the evolution of the stellar mass function and SFR-M$_\star$ relation" MNRAS, 505, 540 (arXiv:2011.13605)
Topping M. W., et al., 2024, "Metal-poor star formation at z > 6 with JWST: new insight into hard radiation fields and nitrogen enrichment on 20 pc scales" MNRAS, 529, 3301 (arXiv:2401.08764)
Tozzi P., et al., 2006, "X-ray spectral properties of active galactic nuclei in the Chandra Deep Field South" A&A, 451, 457 (arXiv:astro-ph/0602127)
Tozzi G., Maiolino R., Cresci G., Piotrowska J. M., Belfiore F., Curti M., Mannucci F., Marconi A., 2023, "Unveiling hidden active nuclei in MaNGA star-forming galaxies with He II λ4686 line emission" MNRAS, 521, 1264

(arXiv:2302.04282)
Trefoloni B., Gilli R., Lusso E., Marconi A., Mazzolari G., Nardini E., Risaliti G., Signorini M., 2024, "The near-infrared SED of blue quasars: what drives the evolution of the dusty torus?" arXiv e-prints, p. arXiv:2410.10941 (arXiv:2410.10941)
Treister E., Urry C. M., 2006, "The Evolution of Obscuration in Active Galactic Nuclei" ApJ, 652, L79 (arXiv:astro-ph/0610525)
Tremonti C. A., et al., 2004, "The Origin of the Mass-Metallicity Relation: Insights from 53,000 Star-forming Galaxies in the Sloan Digital Sky Survey" ApJ, 613, 898 (arXiv:astro-ph/0405537)
Treu T., et al., 2022, "The GLASS-JWST Early Release Science Program. I. Survey Design and Release Plans" ApJ, 935, 110 (arXiv:2206.07978)
Tristram K. R. W., Burtscher L., Jaffe W., Meisenheimer K., Hönig S. F., Kishimoto M., Schartmann M., Weigelt G., 2014, "The dusty torus in the Circinus galaxy: a dense disk and the torus funnel" A&A, 563, A82 (arXiv:1312.4534)
Trouille L., Barger A. J., Cowie L. L., Yang Y., Mushotzky R. F., 2008, "The OPTX Project. I. The Flux and Redshift Catalogs for the CLANS, CLASXS, and CDF-N Fields" ApJS, 179, 1 (arXiv:0811.0824)
Trump J. R., et al., 2023, "The Physical Conditions of Emission-line Galaxies at Cosmic Dawn from JWST/NIRSpec Spectroscopy in the SMACS 0723 Early Release Observations" ApJ, 945, 35 (arXiv:2207.12388)
Übler H., et al., 2023, "GA-NIFS: A massive black hole in a low-metallicity AGN at z ∼ 5.55 revealed by JWST/NIRSpec IFS" A&A, 677, A145 (arXiv:2302.06647)
Übler H., et al., 2024, "GA-NIFS: JWST discovers an offset AGN 740 million years after the big bang" MNRAS, 531, 355
Ueda Y., Akiyama M., Hasinger G., Miyaji T., Watson M. G., 2014, "Toward the Standard Population Synthesis Model of the X-Ray Background: Evolution of X-Ray Luminosity and Absorption Functions of Active Galactic Nuclei Including Compton-thick Populations" ApJ, 786, 104 (arXiv:1402.1836)
Urry C. M., Padovani P., 1995, "Unified Schemes for Radio-Loud Active Galactic Nuclei" PASP, 107, 803 (arXiv:astro-ph/9506063)
Veilleux S., Osterbrock D. E., 1987, "Spectral Classification of Emission-Line Galaxies" ApJS, 63, 295
Venemans B. P., et al., 2015, "The Identification of Z-dropouts in Pan-STARRS1: Three Quasars at 6.5< z< 6.7" ApJ, 801, L11 (arXiv:1502.01927)
Venemans B. P., et al., 2018, "Dust Emission in an Accretion-rate-limited Sample of z ≳ 6 Quasars" ApJ, 866, 159 (arXiv:1809.01662)
Vidal-García A., Plat A., Curtis-Lake E., Feltre A., Hirschmann M., Chevallard J., Charlot S., 2024, "BEAGLE-AGN I: simultaneous constraints on the properties of gas in star-forming and AGN narrow-line regions in galaxies" MNRAS, 527, 7217 (arXiv:2211.13648)
Vignali C., Mignoli M., Gilli R., Comastri A., Iwasawa K., Zamorani G., Mainieri V., Bongiorno A., 2014, "The space density of Compton-thick AGN at z ≈ 0.8 in the zCOSMOS-Bright Survey" A&A, 571, A34 (arXiv:1409.6361)
Vito F., Gilli R., Vignali C., Comastri A., Brusa M., Cappelluti N., Iwasawa K., 2014, "The hard X-ray luminosity function of high-redshift (3 < z ≲ 5) active galactic nuclei" MNRAS, 445, 3557 (arXiv:1409.6918)
Vito F., et al., 2016, "The deepest X-ray view of high-redshift galaxies: constraints on low-rate black hole accretion" MNRAS, 463, 348 (arXiv:1608.02614)
Vito F., et al., 2018, "High-redshift AGN in the Chandra Deep Fields: the obscured fraction and space density of the sub-L_* population" MNRAS, 473, 2378 (arXiv:1709.07892)
Volonteri M., 2010, "Formation of supermassive black holes" A&A Rev., 18, 279 (arXiv:1003.4404)
Volonteri M., Dubois Y., Pichon C., Devriendt J., 2016, "The cosmic evolution of massive black holes in the Horizon-AGN simulation" MNRAS, 460, 2979 (arXiv:1602.01941)
Wada K., Schartmann M., Meijerink R., 2016, "Multi-phase Nature of a Radiation-driven Fountain with Nuclear Starburst in a Low-mass Active Galactic Nucleus" ApJ, 828, L19 (arXiv:1608.06995)
Wang Y., Wang T., Ho L. C., Zhong Y., Luo B., 2024, "The fundamental plane of black hole activity for low-luminosity radio active galactic nuclei across 1 < z < 4" arXiv e-prints, p. arXiv:2402.17991 (arXiv:2402.17991)
Weaver J. R., et al., 2022, "COSMOS2020: A Panchromatic View of the Universe to z∼10 from Two Complementary Catalogs" ApJS, 258, 11 (arXiv:2110.13923)
White R. L., et al., 1997, in American Astronomical Society Meeting Abstracts. p. 103.05
White R. L., Helfand D. J., Becker R. H., Glikman E., de Vries W., 2007, "Signals from the Noise: Image Stacking for Quasars in the FIRST Survey" ApJ, 654, 99 (arXiv:astro-ph/0607335)
Whittam I. H., et al., 2022, "MIGHTEE: the nature of the radio-loud AGN population" MNRAS, 516, 245 (arXiv:2207.12379)
Williams W. L., Röttgering H. J. A., 2015, "Radio-AGN feedback: when the little ones were monsters" MNRAS, 450, 1538 (arXiv:1503.08927)
Willott C. J., Rawlings S., Jarvis M. J., Blundell K. M., 2003, "Near-infrared imaging and the K-z relation for radio galaxies in the 7C Redshift Survey" MNRAS, 339, 173 (arXiv:astro-ph/0209439)
Willott C. J., et al., 2010, "Eddington-limited Accretion and the Black Hole Mass Function at Redshift 6" AJ, 140, 546 (arXiv:1006.1342)
Witstok J., Smit R., Maiolino R., Curti M., Laporte N., Massey R., Richard J., Swinbank M., 2021, "Assessing the sources of reionization: a spectroscopic case study of a 30× lensed galaxy at z 5 with Lyα, C IV, Mg II, and [Ne III]" MNRAS, 508, 1686 (arXiv:2109.03829)

Witten C. E. C., Laporte N., Katz H., 2023, "Evidence for a Low Lyman Continuum Escape Fraction in Three Massive, Ultraviolet-bright Galaxies at z > 7" ApJ, 944, 61 (arXiv:2301.03599)
Wu Q., Shen Y., 2022, "A Catalog of Quasar Properties from Sloan Digital Sky Survey Data Release 16" The Astrophysical Journal Supplement Series, 263, 42
Xie F.-G., Yuan F., 2017, "Fundamental Plane of Black Hole Activity in the Quiescent Regime" ApJ, 836, 104 (arXiv:1701.06143)
Xue Y. Q., Luo B., Brandt W. N., Alexander D. M., Bauer F. E., Lehmer B. D., Yang G., 2016, "The 2 Ms Chandra Deep Field-North Survey and the 250 ks Extended Chandra Deep Field-South Survey: Improved Point-source Catalogs" ApJS, 224, 15 (arXiv:1602.06299)
Yanagisawa H., et al., 2024, "Balmer Decrement Anomalies in Galaxies at z ~6 Found by JWST Observations: Density-Bounded Nebulae or Excited H I Clouds?" arXiv e-prints, p. arXiv:2403.20118 (arXiv:2403.20118)
Yang H., et al., 2017a, "Lyα Profile, Dust, and Prediction of Lyα Escape Fraction in Green Pea Galaxies" ApJ, 844, 171 (arXiv:1701.01857)
Yang H., Malhotra S., Rhoads J. E., Wang J., 2017b, "Blueberry Galaxies: The Lowest Mass Young Starbursts" ApJ, 847, 38 (arXiv:1706.02819)
Yang G., et al., 2020, "X-CIGALE: Fitting AGN/galaxy SEDs from X-ray to infrared" MNRAS, 491, 740 (arXiv:2001.08263)
Yang G., Estrada-Carpenter V., Papovich C., Vito F., Walsh J. L., Yao Z., Yuan F., 2021, "Do Current X-Ray Observations Capture Most of the Black-hole Accretion at High Redshifts?" ApJ, 921, 170 (arXiv:2109.00078)
Yang G., et al., 2023a, "CEERS Key Paper. VI. JWST/MIRI Uncovers a Large Population of Obscured AGN at High Redshifts" ApJ, 950, L5 (arXiv:2303.11736)
Yang G., et al., 2023b, "CEERS Key Paper. VI. JWST/MIRI Uncovers a Large Population of Obscured AGN at High Redshifts" ApJ, 950, L5 (arXiv:2303.11736)
York D. G., et al., 2000, "The Sloan Digital Sky Survey: Technical Summary" AJ, 120, 1579 (arXiv:astro-ph/0006396)
Yue M., Eilers A.-C., Ananna T. T., Panagiotou C., Kara E., Miyaji T., 2024, "Stacking X-ray Observations of "Little Red Dots": Implications for their AGN Properties" arXiv e-prints, p. arXiv:2404.13290 (arXiv:2404.13290)
Yung L. Y. A., Somerville R. S., Finkelstein S. L., Hirschmann M., Davé R., Popping G., Gardner J. P., Venkatesan A., 2021, "Semi-analytic forecasts for JWST - V. AGN luminosity functions and helium reionization at z = 2-7" MNRAS, 508, 2706 (arXiv:2109.13241)
Zajaček M., Tursunov A., Eckart A., Britzen S., 2018, "On the charge of the Galactic centre black hole" MNRAS, 480, 4408 (arXiv:1808.07327)
Zakamska N. L., Greene J. E., 2014, "Quasar feedback and the origin of radio emission in radio-quiet quasars" MNRAS, 442, 784 (arXiv:1402.6736)
Zappacosta L., et al., 2023, "HYPerluminous quasars at the Epoch of ReionizatION (HYPERION): A new regime for the X-ray nuclear properties of the first quasars" A&A, 678, A201 (arXiv:2305.02347)
Zeballos M., et al., 2018, "AzTEC 1.1 mm observations of high-z protocluster environments: SMG overdensities and misalignment between AGN jets and SMG distribution" MNRAS, 479, 4577 (arXiv:1806.10291)
Zeimann G. R., et al., 2015, "Hubble Space Telescope Emission-line Galaxies at z ~2: The Mystery of Neon" ApJ, 798, 29 (arXiv:1410.6159)
Zhang H., et al., 2023, "TRINITY IV: Predictions for Supermassive Black Holes at $zrsim7$" arXiv e-prints, p. arXiv:2309.07210 (arXiv:2309.07210)
Zhu S., Brandt W. N., Zou F., Luo B., Ni Q., Xue Y., Yan W., 2023, "Radio AGN selection and characterization in three Deep-Drilling Fields of the Vera C. Rubin Observatory Legacy Survey of Space and Time" MNRAS, 522, 3506 (arXiv:2304.07864)
da Cunha E., Charlot S., Elbaz D., 2008, "A simple model to interpret the ultraviolet, optical and infrared emission from galaxies" MNRAS, 388, 1595 (arXiv:0806.1020)
de Graaff A., et al., 2024, "Ionised gas kinematics and dynamical masses of z $\gtrsim$ 6 galaxies from JADES/NIRSpec high-resolution spectroscopy" A&A, 684, A87 (arXiv:2308.09742)
de Vries W. H., Morganti R., Röttgering H. J. A., Vermeulen R., van Breugel W., Rengelink R., Jarvis M. J., 2002, "Deep Westerbork 1.4 GHz Imaging of the Bootes Field" AJ, 123, 1784 (arXiv:astro-ph/0111543)
de Vries W. H., Hodge J. A., Becker R. H., White R. L., Helfand D. J., 2007, "Star Formation in Low Radio Luminosity Active Galactic Nuclei from the Sloan Digital Sky Survey" AJ, 134, 457 (arXiv:0704.2074)
van der Vlugt D., et al., 2021, "An Ultradeep Multiband VLA Survey of the Faint Radio Sky (COSMOS-XS): Source Catalog and Number Counts" ApJ, 907, 5 (arXiv:2009.13528)
van der Wel A., et al., 2014, "3D-HST+CANDELS: The Evolution of the Galaxy Size-Mass Distribution since z = 3" ApJ, 788, 28 (arXiv:1404.2844)