ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Dipartimento di Fisica e Astronomia

DOTTORATO DI RICERCA IN
ASTROFISICA

Ciclo XXXV

Tesi di Dottorato

Obscured AGN in deep fields: tracing accretion and star-formation up to $z \sim 5$

Presentata da: **Luigi Barchiesi**

Coordinatore Dottorato:
Prof. **Andrea Miglio**

Supervisore:
Prof. **Cristian Vignali**

Co-supervisore:
Prof. **Francesca Pozzi**

Esame finale anno 2023

Settore Concorsuale: 02/C1 – Astronomia, Astrofisica, Fisica della Terra e dei Pianeti
Settore Scientifico Disciplinare: FIS/05 – Astronomia e Astrofisica

Abstract

During their evolution, galaxies undergo a phase in which the emission from their core (nucleus) can be thousands of times brighter and more energetic that the emission of the entire galaxy. The galaxies in this phase are known as Active Galactic Nuclei (AGN). Their emission ranges from the radio to X-ray wavelengths, and sometimes up to the γ-ray band. Its driving mechanism is the combination of the presence of a Supermassive Black Hole (SMBH), at the center of the galaxy, and gas infalling into it.

The discovery of scaling relations between the mass of the SMBH and some key physical properties of the host galaxy (such as stellar velocity dispersion, stellar mass, luminosity, e.g., Kormendy & Richstone 1995; Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Kormendy & Ho 2013) suggests that the growth of the SMBH and that of the galaxy are coupled, with the AGN activity and the star-formation (SF) processes influencing each other. This led to the formulation of the AGN-galaxy co-evolution paradigms (e.g., Hopkins et al. 2007a; Lapi et al. 2014, 2018). Although the mechanism of this co-evolution and the involved timescales are still a matter of debate, all scenarios agree that a key phase of the co-evolution is represented by the obscured accretion phase, where the high quantity of gas, fueling the star-formation and the AGN accretion, also hides the nuclear region of the galaxy, thus the AGN activity. This phase is of crucial importance in the co-evolution scenarios; however, it is also the least studied, mostly due to the challenge in detecting and recognizing such obscured AGN.

My thesis aims at investigating the AGN-galaxy co-evolution paradigm by identifying and studying AGN in the obscured accretion phase. Sources in this first phase of the coeval growth, characterized by the presence of obscured SF and hidden AGN activity, are the most difficult to identify (section 1.5). Nevertheless, the study of obscured AGN is key for our understanding of the feedback processes and of the mutual influence of the SF and the AGN activity. Moreover, these obscured and elusive AGN are needed to explain the X-ray background (XRB) spectrum (section 2.4.1) and to reconcile the measurements and the theoretical prediction of the BH accretion rate density (BHAD, section 1.4.3).

The review in chapter 1 is aimed at providing a broad overview of the AGN, of their emission mechanisms, and the importance of studying the most obscured ones to unveil the obscured accretion phase of the AGN-galaxy co-evolution paradigm.

In chapter 2, we address the methods to identify and study obscured AGN up to very high redshift. In particular, we investigate the synergies between future IR and X-ray missions in detecting and characterizing AGN, with a particular attention to the most obscured ones. We performed simulations to settle the strategy for a future IR mission, to quickly estimate the number of detectable AGN depending on their luminosity, obscuration, and redshift. We specifically focus on the synergies between the ESA X-ray observatory *Athena*, and a set of template mission based on proposed IR observatories, i.e., *SPICA*, *ORIGINS*, and *PRIMA*.

In chapter 3 and 4, we exploit UV/optical emission lines to select high-redshift obscured AGN in the COSMOS field. In particular, the first method (chapter 3) involves the use of the [Ne v] 3426Å line to select obscured AGN at redshift $z \sim 1$. The second one (chapter 4) exploits the C IV 1550Å line selection to extend the redshift range up to $z \sim 3$. The advantages of selecting AGN using the [Ne v] and C IV emission lines and sampling two different redshift ranges are twofold: on the one hand, it allows us to cover the entire cosmic noon (from $z = 0.6$ to $z = 3$), where the highest SFR density and BH accretion rate density are expected; on the other hand, it allows us to trace the evolution of the obscured AGN population by comparing the results of the two selection methods. We also provide X-ray spectral analysis of their *Chandra* spectra and UV-to-far-IR SED-fitting. These analyses allow us to obtain the AGN obscuration and intrinsic luminosities, as well as the host-galaxies stellar mass (M_*), star formation rate (SFR), and molecular gas mass. We show that our samples host a significant fraction (up to 60 % for the C IV sample) of very obscured sources; many of these are highly accreting (even at over Eddington ratios). By analyzing these properties and comparing them with those of control samples of non-active galaxies and with the expectation from the *in-situ* co-evolution models, we found that the majority of our sources are in the obscured accretion phase.

An in-depth investigation using multi-wavelength approaches can be extremely useful to reveal AGN accretion in apparently normal (i.e., non active) galaxies. In this regard, in chapter 5, we focus on a sample of SF-galaxies at $z \sim 5$ with [C II] 158μm line detection, to investigate the molecular gas content at high redshift, i.e., the fuel for the SF and the AGN activity. We perform a thoughtful spectrophometric investigation of the source with the lowest molecular gas fraction, as it exhibits unusual and peculiar features. The detection of a P-Cygni in the N v emission line profile and the SED-fitting results lead us to identify a second extremely young population of stars, alongside the old one that comprises most of the stellar mass of the galaxy. In addition, the spectral line ratios and the comparison with photo-ionization models and with the literature suggest that this galaxy hosts hidden AGN activity.

Finally, we draw the conclusion of this Thesis in section 6 alongside the future perspectives of our work.

Contents

Contents iv

1 AGN - Active Galactic Nuclei **1**
1.1 AGN structure . . . 1
1.2 AGN classification . . . 8
1.3 AGN emission: X-ray to IR to Radio . . . 10
1.4 AGN-galaxy co-evolution . . . 15
1.4.1 Mergen-driven scenario . . . 15
1.4.2 Alternative: in-situ co-evolution . . . 16
1.4.3 BHAD and SFRD . . . 17
1.5 Obscured accretion . . . 18
1.5.1 The X-ray Background . . . 19

2 Searching for the most obscured AGN: *Athena* – IR synergies **21**
2.1 Selecting obscured AGN . . . 21
2.2 X-ray and IR synergies in detecting obscured AGN . . . 22
2.3 Athena and future IR observatories . . . 25
2.3.1 Athena . . . 25
2.3.2 SPICA-like . . . 26
2.3.3 The *Origins Space Telescope* . . . 27
2.3.4 PRIMA . . . 28
2.3.5 Surveys comparison . . . 29
2.4 Simulations . . . 31
2.4.1 Simulation of the intrinsic AGN counts from XRB synthesis model 31
2.4.2 Simulation of IR detections . . . 33
2.5 Results . . . 35
2.5.1 Predictions of IR source counts . . . 35
2.5.2 Photometric detections . . . 39
2.5.3 Spectroscopic detections . . . 42
2.5.4 SED fitting . . . 48
2.5.5 CT-AGN . . . 52

2.6 Conclusions 55

3 [NeV]-selected obscured AGN at $z \sim 1$ 59
3.1 Introduction 59
3.1.1 The [Ne v] 3426Å selection 60
3.2 The sample 60
3.2.1 X-ray Data 61
3.2.2 Photometric Data 61
3.3 X-ray spectral analysis 62
3.3.1 X-ray detected sources 62
3.3.2 X-ray undetected sources 66
3.4 SED analysis 67
3.4.1 SED-fitting algorithm 68
3.4.2 SED-fitting results 71
3.5 Comparison with the control sample 82
3.6 Interpretation within the *in-situ* co-evolution model 85
3.7 Conclusions 86

4 C IV-selected obscured AGN at $z \sim 2-3$ 89
4.1 Introduction 89
4.1.1 The C IV-selected sample 89
4.1.2 X-ray Data 91
4.1.3 Photometric Data 91
4.2 X-ray spectral properties 91
4.2.1 X-ray undetected sources 93
4.3 X-ray to IR SED-fitting 96
4.3.1 CIGALE 96
4.3.2 X-ray fluxes 98
4.3.3 The control sample 100
4.4 Comparison C IV and control samples 102
4.4.1 SFR-M_* relation 102
4.4.2 SSFR and MS ratio 103
4.5 Comparison C IV and [Ne v] samples 104
4.5.1 Stellar mass 105
4.5.2 SFR 106
4.5.3 SFR-M_* relation 106
4.5.4 SSFR and MS ratio 107
4.5.5 AGN bolometric luminosity 108
4.5.6 M_{BH} and Eddington ratio 110
4.6 Conclusion 111

5 Double stellar population and AGN activity in a galaxy at $z \sim 5.5$ 113
5.1 Introduction . . . 113
5.1.1 ALPINE survey . . . 114
5.2 The ALPINE f_{mol} outliers . . . 114
5.3 GS-14 . . . 117
5.4 Spectral analysis . . . 120
5.4.1 Spectroscopic data . . . 120
5.4.2 Spectral fitting . . . 121
5.4.3 Comparison with literature . . . 122
5.4.4 Comparison with theoretical predictions . . . 123
5.5 Photometric analysis . . . 125
5.5.1 Photometric data . . . 126
5.5.2 SED fitting . . . 126
5.6 Conclusion . . . 129
5.7 GS-14 as archetype of high-z JWST sources . . . 132

6 Conclusions 133

A Appendices 139
A.1 `CIGALE` flux χ distributions . . . 139

Bibliography 145

Chapter 1

AGN - Active Galactic Nuclei

Around 10% of the galaxies are known to host an active nucleus in their core. The nuclear emission can be orders of magnitude higher than the stellar emission of the galaxy and is originated in a very small region at the center of the galaxy, hence the name “active nuclei”. As it is believed that all galaxies harbors a SMBH at their core (with the exception of rare cases in which merger interactions may kick out one of the SMBH), the fraction of active galaxies is reflecting the *duty cycle* of the AGN, i.e. the fraction of time during which the SMBH is active.

AGN are one of the most powerful energy sources of the Universe, and thanks to their extreme luminosity have been the only objects detectable at high redshift. At the present, they still are a fundamental laboratory to investigate and understand the formation and the evolution of both galaxies and SMBH in the early Universe.

In this chapter we provide a brief overview of AGN and the physical mechanism that comes into play, as well as of the open question related to the obscured ones. The AGN structure is introduced in section 1.1; in section 1.2 we will give an overview of the AGN classification, as well as the naming convention we use in this thesis; section 1.3 explains how each part of the AGN contributes to the total emission. The AGN galaxy co-evolution is tackled in section 1.4, while section 1.5 focuses on the obscured accretion phase, the most elusive phase of the co-evolution paradigm and the main goal of this thesis.

1.1 AGN structure

At present, the general structure of AGN is thought to be known; however the exact details are still matter of discussion and investigation.

At the very center of an AGN lies a SMBH, accreting material is swirling around it in the form of an accretion disk or of an accretion flow. Around the disk, and maybe connected to it, there is an obscuring torus, mainly composed of gas and dust. Around the central SMBH and photo-ionized by the central engine, two regions can be identified and classified on the bases of the type of emission lines they produce: The Broad Line Region and the Narrow Line Region. The former is the innermost of the two and can be hidden by the obscuring torus along some lines of sight, the latter is farther than the torus, and always visible. Some AGN ($\sim$ 10%) have also jets of relativistic charged particles, originated near the nucleus and extending almost perpendicular to the disk possibly up to Mpc distance. The jet-medium interaction can form extended radio emission region, known as lobes. The densest and most luminous regions of the lobes are called hot spots. In the following paragraphs more details will be given regarding each AGN component.

Super Massive Black Hole

The SMBH is responsible for the huge emission of the AGN, via the emission of a fraction of the gravitationally energy of the material infalling into it. Its mass is in the range $M_{SMBH} \sim 10^{6-10} M_\odot$ (e.g, Padovani et al. 2017). Several way to estimate the SMBH mass are available: a first distinction is between *direct* and *indirect* methods.

- Direct measurements are those in which the mass is derived from the dynamics of stars and/or gas accelerated by the black hole itself. Direct methods include stellar and gas dynamical modeling and reverberation mapping.

- Indirect methods are those where the black hole mass is inferred from observables that correlate with the black hole mass. This includes masses based on correlations between M_{BH} and the host-galaxy properties, such as stellar velocity dispersion (Ferrarese & Merritt 2000), bulge luminosity (Kormendy & Richstone 1995), and M_* (Kaspi et al. 2000).

As for now, the most accurate and reliable BH mass measurements are based on the study of the motions of individual sources, accelerated by the gravity of the BH. For example, observations of the proper motions and radial velocities of individual stars near Sgr A* (Genzel et al. 2010) or of individual mega-maser sources in M106 (Herrnstein et al. 2005). However, these direct methods rely on resolved sources with measured velocities, hence, can only be used for few nearby SMBHs. On the other hand, the *Reverberation Mapping* method, relies on measuring the delay between the core (continuum emission) and BLR line-emission variability to assert the BLR distance from the core, and can be used for more distant BH. The SMBH mass can be calculated as:

$$M_{SMBH} = f(\frac{\Delta V^2 R_{BLR}}{G}) \tag{1.1}$$

where ΔV is the line width and $R_{BLR} = c\tau$, where τ is the reverberation delay. The effects of everything unknown — the BLR geometry, kinematics, and inclination — are then included into the dimensionless factor f, which will be different for each AGN, but is expected to be of order unity (e.g., Padovani et al. 2017), the most impacting problems in using the *reverberation mapping* method are the following: Firstly, it relies on an unpredictable variability, secondly, due to the fact that it is based on emission lines originated from the BLR, it cannot be used for type 2 AGN.

Using the *Reverberation Mapping*, a scaling relation between the distance of the BLR (from the core) and its luminosity has been found $R_{\rm BLR} \propto L^{\alpha}_{\rm BLR}$ (Kaspi et al. 2000), with $\alpha = 0.67 \pm 0.05$ for the optical continuum and the broad Hβ luminosity (Kaspi et al. 2005). Again this allows us to estimate the SMBH mass from equation 1.1. Despite the higher uncertainties, linked to the dispersion of the scaling relation, this indirect method relies only on the measure of the luminosity and profile of the BLR lines.

We already mentioned that the energy emitted from the AGN comes from the liberation of the gravitational potential energy of the infalling matter. Considering a particle of mass m falling in from infinity, we can write

$$\mathrm{E} = \frac{1}{2} m v^2 = \frac{\mathrm{GM}m}{\mathrm{R}} \tag{1.2}$$

where M is the SMBH mass and R is the radius of innermost stable orbit of the infalling material in the accretion. If the rate at which mass is accreted onto the BH is $\frac{\mathrm{dm}}{\mathrm{dt}} = \dot{M}$, the luminosity of the AGN disk becomes

$$\mathrm{L} = \frac{1}{2} \dot{M} v^2 = \frac{\mathrm{GM}\dot{M}}{\mathrm{R}} \tag{1.3}$$

If we call η the efficiency of the accretion process, i.e. the fraction of the incoming energy emitted by the AGN, $\eta = \frac{\mathrm{L}}{\dot{M}c^2}$, we find that

$$\eta = \frac{\mathrm{G}\dot{M}}{\mathrm{Rc}^2} \tag{1.4}$$

Typically, for AGN a value of $\eta \approx 0.1$ is assumed. We can compare it with the efficiency of H fusion processed, typical of star cores, which results $\eta \approx 0.007$.

With the advent of the Event Horizon Telescope (EHT), SMBHs have been studied with “direct imaging”. The combination of an event horizon and strong lensing near black holes is predicted to produce distinctive feature in their images. In particular, simulated images of black holes typically have a central brightness depression (BH shadow) encircled by a bright emission ring, near the gravitationally lensed photon orbit. Although the BH shadow is not the image of the BH itself (being $\sim$ 10 times larger than the BH gravitational radius $r_g = G\, M_{\rm BH}\, c^{-2}$, Event Horizon Telescope Collaboration et al. 2019), it remains a unique features directly connected with the BH and evidence of its presence. At the present day, the EHT collaboration has produced the images of the two largest (in the sky plane) SMBH (Fig. 1.1): The one in the active galaxies M87 (M87*, Event Horizon Telescope Collaboration et al. 2019, and references therein), and the one in our own Galaxy (Sgr A*, Event Horizon Telescope Collaboration et al. 2022, and references therein). The expanding capabilities of the EHT are aimed at producing new images and even reconstructing the inner motion around the SMBH in the upcoming years (e.g., La Bella et al. 2023).

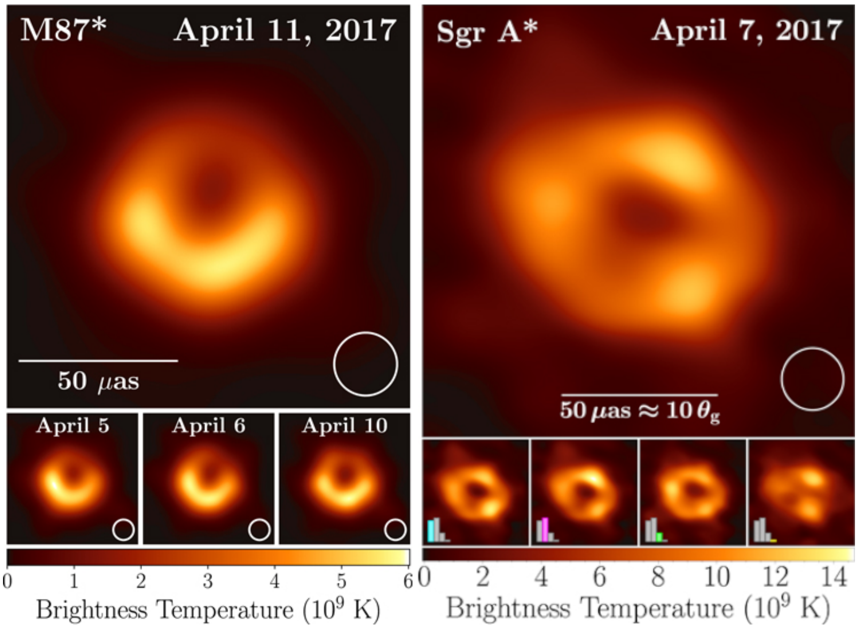

Figure 1.1: EHT images of M87* (*left*) and Sgr A* (*right*). In both images the ring-like structure is clearly visible around a central brightness depressed region (the *BH shadow*). *On the left*, the reconstruction of the M87* images for four different days of the 2017 EHT campaign. The imaging method has multiple parameters that lead to different final images. These images were classified in four groups on the basis of the luminosity distribution and shape. The mean images for each groups for the April 7$^{\text{th}}$ 2017 Sgr A* campaign are displayed *on the right*, alongside the mean of all the images.

Accretion disk

The matter infalling into the SMBH forms an accretion flow around the SMBH, that may take the form of an accretion disk, given the orbital momentum conservation law. The Innermost Stable Circular Orbit (ISCO) "delimits" the inner radius of the disk and it is strictly correlated with the spin of both the SMBH and the disk. One of the first and most widespread disk model is the Shakura & Sunyaev (S&S) model (Shakura & Sunyaev 1973), with a geometrically thin and optically thick disk. The disk temperature is around T$\approx 10^{5-6}$ K and decreases with the distance from the SMBH as $\propto r^{-3/4}$. This disk emits via thermal Black Body (BB) mechanism, mainly in the optical and UV band. Another popular model is the Advection Dominated Accretion Flow (ADAF) (Ichimaru 1977), characterized by a geometrically thick and optically thin disk.

S&S and ADAF disks are respectively characterized by high and low accretion efficiency. The first seems to well reproduce Seyfert and QSO emission, while the latter should be associated with LLAGN (see section 1.2).

Hot Corona

The hot corona is a low density region of gas presumably over the disk, although its exact position and form depend heavily on the model. Recent X-ray reverberation studies (De Marco et al. 2013; Reis & Miller 2013) suggest that the size of the hot corona producing the X-rays is in the range $3 - 20\,r_g$, while dynamical microlensing on lensed quasars restricted the hot corona dimension to a radius $< 30\,r_g$ (Chartas et al. 2016), where r_g is the gravitational radius defined in section 1.1.

The Hot Corona has a temperature around T$\approx 10^{8-9}$ K. Due to the fact that the electrons in this region are so energetic, the disk BB photons are up-scattered via Inverse Compton several times when they reach the Hot Corona, leading to a power law spectrum, that represents the main feature of the AGN X-ray spectra. This power-law spectrum has an exponential cut-off between several tens and few hundreds keV, due to the fact that the photons' energies are similar to the electrons' one and cannot be further up-scattered. The cut-off energy depends on the electron temperature and the optical depth of the corona (e.g., Lanzuisi et al. 2019). Recent results from the Swift-BAT sample (Ricci et al. 2018) show that the average cut-off energy of the sample anti-correlates with the Eddington ratio*.

Broad Line Region

At a distance of about 0.1-1 pc from the central BH there is a region where dense clouds lie (Mathews & Capriotti (1985)), which are are excited by the central engine. These clouds have temperatures around T$\approx 10^4$ K and densities $n \approx 10^{10-11}$ cm^{-3}, high enough to suppress several forbidden lines (due to these states collisionally de-excitating before they have the chance to emits). The motion of these clouds around the SMBH, at velocity around 10^{3-4} km/s, leads to the broadening of the line profiles, from which this region takes its name, that can be used to estimate the velocity of these clouds, hence the SMBH mass.

The BLR is inside the torus so this region can be seen only at high equatorial angles (i.e. close to face-on position), where the line of sight does not intercept the torus. This explains why an obscured AGN does not show broad lines. However, sometimes, also

*The Eddington ratio ($\lambda_{\rm Edd}$) is the ratio between the object luminosity and the Eddington luminosity ($L_{\rm Edd}$). The latter is defined as $L_{\rm Edd} = 4\pi G\,M\,m_{\rm p}\,c\,/\,\sigma_{\rm T}$, where M is the object (i.e, BH) mass, $m_{\rm p}$ is the proton mass, and $\sigma_{\rm T}$ is the Thomson scattering cross-section. The Eddington luminosity is the theoretical maximum luminosity that an object can achieve keeping the balance between the outward force of the radiation and the gravitational pull. It is a theoretical limits and relies on two main assumptions: spherical symmetry and that the protons and electrons are bounded (i.e., the force acting on one affects the other one as well). Due to these assumptions, the Eddington limit can be in fact exceeded (for example, in case of a non spherical accretion, like the one of an accretion disk), however, it remains a useful way of placing a theoretical upper boundary to the AGN emission.

obscured AGN can show broad lines in polarized light (Antonucci & Miller (1985)). This is explained thanks to the circumpolar gas that scatters and polarizes, part of the BLR emission.

Obscuring torus

The obscuring torus is composed primarily of gas and dust and spans from ≈ 1 to 100 pc from the central BH (e.g., Jaffe et al. 2004). The exact nature and form of the torus is still matter of debate, but we can generally view it as composed of dense clouds, with column density $N_H \approx 10^{22-24}\,cm^{-2}$ (Nenkova et al. 2002). Due to high column density, few optical and UV photons can reach us when the line of sight intercepts them. For higher column density even the hard X-ray emission can be heavily dimmed.

This obscuration along certain lines of sight is the basic concept under the unified model, where most of the differences see among AGN types are ascribed to the orientation of the torus along our line of sight.

There are two main families of torus models: *smooth torus* and *clumpy torus* (see Fig. 1.2). The former were the first to be developed, being computationally simpler and, in many aspects, a good approximation when calculating the Spectral Energy Distribution of an AGN. The latter are more recent, more complex, and a more likely representation of the real dust distribution, as a smooth dust would result in collisions that would raise the temperature to levels too high for the dust to survive (e.g., Krolik & Begelman 1988). On the one hand, the *smooth torus* models describe the torus as composed by a smooth distribution of dust and gas (e.g., Pier & Krolik 1992; Stenholm 1994; Efstathiou et al. 1995; Manske et al. 1998; Fritz et al. 2006), with the temperature depending on the distance from the nucleus. On the other hand, in the *clumpy torus* models the torus is composed by a multitude of small clouds, the temperature and density are not directly associated to the distance (e.g., Nenkova et al. 2002; Mor et al. 2009). Moreover, in clumpy models there is not a well defined angle that separates type 1 from type 2 AGN. It is possible to have low equatorial angles with direct view into the core and higher angles obscured by several clouds. The amount of obscuration is, then, a matter of the number of clouds intercepted (Alonso-Herrero et al. 2011), rather than only dependent on the viewing angle. The success of both classes of models in fitting different parts of the observed AGN SEDs keeps the issue of the dust distribution in AGN open, as no conclusions can be drawn from the simple comparison between observed and model SED (Feltre et al. 2012).

With the high spatial resolution provided by the advent of the Atacama Large Millimeter/submillimeter Array (ALMA), resolving the obscuring torus scales for nearby galaxies has become possible. In particular, the CO(3 − 2) observations of 7 Seyfert galaxies with a resolution of 0.1 arcsec (which corresponds to $\sim 4 - 9$ pc) has revealed the existence of small scale circumnuclear structures, that have been associated with molecular tori (Combes et al. 2019). These structures varies in radius between 6 and 27 pc, three show

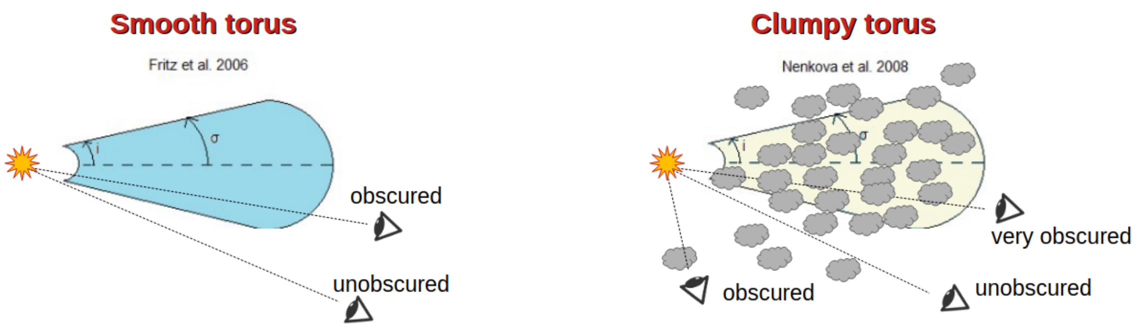

Figure 1.2: Schematic representation of AGN *smooth torus* (*left*) and *clumpy torus* (*right*) models.

a gas hole in the center resembling a torus morphology, and they could be responsible for the nuclear obscuration.

Narrow Line Region

The Narrow Line Region is composed of clouds with low density ($n \approx 10^{3-4}$ cm^{-3}) located between $\approx$ 100 pc and 1 kpc (Capetti et al. 1996) from the central BH and has temperatures around T$\approx$ 10^{3-4} K. Being less dense means that forbidden lines can be emitted by these clouds. Moreover, being these regions external to the torus, the emission lines are never obscured by it, regardless of the line of sight. Due to the fact that the clouds are so far away from the central BH, they have slower speed (a few hundreds-1000 km/s) and narrow emission line profiles, hence the name of the region. As the NLR should be photoionized by the nuclear emission, it should have a (bi)conical morphology, due to the light cones defined by the torus. High-resolution, narrow band imaging (or integral field spectroscopy) has indeed revealed such ionization cones on scales ranging from a few tens pc up to several hundreds pc, in many nearby AGN (Pogge 1988; Evans et al. 1991; Wilson & Tsvetanov 1994; Barbosa et al. 2009).

Jets and Lobes

A fraction ($\approx$ 15%, but depending on the bolometric luminosity, e.g., Urry & Padovani 1995) of AGN are characterized by significant radio emission and defined as *Radio Loud* AGN. Around 10% of them has extended radio emission. This comes from two types of structure: jets and lobes. The jets are composed of plasma of relativistic particles. These charged particles are accelerated up to ultra-relativistic speed by the magnetic field in the innermost regions of the AGN and "expelled" in polar directions (with respect to the disk plane), forming two structures known as jets. The jets emit in the radio band via synchrotron emission and up to the γ-band (in case of Doppler boosting) via Inverse Compton.

The jets can remain ultra-relativistic up to kpc scales, then have strong interaction with the medium and became very luminous (FRI), or up to Mpc scales, in case of FRII sources (see section 1.2 for AGN classification). The latter have less luminous and narrower jets, due to the fact that the jets have fewer interactions with the medium. In case of FRII, the jet-medium interaction produces large radio-lobes and hot-spots. These are regions where the jets collide and interact with the external medium, producing shocks and strong synchrotron emission.

1.2 AGN classification

The complex structure of the AGN resulted in several different classifications being developed, mainly based on the band in which AGN were observed and the resulting properties. The first AGN spectrum dates to 1908, although E. A. Fath classified it as a nebula with strong emission lines.

The first scientist who classified AGN as a different type of object from other known sources was C. Seyfert in 1943, who studied and classified objects with strong emission lines, the broadening of which suggested of velocity up to thousands of km/s. These objects are now called Seyfert galaxies.

In the '50s, the first radio survey identified a peculiar class of object, with optical properties similar to normal galaxies but being point-like and having unknown emission lines. Years later these lines were identified as highly redshifted lines. This meant that these objects, known as QUASAR (QUASi-stellAR radio sources), were in fact far away galaxies, with before-unknown high luminosity.

From the '60s new types of AGN were discovered and this lead to many different classifications, based on the observed properties.

Radio classification

AGN are classified as Radio Loud (RL) or Radio Quiet (RQ), depending on their radio-to-optical[†] flux ratio R_{r-o} (Kellermann 1989). AGN with $R_{r-o} > 10$ are RL, while the others RQ. This classification goes all the way back to Sandage (1965), who realized, soon after the discovery of the first quasar, 3C 273, that there were many similar sources in the sky previously undetected. It was later understood that these quasars were only "radio-faint", but the name stuck (Padovani et al. 2017).

Other flux ratios can also be used to separate these two classes. For example, the ratio between radio and X-ray luminosity, defined as $R_X = \log \frac{\nu L_\nu(5\text{GHz})}{L_X}$. RL AGN have $R_X \leq -4.5$ (Terashima (2005)).

The most used radio AGN classes are:

[†]The optical flux density is measured at 4400 Å and the radio flux density at 5 GHz (rest-frame).

Radio Loud

- Radio galaxies: usually associated to giant elliptical galaxies, for which the most powerful are in cluster cores. Radiogalaxies are dominated by radio non thermal emission. Their bolometric luminosities can reach 10^{47} erg/s. Radio-galaxies can be further divided in FRI and FRII, depending on their radio morphology and radio power (Fanaroff & Riley 1974). FRI have $P_{1.4GHz} < 10^{24}$ W/Hz and the jets dominates their emission . As we already mentioned, these jets are relativistic only up to kpc scales and heavily influenced by their interaction with the medium. FRII have $P_{1.4GHz} \geq 10^{24}$ W/Hz and their emission is dominated by the lobes. Their jets remain relativistic up to Mpc scales.

- Radio Loud Quasars: Very luminous, they can reach $L_{bol} \approx 10^{48}$ erg/s and their optical counterparts appear as point-like sources with strong emission lines.

- BL Lac: Similar to Quasars but without emission lines. Their optical spectra are dominated by a flat continuum. BL Lac are highly variable sources. Together with Optical Violent Variable (OVV) and Flat Spectrum Radio Quasar (FSRQ), they form the Blazar class.

Radio Quiet

- Radio Quiet Quasars: Similar to Radio Loud Quasar but without strong radio emission.

- Seyfert: Usually found in spiral galaxies. They have bright nuclei that produce strong emission lines (Seyfert 1943) but faint radio emission ($\sim 10^{20-23}$ W/Hz). Due to their relatively low luminosity it is usually possible to observe the host galaxy. Seyfert are further divided in Seyfert I or Seyfert II, based on the presence of both broad and narrow lines (Seyfert I) or only narrow lines (Seyfert II).

- LINER: Low Luminosity Narrow Emission-line Region are very low luminosity ($L_{bolo} \approx 10^{40-42}$ erg/s) objects. Found in spiral galaxies, LINERs are characterized by the presence of low-ionization narrow emission lines. Their exact nature is still a matter of debate, in particular, their emission could be explained with low luminosity AGN or Starburst events and SN explosions.

Optical/X-ray classification

The AGN optical classification is based on the presence and properties of the optical emission lines. AGN with both broad emission lines (up to $\Delta v_{FWHM} \approx 10^4$ km/s) and narrow emission lines (up to $\Delta v_{FWHM} \approx 10^3$ km/s) are classified as type 1 or unobscured, whereas AGN with only narrow emission lines are classified as type 2 or obscured. From an X-ray

Table 1.1: A schematic view of AGN optical and radio classification.

	Type 2	Type 1
Radio Quiet	Seyfert 2 LINER	Seyfert 1 RQ Quasar
Radio Loud	NLRG (FR I + FR II)	BLRG Radioquasar (SSRQ + FSRQ) Blazar (BL Lac + OVV)

Table 1.2: Naming scheme used in this work for referring to AGN with different amounts of obscuration and X-ray luminosity. We used the Lusso et al. (2012) bolometric correction (eq 2.1) to compute L_{bol}.

	$\log(N_H/\mathrm{cm}^{-2})$	
	20 – 22	unobscured
	≥ 22	obscured
	22 – 23	extremely obscured
	24 – 26	CT-AGN
$\log(L_{bol}/\mathrm{erg\ s}^{-1})$	$\log(L_x/\mathrm{erg\ s}^{-1})$	
43.0 – 44.6	42.0 – 43.5	low-luminosity
44.6 – 49.7	43.5 – 48.2	high-luminosity

point of view, the X-ray radiation (at least the hard X-ray one) from the core is usually able to pass through the obscuring torus without being significantly absorbed. However, for the most obscured AGN even the hard X-rays can be absorbed. The threshold between unobscured and obscured AGN in the X-ray is a column density of at least $N_H = 10^{22}\,\mathrm{cm}^{-2}$. In these thesis we further divides type 2 AGN in: extremely obscured AGN those with $N_H \geq 10^{23}\,\mathrm{cm}^{-2}$ and Compton Thick (CT) AGN those with $N_H \geq 10^{24}\,\mathrm{cm}^{-2}$.

We report in Table 1.1 a compact radio and optical AGN classification. In Table 1.2, the naming convention we use in this thesis to refer to AGN on the basis of their obscuration and intrinsic luminosity.

1.3 AGN emission: X-ray to IR to Radio

The AGN emission can extend from the radio to the γ-rays. The typical AGN SED of a Radio Quiet AGN is shown in Figure 1.3.

Every band has its own features, that reflects different emission mechanisms and traces different region of the AGN. With few exceptions, the optical and UV emission-line spectra and the infrared to soft X-ray continuum of most RL and RQ AGN are very similar (Sanders et al. 1989) and so must be produced in a similar way. The characteristics of radio-loudness itself may be related in some way to host galaxy type or to black hole spin (Blandford 1990), which might enable the formation of powerful relativistic jets (Urry &

Padovani 1995).

Similarly to the line emission, the broad band features depend heavily on the viewing angle. In type 1 AGN the typical three-bump SED (Fig 1.3) is dominated by the UV/optical emission of the disk (known as Big Blue Bump), while on type 2 AGN almost all the UV/optical emission is hidden by the torus. In case of the most absorbed AGN, the soft, and part of the hard, X-ray emission is also absorbed, leading to quite different spectral shape. In Figure 1.4, we can see a model of AGN FIR-to-UV emission (from Feltre et al. 2012) for different viewing angles.

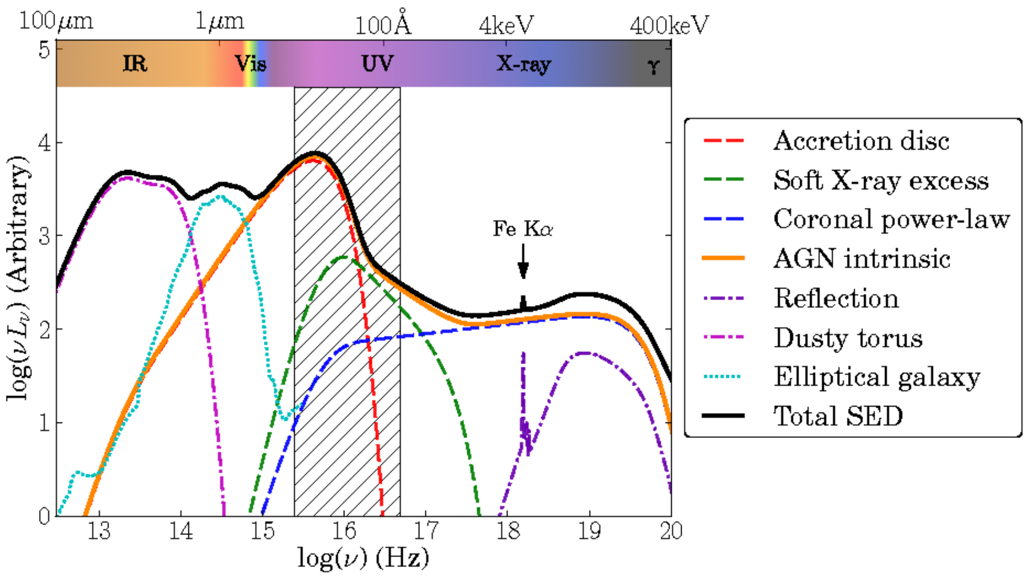

Figure 1.3: Simplified schematic diagram of an AGN SED (Collinson et al. 2017).

X-ray emission

The AGN X-ray emission comes from the inner region of the AGN and gives information about the SMBH, the hot corona and innermost region of the accretion disk. The X-ray emission comes from different sources and processes and can be considered a defining characteristic of the class of AGN. In addition, the X-ray flux shows very fast variability (on timescales as short as few days McHardy 1990; Paolillo et al. 2004) which indicates that it originates in a small region very close to the central object. The intrinsic X-ray emission from AGN is due to processes related to the accretion disk and hot corona (see Mushotzky et al. 1993; Done 2010), however, in jetted AGN the jet contribute to the X-ray emission as well (Padovani et al. 2017).

The primary process is thought to be inverse Compton scattering of the accretion-disk photons to X-ray energies via the hot corona. The X-ray emission is then modified due to the interaction with the matter in the nuclear region (e.g. reflection, scattering, and photo-electric absorption of photons from the accretion disk and/or the obscuring AGN

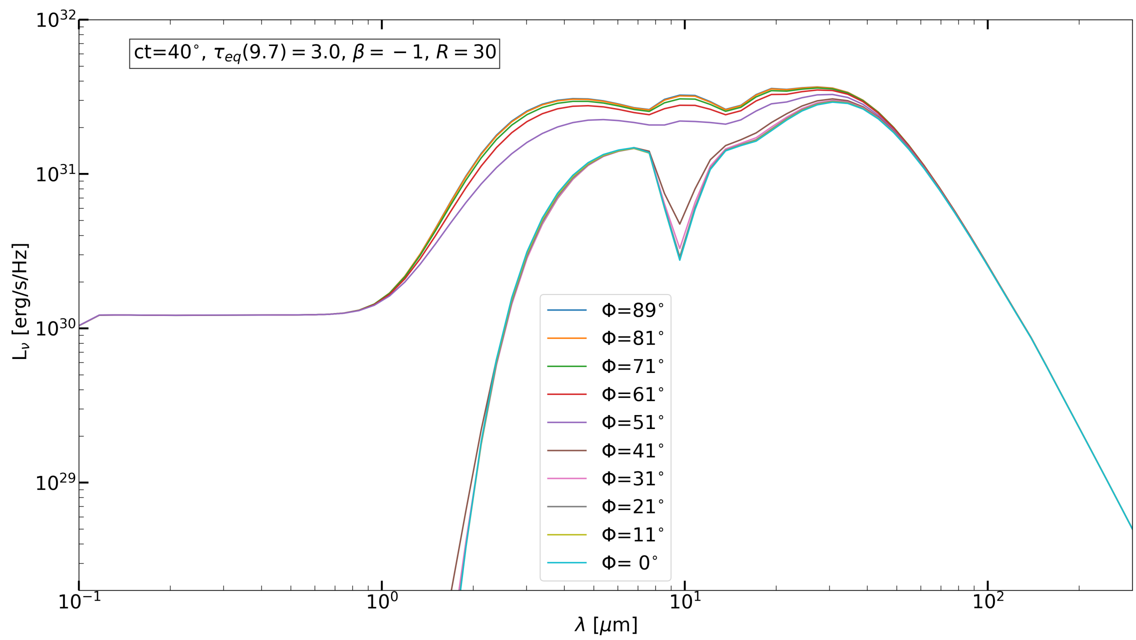

Figure 1.4: Models of AGN SED for different equatorial viewing angle ϕ (left-bottom part of the figure). In this model AGN with $\phi < 50°$ are edge-on, $\phi > 50°$ are face-on. AGN emission derived from Fritz et al. (2006); Feltre et al. (2012) models.

torus). The relative strength of these components can vary quite significantly from source to source, mostly due to differences in the geometry and inclination angle of the torus to the line of sight, leading to a broad range of X-ray spectral shapes.

Hot corona power law This is the dashed blue line in Figure 1.3. The hot corona spectrum is composed by a power law continuum with a thermal exponential cut-off.

The power law component is usually fitted with a spectral index $\Gamma = 1.8 - 1.9$, in the $F_E \propto E^{-\Gamma}$ notation. $\Gamma = \alpha + 1$, where $\alpha = \frac{-\ln\tau}{\ln A}$ is the spectral index, τ is the opacity, $A = e^y$ is the mean energy gained by the photons for each scattering and $y = \frac{4kT}{m_e c^2}$ is the Comptonization parameter. However, the presence of gas can absorb part of the soft X-ray emission and, for highly obscured object, also the hard X-ray emission, leading to flatter photon indexes.

Reflection Bump and Iron Kα line The hot corona emits isotropically, and a fraction of its emission is directed toward the accretion disk. This leads to two different processes.

A part of this emission is reflected by the disk and produce an excess in hard X-ray spectra with respect to the pure power law from the hot corona. This excess is known as Reflection Bump (dark purple line in Figure 1.3).

A part of the hot corona emission is absorbed. These high-energy photons can liberate strong bounded inner electrons from high Z atoms. The most important transition is that involving the Iron Kα line. Photons of energy above $E \geq 7.1$ keV can free a K shell electron, producing a "hole" in the innermost shell. This hole is then filled by an electron of an outer shell. The excess energy can be expelled in two different ways: by unbounding another electron from the outermost shell of the atom (this is known as *Auger effect*), or by the emission of a fluorescence Kα with an energy of 6.4 keV. This 6.4 keV Iron line is very important: on the one hand, it is very luminous and easy to identify, on the other hand, it is generated at few gravitational radii from the SMBH and may show gravitational redshift effects.

Soft excess In the vast majority of AGN, there is an excess at E < 1 keV with respect to the hot corona power law (green dotted line in Figure 1.3), the nature of which is not yet well understood. Early models predict that it could be due to black body (BB) emission from the innermost regions of the accretion disk. However, even the ISCO should not reach sufficient temperatures in the case of SMBH. To the present days, models predict that it can be associated with atomic processes of partially ionized material (Done et al. 2007). Two different geometries of partially ionized material can explain the soft excess: one where the material is optically thick and out of the line of sight, seen via reflection, e.g. from an accretion disc (Crummy et al. 2006). Alternatively, the material can be optically thin and along the line of sight, seen in absorption, e.g. a wind above the disc (Gierliński & Done 2004).

Warm absorber In ~ 50% of local galaxies, photons between 1 ≤ E ≤ 2 keV has been observed to be absorbed. The responsible for this absorption is gas with T≈ 10^{4-5} K, in region smaller than pc scale (e.g., Kaastra et al. 2000). In some cases the absorption lines are blue-shifted by up to thousands $\mathrm{km\,s^{-1}}$ () and it is believed that this absorber can origin from outflowing gas probably connected to AGN winds (Tombesi et al. 2013). The wind-absorber scenario could explain also the rare Ultra Fast Outflows (UFOs), high-ionization absorption lines with velocities (from Doppler shift) typically of 0.1c (Tombesi et al. 2010). These UFOs could be related to AGN winds from the central region, where the high temperatures and high energies are able to accelerate the winds up to quasi-relativistic speed. It has been argued that the absorbers, sometimes considered of different type, could actually be unified in a single, large-scale stratified outflow observed at different locations along the line of sight. The UFOs are likely launched from the inner accretion disc and the Warm absorber at larger distances, such as the outer disc and/or torus (Blustin et al. 2005; Tombesi et al. 2013).

Optical emission

The main feature of the AGN optical emission is the Big Blue Bump (red dotted line in Figure 1.3), a high-luminosity continuum between 300 Å and 3000 Å. It is due to the BB emission of the accretion disk. In particular, the accretion disk has a range of temperatures, decreasing going outward, and the superposition of the various BB, each with its own temperature.

In the optical band there are also emission lines from the BLR (for type 1 AGN) and the NLR (for both type 1 and 2). Common BLR and NLR lines comprehend (among many others) Lyα, Hα, Hβ, [O III], C IV, Mg II, [N II], [O II], and [Ne V](Fig. 1.5). The narrow lines are visible even for the most obscured AGN, thanks to the NLR being uncovered from the torus. Moreover, it is the inner region of the AGN that illuminates the NLR, hence exciting atoms and allowing the production of these lines. Therefore the NLR line (e.g. [OIII], [NeV]) flux can be used to estimate the intrinsic power of the AGN (Schmidt et al. 1998; Gilli et al. 2010; Vignali et al. 2010).

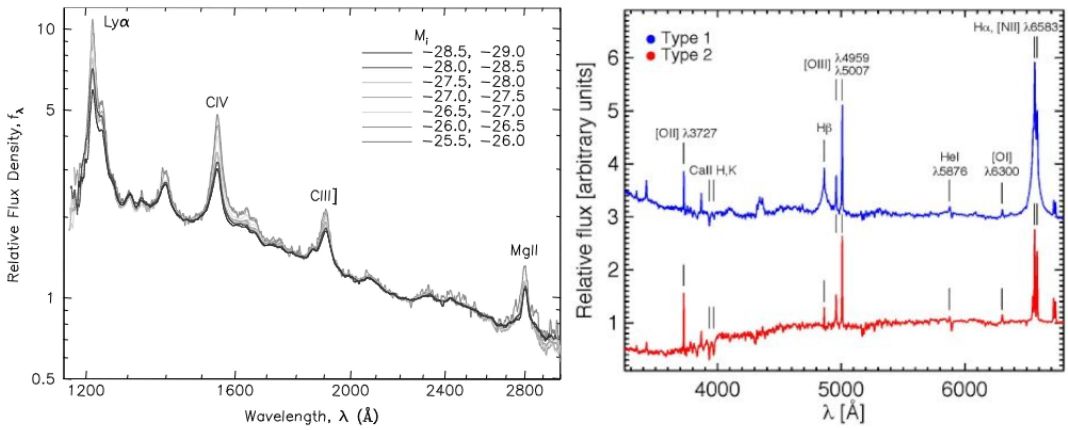

Figure 1.5: *Left panel:* Composite spectra of Broad Line AGN from the Sloan Digital Sky Survey (SDSS), binned by luminosity M_i (Peterson 2006). *Right panel:* Composite optical spectrum of Type 1 (*blue*) and Type 2 (*red*) AGN from SDSS (DiPompeo et al. 2018); the main difference between the spectra are the presence of broad permitted emission lines in Type 1 AGN, while similar narrow emission lines are observed in both spectra.

IR emission

The IR emission of AGN can be split into two different components: one from the dusty torus (light purple line in Figure 1.3) of the AGN and one from the dust of the host galaxy (light blue). The obscuring torus is heated as it absorbs the optical and UV photons emitted from the accretion disk. The dust grains then re-emit the absorbed energy in the mid-IR ($\sim 10{-}20\,\mu$m) given the typical temperatures of the dusty torus, ($\sim 1000{-}2000$ K, at higher temperature the dust grains sublimate). As for the NLR lines, the energy source of this emission are photons from the AGN inner region. As such the mid-IR emission is, again, a proxy of the intrinsic nuclear emission (Gandhi et al. 2009).

However, in the mid-IR there is also a thermal emission from the host galaxy dust. This dust is linked to star-formation (SF) processes and is heated by UV emission from O and B stars. Having two partially overlapping (in wavelength) components in the mid-IR emission means that we need to separate them in order to unveil the AGN.

Radio emission

Except for SF-related processes, the radio emission in AGN is non-thermal and mainly produced by synchrotron processes. The electrons are accelerated up to ultra-relativistic speed by the magnetic field of the central region. With an ultra-relativistic electron of mass m_e and Lorentz factor γ into a magnetic field B, the magnetic field accelerates the electron, producing the emission of a photon of frequency $\nu \propto \mathrm{BE}^2$, where $\mathrm{E} = \gamma m_e \mathrm{c}^2$. If we have a population of electrons with a power law energy distribution N(E)

$$\mathrm{N(E)}d\mathrm{E} \propto \mathrm{E}^{-\delta} d\mathrm{E} \tag{1.5}$$

we obtain a spectra with $\mathrm{F}_\nu \propto \nu^{-\alpha}$, where $\alpha = (\delta - 1)/2$ and typical value around $\alpha \approx 0.8$. In dense region, the synchrotron photons can be absorbed by the same electrons that generated them. We have then a self-absorbed spectra with slopes $\nu^{5/2}$.

In RL AGN the radio emission is mainly located in jets, lobes and hotspot and the nuclear emission is not usually dominant.

1.4 AGN-galaxy co-evolution

In the mid-90's, it was already well established that the galaxy luminosity and the stellar velocity dispersion correlates (Faber–Jackson relation, Faber & Jackson 1976), as well as the effective radius and the luminosity (Kormendy relation, Kormendy 1977). Analogue relations were already available using the stellar mass instead of the luminosity. However, starting from the 1995 new correlations were found that linked the mass of the SMBH with these global properties of the host-galaxy (Kormendy & Richstone 1995; Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Kormendy & Ho 2013). This meant that, somehow, the galactic nucleus is affected by and is having effect on all the galaxy, well over its gravitational influence radius. In particular, the relation between the mass of the SMBH and the stellar mass suggests that the growth of the SMBH and the SF are linked and it led to the formulation of the AGN galaxy co-evolution paradigm.

In all the co-evolution scenarios, the feedback processes are the link between the nucleus-related processes (i.e., accretion into the SMBH, jets, and outflows) and the galaxy-wide SF processes (cloud collapses, stellar winds, SNe). These feedback allow the AGN activity and the SF to regulate each other, thus to produce the observed scaling relations. We will present two different scenarios for the co-evolution: The merger driven scenario (section 1.4.1) and the *in-situ* co-evolution (section 1.4.2).

1.4.1 Mergen-driven scenario

The merger-driven scenario of the co-evolution relies on the presence of two merging galaxies to start the chain of SF and AGN processes. In this scenario, an intense phase of SF is triggered by a wet merger, at least for the most luminous and massive systems (Silk & Rees 1998; Di Matteo et al. 2005; Treister et al. 2012; Lamastra et al. 2013). A fraction of the gas reservoir of the galaxy is funneled towards the SMBH and turns on AGN activity. Thus, this phase is characterized by the growth of both the SMBH and the stellar mass of the galaxy. Due to the large quantity of gas, the majority of the AGN emitted radiation is absorbed and the source appears as an obscured AGN. This phase is likely associated with obscured AGN growth in strongly star-forming submillimeter galaxies (e.g., Archibald et al. 2002; Almaini 2003; Alexander et al. 2005). As the SF and the AGN consume, heat up and expel the gas, the AGN begins to appear less obscured and is identified as a type 1 AGN. Moreover, the decrease in the amount of available gas leads to a decrease of the SF and, eventually, to the end of the AGN activity, leaving behind a "red-and-dead" elliptical galaxy (e.g., Hopkins et al. 2007a; Cattaneo et al. 2009). This scenario is supported by the observations of molecular outflows, extending on kpc scales from the nucleus, in some AGN hosts and ultra-luminous infrared galaxies (ULIRGs) (e.g., Feruglio et al. 2010; Glikman et al. 2012; Cicone et al. 2014). In these systems the mass loss ranges from one to several times the star-formation rate (SFR). Similarly, using various emission line diagnostics, neutral (HI 21 cm, [C II] 158μm) and ionised ([O III] 5007Å, Hα 6564Å) outflows were also detected (e.g., Nesvadba et al. 2008; Alexander et al. 2010; Harrison et al. 2012, see Harrison et al. 2018 for a review). Ionized AGN outflows tend to be more common ($\sim$ 75%, e.g., Brusa et al. 2015, 2016; Perna et al. 2015a,b; Kakkad et al. 2016; Zakamska et al. 2016; LaMassa et al. 2017; Toba et al. 2017 in high luminosity/high Eddington ratio and in red/obscured dusty sources, as expected in the evolutionary model by Hopkins et al. (2007b).

1.4.2 Alternative: in-situ co-evolution

Alternatively to the merger driven scenario, in the *in-situ* co-evolution model (Mancuso et al. 2016a,b, 2017; Lapi et al. 2018), there is no need for a companion galaxy to trigger the SF. Instead, the SF is a local process regulated by energy feedback from supernova explosions (SNe) and from the central SMBH. In the first phase of the galaxy evolution, the dominant feedback is due to SNe, with SFR $\propto t^{1/2}$ and $M_* \propto t^{2/3}$. In this scenario, the galaxy Main-Sequence (MS, see Rodighiero et al. 2011) in the SFR$-$M$_\odot$ plane emerges as the locus where galaxies spend most of their time, thus where it is most probable to find them. In these early stages, the BH mass grows exponentially in gas-rich obscured environment, the AGN emits at mildly super Eddington ratio and its luminosity rises exponentially. After less than one Gyr (for massive galaxies), the AGN luminosity reaches values similar to those of the SF and becomes dominant. Via powerful winds and outflows, the AGN power heats up and removes the gas from the host galaxy, quenching the SF. The galaxy moves below the MS and the stellar populations evolve passively, while the residual gas in the central region accretes into the SMBH at lower sub-Eddington ratios and the AGN luminosity exponentially declines. At the end of this phase, what is left is again, a "read and dead" elliptical galaxy.

1.4.3 BHAD and SFRD

Further indication of the link between the SMBH growth and the SF can be seen when comparing the BH accretion rate density (BHAD) and the star formation rate density (SRFD). The first estimates of the SFRD (Lilly et al. 1996; Madau et al. 1996) were based on the SFR measured with the UV luminosity, and it was later shown that they underestimated the SFR, due to the not sufficient extinction correction in case of dusty environments. More recent UV and IR surveys (e.g., Sanders et al. 2003; Wyder et al. 2005; Cucciati et al. 2012; Gruppioni et al. 2013; Rowan-Robinson et al. 2016a) are now in agreement up to $z < 3$ and shows that the SFRD grows with the redshift up to $z \sim 2$, then rapidly declines (Fig. 1.6). The redshift range between $z \sim 1$ and ~ 3 is called the "cosmic noon".

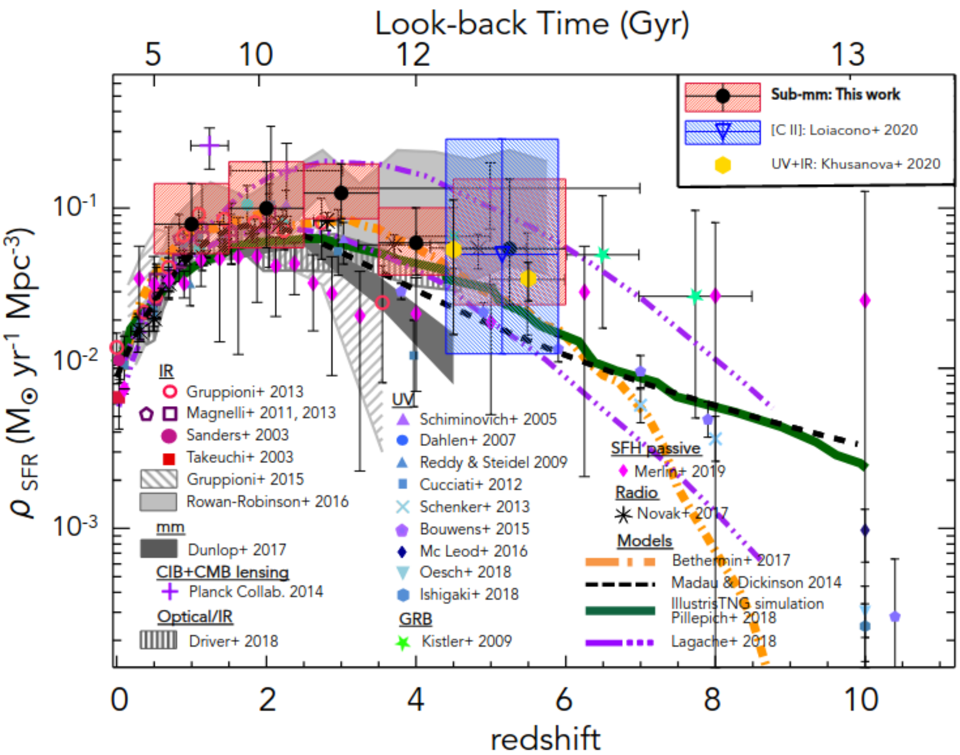

Figure 1.6: Star formation rate density compilation from Gruppioni et al. (2020).

At higher redshift, the reconstruction of the SFRD is more uncertain, with studies showing a steep decline (Bouwens et al. 2015; McLeod et al. 2016; Ishigaki et al. 2018), and other hinting at a flatter SFRD (Gruppioni et al. 2013; Rowan-Robinson et al. 2016b; Novak et al. 2017; Gruppioni et al. 2020; Malefahlo et al. 2022). The reason behind this discrepancy may lie again in the adopted selection method: the studies that found a steep decline are relying mostly on UV-selected sources such as Lyman-break galaxies (LBGs), while the ones that predicts a flatter SFRD are usually exploiting longer wavelengths (i.e., radio or submillimeter bands, like those of Gruppioni et al. (2020); Talia et al. (2021) based on optically-dark dusty galaxies), less prone to extinction-biases.

The reconstruction of the BHAD (e.g., Aird et al. 2010; Shankar et al. 2013; Delvecchio et al. 2014; Sijacki et al. 2015; Volonteri et al. 2016; Vito et al. 2018) has faced similar difficulties: the first studies relied on quasars or type 1 AGN in general, thus missing (and/or underestimating) all the obscured AGN accretion. Mid- to far-IR photometric observations in deep fields (Delvecchio et al. 2014; Schreiber et al. 2015) have provided reliable determination of the SFRD and the BHAD up to redshift ~ 3 (Gruppioni et al. 2013; Magnelli et al. 2013; Delvecchio et al. 2014) using *Herschel PACS* data (observing at 100 and 160 μm; Poglitsch et al. 2010) and have detected high redshift galaxies (up to $z > 6$; e.g., Riechers et al. 2013; Rowan-Robinson et al. 2016b) using *SPIRE*, observing at 250, 350 and 500 μm; Griffin et al. 2010). However, the deepest cosmological surveys performed by *Herschel* at high-z have detected only the most luminous galaxies ($L_{IR} > 10^{12} L_\odot$ at $z \geq 3$; Gruppioni et al. 2013) and suffer from the uncertainties of not being able to accurately disentangle the AGN from the SF emissions, as both contribute to the mid- and far-IR emission. At present, as we can see from Fig. 1.7, agreements has been reached only at $z < 3$, while at higher redshift, most of the estimates comes from X-ray derived BHAD, thus are biased against highly obscured object. Moreover, it is in this regime $3 < z < 6$, where there are still tension with the BHAD derived from simulations. As we will see in section 2, this tension may be alleviated by assuming a higher number of extremely obscured AGN as for now undetected in the X-ray.

Despite all these uncertainties, the shapes of the SFRD and BHAD look remarkably similar. This is expected if we take for true the AGN-galaxy co-evolution. The stellar mass and the BH mass correlate, so the cosmic density of their growths (i.e., their derivative) must have similar shapes.

1.5 Obscured accretion

Despite the differences between the merger-driven and the *in-situ* co-evolution models, both have at their core a phase of obscured accretion, i.e. a phase characterized by AGN accretion in gas- and dust-rich environments. The obscured accretion phase is where most of the stellar mass formation and of the BH accretion should take place. However, the high quantity of gas (and dust) that are fueling the SF and AGN processes, are also hiding them. Therefore, the search for and the characterization of objects in this phase in the optical and UV is challenging due to the heavy reddening that affects them (Rowan-Robinson et al. 1997; Hughes et al. 1998). A solution to this problem is to study indirectly the primary emission, by measuring the dust-reprocessed radiation in the IR. However, in the IR we have both the emission of the AGN torus and that due to SF (ascribed to dust heated by young stars in the galaxy) and it is difficult to separate the two components. This is

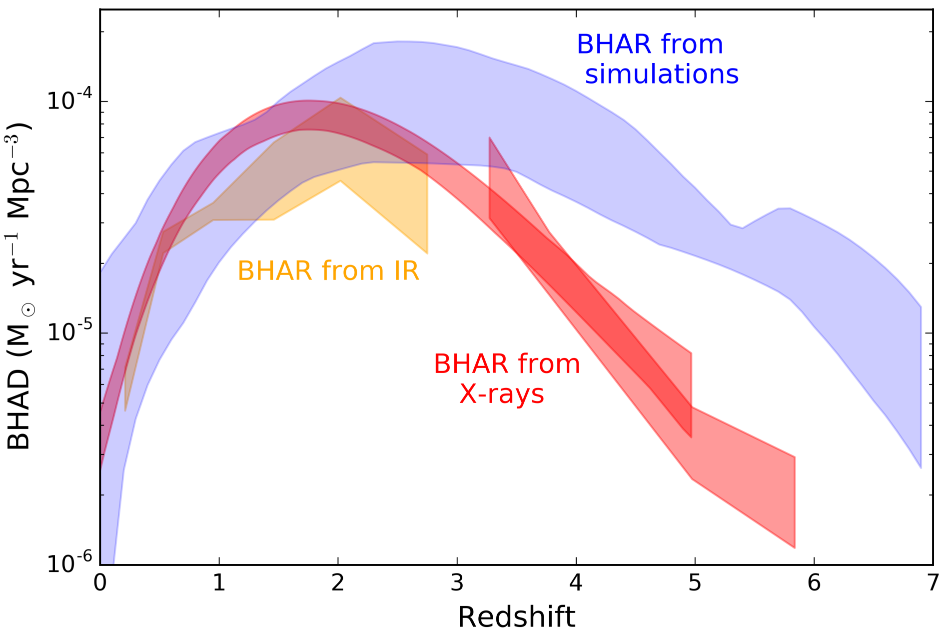

Figure 1.7: Redshift evolution of the BHAD adapted from Vito et al. (2018). The blue area is a theoretical BHAD curve obtained from Volonteri et al. (2016); Sijacki et al. (2015); Shankar et al. (2013). The red areas are the X-ray-derived BHADs, obtained from observations, from Aird et al. (2015); Ueda et al. (2014); Vito et al. (2014). The yellow area is the IR-derived BHAD from Delvecchio et al. (2014).

particularly true for low-luminosity AGN (LLAGN): their mid-IR emission is diluted and overshadowed by the host-galaxy luminosity (i.e., Gruppioni et al. 2013).

Despite the significant challenges in identifying sources in the obscured accretion phase, we must study this phase, as it is the key moment of the co-evolution, where most of the accretion takes place and the feedback processes start to kick-in.

1.5.1 The X-ray Background

Obscured and accreting AGN are also fundamental to explain the spectra of the X-ray background (XRB).

That a diffusive X-ray glow pervades the sky was first discovered by Giacconi et al. in 1962, in the first X-ray astronomy experiment. In Figure 1.8 the spectrum of the XRB is shown. As far as today more than 90% of the XRB under a few keV has been resolved in hundreds of millions of individual X-ray sources distributed across the entire universe. Most of them are AGN (e.g., Moretti et al. 2012). However, at energies E $\sim$ 30 keV, where the XRB spectrum peaks, only 30%-40% of its emission has been resolved so far (Harrison et al. 2016) and AGN are the main contributors to it. The integrated emission of the AGN, obtained from the X-ray deepest surveys, can account for most of the XRB surface brightness below 10 keV (Xue et al. 2011; Moretti et al. 2012). However, where the XRB spectra peaks (20 – 40 keV), most of the XRB is still unresolved. Extrapolating at these energies the spectrum from the deep surveys showed that an additional large population of heavily obscured, Compton Thick (CT) AGN with $N_H > 10^{24}$ cm^{-2} is required to fully reproduce the XRB spectra. XRB population synthesis models predict that the maximum contribution to the XRB of the "missing" CT objects in the $L_x^{intr} \approx 10^{42} - 10^{44}$ erg/s luminosity range should peak at $z \approx 1$ (Gilli et al. 2007; Ueda et al. 2014; Ananna et al. 2019). As it is difficult to constrain the abundance and the properties of these CT AGN using only XRB synthesis models (e.g., Gilli et al. (2007) estimates that the CT AGN can be up to ~ 50 of all the obscured AGN, while Akylas et al. (2012); Georgantopoulos & Akylas (2019) lower this fraction to 10 – 20 % in the local Universe), direct selections and studies of them are needed (see e.g., Alexander et al. 2005; Daddi et al. 2007; Georgantopoulos et al. 2008, 2009; Gilli 2013; Marchesi et al. 2022).

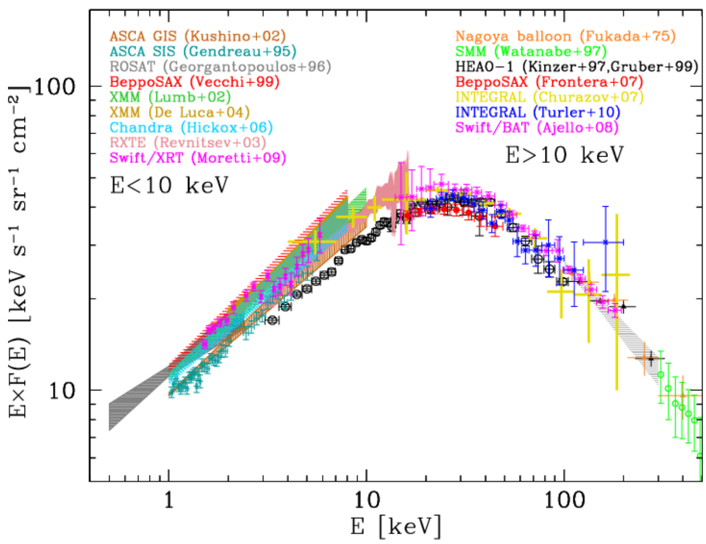

Figure 1.8: Compilation of the measurements of the cosmic X-ray background spectrum in the 0.5 - 400 keV energy range. Data points with different colors come from different combinations of missions and instruments as labeled and referenced (left labels: E < 10 keV; right labels: E > 10 keV) (Gilli 2013).

Chapter 2

Searching for the most obscured AGN: *Athena* – IR synergies

The goal of this chapter is to study the synergies between the X-ray observatory *Athena* and future IR facilities in searching and identifying AGN, with a particular focus on the most obscured ones. We will show how the search for these elusive CT-AGN, very difficult to identify and characterize with the current instruments, can benefit from a combined use of IR and X-ray data.

2.1 Selecting obscured AGN

To study the link between star formation and black hole accretion we need, in particular, to investigate the epoch in which the bulk of these growths seems to take place (i.e. the cosmic noon, see section 1.5). However, the high quantity of gas and dust that fuels both processes absorbs most of the energy emitted by stars and accreting SMBHs and re-emits it at longer wavelengths, mostly in the IR. This make selecting AGN, and in particular obscured AGN, not a trivial task.

Although several methods for selecting type 2 AGN are available (see Vignali 2014 for a review), a complete census of these objects cannot be achieved using only a single observing-band. The X-ray radiation, originated by the hot corona in the innermost region of the AGN (section 1.3), is a good tracer of the AGN intrinsic emission, however, when the nucleus is obscured by column densities as large as $\sim 10^{24-25}\mathrm{cm}^{-2}$, even hard X-rays are severely depressed. Mid-infrared wavelengths can be effectively used as an X-ray complementary selection method, as the optical-to-X-ray absorbed radiation is re-emitted at these wavelengths after being thermally reprocessed by the obscuring torus. However, distinguishing between the AGN and the SF contributions to the total IR emission is not trivial, especially for galaxies with high SFR and strong PAH features. Finally, obscured sources can also be identified using mm-observations, as they can be sensible up to $N_H \sim 10^{26}\,\mathrm{cm}^{-2}$ (e.g., Behar et al. 2015; Kawamuro et al. 2022). Type 2 AGN can be also selected using narrow, high-ionization emission lines, such as the [O III]5007Å, the [Ne V]3426Å and the C IV1549Å narrow emission lines. We will focus on these lines in chapter 3 and 4.

2.2 X-ray and IR synergies in detecting obscured AGN

As already mentioned, the search for and the characterization of obscured AGN in the optical and UV is challenging due to the heavy reddening that affects them (Rowan-Robinson et al. 1997; Hughes et al. 1998). A solution to this problem is to study indirectly the primary emission, by measuring the dust-reprocessed radiation in the IR. A study at such wavelengths allows us to observe the physical processes at work in the obscured regions and, therefore, to derive the SFRD and the BHAD if we are capable of properly separating them. A direct measure of the dust-obscured SF activity is possible with space-based IR observatories, whose data are not affected by to the dust attenuation.

The mid- to far-IR photometric observations in deep fields (Delvecchio et al. 2014; Schreiber et al. 2015) have provided extremely reliable determination of the SFRD and the BHAD up to redshift $\sim 3-4$ (Gruppioni et al. 2013; Magnelli et al. 2013; Delvecchio et al. 2014) using *Herschel PACS* data (observing at 100 and 160 μm; Poglitsch et al. 2010) and have detected high redshift galaxies (up to $z > 6$; e.g., Riechers et al. 2013; Rowan-Robinson et al. 2016b) using *SPIRE*, observing at 250, 350 and 500 μm; Griffin et al. 2010). However, the deepest cosmological surveys performed by Herschel at high-z have detected only the most luminous galaxies ($\mathrm{L_{IR}} > 10^{12}\,L_{\odot}$ at $z \geq 3$; Gruppioni et al. 2013) and suffer from the uncertainties of not being able to accurately disentangle the AGN from the SF emissions, as both contributes to the mid- and far-IR emission. Even in the case when Spectral Energy Distribution (SED) fitting allows us to separate the two components, the results usually depend on the adopted technique and modeling. Therefore, the initial phase of the BH-galaxy co-evolution paradigm has been, so far, the most elusive and difficult to track, primarily due to the the large amounts of cold gas and dust (Fabian 1999). According to the co-evolution scenario, this heavily obscured phase should be characterized by the presence of a strong starburst and an obscured AGN, both giving a significant contribution to the mid- and far-IR emission (e.g. Page et al. 2004; Vignali et al. 2009; Lapi et al. 2018).

Compton-thick (CT) AGN, i.e. highly-obscured AGN with column density $\log(\mathrm{N_H}/\mathrm{cm}^{-2}) \geq 24$, are also required to explain the spectra of the extragalactic XRB (see section 2.4.1): their contribution is needed to reproduce the hump observed at $20-30\,\mathrm{keV}$ in the XRB (e.g. Comastri et al. 1995; Gilli et al. 2007; Ballantyne et al. 2011). However, the fraction of CT-AGN derived from XRB synthesis models ranges from 10 % to 30 – 50 % (Treister et al. 2009; Gilli et al. 2007; Ueda et al. 2014; Ananna et al. 2019) and suffers from large uncertainties due to degeneracies between several model parameters. Although the expected number of CT-AGN is large, their detection and identification beyond the local universe are challenging and, even in the local universe, their observed fraction ($\sim 5-10\,\%$, e.g. Burlon et al. 2011; Ricci et al. 2015) is significantly lower than models predictions, but X-rays detection bias may be responsible of such low numbers (Burlon et al. 2011).

To select unobscured or mildly-obscured AGN, one of the most effective tools is the observation in the X-ray band, thanks to the radiation coming directly from the AGN innermost regions. However, this selection loses effectiveness with the increasing column density (at $N_H \geq 10^{23}\,\mathrm{cm}^{-2}$, the soft continuum is severely depressed) and, even in deep fields, only a limited fraction of the most obscured AGN has been revealed in the X-ray (e.g. Tozzi et al. 2006; Lanzuisi et al. 2013; Marchesi et al. 2016a; Del Moro et al. 2017). Moreover, the BHAD estimation from the X-ray luminosity is strongly dependent on the bolometric correction (k_{bol}), that allows to estimate the “total” bolometric power (L_{bol}) of the AGN from the AGN emission in a specific band, but suffers from significant uncertainties in the $k_{bol}-L_{bol}$ relation (e.g. Hopkins et al. 2007b; Lusso et al. 2010, 2012; Duras et al. 2020). Likewise for obscured SF galaxies, the IR is an effective band in detecting these obscured AGN, thanks to the energy absorbed by the obscuring material and re-emitted in the mid- far-IR wavelength range (e.g. Pozzi et al. 2007). The effectiveness of identifying highly obscured AGN by selecting bright mid-IR sources with faint optical or near-IR emission has been shown in the past by *Spitzer* (e.g. Houck et al. 2005; Weedman et al. 2006; Polletta et al. 2008) and *WISE* (e.g. Mateos et al. 2012; Assef et al. 2013).

The wavelength range is which the AGN torus emits is more limited with respect to the one from the host-galaxy SF (the AGN emitting mostly in the $5-100\,\mu$m rest frame range, while the galaxy emission is still significant up to $1000\,\mu$m), however we need to probe the entire IR band to discern and separate the two components. Therefore, to be able to photometrically identify highly-obscured and CT-AGN, both mid-IR and far-IR instruments are necessary, as we need to distinguish between the emission from hot dust (T>100 K), mostly due to reprocessed AGN emission from the dusty torus and galactic dust, and the emission from warm/cold dust (T<100 K) dominated by the reprocessed SF emission.

A future IR cryogenic cooled space telescope may be able to deliver all of this. We investigated three different IR mission, using as ‘templates’ the proposed, now withdrawn, ESA *SPICA* mission, the NASA concept study *Origins Space Telescope* (*OST*), and the proposed NASA mission *PRIMA*. A *SPICA*-like mission (described extensively in section 2.3.2) with a mid- and far-IR photometric coverage will be able to sample the band where the peak of the CT-AGN emission is up to $z \approx 6$ and provide also low-resolution spectra, while far-IR photometric cameras will cover the wavelength range where the SF is the main contributor, thus helping in separating the two emission mechanisms. Detected sources may be followed-up with SPICA SAFARI-like instruments (such as the one proposed for *PRIMA*) to provide low resolution spectra in the mid- and far-IR wavelength range up to $z \approx 3.5$. In these spectra high ionization lines, such as [Ne v]at 14.32 μm and 24.32 μm, [Ne vi] at 7.7 μm and [O iv] at 25.89 μm, would be signatures of AGN presence and their fluxes could be used as proxy of the AGN bolometric luminosities (e.g. Spinoglio & Malkan 1992; García-Bernete et al. 2016; Fernández-Ontiveros et al. 2016). Alternatively, a highly capable grating mid- and far-IR spectrometer, as the *OSS* in the *OST* (see section 2.3.3), will be able to cover instantaneously both the hot and warm/cold dust wavelength regime, thus allowing us to distinguish between the torus and the host-galaxy emission and to estimate, in a reliable way, the AGN bolometric power, the host SF and the BH accretion rate, without the need of additional constraints from other optical-IR instruments or surveys (at least for the mid- and high-luminosity AGN).

To confirm the CT nature of sources, X-ray spectroscopy has been the most reliable method so far. Unfortunately, with the current deepest X-ray surveys, we have only scratched the surface of the obscured AGN population at $z \sim 1-4$. Moreover, the identification of an AGN as CT depends strongly on the adopted analysis techniques and modeling, with different techniques giving sometimes different classifications of the same sources (e.g. Castelló-Mor et al. 2013). The best results have come from X-ray deep and ultradeep surveys (e.g. COSMOS, Elvis et al. 2009; Civano et al. 2016, CDF-S, Giacconi et al. 2002; Luo et al. 2008; Xue et al. 2011; Xue 2017) that allowed to detect CT-AGN even at high redshift, but with the disadvantage of having small fields, thus low source statistics. *XMM-Newton* and *Chandra* surveys unveiled only up to few tens of candidates of CT-AGN (e.g. Comastri et al. 2011; Brightman & Ueda 2012; Georgantopoulos et al. 2013; Lanzuisi et al. 2017), while an hard X-ray mission like *NuSTAR* has difficulties in detecting objects beyond $z \sim 1$, due to its sensitivity (Lansbury et al. 2017a,b, the notable exceptions of CT-AGN at $z \sim 2$ detected by Del Moro et al. 2014; Zappacosta et al. 2018).

A breakthrough in selection and identification of CT-AGN is expected with the new ESA X-ray observatory *Athena*, due for launch in early 2030s. *Athena*, with its combination of angular resolution, field of view and collecting area at ~ 1 keV (see section 2.4.1), is the ideal instrument to perform X-ray surveys and will be more than two orders of magnitude faster than *Chandra* or *XMM-Newton* (Nandra et al. 2013). The excellent spectral capabilities of *Athena WFI* for surveys will yield samples of the most heavily obscured AGN up to redshifts $z = 4$ and up to 100 times larger than is currently possible. Since $L_{\text{X-ray}} \gtrsim 10^{42}\,\mathrm{erg\,s^{-1}}$ are commonly associated with AGN activity, *Athena* will also be invaluable in quickly identifying moderate- and high-luminous AGN, especially in cases where we can not distinguish between the torus mid-IR emission and the host-galaxy contribution from SED fitting analysis.

The planned launch date for *Athena* is $\gtrsim 2030$. Having, in the same years, a new IR mission dedicated (also) to survey high-z galaxy and AGN would be an extraordinary opportunity to study obscured accretion and to characterize AGN up to high redshift. A possible example of such mission would have been *SPICA*, but also the NASA concept *OST* for an 4.5 K cooled IR space observatory, which had a proposed launch date of

$\sim$ 2035 and similar survey strategy and wavelength coverage as *SPICA*. Having data at X-ray and IR wavelengths will be the key to accurately reconstruct the accretion luminosity across cosmic time up to very high redshift ($z \approx 6$). The search for the elusive CT-AGN may decisively benefit from the synergies between *Athena* and future IR mission: while a *SPICA*-like telescope and the *OST* should be able to effectively detect the majority of them (even those too obscured to be detected in X-rays), the X-ray spectra provided by *Athena* will be invaluable to confirm their CT-nature, even in paucity of photons, for a large fraction of them.

2.3 Athena and future IR observatories

2.3.1 Athena

Athena (Advanced Telescope for High ENergy Astrophysics [*] Barcons et al. 2017) is the X-ray observatory mission selected by ESA, within its Cosmic Vision programme, to address the *Hot and Energetic Universe* scientific theme. It is the second L(large)-class mission within that programme and is due for adoption in late 2022 and launch around the 2035s. *Athena* will consist of a single large-aperture grazing-incidence X-ray telescope, utilizing a novel technology (High-performance silicon pore optics) developed in Europe, with 12 m focal length and 5 arcsec Half Energy Width (HEW) on-axis angular resolution, degrading gradually to less than 10 arcsec at 30 arcmin off-axis (Bavdaz et al. 2018). There will be two instruments in the focal plane. One is the Wide Field Imager (WFI)[†] (Meidinger et al. 2018) providing simultaneous sensitive wide-field imaging and spectroscopy (FWHM$\leq$ 170 eV at 7 keV) and high count-rate capability (>90% throughput and <1% pile-up for 1 Crab) over the 0.2 to 15 keV energy range. This is achieved through two sets of Silicon-based detectors using DEPFET Active Pixel Sensor technology: the Large Detector Array is a mosaic of 2×2 arrays spanning a $\sim 40 \times 40\,\mathrm{arcmin}^2$ Field of View (FoV) oversampling the PSF by more than a factor of two and the Fast Detector is a single array optimised for high count rate applications. The other instrument is the X-ray Integral Field Unit (X-IFU)[‡] (Barret et al. 2018) delivering simultaneous spatially resolved (5 arcsec pixels) high-resolution X-ray spectroscopy (FWHM<2.5 eV below 7 keV) over a limited field of view (~5 arcmin equivalent diameter) over the 0.2 to 12 keV energy range, with high count-rate capability (10 eV spectral resolution at 1 Crab intensities with low pile-up and >50% throughput). This performance is based on a large format array of superconducting molybdenum-gold Transition Edge Sensors coupled to absorbers made of gold and bismuth, cooled at ~90 mK inside a nested set of cryostats.

[*] https://www.the-athena-x-ray-observatory.eu/
[†] http://www.mpe.mpg.de/ATHENA-WFI/
[‡] http://x-ifu.irap.omp.eu/

Surveys with *Athena*

Some of the core science objectives of *Athena*[§] (e.g. finding distant evolved groups of galaxies at $z > 2$, complete the census for high-redshift and heavily obscured AGN) require performing surveys of the X-ray sky with the WFI. These surveys are expected to take a significant part of the observation time during the four years of the nominal mission lifetime. At the time of writing, the nominal number of pointings and exposure times (the so-called "Tier 2 post-CORE strategy") includes an "ultradeep" layer (4×1400 ks), a "deep" layer (3×980 ks+7×840 ks) and a "shallow" layer (106×84 ks) (see Table 2.1). Each pointing covers $\sim 0.4\,\mathrm{deg}^2$, so the total area of each layer is 1.6, 4 and 42.5 deg^2, respectively. The exact places and geometries (single pointings vs. mosaics) of those layers will of course be determined at a later stage by the international astronomical community through the guaranteed time and open time, but they are expected to include regions of the sky with substantial multi-wavelength coverage. An initial estimate is encapsulated in the *Athena* Mock Observing Plan (MOP).

2.3.2 SPICA-like

SPICA was a joint European-Japanese project to develop a new generation cryogenic infrared space telescope, with the early design dating back to more than two decades ago (e.g. Nakagawa et al. 1998, 2014; Swinyard et al. 2009). Here we discuss the *SMI* and *B-BOP* instruments as it were designed during the ESA M5 selection process and a *SPICA*-like survey strategy.

SMI

SMI is composed of four channels, SMI-CAM, LR, MR and HR. In this work, we consider *SMI-CAM* and *SMI-LR*, that work simultaneously. *SMI-CAM* is a mid-IR slit viewer photometric camera that covers the $30 - 37\,\mu$m range and provides $34\,\mu$m broad-band images with a field of view of $10' \times 12'$ (with the exclusion of the four slits used by the spectrometer). *SMI-LR* is a wide field-of-view multi-slit prism spectrometer, composed of four long slits of $10'$ in length and $3.7''$ in width. It has a resolution of $R = 50 - 120$ and allows spectroscopic surveys in the $17 - 36\,\mu$m wavelength range. The *SMI-CAM* provides photometric scientific data, and is invaluable in accurately determining the position of the slits on the sky for pointing reconstruction in creating spectral maps. We refer to Kaneda et al. (2017) for further instrument specifications and survey strategies.

[§]The full list of science requirements can be found under `https://www.cosmos.esa.int/web/athena/study-documents`

B-BOP

The B-BOP polarimetric instrument allows simultaneous imaging observations in three bands, centered at 70 μm, 200 μm and 350 μm, over an instantaneous field of view from $1.8' \times 1.8'$ to $2.7' \times 2.7'$ at FWHM resolutions of $6''$, $17''$ and $30''$, respectively. B-BOP is two to three orders of magnitude more sensitive than current or planned far-IR/submillimeter polarimeters, e.g. *SHARP* (Li et al. 2008), *HAWC+ (SOFIA)* (Dowell et al. 1998), *BLAST-TNG* (Galitzki et al. 2014). It provides wide-field $70-350\,\mu$m polarimetric images in Stokes Q and U of comparable quality (in terms of resolution, signal-to-noise ratio, and both intensity and spatial dynamic ranges) to *Herschel* images in Stokes I. More details about the B-BOP instrument are in Rodriguez et al. (2018) and André et al. (2019).

Surveys with SMI

Hereafter, we consider two reference blind spectroscopic surveys:

- a **1 deg^2 Ultradeep survey** with a total observational time of 600 hr, an SMI-CAM 5 σ sensitivity of $\approx 3\,\mu$Jy and an SMI-LR 5 σ sensitivities (high-background case) of $\approx 50\,\mu$Jy at 20 μm, $\approx 110\,\mu$Jy at 30 μm and a line sensitivity of $\approx 9.2 \times 10^{-20}$ W/m^2 (this survey configuration corresponds to the Kaneda et al. 2017 deep survey);

- a **15 deg^2 Deep survey** with a total observational time of 600 hr, an SMI-CAM 5 σ sensitivity of $\approx 13\,\mu$Jy and an 5 σ SMI-LR sensitivities (high-background case) of $\approx 160\,\mu$Jy at 20 μm, $\approx 380\,\mu$Jy at 30 μm and a line sensitivity of $\approx 33 \times 10^{-20}$ W/m^2 (this survey has the same depth of the Kaneda et al. 2017 shallow survey, although over a larger area).

2.3.3 The *Origins Space Telescope*

The *OST* (Battersby et al. 2018; Meixner et al. 2019) is a concept study for a Far-Infrared Surveyor mission, the subject of one of four science and technology definition studies supported by NASA for the 2020 Astronomy and Astrophysics Decadal Survey. The *OST* is being designed with the aim of covering a large area of the sky, thus allowing to search for rare objects at low and high redshifts.

OST is composed of a 5.9 m-diameter telescope with a *Spitzer*-like architecture, actively cooled down to 4.5 K. It is designed with three onboard instruments: the *Origins Survey Spectrometer* (*OSS*) to cover the $25-588\,\mu$m wavelength range instantaneously at R= 300, *MISC-T* (*Mid-Infrared Spectrometer Camera Transit*) a $2.8-20\,\mu$m transit spectrometer and the *Far-infrared Imager and Polarimeter (FIP)*, that delivers imaging or polarimetry at 50 μm and 250 μm. In this work, we take into consideration primarily the

OSS instrument, as it is perfectly suited for detecting and characterizing AGN in blind surveys.

OSS

The *Origins Survey Spectrometer* is a R=300 spectrometer covering the full 25 – 588 μm wavelength range instantaneously, using six logarithmically-spaced grating modules. True 3D spectral mapping is performed thanks to the 30 to 100 spatial beams that each module couples. Higher spectral resolution can be achieved by inserting into the light-path two mirrors that divert the light into interferometer optics allowing a $R = 43\,000$ resolution, keeping the full spectral range but working only on single pointing (i.e. losing the survey capabilities). Even higher $R = 300\,000$ spectral resolution can be achieved using an insertable etalon, further restricting the FoV and reducing the spectral range to 100 – 200 μm. As in this chapter we focus on detecting AGN in surveys, we take into consideration only the base R=300 spectroscopy mode.

The six *OSS* bands cover the 25 – 44, 42 – 74, 71 – 124, 119 – 208, 200 – 350, 336 – 589 μm wavelength ranges, with beam sizes of 1.41, 2.38, 4.0, 6.78, 11.3, 19.0 arcsec and 1 hr 5 σ sensitivities (R=300) of 22, 28, 40, 104, 104, 338 μJy. With the current *OSS* specs, an improvement of more than a factor 1 000 in sensitivity is expected with respect to already flown far-IR observatories (*SOFIA*, *Herscel-SPIRE*, and *Herschel-PACS*). We refer to Meixner et al. (2019) for further instrument specifications.

Survey with *OSS*

For *OST* we consider two blind blank-field spectroscopic surveys as described in Meixner et al. (2019):

- a **0.5 deg^2 deep survey** with an observational time of ≈ 1000 hr, a R=4, 5 σ sensitivity of 4.5, 5.8, 8.4, 21.8, 21.8, 70.9 μJy (for channel 1 to 6) and a R=300 5 σ sensitivity of 39, 50, 72, 189, 189, 614 μJy;

- a **20 deg^2 wide survey** with an observational time of ≈ 1000 hr, a R=4, 5 σ sensitivity of 30, 39, 56, 145, 145, 473 μJy (for channel 1 to 6) and a R=300 5 σ sensitivity of 262, 336, 483, 1260, 1260, 4094 μJy.

2.3.4 PRIMA

PRIMA is a concept for a cryogenically-cooled, far-infrared observatory for the next decade. PRIMA combines a passive thermal design in an earth-sun L2 halo orbit with closed-cycle coolers to support a 2-3 meter telescope cooled to 4 K, and instruments and focal planes cooled to 1 K and below. The primary instrument is a R=100 – 300 spectrograph. It will cover from 25 to at least 200 μm, with a goal of reaching 330 μm. This full range will be covered with 4 or 5 wideband grating spectrometer modules operating simultaneously, each fed with a long slit; this architecture provides both optimal pointed object sensitivity as well as excellent spectral mapping speed. The spectrometer will also be equipped with a mode providing spectral resolving power of several thousand, also covering the full band with incurring only modest sensitivity penalty. A multi-purpose imaging experiment (*PRIMAGER*) developed by an European consortium is also under study for PRIMA. According to its latest specifications it should cover the $25-80\,\mu$m range with 12 filters at R$\sim$ 10 resolution and the $80-300\,\mu$ band with 4 filters at R$\sim$ 4.

Survey with *PRIMAGER*

Due to the PRIMA mission being in a early state of development and subject to almost daily changes, the survey strategy is not defined yet. For our work, we assumed a 600 hr deep survey of 3 deg^2, adopting the specification from the *PRIMA* factsheet version 1.1, 22 Feb 2022.

2.3.5 Surveys comparison

Athena WFI, *OST OSS*, and *SPICA SMI* have planned survey with different layers of depth and area. The same is true for *PRIMA*, with the added difficulty of daily changing specifications. In this work, we compared the different layers, matching them to have similar areas as much as possible. As we compared the expected number of sources per deg^2 for all the surveys and investigated the fraction of sources that may be detected, the exact area coverage is not critical. We investigate the capabilities of the observatories considering a **DEEP** survey (comparing the *SPICA* ultradeep survey with the *Athena* ultradeep survey, the *OST* deep survey, and the *PRIMA* deep survey), and a **WIDE** survey (composed of the *SPICA* deep survey, the *Athena* shallow survey and the *OST* wide survey). Table 2.1 summarizes the main parameters of all the *SPICA*, *OSS*, *PRIMA*, and *Athena* surveys, while Table 2.2 reports the naming scheme used in this work.

Table 2.1: Parameters of the *SPICA*, *OST*, *PRIMA*, and *Athena* surveys. The time per field $t_{\rm field}$ are pure integration times, without over-heads. The sensitivities are computed at 5σ. The *SPICA* R=150 sensitivities are reported at $20\,\mu$m. The B-BOP sensitivities refer only to the $70\,\mu$m channel, while those for the *OST* refer to channel 1 and channel 6. In this table, the *Athena* survey strategy is simplified, as the deep survey encompasses also the ultradeep pointings and the shallow survey will comprehend also the deep pointings. The *Athena* sensitivities refer to the $2-10\,$keV band. The surveys investigated and the naming scheme used in this work are summarized in Table 2.2.

Instrument		Survey	$t_{\rm field}$ (ks)	Sensitivity (μJy)	$t_{\rm tot}$ (hr)	Area (deg^2)
SPICA	SMI	ultradeep	67.86	3 (R=4), 50 (R=150)	605	1
	B-BOP		-	60	70	
	SMI	deep	5.22	13 (R=4), 160 (R=150)	482	15
	B-BOP		-	100	350	
OST	*OSS*	deep	-	4.5 – 79 (R=4), 39.4 – 614 (R=300)	1000	0.5
		wide	-	30 – 470 (R=4), 260 – 4090 (R=300)	1000	20
PRIMA	*PRIMAGER*	deep	-	4 – 6 (R=10), 10 – 50 (R=4)	600	3
Athena	*WFI*	ultradeep	4×1400	$\sim 3\times10^{-17}$ erg/s/cm^2	$\sim$ 1600	1.6
		deep	3×980+7×840	$\sim 4\times10^{-17}$ erg/s/cm^2	$\sim$ 2500	4
		shallow	106×84	$\sim 1.2\times10^{-16}$ erg/s/cm^2	$\sim$ 2500	42.5

Table 2.2: Reference surveys used in this work. The surveys composing the reference surveys were matched in similar area coverage.

Instrument	Survey	Area (deg^2)	ref. survey
SPICA	ultradeep	1.0	DEEP 1 deg^2
OST	deep	0.5	
PRIMA	deep	3.0	
Athena	ultradeep	1.6	
SPICA	deep	15	WIDE 15 deg^2
OST	wide	20	
Athena	shallow	42.5	

2.4 Simulations

To studies the synergies between the X-ray and IR instruments, we developed our own tool, able to quickly estimate the number of AGN detected at both wavelength. The tool uses XRB models to compute the intrinsic number of AGN and needs just the instruments capabilities to estimate the fraction of AGN detected as function of z, L_x, and N_H.

2.4.1 Simulation of the intrinsic AGN counts from XRB synthesis model

We have estimated the 'intrinsic' AGN number as a function of the redshift z, 2-10 keV intrinsic X-ray luminosities L_x and intrinsic column density N_H based on the Gilli et al. (2007) X-ray background synthesis model. Recent works (e.g., **?**) suggest that the number of AGN at $z > 3$ may be higher than the predicted by the Gilli et al. (2007) model, however, considering the uncertainties involved in these predictions, we can safely assumed that the number of AGN we provide is at least a lower limit. We have started by dividing the (z, L_x, N_H) space in bins, with 21 z bins $\in [0, 10]$, 18 L_x bins with $L_x \in [10^{42}, 1.5 \times 10^{48}]$ erg/s and 6 N_H bins with $\log(N_H/\mathrm{cm}^{-2}) \in [20, 26]$. We, then, make use of the software POMPA[¶] to get the $\log N(> S)$ source density in each of those bins, with S being the 2-10 keV source flux.

From the $N(> S)$ we also obtained the total intensity of each bin (energy received per unit area, unit time and unit sky area). Instead of the composite power law plus reflection spectral model in Gilli et al. (2007), we have adopted here a torus model from Brightman & Nandra (2011) with photon index $\Gamma = 1.9$, aperture angle 30 deg and inclination 80 deg. We have also included an additional scattering component modeled with a power-law with the same photon index and normalization equal to one percent of that of the primary emission.

In addition, we also included "normal" galaxies using the galaxy $\log N - \log S$ curve of Lehmer et al. (2012) and a power-law spectral shape with $\Gamma = 1$ (the median value for

[¶] http://www.bo.astro.it/~gilli/counts.html the 332 sources with spectral slope not based on upper limits in Lehmer et al. 2012) and $S \in [5 \times 10^{-20}, 5 \times 10^{-16}]$ erg cm^{-2} s^{-1}.

Physical fluxes are converted to counts using the latest (at the time of writing) set of FoV-averaged without filter WFI matrices[‖] using `xspec`[**] (Arnaud 1996).

Predictions of *Athena* source counts

When observed by the WFI on board *Athena*, all the sources above are seen against a "background" which includes several components[††,‡‡]:

- Particle background: due to cosmic rays impacting the detector and the surrounding structure. It is assumed to contribute 6×10^{-4} cts/keV/s/arcmin^2 with a flat power-law spectrum characterized by a spectral index of $\Gamma = 0$.

- Diffuse Galactic X-ray background: it is assumed to be uniform and to have the spectral shape and intensity from McCammon et al. (2002), which includes two thermal components, one unabsorbed representing emission from the Local Hot Bubble and one with foreground absorption by our own Galaxy representing emission from the hot Galactic halo, in `xspec` parlance: $\texttt{apec}_1$ + `phabs` $(\texttt{apec}_2)$ with kT_1=0.099 keV and kT_2=0.225 keV and respective `apec` normalizations of 1.7×10^{-6} and 7.3×10^{-7}, corresponding to 1 arcmin^2; the column density for the Galactic absorption is assumed to be the same as for extragalactic sources (see below).

- XRB: due to the integrated emission of the above AGN and galaxy populations. It is assumed to have intensity and power-law spectral shape from McCammon et al. (2002) but with $\Gamma = 1.45$ and normalisation 10^{-6} photons s^{-1}cm^{-2}keV^{-1}, again normalized to 1 arcmin^2.

- Stray light: X-ray photons from the two previous components, coming from outside the FoV but impacting the WFI after a single reflection on the mirror. We have modeled it by averaging the contribution of the two previous components over the full WFI FoV and fitting it with two thermal components and a power-law, with the same temperatures and photon index but with free relative normalizations.

All extragalactic components are assumed to undergo absorption by gas in our own Galaxy with hydrogen column density of $N_H = 1.8 \times 10^{20}$ cm^{-2}, which is the average Galactic column density of the estimated survey pointings in the MOP and coincides with the foreground column density in McCammon et al. (2002).

[‖]`http://www.mpe.mpg.de/ATHENA-WFI/response_matrices.html`

[**]`https://heasarc.gsfc.nasa.gov/xanadu/xspec/`

[††]`https://www.mpe.mpg.de/ATHENA-WFI/public/resources/background/WFI-MPE-ANA-0010_i7.1_Preparation-of-Background-Files.pdf`

[‡‡]Exact matrix used: `athena_wfi_rib2.3_B4C_20190122_wo_filter_FovAvg.rsp`

Depending on the (z, L_x, N_H) and the exposure time, a fraction f_{det} of the sources in each of the bins, described above, will be detected, contributing to resolve a fraction of the full intensity of the extragalactic XRB, thus lowering the remaining (unresolved) background. For each of the exposure time values of the survey layers, we have determined the resolved fraction iteratively starting from an initial fiducial value of 80% and f_{det} is also determined for each bin in each iteration. To decide whether a source is detected above the background, we used Cash statistic C (Cash 1979): calling T the total source+background counts, B the background counts, and t the exposure time, the mode (most probable value) of the source counts is $\hat{s} = \max(0, T - B)$. For a 5σ detection for a single free parameter (s), we required that the Cash statistic improves by $\Delta C = 25$ between $s = 0$ (i.e. no source) and $s = \hat{s}$. For each bin, 100 sources are simulated at the bin centre and f_{det} is the fraction of those sources that are detected. It turns out that the resolved fraction is around 80% and varies a few percent for exposure times above ~10 ks. In order to estimate the source+background and the background counts we have assumed a circular extraction area of radius 5.7 arcsec, which is the WFI FoV-weighted average of the *Athena* mirror HEW.

Athena source characterization

By their own nature, X-ray detectors work in photon-counting mode, thus they are able to measure the energy, the position on the detector, and the time of each photon (with various degrees of accuracy, depending on the detector technology). One of the consequences of this is that they all are "Integral Field Units", providing spectra for all detected sources. For *Athena*, this ranges from the superb ~2.5 eV resolution of X-IFU to the modest (but good for a Si-based detector) 170 eV resolution at 7 keV of WFI. We have simulated 100 spectra at the "centres" of each of the $1 \leq z \leq 4$, $10^{44} \leq L_x(\mathrm{erg/s}) \leq 5 \times 10^{45}$, $24 \leq \log(N_H/\mathrm{cm}^{-2}) \leq 26$ bins and fitted them. If both the fitted luminosity and the column density, obtained from the spectral fitting, are within 30% of the input values, we say that we have “characterized” the source. Finally, by dividing the number of characterized sources by 100 (the number of spectra we simulated) we have estimated the fraction of characterized sources in each bin.

2.4.2 Simulation of IR detections

To compute the AGN and host galaxy expected fluxes at different redshifts in IR bands, we need a proper set of SEDs to model the AGN as a function of different parameters: the intrinsic rest-frame 2 – 10 keV luminosity L_x, the equivalent hydrogen column density N_H and the redshift z.

We started with the AGN and host galaxy SED compilation from Lanzuisi et al. (2017). They used a sample of 2 333 X-ray selected (either *XMM-Newton* or *Chandra*) AGN in the COSMOS field (Scoville et al. 2007) with at least 30 X-ray counts from

Lanzuisi et al. (2013, 2015), Marchesi et al. (2016a) and then selected only those with a $> 3\sigma$ IR detection in one of the *Herschel* bands. The X-ray spectral analysis provided us with column density N_H and absorption-corrected $2-10$ keV luminosity for each source (see Lanzuisi et al. 2017 and references therein for more details). The final sample was composed of 692 sources X-ray and FIR detected, all with both X-ray spectral properties and SED decomposition. The SEDs were obtained by Delvecchio et al. (2014, 2015), following the recipes described in Berta et al. (2013), using the `sed3fit` code (with a "smooth" torus model), which allowed us to disentangle the AGN and host-galaxy contribution using three components – stellar emission, AGN torus emission and SF-heated dust emission – with photometric points from the UV to sub-mm. We excluded from our sample the sources with 2-10 keV rest-frame luminosities $\log(L_x/\mathrm{erg/s}) < 42$, as SF galaxies without AGN should not be able to exceed this threshold. Moreover, to overcome the degeneracy between the AGN and the SF contribution during the SED fitting, we selected only the SEDs with AGN significance $S_{AGN} \geq 99\%$, evaluated through the F-test between the best-fit reduced χ^2 with and without the AGN component (see Delvecchio et al. 2014). Thus, the final sample is composed of 422 AGN in the redshift range $0 < z \leq 4$, 2-10 keV rest-frame luminosities in the $42 \leq \log(L_x/\mathrm{erg/s}) \leq 45.5$ range and host-galaxy masses $7.9 \leq \log M/M_\odot \leq 12.2$.

With the aim of obtaining template SEDs representative of AGN with different intrinsic rest-frame $2-10$ keV luminosity and obscuration, we created four L_x bins and four N_H bins. The binning was chosen to maximize the number of sources in each bin, whilst maintaining a good sampling of the L_x and N_H parameter space. The binning used and the number of SEDs in each bin are shown in Table 2.3. Two of the most extreme bins had no sources in them; for these bins we chose to use SEDs randomly extracted from the two nearest bins (one with the same L_x and one with the same N_H). The binning choice was also driven by the binning we adopted to obtain the number of expected AGN (see Sec 2.4.1) and we used the same binning edges, although we merged some of the less populated bins.

Throughout this thesis, we assumed a bolometric correction of $k_{bol} = L_{bol}/L_x$, eq 2.1, from Lusso et al. (2012).

$$\begin{aligned}\log(L_x/L_\odot) = 0.230x + 0.050x^2 + 0.001x^3 + 1.256 \\ \text{with } x = \log(L_{bol}/L_\odot) - 12\end{aligned} \tag{2.1}$$

In this work, we will refer to the predicted AGN in the simulated redshift range $0 < z < 10$ with $20 \leq \log(N_H/\mathrm{cm}^{-2}) < 26$ and $42 \leq \log(L_x/\mathrm{erg/s}) < 48.2$ ($43 \leq \log(L_{bol}/\mathrm{erg/s}) < 49.7$), thus covering the whole simulated parameter space, as all the AGN. The naming scheme used for referring to AGN with different properties (namely, N_H and luminosity) is reported in Table 1.2.

Table 2.3: Number of binned sources, using the compilation of Lanzuisi et al. (2017) as explained in section 2.4.2, in each 2-10 keV rest-frame luminosity L_x and amount of obscuration N_H bin. The used N_H bins are: $20 \le \log(N_H/cm^{-2}) \le 22$; $22 < \log(N_H/cm^{-2}) \le 23$; $23 < \log(N_H/cm^{-2}) \le 24.18$; $24.18 < \log(N_H/cm^{-2}) \le 26$. We refer to the sources in the first two L_x bins as low-luminosity ones and to those in the other two as high-luminosity ones. We consider the sources in the first N_H bin as unobscured AGN, those in the second and third bins as obscured and those in the last N_H bin as CT-AGN. For the two low-luminosity CT bins we had no SED available and we chose to use the SED of the nearest bins.

$\log(N_H/cm^{-2})$	21.5	22.5	23.5	24.5
$42.0 \le \log(L_x/erg/s) < 42.9$	21	14	5	-
$42.9 \le \log(L_x/erg/s) < 43.5$	43	54	10	-
$43.5 \le \log(L_x/erg/s) < 44.2$	48	76	53	3
$44.2 \le \log(L_x/erg/s) < 48.2$	23	25	38	6

2.5 Results

2.5.1 Predictions of IR source counts

In Sec 2.4.1, we predicted the number of expected AGN (per deg^2) as function of z, L_x and N_H. For each z, L_x and N_H bin we randomly extracted 20 template SEDs, from the corresponding bins in Section 2.4.2), and assigned them a weight equal to one-twentieth of the total number of predicted AGN. We measured their total (AGN+host) flux densities in the SMI 34μm and B-BOP 70μm, 200μm 350μm bands (for *SPICA*), in the six *OSS* bands (for *OST*), and in all the *PRIMAGER* bands. We obtained the number sources detectable at 5 σ by summing the weights of the sources with flux larger than the survey sensitivity.

In case of detection, we also differentiated if we were primarily detecting either the AGN or the host galaxy, on the basis of which of the two components had the highest flux in the band we were considering. In the case of a *SPICA* SMI 34μm detection, we considered whether we could also detect the source with the SMI-LR mode (as the SMI-CAM photometric channel operates at the same time with the SMI-LR spectroscopic channel). We used, as SMI-LR sensitivity, a mean value between the sensitivities at the boundaries of the SMI-LR wavelength range. For *OST OSS* we always considered both the detection in photometric mode (R=4) and in spectroscopic mode (R=300). For *PRIMA* we consider just the photometric detection.

We iterated this whole process forty times (different SEDs were extracted at each iteration), each time obtaining an estimate for the number of detected sources, AGN/torus and spectroscopic detection. We computed the median of these numbers and used the 16th-84th percentiles as uncertainties. Extracting 20 SEDs for each bin (and not one SED for each of the expected detection in that bin) allowed us to save computational time while maintaining a result resolution of 5%. We chose to iterate the process forty times to have good statistical significance of the median and of the 16^{th}-84^{th} percentiles.

In conclusion, from the IR simulations, we derive, for each bin, the percentage of sources photometrically detected, spectroscopically detected, and for which it is the AGN emission to be primarily detected. From the X-ray simulations, we obtained the total number of AGN in each bin, as well as the fraction of sources photometrically detected and spectroscopically characterized using *Athena WFI*. In each bin, we computed the fraction of sources detected both in the IR and in the X-rays, as the minimum between the detection fraction in the two wavelength ranges. Although, in some bins, there may be sources detected in the IR and not in the X-rays, alongside those detected in the X-rays and not in the IR, usually the detection fraction in one of the wavelength ranges is much higher than the other, and we can safely consider the minimum as a reliably estimate of the fraction of sources detectable with both instruments.

We summarize the main results of our prediction in Tables 2.4 and 2.6. Figures 2.1, 2.2, 2.3, and 2.4 illustrate the number of photometric detections per deg^2 for *SPICA-SMI* DEEP, *SPICA-SMI* WIDE, *OSS* DEEP, and *OSS* WIDE, respectively. All the figures also report the number of sources detected by *Athena WFI*. The black lines are the total number of AGN expected, the red areas represent those which can be detected using *SPICA* SMI-CAM at 34μm (or band 1 of *OSS*), the blue areas those which can be detected in the X-rays by *Athena*, the purple areas are the AGN that can be detected both by *SPICA* SMI-CAM (or *OSS*) and by *Athena*. The red uniform area represents the IR detected sources in which the main component of emission is due to the host-galaxy, while for those represented with the starry red area, the AGN is the main contributor to the detected emission. Figures 2.6 and 2.7 have a similar color code, but with the dashed areas representing the sources for which we will have spectroscopic detection in a DEEP survey with a *SPICA* SMI-LR-like instrument or with the *OSS*. Finally, figure 2.11 (figure 2.12) shows a comparison of the number of sources that we will be able to detect in a DEEP (WIDE) survey with *SPICA* and *OST*, with the color code indicating the expected number of photometric bands in which we may detect the source (considering SMI-CAM and B-BOP for *SPICA* and the six *OSS* bands for the *OST*).

Table 2.4: Percentage of all the AGN, CT-AGN, CT-AGN at $z \leq 4$, and CT-AGN at $z \leq 2$ photometrically detected with various configurations of instruments and surveys. SMI refers to *SPICA SMI-CAM*. OSS refers to the sources detected with the *OST OSS* in photometric mode (R=4) in at least one of the bands, while the number within the parenthesis to those with detection in all the six *OSS* bands. SMI AGN and OSS AGN refer to the direct detection of AGN emission (thus the cases where the AGN is more luminous than the host-galaxy in the considered band) for respectively the *SPICA SMI-CAM* and the *OST OSS* instruments. WFI refers to the source photometrically detected by *Athena WFI*, while SMI+WFI (OSS+WFI) refers to the sources with both *SMI-CAM* (*OSS*) detection and *Athena* photometric detection.

	Survey	SMI	SMI AGN	OSS	OSS AGN	WFI	SMI+WFI	OSS+WFI
		%	%	%	%	%	%	%
All AGN	DEEP	88 ± 3	22^{+8}_{-6}	95^{+1}_{-2} (84 ± 3)	21^{+6}_{-7} (< 1)	52	51	51 (50)
	WIDE	72 ± 5	16 ± 6	83^{+3}_{-4} (60^{+6}_{-5})	15 ± 5 (< 1)	29	27	28 (25)
CT-AGN	DEEP	87 ± 3		96 ± 1 (84 ± 4)		20	20	20 (20)
	WIDE	71 ± 5		84 ± 3 (59^{+5}_{-6})		4	4	4 (4)
CT $z \leq 4$	DEEP	94 ± 2		99 ± 1 (89^{+4}_{-3})		22	22	22 (22)
	WIDE	80 ± 5		93^{+2}_{-3} (68^{+6}_{-7})		5	5	5 (5)
CT $z \leq 2$	DEEP	99 ± 1		99 ± 1 (92^{+2}_{-3})		28	28	28 (28)
	WIDE	84^{+5}_{-4}		98 ± 1 (78 ± 6)		6	6	6 (6)

Table 2.5: Same as 2.4 for PRIMA.

	Survey	PRIMA	PRIMA AGN	WFI	PRIMA+WFI
		%	%	%	%
All AGN	DEEP	82 ± 5	19 ± 7	52	50
CT-AGN	DEEP	84 ± 4		20	20
CT $z \leq 4$	DEEP	90 ± 3		22	22
CT $z \leq 2$	DEEP	91 ± 2		28	28

Table 2.6: Percentage of all the AGN, CT-AGN, CT-AGN at $z \leq 4$, and CT-AGN at $z \leq 2$ spectroscopically detected with various configurations of instruments and surveys. SMIsp refers to *SPICA SMI-LR*. OSSsp refers to the sources detected with the *OST OSS* in spectroscopic mode (R=300) in at least one of the bands, while the number within the parenthesis to those with detection in all the six *OSS* bands. WFIsp refers to the source spectroscopically detected by *Athena WFI*, while SMIsp+WFI (OSSsp+WFI) refers to the sources with both *SMI-LR* (*OSS*) spectroscopic detection and *Athena* photometric detection.

	Survey	SMIsp %	OSSsp %	WFIsp %	SMIsp+WFI %	OSSsp+WFI %
All AGN	DEEP	38 ± 6	81 ± 3 (55 ± 6)	20	28	47 (38)
	WIDE	18 ± 5	53 ± 5 (20^{+5}_{-4})	6	10	22 (11)
CT-AGN	DEEP	37 ± 6	82^{+3}_{-2} (55^{+6}_{-7})	1	16	20 (18)
	WIDE	17 ± 5	54 ± 6 (22^{+5}_{-6})	<1	3	4 (4)
CT $z \leq 4$	DEEP	43^{+8}_{-7}	92^{+6}_{-7} (63 ± 7)	2	18	22 (21)
	WIDE	18 ± 5	63^{+7}_{-6} (26^{+7}_{-6})	<1	4	5 (4)
CT $z \leq 2$	DEEP	55^{+9}_{-7}	98 ± 1 (75^{+6}_{-7})	2	22	28 (27)
	WIDE	20 ± 6	77^{+5}_{-6} (34^{+8}_{-7})	<1	5	6 (6)

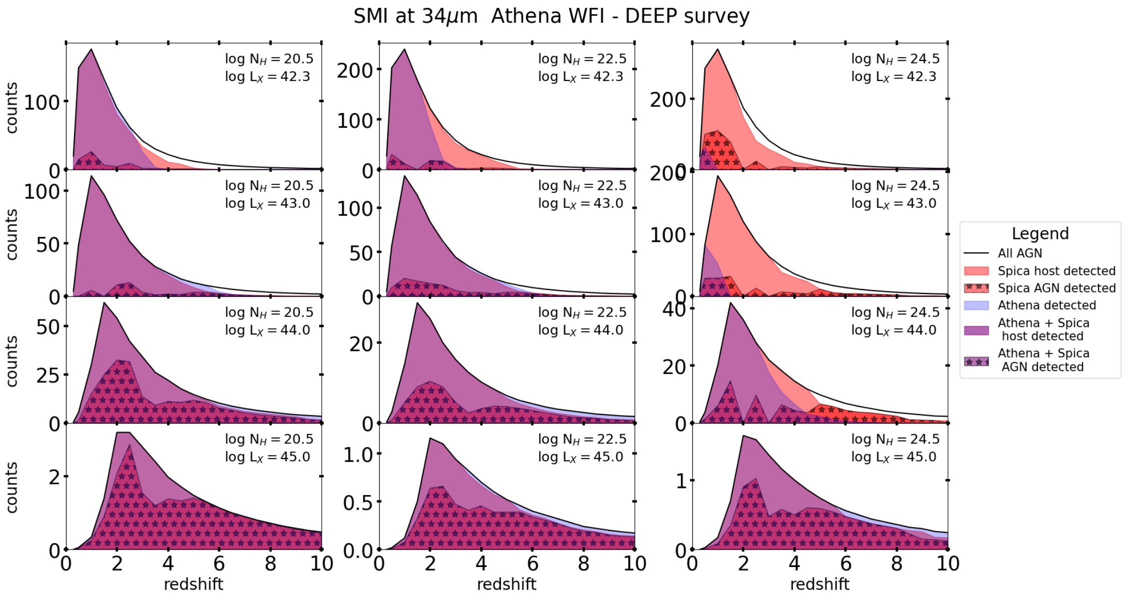

Figure 2.1: Number of AGN expected per deg^2 and per $\Delta z = 1$ for a *SPICA*-like DEEP survey. The black lines are the total number of expected AGN, the red areas represent those which can be detected with the *SPICA* SMI-CAM at 34μm, the blue areas those which can be detected in the X-rays with the *Athena WFI*, the purple areas are the AGN that will be detected both by *SPICA* SMI-CAM and *Athena*. The red uniform area represents the sources detected by *SPICA* in which the main component is due to the host-galaxy emission, while for those represented with the starry red area the AGN is the main contributor to the detected emission. The columns refer to AGN with different amount of obscuration ($20 \leq \log(\mathrm{N_H/cm^{-2}}) \leq 21$; $22 < \log(\mathrm{N_H/cm^{-2}}) \leq 23$; $24.18 < \log(\mathrm{N_H/cm^{-2}}) \leq 25$, from left to right), the rows to different AGN luminosity ($42.0 \leq \log(\mathrm{L_x/erg/s}) < 42.3$, $42.9 \leq \log(\mathrm{L_x/erg/s}) < 43.2$, $43.9 \leq \log(\mathrm{L_x/erg/s}) < 44.2$, $44.9 \leq \log(\mathrm{L_x/erg/s}) < 45.2$, from top to bottom). For forty times we extracted 20 SED for each bin, measured the flux in each *SPICA* band and compared them with the $5\,\sigma$ sensitivities to compute the number of detectable sources. The median of these forty values are the numbers used in this figure, while the $84^{th} - 16^{th}$ percentiles are used as uncertainties (not reported in the figure).

2.5.2 Photometric detections

Considering the DEEP survey, we will be able to detect in the IR $\gtrsim 90\,\%$ of all the AGN, approximately half of them will have photometric detection both in the IR and in the X-ray bands; this synergy will allow us to identify the source as an AGN and will help in placing better constraints to its properties. By a DEEP survey, working jointly, a IR cryogenic observatory (like *SPICA*, *OST*, and *PRIMA*) and *Athena* can detect all the AGN up to $z = 4$, *de-facto* completely covering the "cosmic noon" and bringing important insights about the AGN density evolution at higher redshifts. At the same time, the bulk of the intrinsic X-ray emission is produced at the "knee" of the luminosity function ($\mathrm{L_x} \sim 5 \times 10^{44}$ erg/s for $z \sim 1 - 4$, Aird et al. 2010) and these sources will be easily detected and characterized by *Athena* up to $z \sim 3$. Moreover, *Athena* will be of

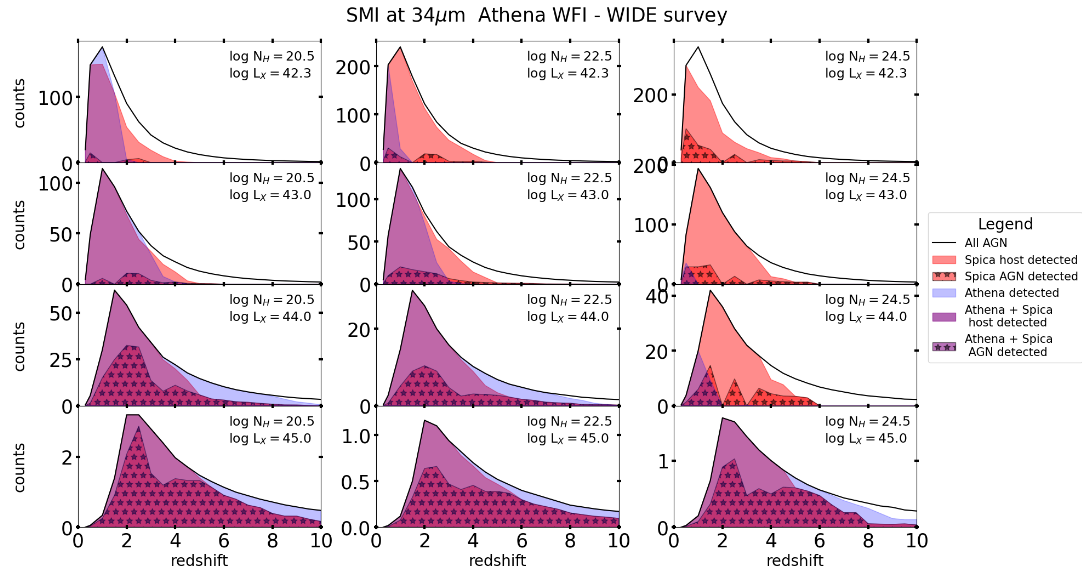

Figure 2.2: Number of AGN expected per deg^2 and per $\Delta z = 1$ for a *SPICA*-like DEEP survey. The lines and areas are coded as in Figure 2.1.

paramount importance to identify as AGN even sources for which the IR torus emission is diluted and hidden behind the more powerful host-galaxy contribution. The combined used of data in both bands will be fundamental in obtaining a more robust estimate of the AGN bolometric power. In particular, on the one hand, obtaining the AGN bolometric power with only X-ray data requires to assume an AGN X-ray bolometric correction, thus introducing large uncertainties in the derived values, due to our limited knowledge of the $k_{bol} - L_{bol}$ relation (Hopkins et al. 2007b; Lusso et al. 2010, 2012; Duras et al. 2020). On the other hand, using only an instrument similar to *SPICA* SMI-CAM, it will not be possible to distinguish between the AGN and the galaxy emission, as *SPICA SMI-CAM* can measure only the total (AGN+host-galaxy) emission.

Considering a WIDE survey, the fractions of sources that will be detected are lower by $\sim 20\,\%$ with respect to the DEEP survey, but the (at least) fifteen times wider area will allow us to detect at least 10 times more sources. Similarly to the DEEP, the WIDE survey may allow us to perfectly cover the “cosmic noon” with the advantage of a ten-folded statistic (at least). In a WIDE survey, *Athena WFI* may loose a significant fraction of the most obscured AGN or of the low-luminosity ones. However, a *SPICA*-like (or *OST*-like) mission should be able to effectively recover these sources, although we may need an effective way (e.g. *SPICA-SAFARI* or *OST* and *PRIMA* spectroscopic follow-up, see sec 2.5.3 and 5.5.2) to characterize their AGN properties (intrinsic bolometric luminosity and amount of obscuration). For this shallower survey it is still valid our statement that, for most of the detected sources at $\log(\mathrm{N_H/\,cm^{-2}}) \leq 23$, $\log(\mathrm{L_x/erg/s}) > 43$ and $z \leq 3$ (and higher z for higher $\mathrm{L_x}$), *Athena* detections would provide essential evidence of the AGN nature of those sources.

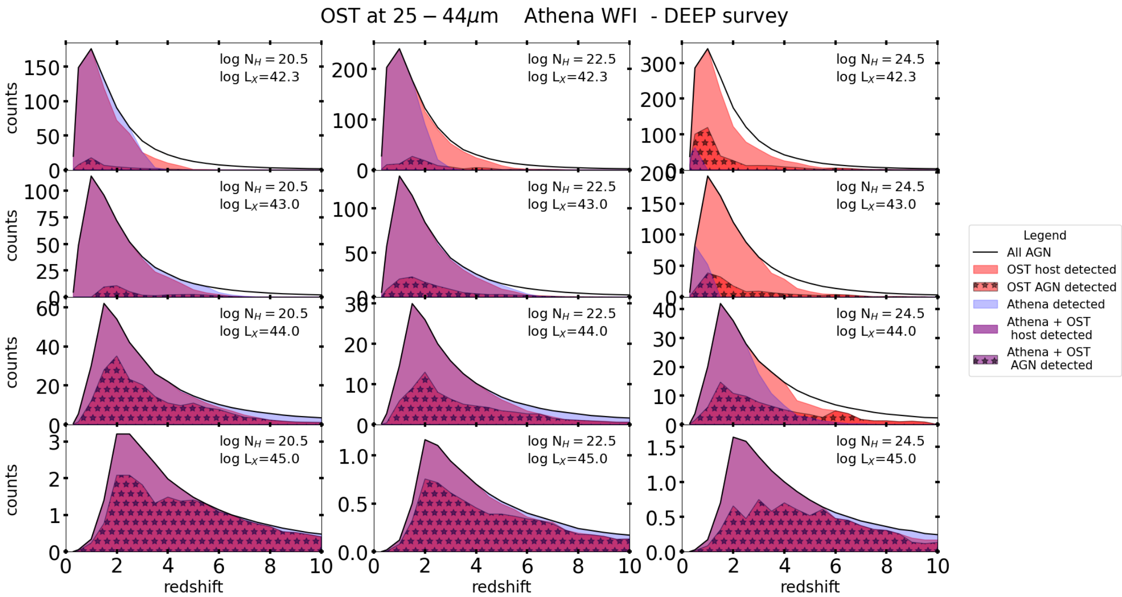

Figure 2.3: Number of AGN expected per deg^2 and per $\Delta z = 1$ for a *OST OSS* DEEP survey. The black lines are the total number of expected AGN, the red areas represent those which can be detected in the band 1 by *OST OSS*, the blue areas those which can be detected in the X-rays with the *Athena WFI*, the purple areas are the AGN that will be detected both by *OSS* and *Athena*. The red uniform area represents the sources detected by *OSS* in which the main component is due to the host-galaxy emission, while for those represented with the starry red area the AGN is the main contributor to the detected emission. The columns refer to AGN with different amount of obscuration ($20 \leq \log(\mathrm{N_H/cm^{-2}}) \leq 21$; $22 < \log(\mathrm{N_H/cm^{-2}}) \leq 23$; $24.18 < \log(\mathrm{N_H/cm^{-2}}) \leq 25$, from left to right), the rows to different AGN luminosity ($42.0 \leq \log(\mathrm{L_x/erg/s}) < 42.3$, $42.9 \leq \log(\mathrm{L_x/erg/s}) < 43.2$, $43.9 \leq \log(\mathrm{L_x/erg/s}) < 44.2$, $44.9 \leq \log(\mathrm{L_x/erg/s}) < 45.2$, from top to bottom). For forty times we extracted 20 SED for each bin, measured the flux in each *OSS* band and compared them with the $5\,\sigma$ sensitivities to compute the number of detectable sources. The median of these forty values are the numbers used in this figure, while the $84^{th} - 16^{th}$ percentiles are used as uncertainties (not reported in the figure).

Focusing on *OST*, the six spectrophotometric bands of *OSS* will give us a good coverage at and beyond the "cosmic noon" even in photometric mode. Moreover, thanks to the capabilities of *OSS*, for the majority of the sources detected by *Athena*, we will have six photometric points (and maybe up to six spectra) in the mid- and far-IR.

In Fig. 2.5, we show the number of expected AGN detected with the *PRIMAGER* filter 1A ($\sim 25\,\mu$m). As we can see, the capabilities of PRIMAGER do not differ significantly from that of a *SPICA*-like (or *OST*-like) mission. While the considered survey is shallower than those of *SPICA* or *OST*, the larger area will increase the total number of detected sources. Furthermore, the huge number of photometric filters will play a fundamental role in recovering the AGN and host-galaxy properties (see section 5.5.2).

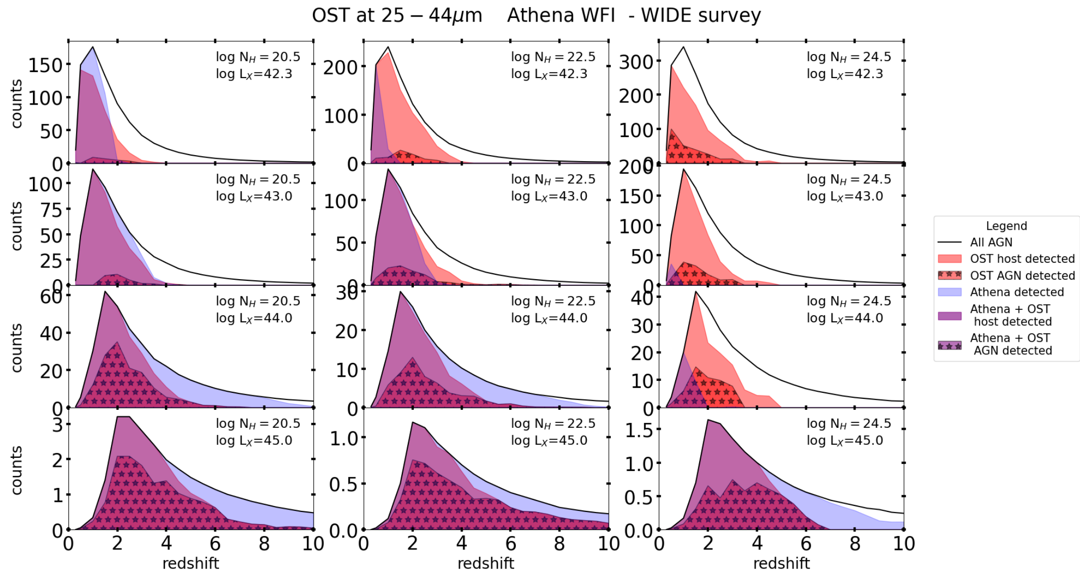

Figure 2.4: Number of AGN expected per deg^2 and per Δz = 1 for a *OST OSS* WIDE survey. The lines and areas are coded as in Figure 2.3.

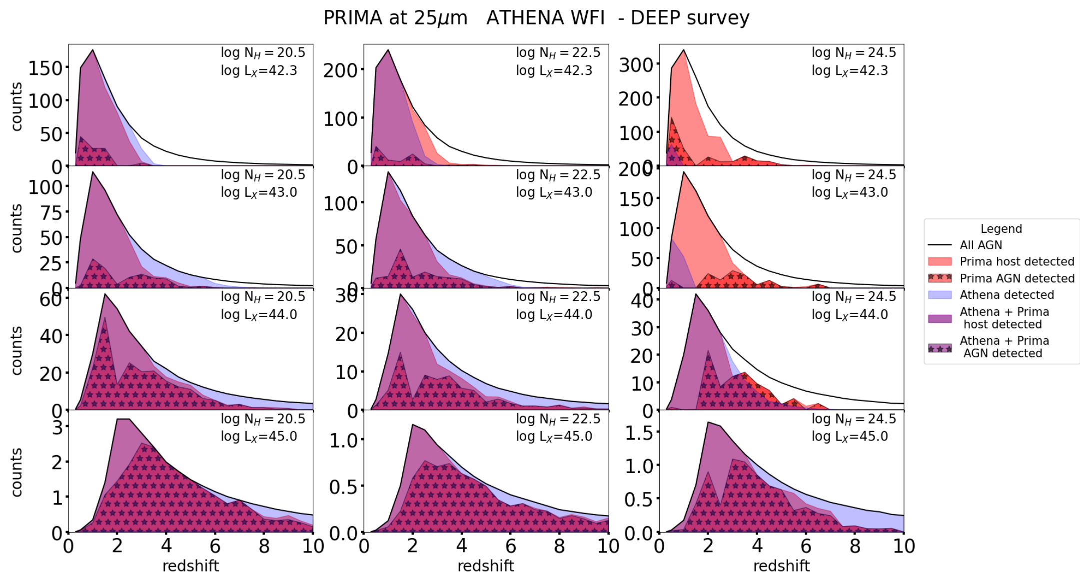

Figure 2.5: Number of AGN expected per deg^2 and per Δz = 1 for a *PRIMA* DEEP survey with the *PRIMAGER* filter 1A ($\sim$ 25 μm). The black lines are the total number of expected AGN, the red areas represent those which can be detected in the band 1A by *PRIMA*, the blue areas those which can be detected in the X-rays with the *Athena WFI*, the purple areas are the AGN that will be detected both by *PRIMA* and *Athena*. The red uniform area represents the sources detected by *PRIMA* in which the main component is due to the host-galaxy emission, while for those represented with the starry red area the AGN is the main contributor to the detected emission. The columns refer to AGN with different amount of obscuration.

2.5.3 Spectroscopic detections

We estimated that, exploiting *Athena WFI* spectra, we will be able to recover the intrinsic AGN luminosity and obscuration (with 30 % uncertainties) for $\approx$ 20 % of all the AGN and up to $\approx$ 50 % of the high-luminosity ones. Even low-count X-ray spectra will allow reliable N_H characterization, as for obscured and CT-AGN, we expect simple spectra with a flat continuum and, possibly, a strong iron Kα line. Furthermore, an X-ray detection at a level of $\sim 2\times10^{-16}$ erg cm^{-2} s^{-1} in the 2 – 10 keV band at $z \geq 1$, or simply any detection at $z \gtrsim 2.3$ in the DEEP survey, imply X-ray luminosities $\gtrsim 10^{42}$ erg/s in that band, revealing almost unequivocally an AGN.

Considering a SMI-LR-like instrument (figure 2.6), we find that with a DEEP survey we will be able to detect spectroscopically more than half of all the AGN within $z = 2$. For the high-luminosity objects, we expect to have some spectroscopic detections even up to $z \approx 4$. Considering a WIDE survey, the fraction of AGN that we should be able to detect spectroscopically drops considerably. For the sources for which a fast low-resolution spectroscopic characterization will not be possible, we may make use of photometric data (such as those produced by SMI-CAM- or B-BOP-like instruments) and of SED-fitting technique (see section 5.5.2) to disentangle the AGN and SF emissions.

Considering the *OSS*, we find that, with a DEEP survey, we will be able to detect spectroscopically between 55 $\pm$ 6 % and 81 $\pm$ 3 % of all the AGN, depending on the *OSS* band considered. Thus, we will have R=300 full IR spectra for more than half of all the predicted AGN at $z \leq 10$. For the high-luminosity objects, we expect to have some spectroscopic detections even up to $z > 4$. In the WIDE survey, the fraction of AGN that we should be able to detect spectroscopically lowers by $\approx$ 30 % with respect to the DEEP survey.

Several important galaxy and AGN mid-IR spectral features fall in the wavelength range of a *SPICA SMI-LR*-like (or *OSS*) instrument and can be of fundamental importance in both characterizing AGN and host-galaxy properties and measuring the redshift. In particular, we expect AGN-related high-ionization lines[§§], such as [Ne v]14.3μm and 24.3μm, [O iv]25.9μm (e.g. Tommasin et al. 2008, 2010; Feltre et al. 2016), but also SF-related Polycyclic Aromatic Hydrocarbons (PAHs) features (e.g. Leger et al. 1989) and the 9.7μm and 18μm silicate features. The 9.7μm feature is typically associated with type 1 AGN when observed in emission and with type 2 when it is in absorption (with the few notable exceptions of the type 2 quasars of Sturm et al. 2006; Teplitz et al. 2006, which show the silicate feature in emission); moreover, more obscured AGN usually have deeper silicate features, with depths that correlate with the X-ray derived N_H (e.g. Wu et al. 2009; Shi et al. 2006, but see also Goulding et al. 2012).

Considering the AGN spectroscopic properties, the most evident mid-IR AGN lines ([Ne v]14.3μm and 24.3μm, [O iv]25.9μm) are redshifted out of a *SMI-LR* spectral range at modest redshift ($z < 1$), and would require follow-ups with either *SPICA-SAFARI* or *SMI-MR* to be properly detected. In particular, exploiting 5 hours pointed observations with a *SPICA SAFARI*-like instruments, for low-luminosity (high-luminosity) un-

[§§]The high-ionization potential assures us that the emission can not be associated with star-formation activity and has to be, instead, linked to AGN activity.

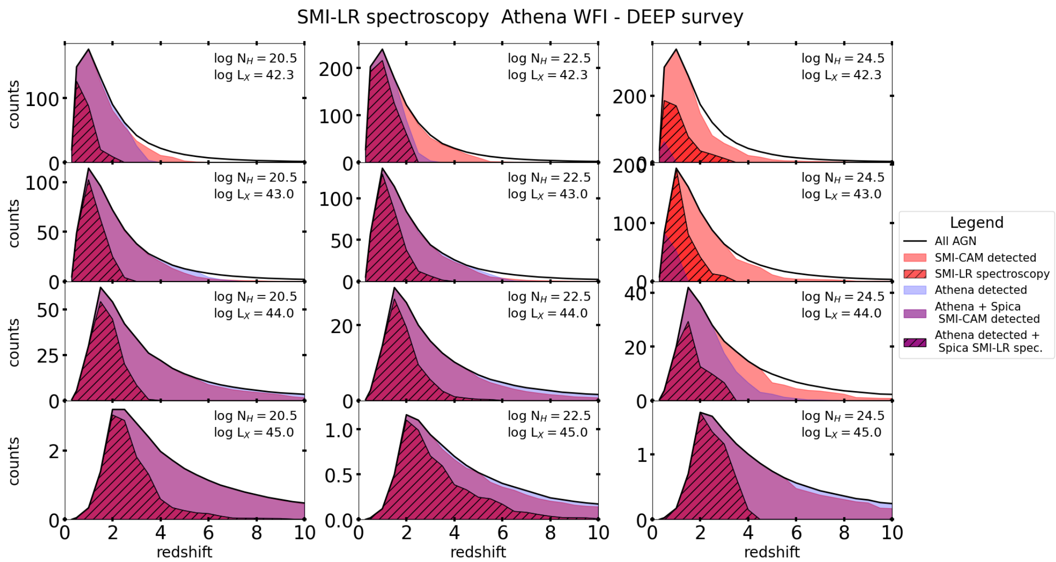

Figure 2.6: Number of AGN expected per deg^2 and per Δz = 1 for a *SPICA*-like DEEP survey. The black lines are the total number of AGN expected, the red areas represent those which can be detected with the *SPICA* SMI-CAM at 34μm, the blue areas those which can be detected in the X-rays by *Athena SMI*, the purple areas are the AGN that will be detected both by *SPICA* SMI-CAM and *Athena*. The dashed areas represent the sources which could be spectroscopically detected with SMI-LR (5 σ detection with a resolution R= 50 – 120 in the 17 – 36 μm wavelength range). The columns refer to AGN with different amount of obscuration ($20 \leq \log(\mathrm{N_H/cm^{-2}}) \leq 21$; $22 < \log(\mathrm{N_H/cm^{-2}}) \leq 23$; $24.18 < \log(\mathrm{N_H/cm^{-2}}) \leq 25$, from left to right), the rows to different AGN luminosity ($42.0 \leq \log(\mathrm{L_x/erg/s}) < 42.5$, $42.9 \leq \log(\mathrm{L_x/erg/s}) < 43.2$, $43.9 \leq \log(\mathrm{L_x/erg/s}) < 44.2$, $44.9 \leq \log(\mathrm{L_x/erg/s}) < 45.2$, from top to bottom).

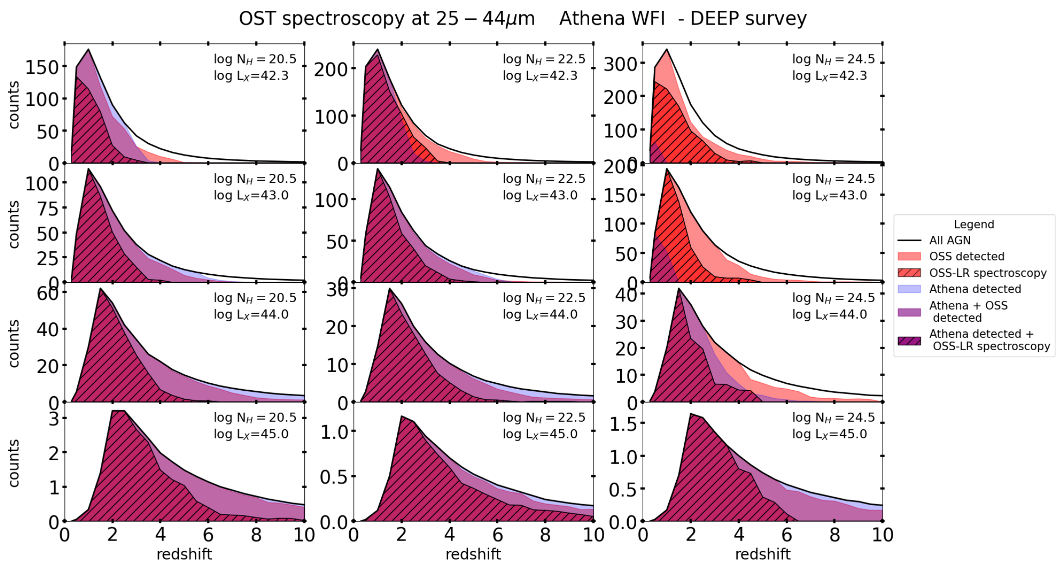

Figure 2.7: Number of AGN expected per deg^2 and per Δz = 1 for a *OSS* DEEP survey. The lines and areas are coded as in Figure 2.6, but with the red areas referring to the band 1 of the *OST OSS* instrument.

obscured AGN the [Ne v]24.3μm and [O iv]25.9μm lines can be studied up to $z \approx 1$ and $z \approx 2$ ($z \approx 3$ and $z \approx 4$). The same lines can be detected up to $z \approx 1.5$ and $z \approx 2.5$ ($z \approx 4$ and $z \geq 4$) for low-luminosity (high-luminosity) obscured AGN (Spinoglio et al. 2017, 2021). For the majority of the sources, the brightest and most recognizable features may be due to the galaxy emission, in particular the dust continuum and the PAH emission lines (e.g. Lutz et al. 1998; Hollenbach & Tielens 1999; Yan et al. 2007; Fadda et al. 2010; Spinoglio et al. 2021). However, for most cases, we expect the 9.7μm silicate feature, to be easily recognizable. The silicate feature can be helpful in estimating the AGN amount of obscuration (e.g. Vignali et al. 2011; La Caria et al. 2019) and can be entirely within a SMI-LR-like instrument wavelength range up to $z \approx 3$. This is particularly important for the CT-AGN, for which, with the exception of the most luminous ones ($L_x \approx 10^{44}$ erg/s for $z < 3$), it will be difficult to obtain the AGN properties using only X-ray spectroscopy. Thus, the analysis of both X-ray and IR spectroscopic data will play a fundamental role in discerning their characteristics.

One of the advantages of using a *SPICA SMI*-like (or *OSS*-like) instrument is that it can provide, at the same time, photometry and spectroscopy, and it can quickly produce low-resolution spectroscopic data for a large number of sources. X-ray spectral analysis is, so far, the best way to characterize the AGN (e.g. to determine the intrinsic luminosity), but it requires a redshift determination. Fortunately, there is a good prospect that redshifts can be determined solely from X-ray spectra (e.g. Iwasawa et al. 2012; Simmonds et al. 2018; Peca et al. 2021). We simulated mid-IR spectra to evaluate the possibility of measuring the redshift from a *SMI-LR*-like spectra. In particular, we used a code that exploits the most prominent PAH features to recover the redshift. This implies that it is most suited for sources with sufficient star-formation to produce strong PAH features and low- or moderate-luminosity AGN, as the equivalent width of the PAH features is greatly reduced when the AGN luminosity is high (Voit 1992; Genzel et al. 1998; O'Dowd et al. 2009). We chose ten sources among the ~ 500 of our sample and use their SEDs as spectral templates. The ten sources were chosen so that in our sample there were un-obscured, obscured and CT-AGN, high- and low- luminosity sources, AGN- and host-dominated sources, with emission and absorption 9.7 μm features. For each template, we simulated sixteen spectra in the redshift range $0.5 \leq z \leq 6$ (as we can see from Fig 2.6, at $z > 6$ the fraction of spectroscopically detected sources is very small). The simulated spectra were created accordingly to the latest *SMI* specifications. We used a spectral resolution of $R = 50$ at 17 μm and R= 120 at 36 μm and assumed a linear regime between these two values. We used the LR continuum sensitivity (see section 2.3.2), rescaled to the Ultradeep survey, as standard deviation of the Gaussian noise that we included in the spectra. Figure 2.8 shows two of the simulated spectra. Finally, we exploit a modified version of the PAHFIT code (Smith et al. 2007) that uses a power-law continuum, several PAH emission lines and silicate absorption, to fit the spectra and provide the source redshift (Negrello et al. in prep). As shown in figure 2.9, we find that we should be able to effectively recover

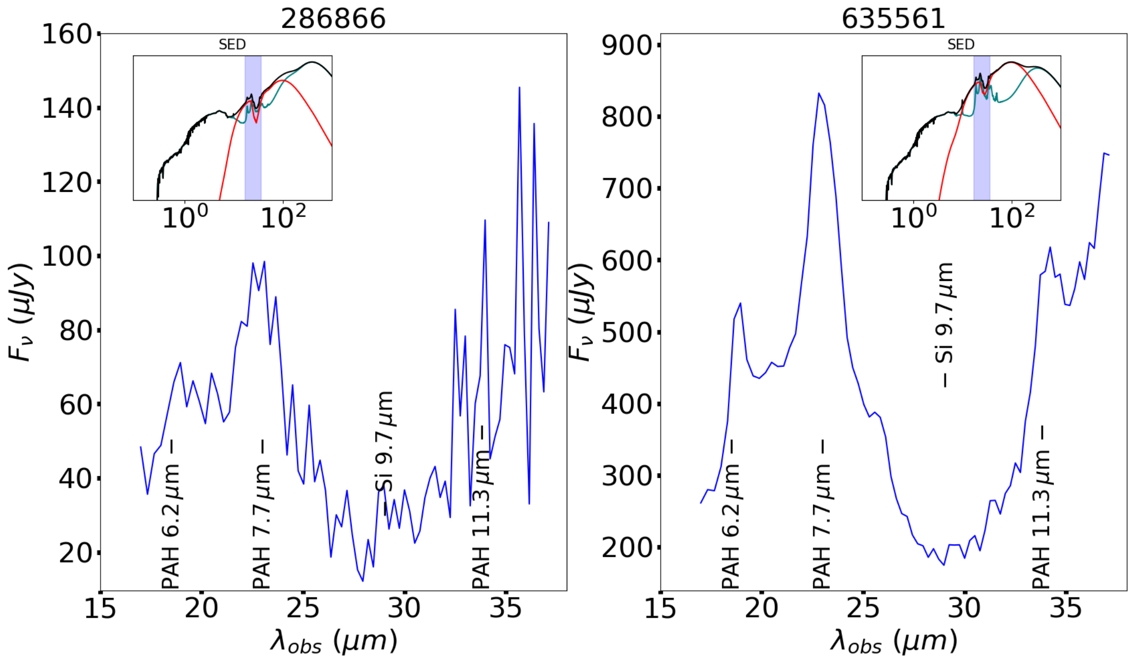

Figure 2.8: Simulated SMI-LR spectra for a *SPICA*-like DEEP survey. Left panel is for $\log(\mathrm{L_x/erg/s}) = 43.9$, $\log(\mathrm{N_H/cm^{-2}}) = 23.4$ (i.e. moderate-luminous obscured) AGN at $z = 2$; right panel for $\log(\mathrm{L_x/erg/s}) = 44.3$, $\log(\mathrm{N_H/cm^{-2}}) = 24.1$ (i.e. high-luminous) CT-AGN $z = 2$. We used the Lanzuisi et al. (2017) SEDs (inset in the upper left of the plots, where the red line is the AGN component, the blue one is related to the host-galaxy and the black one is the total AGN+galaxy emission) as spectral templates and added a white noise with amplitude based on the LR continuum sensitivity expected for a DEEP survey. The spectra were sampled with a resolution of $R = 50$ at 17 μm and R= 120 at 36 μm.

the redshift in case of moderate and luminous type 2 AGN up to $z \sim 3-4$. However, the code currently does not use AGN silicate 9.7μm emission feature in finding the z and is prone to mis-interpret it as a PAH feature when it is strong; however, the code is still able to recover the redshift in case of faint AGN silicate emission feature, as long as the silicate line is not mistakenly interpreted as a PAH feature. In particular, excluding three sources with strong Si 9.7μm emission features and one source with very low signal-to-noise ratio, we recovered the redshifts with a median error $|z_{best} - z_{true}| / (1 + z_{true}) = 0.02$. In conclusion, a *SPICA SMI*-like instrument can provide us with low-resolution spectra for more than 6 000 AGN in a DEEP survey and, in the cases of sources with prominent host-galaxy PAH lines and low AGN activity (i.e. unable to destroy these lines), measures of their spectroscopic redshifts. Moreover, it should be possible to obtain an estimate of the AGN obscuration from the depth of the silicate 9.7μm absorption feature (Wu et al. 2009; Shi et al. 2006 had shown that there is a correlation between these two quantities), that may provide support to the outcomes of X-ray spectroscopy in terms of obscuration (especially for low-statistics obscured AGN).

Focusing on the *OSS* capabilities, this instrument is able to exploit the above men-

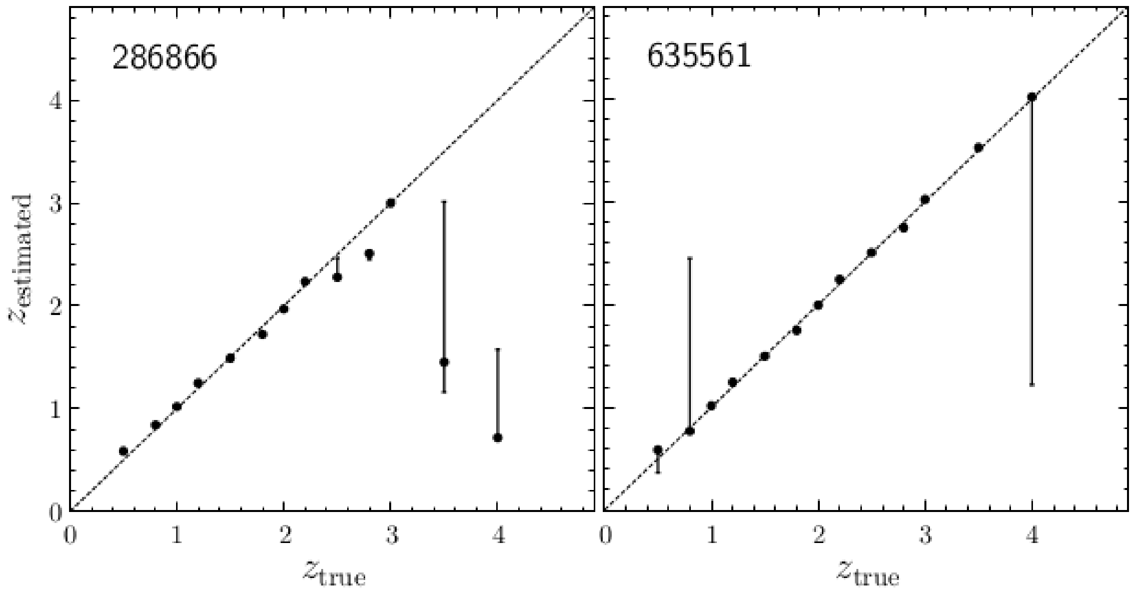

Figure 2.9: Redshift estimate from simulated *SMI-LR*-like spectra using a modified version of the PAHFIT code (Negrello et al. in prep). Left panel is for $\log(L_X/\text{erg/s}) = 43.9$, $\log(N_H/\text{cm}^{-2}) = 23.4$ moderate-luminous obscured AGN; right panel for $\log(L_X/\text{erg/s}) = 44.3$, $\log(N_H/\text{cm}^{-2}) = 24.1$ high-luminous CT-AGN. We find that we should be able to effectively recover the redshift up to $z \approx 3-4$; at $z > 4$, there are too few strong PAH features in the *SMI-LR* wavelength range to allow a proper determination of the redshift.

tioned mid-IR lines, but with the advantages of higher spectral resolution and larger wavelength range coverage than a *SPICA*-like instrument. Lines as the [Ne v]14.3μm and 24.3μm, [O iv]25.9μm, which are redshifted out of *SMI-LR* spectral range at modest redshift ($z < 1$), thus requiring follow-ups with other instruments, can be detected up to very high-redshift, their detection being a matter of instrument sensitivity and not of limited wavelength range. Fine structure lines of O, C, Ne, S, N, Fe, Ar, and Si are keys to probe neutral and ionized gas phases and can be used to estimate redshift, amount of gas, its ionization, presence of outflows and the contribution of SF and AGN to the emission (Meixner et al. 2019; Mordini et al. 2021). The *OSS* DEEP survey will allow us to detect the PAH up to $z \approx 4$ for galaxies with SFR $\sim 10 M_\odot$/yr and over $z > 6$ for those with SFR $\sim 100 M_\odot$/yr (Meixner et al. 2019). Regarding the estimate of the BHARD, the high sensitivity and large spectral range of the *OSS* survey will allow the detection of the [Ne v] line (whose flux correlates with the AGN intrinsic emission), as well as of the [O iv] and [Ne ii] line, whose line flux ratio can be used as a diagnostic of AGN power (e.g., Genzel et al. 1998; Lutz et al. 2003; Armus et al. 2006, 2007; Mordini et al. 2021). Gruppioni et al. (2016) well calibrated the relations between [Ne v] (or [O iv]) and AGN bolometric emission in samples of local galaxies. Bonato et al. (2019) explored the spectroscopic survey capabilities of the *OSS* and find out that, in a DEEP survey, it can detect emission lines associated to SF up to $z \approx 8.5$ and those of AGN emission, in particular [Ne v]24.3μm and [O iv]25.9μm, up to $z \approx 3$ and $z \approx 5$, respectively.

2.5.4 SED fitting

Working jointly, an IR cryogenic observatory, like *SPICA* or *OST*, and *Athena* can detect a huge fraction of the the AGN up to $z = 10$. The combined used of data from both wavelength bands will be fundamental, as with only IR photometric detections, it will not be possible to distinguish between the AGN and the galaxy emission, as *SPICA SMI-CAM* or *OSS* can measure only the total (AGN+host-galaxy) emission, this in the lucky case of having already identified the source as an AGN. In this regard, a spectrometer similar to *SPICA* SMI-LR or *OSS* may surely help, but we will not be able to spectroscopically detect all the AGN, as the spectroscopic recovery fraction decreases rapidly for $z > 2$. The additional use of other photometric data (such as those of B-BOP-like instruments, or from all the *OSS* bands), along with SED decomposition technique, could be used to overcome this limitation. For example, considering all the 16 *PRIMAGER* bands, we found that with a DEEP survey, we will have an extremely good coverage, being able to detect the majority of the sources in all bands over the entire "cosmic noon" (Fig. 2.10). Moreover, for the most luminous sources we will have detection in a significant number of bands even at $z > 6$. Even in photometric mode, having multiple far-IR detections will be invaluable in disentangling the AGN and host-galaxy contribution: indeed, these bands cover both the regime where the torus has its peak emission and the regime where the SF is the main contributor to the SED. But we expect that the real game-changer in obscured AGN studies will come from a multi-wavelength approach.

The use of *Athena* data, optical and near-IR photometric datapoints (in fields where these are already available, such as COSMOS and the Euclid Deep Fields), mid-IR and far-IR photometric instruments (similar to those designed for *SPICA*) can provide a never-reached-before coverage of all the AGN and galaxy emission up to high redshift. In addition, with follow-up pointed spectroscopic observations of these targets (photometrically detected, but without a clear separation between galaxy and AGN components) we will be able to distinguish the AGN and host-galaxy emissions using fine-structure lines and PAH features up to $z \sim 4$ (Spinoglio et al. 2021). Therefore, we expect to be able to improve our knowledge of the AGN $k_{bol} - \mathrm{L}_{bol}$ relation, similarly to what was done in Lusso et al. (2012), but with a larger sample and to higher redshift. This should enable us to extend the $k_{bol} - \mathrm{L}_{bol}$ dynamical range and, having more sources, to reduce its uncertainties.

In figure 2.11, we show the comparison of the total number of AGN detected by a *SPICA*-like observatory and by the *OST OSS* for a DEEP survey. The color code indicates the number of bands for which we can have photometric detection, considering SMI-CAM and the three B-BOP channel for *SPICA* and the six *OSS* bands for *OST*. For $z < 2$ both mission are able to detect almost all the sources in at least one band. While at $z > 2$ *SPICA* begins to lose a fraction of the sources, *OSS* continues to detect them all up to $z \approx 3.5$.

Moreover, having a sufficient number of photometric points, would allow us to break the SF-AGN emission degeneracy and to estimate the AGN properties using SED-fitting

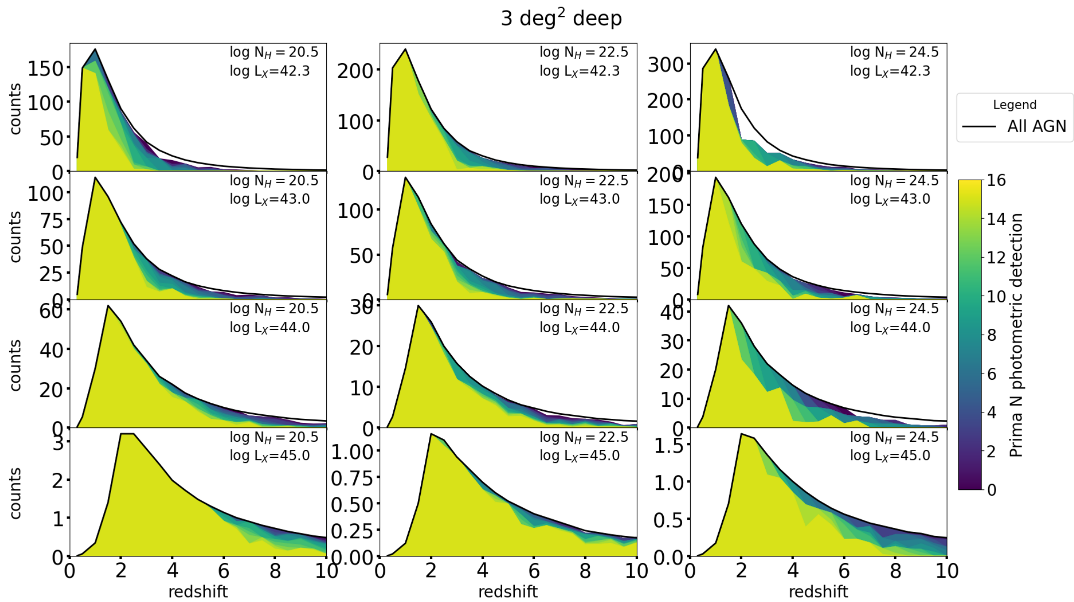

Figure 2.10: Number of AGN expected per deg^2 and per $\Delta z = 1$ for a *PRIMA* DEEP survey. The color code indicates our predictions for the number of photometric detections.

decomposition techniques (jointly with additional instruments at different wavelengths) when spectra are not available. Although X-ray photometry and/or low-resolution spectroscopic data may not be used in SED fitting (but see X-CIGALE, Yang et al. 2020), X-ray information could be crucial to select the templates to fit in the IR.

As a quick test for this statement, we simulated *SPICA*-like observations and used `sed3fit` to recover the stellar mass, SFR and AGN bolometric luminosity. In particular, we selected a random sample of seventy sources among the $\approx$ 550 of the Lanzuisi et al. (2017) compilation that we used in this work and exploit `sed3fit` with up to 33 Laigle et al. (2016) photometric data (from optical to sub-mm including *Spitzer* and *Herschel*, although not all the photometric points were available for all the sources) to obtain the AGN, host-galaxy and total SED. We used the total SED to estimate the *SPICA SMI* and *B-BOP* fluxes, added these four photometric points to the previous 33 and used all of them in new runs of `sed3fit`. We focused on the width of the $84^{th} - 16^{th}$ percentile of the probability distribution functions (PDF) from `sed3fit` to check for the improvement provided by the use of *SPICA*-like data. We found that for, respectively, 50 and 45 of the 70 sources investigated we have narrower AGN bolometric luminosity and SFR PDFs. The median SF relative error (i.e. $0.5 \times (SFR_{84^{th}} - SFR_{16^{th}})/SFR_{best}$) improves from 11 % to 1 % and that of the AGN L_{bol} from 61 % to 29 %. The improvements in the stellar mass are smaller, since `sed3fit` estimates this parameter exploiting the optical wavelength band, therefore we get only a minor advantage in using more IR data. The major improvements come from the cases where the AGN-SF degeneracies are stronger; as we can see from Figure 2.13, the additional mid- and far-IR points provided by a *SPICA*-like observatory help in disentangling the two contributions and allow us to obtain

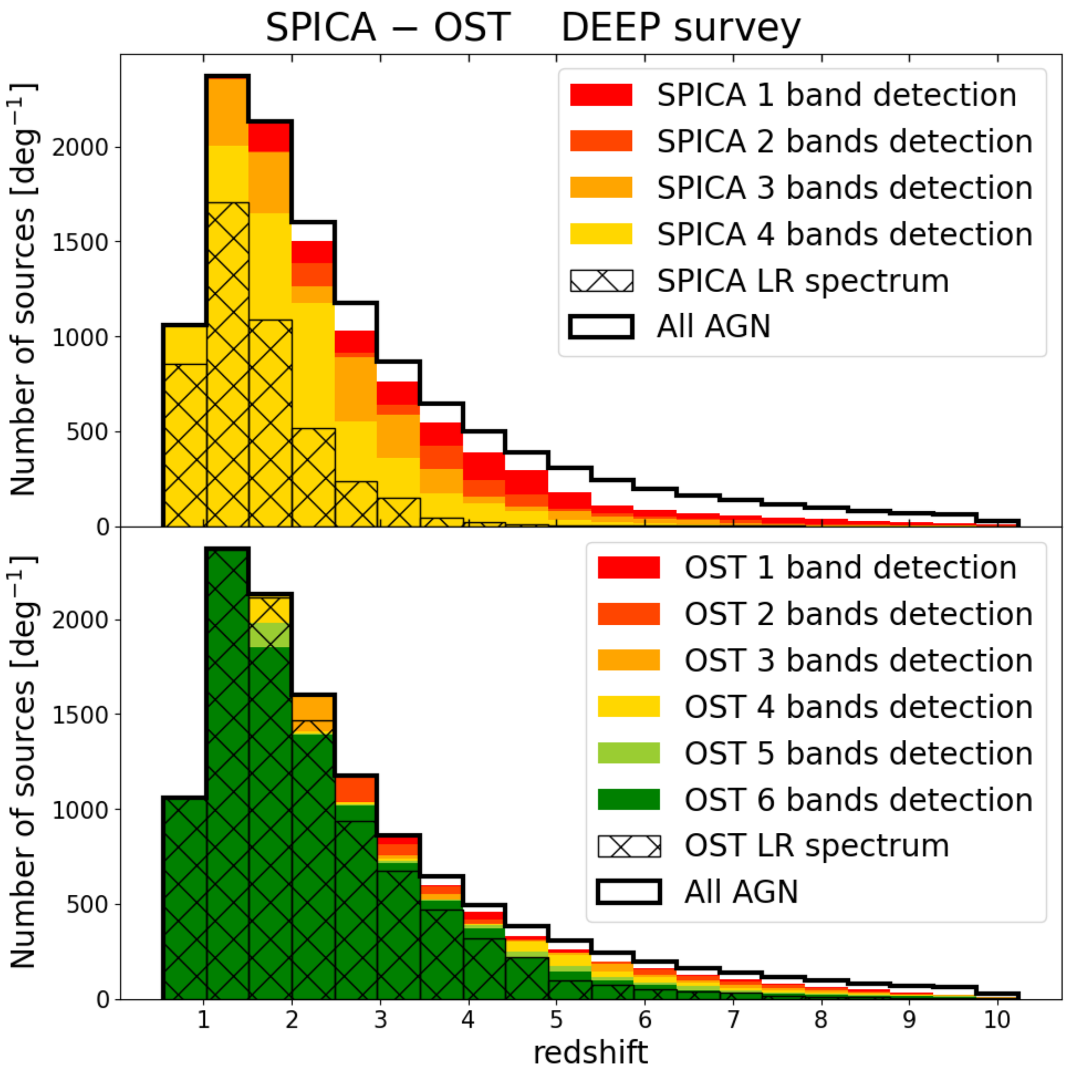

Figure 2.11: Distribution of the number of AGN expected per deg^2 for a *SPICA*-like (*top panel*) and for the *OST OSS* (*bottom panel*) from a DEEP survey. The color code indicates our predictions for the number of photometric detections. For the *SPICA* panel we took into consideration the SMI-CAM and the three B-BOP channel, while the dashed area represents the sources which can be spectroscopically detected with a SMI-LR-like instrument (5 σ detection with a resolution R= 50 – 120 in the 17 – 36 μm wavelength range). For the *OST* panel, we consider the six *OSS* bands in photometric mode (R=4), while the dashed line indicates the sources for which we can have R=300 spectra in at least one of the *OSS* bands.

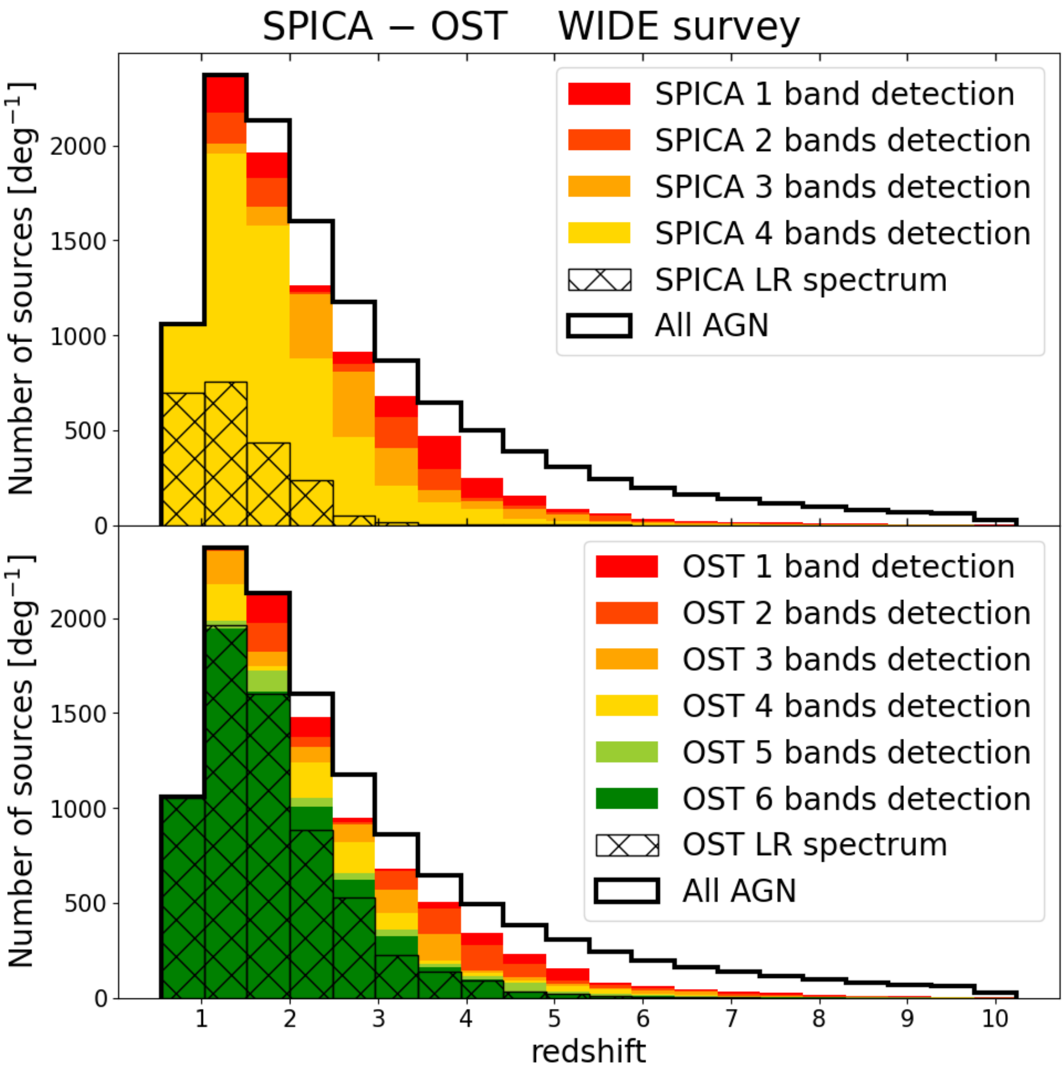

Figure 2.12: Distribution of the number of AGN expected per deg^2 for a *SPICA*-like (*top panel*) and for the *OST OSS* (*bottom panel*) from a WIDE survey. The lines and areas are coded as in Figure 2.11.

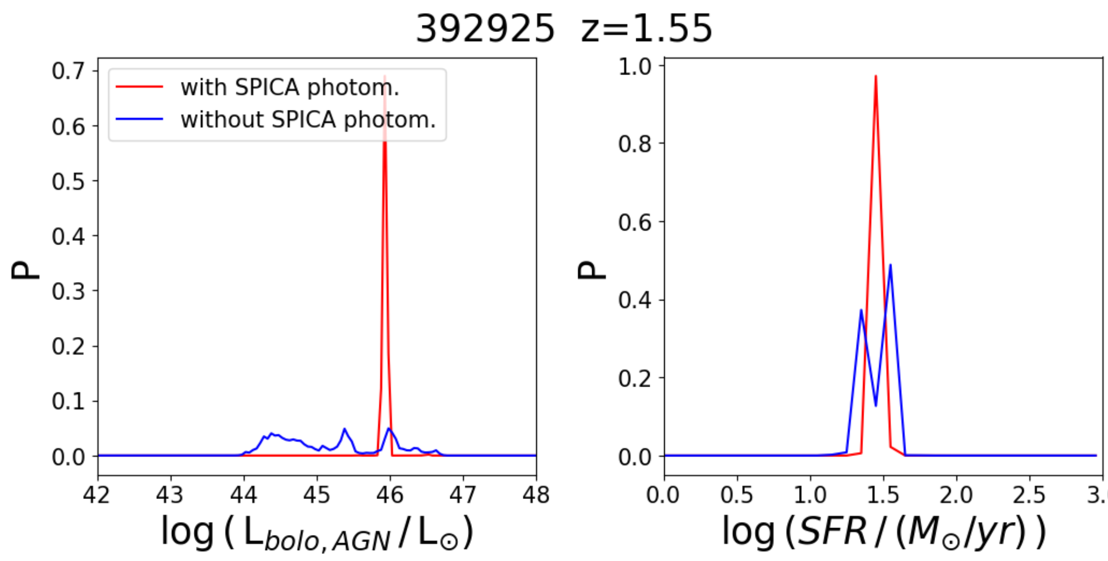

Figure 2.13: SED-fitting probability distribution functions (PDFs) for AGN 392925 ($\log(L_x/\mathrm{erg/s})$ = 45.0, $\log(N_H/\mathrm{cm}^{-2})$ = 24.1) at z = 1.55. *Left panel*: AGN bolometric luminosity; *right panel*: SFR in the last 0.1 Gyr. The blue lines are obtained from the SED-fitting using 33 photometric bands from optical to far-IR (in particular, from IR telescope *Spitzer* and *Herschel*). The red lines are the PDF obtained using the same 33 photometric bands plus the four simulated photometric observations of *SPICA SMI-CAM*, *B-BOP1*, *B-BOP2* and *B-BOP3*. The use of the additional *SPICA* photometric points provides better constraints on the AGN bolometric power and, overall, allows us to properly disentangle the AGN and the SF emission.

better estimates of both the host-galaxy and AGN parameters (e.g. stellar mass, SFR, AGN L_{bolo}).

2.5.5 CT-AGN

The search for CT-AGN is of fundamental importance in view of the still open questions of the missing sources responsible for the XRB and the galaxy co-evolution paradigm. About the former issue, X-ray background synthesis models predict a putative population of CT-AGN challenging to detect with current facilities (Gilli et al. 2007; Gilli 2013; Harrison et al. 2016). In addition, the black-hole-galaxy co-evolution paradigm states that the scarcely known first phases of BH (and galaxy) growth should be associated with very obscured AGN activity and obscured SF, therefore these elusive AGN are of the foremost importance to search for (e.g. Silk & Rees 1998; Di Matteo et al. 2005; Lamastra et al. 2013).

The adopted model for the XRB predicts that low-luminosity CT-AGN form the bulk of the very obscured AGN population. However, due to their large obscuration, it is difficult to detect them with X-ray instruments, while a *SPICA*-like (or *OST*) observatory should be able to effectively sample them with a detection fraction as high as 90% in the DEEP survey and 70 % in the WIDE survey.

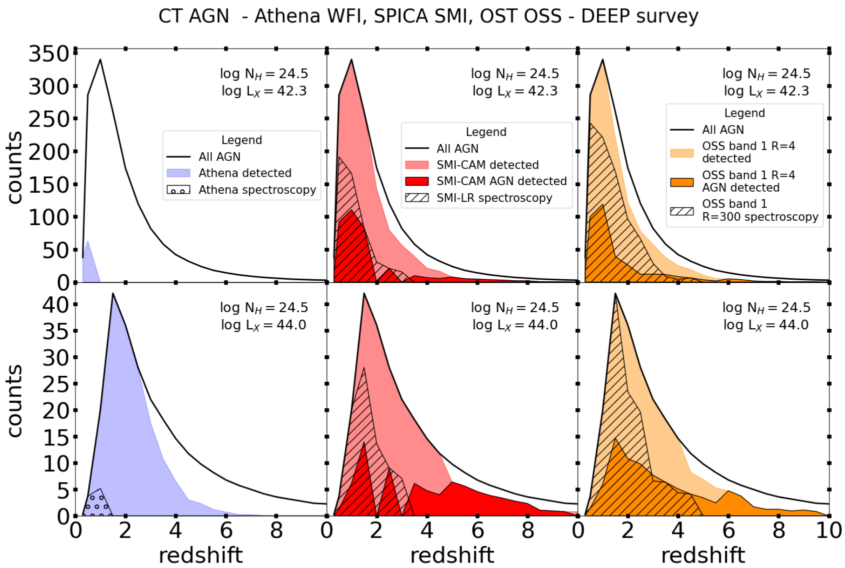

Figure 2.14: Number of CT AGN expected per deg^2 and per Δz = 1 from a DEEP survey with *Athena* (*left panels*), with a *SPICA*-like observatory (*middle panels*), and with the *OST OSS*(*right panels*). The black lines are the total number of expected CT-AGN, the blue areas represent those which can be detected in the X-ray by *Athena WFI* and the dotted areas those for which we will be able to recover the L_x and N_H (with 30 % uncertainties) using *Athena WFI* spectroscopy; the red areas represent those which can be detected in the IR with the *SPICA* SMI-CAM at 34μm. The light red areas indicate that the host galaxy has the largest contribution to the flux in the SMI-CAM detection, while the dark red areas that the AGN is the dominant component. The dashed areas indicate that we will be able to have also mid-IR spectroscopic 5 σ detection with *SPICA* SMI-LR with a resolution R= 50 – 120 in the 17 – 36 μm wavelength range. The light orange areas are the sources for which we expect detection in the *OSS* band 1 (25 – 44 μm) in photometric mode (R=4), the dark orange area those for which the AGN is the dominant component of the detected emission. The dashed areas refer to AGN for which we will have enough flux to have low-resolution (R=300) spectra in *OSS* band 1. For the *OST*, we showed only the prediction for the band 1 of the *OSS*, but the instrument has five other bands (up to 588 μm), for which we predict a larger number of AGN detected both in photometry and in spectrometry. The upper row refers to AGN with luminosity $42.0 \leq \log(L_x/\text{erg/s}) < 42.5$, while the lower to those with $43.9 \leq \log(L_x/\text{erg/s}) < 44.2$.

As we can see from Figure 2.14, low-luminosity obscured AGN are difficult to detect even with the planned *Athena* facility. In fact, we estimated that *Athena* will be able to detect (light blue) more than half of all the high-luminosity CT-AGN, while the major part of the CT-AGN will be missed (∼ 20 % detection fraction of all the CT-AGN), being low-luminosity sources (∼ 4260 of the 5200 CT-AGN per deg^2 have $\log(L_x/\text{erg s}^{-1}) < 43.5$). The characterization of these sources with *Athena WFI* (dotted area) will be challenging due to their low statistics: we may be able to recover L_x and N_H only for ∼ 6 % of the high-luminosity ones and < 1 % for the low-luminosity. Sufficiently powerful IR surveys will be capable of detecting the reprocessed emission due to accretion at IR wavelengths[¶¶]. For

[¶¶]The effective capabilities of generic IR surveys truthfully depend on a variety of factors, e.g. spectral and spatial resolution, sensitivities, detector and mirror noise, confusion limits. However, assuming a example, a *SPICA*-like DEEP survey is capable of detecting (red areas) ≈ 90 % of the CT-AGN and provide low-resolution spectra (shaded areas) for one third of them. The even more powerful *OST* will allow us to detect photometrically (R=4) almost all the CT-AGN and to have R=300 spectra for at least half of them. Spectroscopic follow-up with other IR instruments will be fundamental in identifying CT-AGN with photometric detection but without a clear detection with *SMI-LR*-like or *OSS* instruments. Low-luminosity (high-luminosity) CT-AGN can be spectroscopically identified with a *SAFARI*-like instrument up to $z \sim 1.5$ ($z \sim 2.5$) and $z \sim 2.5$ ($z > 4$) via the [Ne v]24.3μm and the [O iv]25.9μm lines with 5 hours observations (Spinoglio et al. 2021). Longer exposures (>10 hours) will allow to identify CT-AGN beyond $z \sim 4$. Despite the fact that the X-ray detection of these CT objects will be difficult, the *Athena* contribution to their study remains essential. In fact, only X-ray spectra (for individual sources or at least stacked X-ray data in case of no detection) can confirm the AGN nature of these sources (i.e., a flat X-ray continuum plus a strong iron line are clear indicators of heavy obscuration).

For the most luminous ones, the combined use of *Athena* X-ray spectroscopy, mid-, and far-IR spectrophotometry (with SMI-CAM- and B-BOP-like instruments or with the *OSS*) will provide never achieved before insights of the intrinsic AGN emission and bolometric power and opportunities to sample them at redshift higher than the cosmic noon. As shown in Figure 1.7 and discussed in section 1.4.3, at present, there are tensions between BHAD traced by X-rays, and those obtained by simulations, especially at $z > 2$. At $z \sim 4$, the BHAD expected from simulations (Aird et al. 2015; Vito et al. 2018) are ~ 5 times higher than those computed via X-ray surveys, while actual IR-surveys do not go beyond $z \sim 3$. Different arguments, mostly driven by observations in the IR (Gruppioni et al. 2016; Bisigello et al. 2020), predict a higher number of deeply obscured AGN, as for now undetected in the X-ray, that may alleviate this tension.

Exploiting all the capabilities of new generation IR cryogenic observatories (i.e. fast LR spectrophotometry to detect and characterize low-redshift CT-AGN, spectroscopic follow-ups to identify those with only photometric detection and multi-wavelength SED-fitting exploiting already observed deep fields to separate the AGN and host-galaxy emission) will allow us to discover the deeply obscured CT-AGN population up to very high redshift ($z > 4$). The presence of such an obscured population does not violate the limit imposed by the spectral energy density of the XRB, since these sources will provide an almost negligible contribution in the X-ray band (Comastri et al. 2015). The updated picture, in terms of obscured accretion over cosmic time, that the IR missions discussed in this chapter will lay out is fundamental in providing inputs to the simulations, which are still affected by considerable uncertainties and depend strongly on the adopted assump-

SPICA-like telescope, the driving factor will be the sensitivity and spectral resolution. Higher than SPICA spatial resolution, as well as lower noise, will mostly benefit the surveys by lowering the confusion limits, thus allowing us to perform deeper surveys. While constraining the emission region would help in distinguish AGN and SF emission only at low*-redshift, due to the fact that as we move to high-redshift the galaxies tend to become more compact.

tions (e.g., González et al. 2011; Jaacks et al. 2012; Sijacki et al. 2015; Shankar et al. 2013; Volonteri et al. 2016; Thomas et al. 2019).

2.6 Conclusions

In this chapter, we investigated the capabilities of IR deep and wide surveys to detect and characterize AGN, as well as the advantages resulting by the synergies with the future X-ray observatory *Athena*. We used XRB synthesis models to predict the total number of AGN as a function of intrinsic X-ray luminosity, amount of obscuration and redshift. A sample of more than 500 AGN from the COSMOS field with both X-ray spectra and optical-to-FIR SED-fitting has been used to investigate the AGN mid- and far-IR emission and to predict the fraction of AGN that will be detected by instrument similar to *SPICA SMI*, *B-BOP*, the *OST OSS*, and the *PRIMA PRIMAGER*. We compared the estimated total number of AGN with the fraction of those that will be detected by *Athena*, by a *SPICA*-like observatory, by the *OST*, and by the *PRIMAGER*, with the goal of enlightening the synergies between X-ray and IR surveys in terms of detection and characterization of AGN (in particular, of the most obscured ones) up to high redshift (i.e. $z \sim 10$) and down to low luminosity (i.e. $\log(\mathrm{L_x}/\mathrm{erg/s}) \sim 42$). In the IR bands, we also differentiated if we were primary detecting either the AGN or the host-galaxy emission. Moreover, we investigated if we were able to spectroscopically identify and characterize the AGN emission with *SMI-LR*-like instruments or exploiting the full R=300 spectral resolution of the *OSS*. In particular, we put emphasis on the elusive CT-AGN, a large fraction of those being missed by the current X-ray surveys, and that will require deep IR surveys for their detection and *Athena* spectra for their characterization. Our main results are the following:

1. An *OST* WIDE survey will be able to photometrically detect $\sim 60\,\%$ of all the AGN in all the *OSS* bands and more than $\sim 80\,\%$ in at least one of the bands, while the DEEP survey will detect at least 95 % of all the AGN at $z \leq 10$ in at least one of the *OSS* bands and more than $\sim 80\,\%$ in all of them.

2. Considering an *Athena* and *SPICA*-like DEEP survey, we estimate that we can have both IR and X-ray detections for half of all the AGN and for more than the 80 % of the high-luminosity ($\log(\mathrm{L_x}/\mathrm{erg/s}) \geq 43.5$) AGN; for these sources just the *Athena* detection will assure us of the presence of AGN activity, even when the mid-IR torus contribution is diluted into the host-galaxy emission. The *Athena WFI* spectral capabilities should also provide X-ray spectra, hence a reliable way to characterize the AGN, for thousands of AGN. This will allow us to estimate the AGN bolometric power and, having such a large number of sources with data in both the wavelength bands, to improve the AGN $\mathrm{L_{mid\text{-}IR}} - \mathrm{L_x}$ and $k_{bol} - \mathrm{L}_{bol}$ relations. Moreover, the use of low-resolution spectra (such as those produced by a SMI-LR-like spectrometer),

or alternatively, of mid- and far-IR photometric points (available with SMI-CAM- and B-BOP-like instruments) and SED-fitting decomposition technique, will help in constraining the AGN properties even in case of low-statistic X-ray spectra, like those we expect for very obscured and CT-AGN. However, even for sources with low photon statistics, the detection of a flat continuum and/or a strong iron emission line in the *Athena WFI* spectra will be invaluable clues to identify these sources as very obscured AGN.

3. The *PRIMAGER* instrument will have similar performance to *SPICA*, with the slightly lower sensitivity balanced by the larger FoV. In addition, the huge number of photometric filters will allow us to better recover the AGN and host-galaxy properties.

4. The synergies of *Athena* with the *OST* are similar to those with a *SPICA*-like mission, but with the advantages of having up to six photometric point in the far-IR, a higher fraction (with respect to a *SPICA*-like mission) of sources with mid-IR spectra and providing R= 300, $25-588\,\mu$m spectra for $\sim 50\,\%$ of all the AGN.

5. The CT-AGN, for which the X-ray detection and characterization are challenging due to their high amount of obscuration, will be very well sampled by deep IR surveys. A *SPICA*-like DEEP survey should detect 86 % of the low-luminosity ($\log(L_x/\mathrm{erg/s}) < 43.5$) CT-AGN at $z \leq 10$, while, *OST* can provide with six mid- and far-IR photometric detections for a similar fraction of the AGN, and at least one band detection for more than $\sim 95\,\%$ of all the CT AGN at $z < 10$. These kinds of missions would allow us to sample higher redshifts than what we are actually capable of at IR wavelengths, and detecting very obscured CT-AGN missed by X-ray surveys, to provide clues and possibly inputs to BHAD simulations. However, for most of these sources, the major contributor to the flux comes from the host galaxy emission, and the challenge will be to recognize them as AGN. In this context, pointed observations, such as those provided by *SPICA SAFARI*-like or *OSS* follow up, can be decisive in detecting the [Ne V]24.3μm and the [O IV]25.9μm lines, which would allow us to unambiguously classify the source as an AGN.

6. For high-luminosity CT-AGN, we will have X-ray detection with *Athena* up to $z \sim 3$. This will allow us to to ascertain the AGN nature of these obscured sources and will provide vital input for the selection of IR SED templates and for their spectroscopic characterization.

7. The use of mid- and far-IR photometric observations (such as those produced by SMI- and B-BOP-like instruments or by the *OST OSS* in R=4 photometric mode) can provide up to four (six in the case of the *OSS*, and an amazing 16 for *PRIMA*) photometric data points, which will be unique to improve the reliability of SED-fitting techniques to estimate the AGN properties up and beyond the cosmic noon.

8. Mid-IR spectra are very helpful to recognize AGN and to separate their emission from the SF-related one. A *SPICA*-like observatory can provide LR spectra (with the SMI-LR instrument) for the majority of the detected objects at $z \leq 2$, and up to $z \approx 4$ for the most luminous sources, as well as follow-up deep spectroscopy with SAFARI for galaxies up to and beyond $z \sim 4$. On the other hand, the *OST OSS* can deliver R= 300, $25 - 588\,\mu$m spectra for $\sim 50\,\%$ and $\sim 80\,\%$ of all the AGN with, respectively, its WIDE and DEEP surveys, the last one up to and beyond $z \approx 4$. The main spectral features to focus on will be the host-galaxy PAH lines and the AGN $9.7\,\mu$m silicate feature. We showed that, for what concerns the PAHs, those can be used to quickly recover the source spectroscopic redshift, without requiring specific higher-resolution follow-ups, while the silicate feature, when it is in absorption, jointly with X-ray spectral information, will allow us to assess and quantify the amount of obscuration. An *Athena* DEEP (WIDE) survey will provide X-ray spectral characterization for $\approx 20\,\%$ ($\approx 6\,\%$) of all the AGN and up to $\approx 50\,\%$ ($\approx 21\,\%$) of the high-luminosity ones, and, with the help of a multi-wavelengths approach, we will be able to characterize the properties even for X-ray low-statistic sources, such as CT-AGN.

Chapter 3

[NeV]-selected obscured AGN at $z \sim 1$

In this chapter, we investigated the properties of 94 [Ne v]3426Å-selected type 2 AGN in COSMOS at $z = 0.6 - 1.2$, performing optical-to-FIR SED-fitting of COSMOS2020 photometric data to estimate the properties of the AGN (bolometric luminosity, obscuration, Eddington ratio) and of the host galaxy (stellar mass, star formation rate, age, molecular gas mass). In addition, we performed X-ray spectral analysis of the 36 X-ray detected sources to obtain reliable values of the AGN obscuration and intrinsic luminosity, and to constrain the AGN properties of the X-ray undetected.

3.1 Introduction

As we discussed in chapter 1, highly accreting obscured AGN are the primary target to study the obscured accretion phase of the BH-galaxy co-evolution. However, the main difficulty lies in finding sources in this phase (see section 2.1).

Type 2 AGN can be also selected using narrow, high-ionization emission lines, such as the [O III]5007Å, the [Ne v]3426Å and the C IV1549Å narrow emission lines (section 1.3). These lines, being produced in the Narrow Lines Region (NLR), do not suffer from the nuclear extinction and their flux is a better proxy of the AGN intrinsic emission. Several works showed that pairing AGN optical lines selection with X-ray data is an effective method to find obscured and CT AGN (e.g., (e.g., Maiolino et al. 1998; Cappi et al. 2006; Vignali et al. 2006, 2010; Gilli et al. 2010; Mignoli et al. 2013, 2019, see also chapter 4)). In particular, the [Ne v] line, despite being $\sim$ 9 times fainter than the [O III] line, allows AGN selection up to $z \approx 1.5$ (using optical spectroscopy), whereas the [O III] is redshifted out of the optical range at $z \approx 0.8$. Moreover, due to its high-ionization potential of 97 eV (vs. 54 eV of [O III]), it is an unambiguous marker of AGN activity (e.g., Gilli et al. 2010; Mignoli et al. 2013; Cleri et al. 2022).

The use of different emission lines to select AGN at various redshifts (i.e., $z \lesssim 0.8$ for [O III], $0.6 \lesssim z \lesssim 1.5$ for [Ne v] and $1.5 \lesssim z \lesssim 3$ for C IV) allows the study of the redshift evolution of both the AGN and the host properties (Vignali et al. 2010; Gilli et al. 2010; Mignoli et al. 2013; Vignali et al. 2014; Mignoli et al. 2019).

3.1.1 The [Ne v] 3426Å selection

In addition to be a tracer of AGN activity, thus used to select AGN, the [Ne v] emission line is also a proxy of the AGN intrinsic luminosity. As such, the X/[Ne v] flux ratio was used by Vignali et al. (2014), hereafter V14, to trace the obscuration of a sample of [Ne v]-selected type 2 AGN in the C-COSMOS field. In fact, both the observed X-ray and the [Ne v] fluxes are linked to the intrinsic AGN emission, but the observed X-ray flux also suffers from the source obscuration. Gilli et al. (2010) calibrated a relation between X/[Ne v] and the N_H using a sample of 74 bright, nearby Seyferts with both X-ray and [Ne v] data and for which the column density was determined unambiguously. They found that the mean X/[Ne v] ratio for unobscured Seyferts is about 400, about 80% of local Seyferts with X/[Ne v]< 100 are obscured by column densities above 10^{23} cm^{-2} and essentially all objects with observed X/[Ne v]< 15 are CT (but see also Cleri et al. 2022).

In this chapter we present the properties of a sample of 94 [Ne v]-selected type 2 AGN, obtained via X-ray spectral analysis and from optical-to-FIR spectral energy distribution (SED) fitting. The use of the [Ne v] selection method restricted the sample to AGN in the $0.65 < z < 1.2$ redshift range, where most of the XRB "missing" sources should lie. This work is an extension of previous works (Mignoli et al. 2013, V14) for the AGN in the C-COSMOS field. We made use of newer X-ray data from the Chandra COSMOS Legacy catalogue (Civano et al. 2016; Marchesi et al. 2016b), that extended the X-ray coverage of the COSMOS field from 0.9 to 2.2 deg^2 and provided a more uniform coverage. We also studied the hosts of type 2 AGN to characterize their parameters, stellar mass and star formation rate, and investigated whether these galaxies are different from "normal" galaxies due to the AGN influence. The optical selection of the sample is presented in § 3.2, with the X-ray data in § 3.2.1 and the photometric data in § 3.2.2. The X-ray analysis results are reported in § 3.3, along with the fraction of CT objects. The SED-fitting algorithm and the results obtained with it are presented in § 3.4. In § 3.5 we present the comparison with a stellar mass- and redshift-matched control sample of non-active galaxies and the interpretation of our results in the light of the *in-situ* co-evolution scenario is in § 3.6. Conclusions are then reported in § 5.6.

3.2 The sample

We studied the [Ne v] type 2 AGN sample described in Mignoli et al. (2013). It was derived from the zCOSMOS-Bright spectroscopic survey (Lilly et al. 2007, 2009), which provided the 5500 – 9700 Å spectra of $\sim$ 20000 objects in the COSMOS (Scoville et al.

2007) field.
From the the zCOSMOS-Bright catalog, 94 type 2 AGN in the redshift range $\sim 0.65-1.20$ were selected on the basis of their spectral properties. The redshift range assured that both the [Ne v]3346Å and the [Ne v]3426Å emission line fall within the spectral coverage. Sources previously identified as type 1 AGN due to broad (>1000 km s^{-1}) emission lines or a blue underlying continuum in their spectra were excluded. The sample was composed of sources from redshift $z = 0.6606$ to $z = 1.1767$, with mean $z = 0.85 \pm 0.13$ and a median $z = 0.86$. The mean (aperture corrected) [Ne v] flux was $F_{[\mathrm{Ne\,v}]} = (1.81 \pm 1.23) \cdot 10^{-17}$ erg cm^{-2} s^{-1}, with median $F_{[\mathrm{Ne\,v}]} = 1.44 \cdot 10^{-17}$ erg cm^{-2} s^{-1}; the mean [Ne v] Equivalent Width is $\mathrm{EW}_{[\mathrm{Ne\,v}]} = 18.2 \pm 15.8$ Å with median $\mathrm{EW}_{[\mathrm{Ne\,v}]} = 13.9$ Å (see Mignoli et al. 2013 for further details).

3.2.1 X-ray Data

All 94 [Ne v]-selected sources fall in the *Chandra*-COSMOS Legacy mosaic, which is composed by data from the C-COSMOS survey (the central ~ 0.9 deg^2; Elvis et al. 2009) and from the COSMOS Legacy survey (covering the external ~ 1.7 deg^2 with a similar depth of the C-COSMOS survey; Civano et al. 2016). The whole mosaic covered ~ 2.2 deg^2 with a total exposure time of ~ 2.8 Ms. Thanks to the broader coverage of the COSMOS Legacy survey, we had twenty-three more sources with X-ray coverage with respect to V14; moreover, we included two sources which were previously excluded because they fell in bad positions of the C-COSMOS mosaic.
Thirty-six [Ne v]-selected type 2 AGN were detected by *Chandra* within 1.5" from the optical position. The majority of them having a X-ray-optical separation ≤ 0.5". The X-ray spectra were extracted as described in Marchesi et al. (2016b), using the *CIAO* (Fruscione et al. 2006) tool `specextract` from circular regions of radius r_{90} (i.e., the radius that contains 90% of the PSF in the 0.5 - 7 keV observed-frame band). The background spectra were extracted from annuli centered on the source position with inner radius $r_{90} + 2.5$" and outer radius of $r_{90} + 20$", paying attention to avoid the inclusion of X-ray sources. For each source, the spectra of its observations were combined in a single spectrum via the *CIAO* tool `combine_spectra`.

3.2.2 Photometric Data

The optical and IR data used to identify the [Ne v]-selected sources are taken from the COSMOS 2020 catalog (Weaver et al. 2022) (improved version of the previous COSMOS 2015 catalog Laigle et al. 2016), which contains photometry in 44 bands (from 1526Å to 8μm) for $\sim$ 1.7 million objects in the 2 deg^2 COSMOS field, along with matches with X-ray, near ultraviolet, and Far-IR data.
We used 3" aperture fluxes from 20 photometric bands and the COSMOS2020 matches with the 24μm band from the MIPS (Multi-Band Imaging Photometer) detector onboard

Spitzer, with the 100 μm and 160μm bands and the 250 μm, 350 μm and 500 μm bands from the PACS and SPIRE detectors of *Herschel*, and with the 850 μm from the SCUBA instrument at *JCMT* (we refer to Weaver et al. 2022 for a complete description of the COSMOS2020 catalogue and source associations). In total, we used data from 31 filters (see Table 3.1).

Two sources of the NeV sample were not in the latest COSMOS 2020 catalog (at the time of writing); for these sources we used the photometric information from the COSMOS 2015 catalog. As we analyzed all the sample also with the COSMOS 2015 catalog and we find no systematic between the values obtained using one or the other catalog, we are confident in reporting the results of these two sources along with those obtained with the newest COSMOS 2020 catalog.

3.3 X-ray spectral analysis

3.3.1 X-ray detected sources

Using the *XSPEC* software (Arnaud 1996), we performed the X-ray spectral analysis of 36 type 2 AGN with X-ray detection. The median number of net (i.e., background-subtracted) counts is 85, four sources have 15 counts or less. We divided the 36 sources into two sub-samples, on the basis on their net-counts. The high-counts sample is composed of 17 sources with at least 90 net-counts, the low-counts sample of 19 sources with less than 90 net-counts. For the low-counts sample, we used unbinned data and *C*-statistic (Cash 1979); for the high-counts sample, we rebinned the data at 25, 15 and 10 counts per bin (respectively for sources with net-counts > 500, > 200 and > 100) and used Gaussian statistic.

We first fit the sources with a power-law model, modified by the Galactic absorption (computed via the nh tool at the source positions, which derives it from the HI map by Kalberla et al. 2005) obtaining a mean photon index value of $\Gamma = 0.82$ with a standard deviation of 0.90. Six sources have $\Gamma \geq 1.6$, typical of unobscured AGN, sixteen objects have $\Gamma \leq 1.0$, seven of which with negative spectral index, usually found in very obscured sources.

Due to the low number of net-counts for the sources in our sample, we chose to adopt a simple phenomenological model to characterize the obscuration. Although we are aware that the real emission is likely more complicated, a model composed by a fixed-index power law ($\Gamma = 1.8$) plus an absorption component, allows us to characterize the AGN obscuration even for sources with net-counts ~ 10. The mean N_H of the high-count sample is $\approx 7.5 \cdot 10^{22}$ cm^{-2}, while for the low-count sample we obtained $\approx 32.2 \cdot 10^{22}$ cm^{-2}. On the basis of their obscuration, we classified nine sources as highly obscured ($N_H > 10^{23}$ cm^{-2}) and two as CT objects. Twelve sources have low values of obscuration ($N_H < 10^{22}$ cm^{-2}). We computed the intrinsic (i.e., absorption-corrected) 2 − 10 keV rest-frame luminosity

Table 3.1: Summary of COSMOS2020 photometric bands used in the [Ne v] sample analysis. The effective wavelength is the median wavelength weighted by transmission and the widths are defined as the difference between the maximum and the minimum wavelengths (calculated as the first and the last wavelengths with a transmission of at least 1%).

Instrument /Survey	Filter	Effective λ [Å]	Width [Å]
MegaCam/CFHT	$u*$	3823.3	670
Suprime-Cam /Subaru	IB427	4256.05	305.6
	B	4400.33	1399.1
	IB464	4633.48	330.5
	IB505	5060.57	378.1
	IB527	5261.1	242
	V	5477.8	955
	IB574	5764.8	271.5
	r	6136.24	1918
	IB679	6778.75	555.3
	IB709	7070.67	511.5
	IB738	7358.64	490.2
	i^+	7630.05	1872.5
	IB827	8241.69	514.3
	z^{++}	9020.18	1960.4
	Y	9759.16	1752.8
VIRCAM /VISTA (UltraVISTA-DR2)	Y^{UD}	10214.2	970
	J^{UD}	12534.6	1720
	H^{UD}	16453.4	2900
	K_S^{UD}	21539.9	3090
IRAC/*Spitzer* (SPLASH)	ch1	35634.3	7460
	ch2	45110.1	10110
	ch3	57593.4	14140
	ch4	79594.9	28760
MIPS/*Spitzer*	24μm	232096	110494
PACS/*Herschel*	green	979036	558974
	red	1.54×10^6	1.26×10^6
SPIRE/*Herschel*	PSW	2.43×10^6	1.26×10^6
	PMW	3.41×10^6	1.46×10^6
	PLW	4.82×10^6	2.92×10^6
SCUBA/*JCMT*	2.450 GHz	4.48×10^6	1.04×10^6

of the sample, obtaining a mean value of $8.3 \cdot 10^{43}$ erg/s for the high-counts sample and $2.7 \cdot 10^{43}$ erg/s for the low counts sample. The mean intrinsic $2-10$ keV rest-frame luminosity of the whole sample is $5.4 \cdot 10^{43}$ erg/s, with a standard deviation of $5.7 \cdot 10^{43}$ erg/s.

As the sources have low numbers of counts we preferred using simple phenomenological models. However, few sources display more complex spectra that called for deeper investigations. Source *lid*1840 (COSMOS Legacy ID) with 238 net-counts and $N_H < 0.5 \cdot 10^{22}\,\mathrm{cm}^{-2}$ shows the possible presence of a 6.9 keV line with a significance of $\approx 2.6\,\sigma$ and an equivalent width of $EW_{6.9} = 0.51 \pm 0.38$ keV. Source *cid*1508 (C-COSMOS ID) has a spectrum which shows the presence of a possible soft excess with respect to the absorbed power law model. We substituted the absorption component with a partial covering fraction absorption (`zpcfabs`) and found a lower limit on the covering fraction $f > 0.77$, while the value of N_H being compatible with the one obtained using the simple absorption component. For source *cid*138 (99 net-counts) we found a variation of the net-count rate of more than a factor 6 (with a significance of 6.7σ) between the 2007 and 2014 observations. We did not find any significant variation in the hardness ratio ($HR = (H-S)/(H+S)$, where H and S are the net-counts in the $2-7$ and $0.5-2$ keV energy range), spectral index or N_H. As the flux variation is not accompanied by a variation of the spectral properties, we hypothesized it may be linked to differences in the AGN accretion rate.

We computed the 2-10 keV rest-frame flux (not corrected for obscuration) of the sample, obtaining a median flux of $F_{2-10} = 0.61 \cdot 10^{-14}$ erg/s/cm^2. We compared the fluxes with those from V14, which were obtained via an *Xspec* spectral fitting with a powerlaw model with $\Gamma = 1.4$. Except for *cid*138, the values are in agreement within their errors. We used the 2-10 keV rest-frame flux and the [Ne v] flux from Mignoli et al. (2013) to compute the X/[Ne v] ratio. The mean X/[Ne v] is 313 with a standard deviation of 321. Ten sources had X/[Ne v] < 100. No source has X/[Ne v] lower than 15. In Fig. 3.1, we compared our data with the X/[Ne v] vs N_H diagram obtained by Gilli et al. (2010). The plot was produced using the spectral templates of Gilli et al. (2007). These are AGN X-ray spectral models with a primary power-law with $\Gamma = 1.9$, cut-off energy $E_C = 200$ keV, a variety of absorptions ($\log N_H$) $= 21.5, 22.5, 23.5, 24.5, > 25$, a 6.4 keV emission line and, in case of obscured spectra, a 3% of soft scattered component (see section 1.3). The blue solid line was obtained using the mean X/[Ne v] ratio of a sample of 74 unobscured Seyfert galaxy in the local Universe and, starting from it, computing the expected X/[Ne v] ratio at increasing levels of absorption, using the spectral templates. The same computation was carried-out starting from the mean X/[Ne v] ratio $\pm 1\sigma$ and $\pm 90\%$, to produce the 1σ and 90% limits. The procedure is extensively described in Gilli et al. (2010). As we can see from Fig. 3.1, the X-ray detected sources of [Ne v] sample populate the obscured AGN region of the diagram, i.e., $10^{22} < N_H < 10^{24}$ cm^{-2}, with a few sources in the unobscured region ($N_H < 10^{22}$ cm^{-2}); the majority of the sources lie within

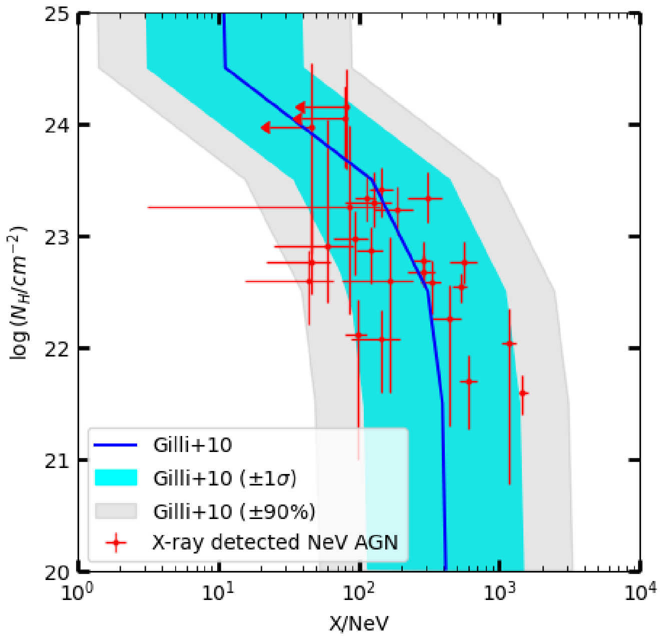

Figure 3.1: X/[Ne v] vs N_H diagram. Red dots are the X-ray detected sources of the current [Ne v] sample, with N_H obtained from the X-ray spectral analysis. The solid line shows the expected X/[Ne v] values as a function of absorption, as computed by Gilli et al. (2010) (but see also Cleri et al. 2022) using spectral templates with different N_H, starting from the mean X/[Ne v] obtained from a sample of unobscured Seyfert galaxies. The cyan shaded region was computed in the same way, but starting from the mean X/[Ne v] $\pm 1\sigma$ and the grey shaded region starting from the mean X/[Ne v] $\pm 90\%$. The [Ne v] sample populates the obscured quasar region of the diagram, i.e., $N_H \gtrsim 10^{22}$ cm^{-2}.

the 1 σ limit.

We compared the results of our spectral analysis with those presented by Marchesi et al. (2016b), hereafter M16, and Lanzuisi et al. (2018, who studied CT-candidate from M16). M16 performed an X-ray spectral analysis of the $\sim$ 1850 extragalactic sources in the *Chandra* COSMOS-Legacy survey with more than 30 net-counts in the 0.5 – 7 keV band. Spectra were fitted with a fixed photon-index $\Gamma = 1.9$ power-law, an absorption component, an optional un-absorbed second power-law to model the scattered emission, and a 6.4 keV (rest frame) Gaussian feature to reproduce the iron Kα emission line. Out of our 36 X-ray detected sources in our sample, 6 of them are not in the M16 work, as they are below the 20 net-counts threshold. For the reaming 30 sources, we found an excellent agreement in $L^{intr}_{2-10keV}$ and N_H: the median (16^{th} and 84^{th} percentile) ratio between M16 and ours $L^{intr}_{2-10keV}$ and N_H are $1.000^{+0.004}_{-0.003}$ and $1.02^{+0.10}_{-0.02}$, respectively. Regarding the two CT X-ray detected sources, *cid*1019 is classified as CT in Lanzuisi et al. (2018) as well, while *cid*1706 has 15 net-counts and was not included in M16. The other 5 sources not in M16, are all obscured AGN, two of them with $N_H > 10^{23}$ cm^{-2}.

3.3.2 X-ray undetected sources

Out of the 94 sources of the [Ne v] sample, 58 have no X-ray detection. We run a two-samples Kolmogorov–Smirnov (KS) test in order to investigate if X-ray undetected sources are such because they fall in regions with shorter exposure-map derived time* than those of the X-ray detected sources. We built two empirical distribution functions of the exposure time for the X-ray detected sources and for the X-ray undetected, using the exposure maps derived from the entire COSMOS Legacy mosaic. We found that, with a confidence of 89% (KS statistic of $D = 0.12$), the undetected sources are not associated with lower exposure times.

We used the *CIAO* tool `srcflux` to calculate the net count-rate limit for each undetected source, and the tool `modelflux` to calculate the flux upper limit. Count-rates were computed using all the observations for which the source fall in the field of view. Using `srcflux`, we calculated the $0.5 - 7$ keV (observed-frame) net count-rate in a circular region centered on the source position, that contained 90% of the PSF at 1 keV. Background counts were extracted in an annular region centered on the source position, using as inner and outer radii of 1 and 5 times the radius of the source region. `srcflux` computed the net count-rate by dividing the net-counts by the effective (i.e., vignetted-corrected) exposure time at the source position. Because these sources were not detected in the X-ray band, these count-rate values can be considered as upper limits with a 90% confidence interval. To correctly compute the net counts in case of low photon statistics and upper limits when the source is not detected, we adopted `srcflux`, based on the tool `aprates`†, which finds credible intervals by computing the Bayesian background-marginalized posterior probability distribution function (PDF), assuming non-informative priors for the intensities in the source and background apertures. The value at the peak of this PDF is determined, and upper bounds of the credible interval are determined by summing values of the PDF above the peak until the desired confidence level is attained. We refer to Primini & Kashyap (2014) for more details about this statistic technique. We used `modelflux` to calculate, from the net count-rates, the upper limit for the flux in the $2 - 10$ keV rest-frame energy range. We used a power-law model with spectral index $\Gamma = 0.4$ (the average spectral index of the low count sample), modified by Galactic absorption. This model takes into account the source obscuration via a flatter photon index than the intrinsic one $\Gamma = 1.8 - 1.9$.

We used the rest-frame $2 - 10$ keV flux upper limits and the [Ne v] fluxes to compute the X/[Ne v] ratio upper limits. 54 sources have an upper limit < 100, and 16 sources have X/[Ne v] ratios < 15. This means that 93% of the X-ray undetected sources are candidate to be AGN with $N_H > 10^{23}$ cm^{-2}, and 28% to be CT AGN. We used the X/[Ne v] upper limits to compute lower limits on the N_H. Using the net-count rates, the N_H lower limits,

*In this section, with "time" we refer to *vignetting*-corrected time, i.e. taking into account the lower effective exposure off-axis.

†`https://cxc.cfa.harvard.edu/ciao/ahelp/aprates.html`

and a fixed $\Gamma = 1.8$ spectral index, we estimate the intrinsic (i.e., absorption corrected) X-ray flux (and luminosity) upper limit via `modelflux`.

Testing this procedure on the X-ray detected sources, we found a strong correlation between the intrinsic luminosities obtained in this way and those from the X-ray spectral analysis for the high-counts sample. The correlation is loose for the low-count sample, but for the majority of these sources, the two luminosities do not differ more than $2\,\sigma$. We did not find any particular trend between the accuracy of our method and the amount of obscuration, thus we concluded that it is a reliable way to estimate an upper limit on the X-ray intrinsic luminosity and that the major source of error is linked to the uncertainties of estimating the net-counts upper limit in case of low statistic.

Considering the whole (both X-ray detected and undetected sources) [Ne v] sample, at least 67% of the sources have X/[Ne v] ratios compatible with absorption $N_H > 10^{23}$ cm^{-2}, and at least 19% of the sources are likely CT AGN. The X/[Ne v] flux ratios are plotted in Fig. 3.2, the vertical dashed line shows the threshold, defined by Gilli et al. (2010), between Compton-thick (leftward direction) and Compton-thin (right-ward direction) sources.

It is worth mentioning that, for some of the undetected sources with low X/[Ne v] ratio, their low X-ray flux could be a consequence of the AGN duty cycle rather than an indication of obscuration. In fact, Saade et al. (2022), exploiting *NuSTAR*, *Chandra*, *XMM-Newton*, and *Swift* observations, found that in their sample of nine low X/[O III] flux ratio AGN, one of them was not obscured, and was more likely a recently de-activated AGN. For this source, the interpretation is that the X-ray emission from the core has already faded, but the [O III] flux from the NLR has a delay in the order of tens or hundreds of years (Ichikawa & Tazaki 2017). As the [Ne v] and the [O III] come both from the NLR, it is possible that the same could apply to some of our undetected sources. However, a study similar to the one of Saade et al. (2022) is challenging and requires hard X-ray data (such as those provided by *NuSTAR* that, being not significantly effected by the obscuration, trace the AGN intrinsic luminosity) that we do not have at our disposal, therefore this investigation is beyond the scope of this chapter.

3.4 SED analysis

We used the SED fitting technique to derive AGN and galaxy properties from photometric data. In Sect. 3.4.1 we briefly present the SED-fitting algorithm as well as the AGN torus models we used. The results of the SED-fitting, along with the comparison with those obtained from the X-ray analysis, are presented in Sect. 3.4.2.

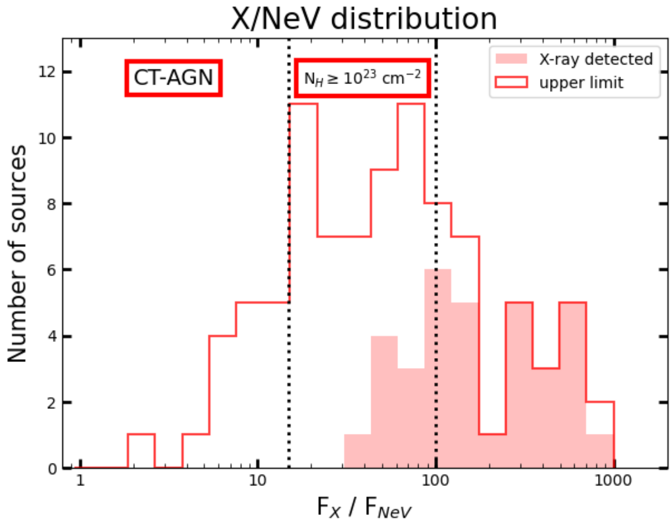

Figure 3.2: X/[Ne v] ratio distribution. Filled and empty histograms refer to X-ray detections and upper limits, respectively. The leftward area (X/[Ne v] $\leq$ 15) is the region defined by Gilli et al. (2010) for Compton-thick AGN, while the central area (15 < X/[Ne v] $\leq$ 100) is where very obscured ($N_H > 10^{23}$ cm^{-2}) and possible CT-AGN should be located.

3.4.1 SED-fitting algorithm

We made use of the SED-fitting algorithm *SED3FIT* (Berta et al. 2013), based on the *MAGPHYS* (da Cunha et al. 2008) code, that performs SED-fitting with a combination of three components: stellar emission, dust emission from star formation and a possible dusty torus/AGN component. The stellar and dust emission are linked by energy balance arguments. Torus emission is independently included.

MAGPHYS (Multi-wavelength Analysis of Galaxy Physical Properties) is a model package to interpret observed SEDs of galaxies (at rest wavelengths in the range 912Å < λ < 1mm) in terms of galaxy-wide physical parameters pertaining to the stars and the interstellar medium, following the approach described in da Cunha et al. (da Cunha, Charlot, & Elbaz (2008)). The analysis of the SED of an observed galaxy with MAGPHYS is carried out in two steps:

1. The creation of a library of model spectral energy distributions at the same redshift and in the same photometric bands as the observed galaxy, for wide ranges of plausible physical parameters.

2. The build-up of the marginalized likelihood distribution of each physical parameter of the observed galaxy, through the comparison of the observed spectral energy distribution with all the models in the library.

The code uses two libraries of models: one that takes into account the stellar emission and the effects of dust attenuation (we will refer to these models as "optical models"), the other that include the IR emission of the dust (we will refer to them as "IR models"). The optical and infrared libraries are linked together to provide the full SED of model galaxies from the far ultraviolet to the far-infrared wavelengths. The optical models store 50 000 stellar population spectra, with both the dust-free spectrum and the dust-attenuated spectrum for each galaxy template. These spectra were generated using the Bruzual (Bruzual (2007)) stellar population synthesis code, assuming a Chabrier (2003) IMF.

The SED at time t of a stellar population characterized by a star formation rate $\psi(t)$ is given by:

$$L_\lambda(t) = \int_0^t dt' \, \psi(t-t') \, S_\lambda(t', Z) e^{-\tau_\lambda(t')} \tag{3.1}$$

where $S_\lambda(t', Z)$ is the power radiated per unit wavelength and per unit initial mass by a simple stellar population (SSP) of age t' and metallicity Z, and $\tau_\lambda(t')$ is the 'effective' absorption optical depth of the dust seen by stars of age t'.

The main adjustable parameters of these models are:

- **Star formation history:** the star formation rate as a function of time $\psi(t)$. It is build as a continuous star-formation (characterized by an age t_{form} and a star formation timescale parameter γ, with $\psi(t) \propto e^{-\gamma t}$) and random bursts superimposed to this continuous model.

- **Metallicity:** uniformly distributed between 0.02 and 2 times solar metallicity.

- **Dust attenuation:** computed using the simple, angle-averaged model of Charlot & Fall (Charlot & Fall (2000)). This accounts for the fact that stars are born in dense molecular clouds, which dissipate typically on a timescale of 10^7 yr.

The IR models store 50 000 dust emission spectra. The mid- and far-infrared emission from dust in galaxies is computed using the model of da Cunha et al. (da Cunha, Charlot, & Elbaz (2008)). The total dust emission from a galaxy is the sum of the dust emission originating from the stellar birth clouds and the dust emission originating from the ISM (Inter Stellar Medium).

- **Birth clouds:** The SED of the power re-radiated by dust in the stellar birth clouds is computed as the sum of three components: a component of polycylic aromatic hydrocarbons (PAHs); a mid-infrared continuum characterizing the emission from hot grains at temperatures in the range 130-250 K; and a component of grains in thermal equilibrium with adjustable temperature in the range 30-60 K.

- **Ambient ISM:** In the ambient ISM, the relative proportions of the three components are fixed, for simplicity, to reproduce the spectral shape of diffuse cirrus emission in the Milky Way, and a fourth component of cold grains in thermal equilibrium with adjustable temperature in the range 15–25 K is included.

MAGPHYS code provides a consistent interpretation of ultraviolet, optical and infrared SEDs of galaxies. This is achieved by accounting consistently for the total energy absorbed by dust in stellar birth clouds and in the ambient ISM, and for the re-distribution of this energy at far-infrared wavelengths. The main underlying assumptions are that the energy re-radiated by dust is equal to that absorbed (i.e. the energy is conserved), and that starlight is the only significant source of dust heating in the galaxies under study.

Different combinations of star formation histories, metallicities and dust content can lead to similar amounts of energy absorbed by dust in the stellar birth clouds, and these energies can be distributed in wavelength using different combinations of dust parameters. Consequently, in the process of fitting, a wide range of optical models is associated with a wide range of infrared spectra and compared to observed photometry, seeking for χ^2 minimization.

One of the main assumptions of the *MAGPHYS* code is that the only significant source of dust heating is the starlight, thus ignoring any possible contribution of the AGN to the SED. The *SED3FIT* code solves this limitation by adding a warm dust component to the modeled SED emission. It represents dust surrounding the active nucleus, assumed to be distributed in a toroidal region. The code uses χ^2 minimization to find the best-fit model. Allowing the normalization of stars+dust to be free, i.e., not strictly anchored to the observed photometry but simply randomly picked from a grid of values, the torus is effectively fit to the data in a simultaneous 3-component mode.

The torus library we used to model the AGN contribution to the SED assumes that the AGN dust and gas are distributed in a toroidal shape, i.e., "smooth-torus" model. It was developed by Fritz et al. (2006) and updated by Feltre et al. (2012). The geometry of the torus is a *flared disc*. Its size is defined by the outer radius $R_{\rm max}$ - the inner radius being defined by the sublimation temperature of dust grains under the influence of the strong nuclear radiation field - and by the angular opening angle Θ of the torus itself. The main dust components are silicate and graphite grains, in almost equal percentages. The torus density law adopted is:

$$\rho(r,\theta) = \alpha \cdot r^{\beta} \cdot e^{-\gamma|\cos\theta|} \tag{3.2}$$

where α is a normalization constant and the parameters β and γ allow to create density gradients both in radial (r) and in polar (θ) directions. The models assume that the torus is illuminated by a central point-like energy source with isotropic emission. Its spectrum is described as a composition of power-laws with variable indices (see Feltre et al. 2012). The radiation emitted is given by the sum of the primary source located in the torus center and a secondary contribution given by thermal and scattering dust emission.

To reduce the calculation time, we selected only a sub-sample of the 24 000 elements torus library. We choose:

- $\mathbf{\Phi = 1^\circ, 21^\circ, 41^\circ, 61^\circ, 89^\circ}$: to be able to model different inclination angles between the line of sight and the torus equatorial plane (i.e., to model both type 1 and type 2 objects).

- **R = 30** : this value limits the models to compact tori of a few parsec (given that $R_{\min}$ is directly connected to the sublimation temperature and to the accretion luminosity of the central BH), as done in Pozzi et al. (2010). In fact, high-resolution IR and recent ALMA observations support a compact dust distribution in nearby luminous AGN (i.e., (Jaffe et al. 2004; Elitzur 2008; Combes et al. 2019)).
- **ct** $= \mathbf{20^\circ, 40^\circ, 60^\circ}$: all the possible values of the half-width of the torus apertures.
- $\beta = \mathbf{0, -1}$: the first is linked to an homogeneous density distribution, the second to a density decreasing exponentially with the distance from the nucleus.
- $\gamma = \mathbf{0}$: we considered only torus with an homogeneous distribution of density in polar direction.
- $\tau_{\rm eq} = \mathbf{0.1, 0.3, 0.6, 1, 3, 6}$: as suggested by Feltre et al. (2012), we avoided extreme optical depths.

Reducing the torus parameter space allowed us to reduce the calculation time to an acceptable level while maintaining 180 different torus models. We instructed the SED fitting algorithm to run 100 normalizations for each torus model, for a total of 18 000 torus SED available.

3.4.2 SED-fitting results

Fig. 5.10 shows the SED-fitting results for one of the [Ne v] AGN, source zCOSMOS 380027, undetected in the X-ray, at $z = 0.9307$.

IR8 relation

To confirm the goodness of the galaxy SED-fitting, we investigated if the best-fit galaxy SED is able to reproduced the "IR8" relation (Elbaz et al. (2011)). This is a scaling relation between the $8 - 1000\,\mu$m galaxy luminosity and the $8\,\mu$m galaxy luminosity, which is defined as $L_8 = \nu L_\nu\,(8\,\mu m)$. The IR8 ratio is

$$\mathrm{IR8} = \frac{L_{\rm IR}^{\rm gal}}{L_8^{\rm gal}} \tag{3.3}$$

and for star forming galaxies its mean value is IR8 = 4.9 [−2.2, +2.9], with $1\,\sigma$ confidence (Elbaz et al. (2011)). This correlation is based on the SF origin of both the $8\,\mu$m and the $8 - 1000\,\mu$m luminosities. In fact, the first is linked to PAH emission mostly excited by UV radiations, the second to the UV and the more numerous optical photons that heat

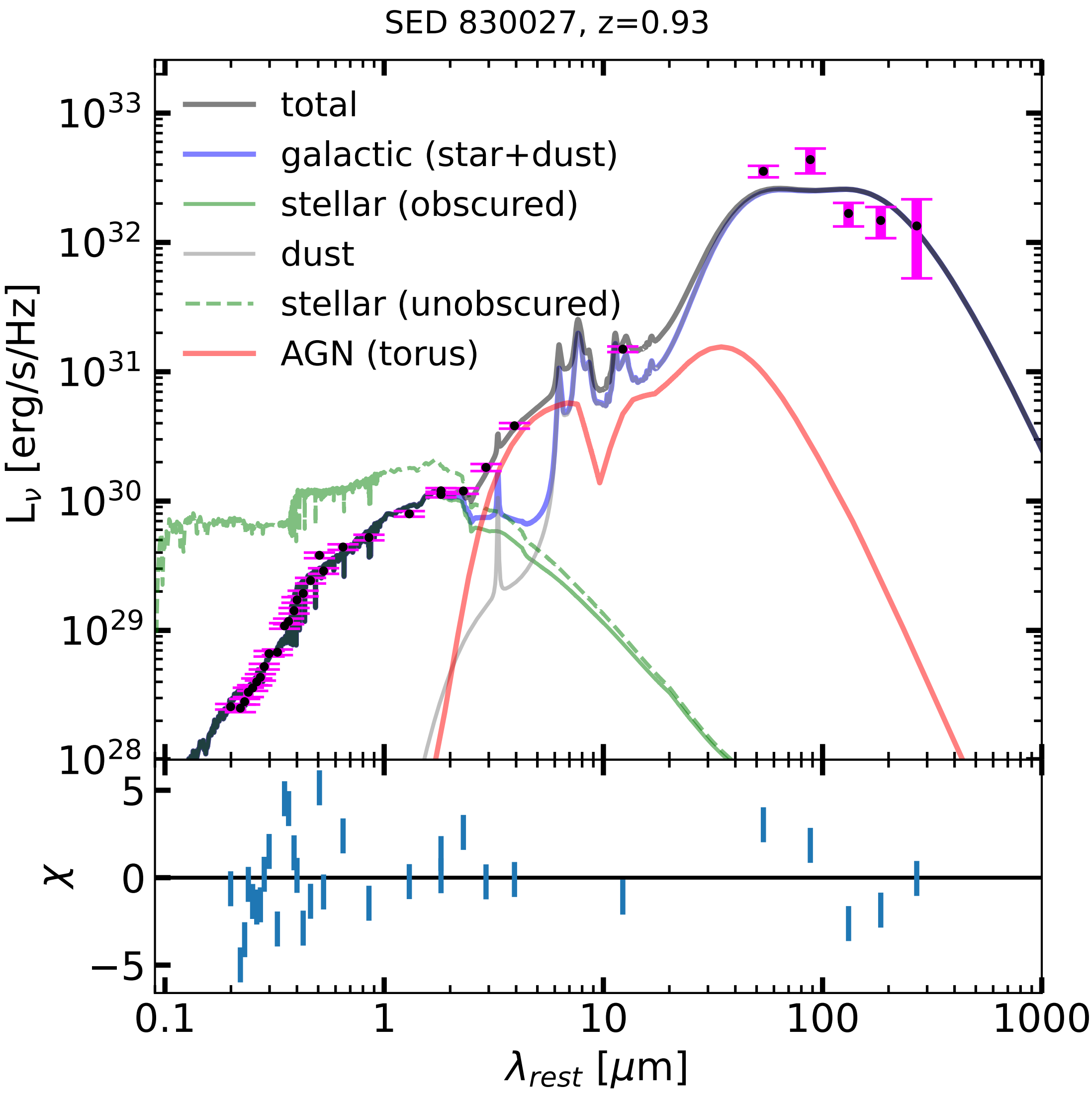

Figure 3.3: Example SED fit for a [Ne v]-selected AGN of the current sample. The black points with purple error bars are the photometric data, the dark gray line is the best-fit SED, composed of the galactic emission (blue line) and the AGN component (red line). The light grey line is the host galaxy dust emission, the continuous (dashed) green line refers to the dust-obscured (intrinsic) stellar emission. The fitting provides a stellar mass of $\log(M_*/\mathrm{M}_\odot) = 10.64 \pm 0.05$, a SFR of $51 \pm 10\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$, and an AGN bolometric luminosity of $\log(L_{\mathrm{bol}}^{\mathrm{sed}}/\mathrm{erg\,s}^{-1}) = 45.6^{+0.3}_{-0.1}$. The bottom panel shows the residuals, where $\chi = (\,\mathrm{observation} - \mathrm{model}\,)/\mathrm{error}$.

up the dust, that consequently emits via grey-body in the far-IR bands. Both can be used as SFR tracers.

In figure 3.4 we show that the majority of our sample lies within the 1 σ limit. Moreover, the most distant sources are those with few FIR datapoints.

These results confirm the meaningfulness of the galaxy SED-fitting and underline the importance of FIR detections in constraining the IR SED and thus in separating the AGN emission from that of the galaxy.

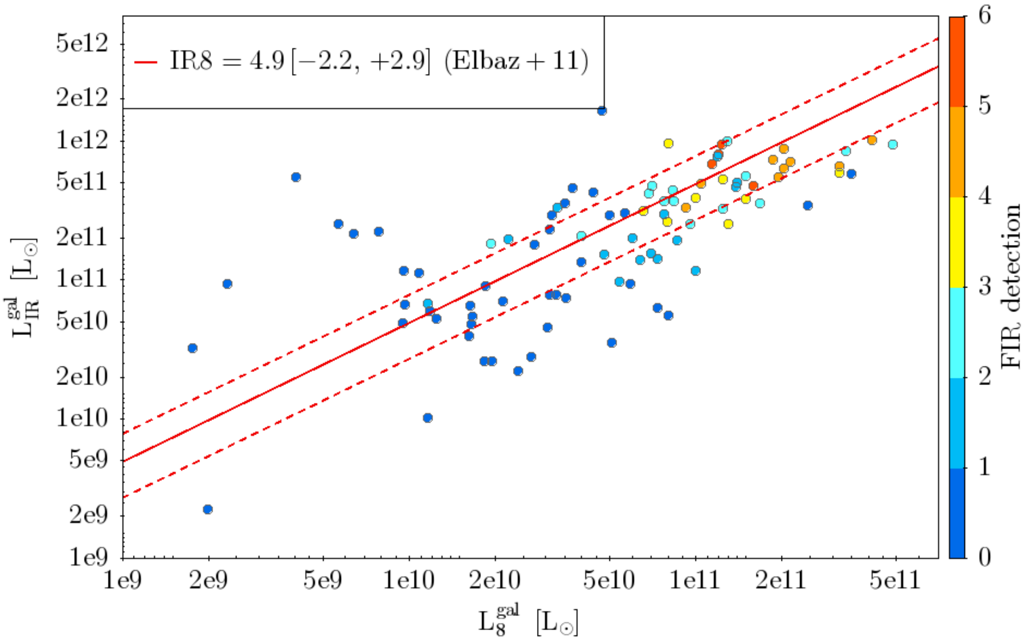

Figure 3.4: Comparison between the galaxy 8 – 1000 μm luminosity and the L_8 luminosity. L_8 is the galaxy 8 μm luminosity ($L_8 = \nu L_\nu (8\,\mu m)$). The color code indicates the number of photometric points available for each source in the FIR band. The solid red line is the mean IR8 (IR8= L_{IR}^{gal}/L_8^{gal}) for star-forming galaxy (Elbaz et al. 2011); the dashed red lines are its 1 σ limit.

AGN bolometric luminosity

Unlike type 1 objects, for obscured AGN it is not possible to obtain the AGN bolometric luminosity from the optical emission. However, as the SED-fitting algorithm allows us to disentangle the AGN and the host contributions, one of the direct outputs of *SED3FIT* is the AGN bolometric luminosity, derived via the integration of the spectrum of the central source that illuminates the torus.

We found that for 18 sources the SED fitting shows a low contribution of the AGN to the total emission, hence for these sources we have only upper limits on their AGN bolometric luminosity. The remaining 76 sources have a median $\log(L_{bol}^{sed}/\mathrm{erg\,s^{-1}}) = 44.4 \pm 0.7$.

To further support these values, we compared them with the bolometric luminosities

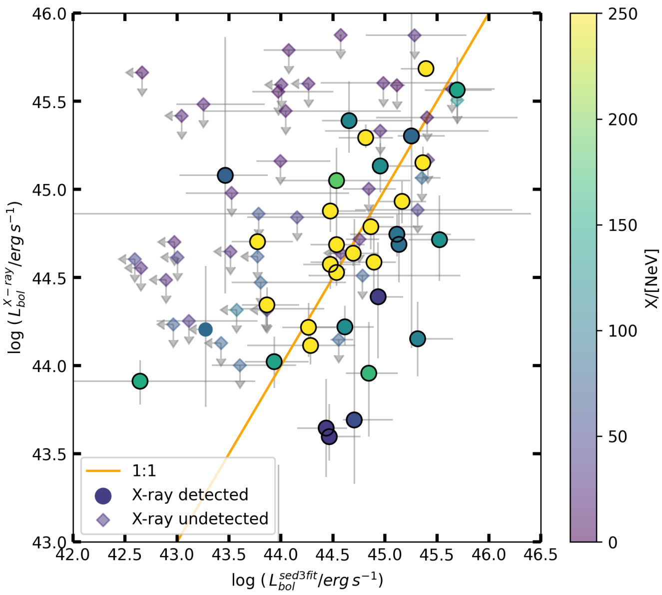

Figure 3.5: Comparison of the AGN bolometric luminosities obtained from the SED-fitting (L^{sed}_{bol}) and from X-ray spectral analysis (L^{x}_{bol}). The L^{x}_{bol} were computed from the 2 – 10 keV rest-frame intrinsic luminosities using the Lusso et al. (2012) bolometric correction. Circles indicate X-ray detected sources, while diamonds are the X-ray undetected. For these sources we have only an upper limit on their L^{x}_{bol}. The color code indicates the X/[Ne v] ratio for the X-ray detected sources and upper limits for the X-ray undetected sources. The orange line is the 1:1 correlation.

obtained from the X-ray spectral analysis. We will refer to the bolometric luminosities computed from the SED-fitting as L^{sed}_{bol} and to those obtained from X-ray analysis as L^{x}_{bol}. For the sources with X-ray detection we were able to compute the intrinsic 2–10 keV rest-frame luminosity, as reported in § 3.3.1. Using the bolometric correction K_{bol} from Lusso et al. (2012) (see also Duras et al. 2020), we obtained the AGN bolometric luminosities L^{x}_{bol} of these sources, with a median value of $\log(L^{x}_{bol}/\mathrm{erg s^{-1}}) = 44.7 \pm 0.5$. Considering the X-ray detected sources, the median $L^{sed}_{bol}/L^{x}_{bol}$ ratio is 1.01 ± 0.02. In Fig. 3.5 we show the comparison of the bolometric luminosities: except for two sources, the two bolometric luminosities are compatible within twice their uncertainties. This correlation is extremely important, as it allows an estimate of the AGN bolometric luminosities using only optical-to-FIR photometric data. Moreover, it confirms the reliability of both the SED-fitting procedure and the X-ray analysis.

Estimate of the 2 – 10 keV intrinsic luminosities

We obtained an independent estimate of the AGN 2 – 10 keV rest-frame intrinsic luminosity from the AGN 12 μm luminosity using the Gandhi et al. (2009) relation:

$$\log\left(\frac{L^{sed}_{2-10keV}}{10^{43}\mathrm{erg/s}}\right) = \frac{1}{0.97}\left(\log\left(\frac{L^{sed}_{12\mu m}}{10^{43}\mathrm{erg/s}}\right) - 0.33\right) \tag{3.4}$$

This relation was found via a sample of 42 Seyfert, including both type 1 and type 2 AGN, observed with the *VLT*/VISIR with sufficient angular resolution to isolate the AGN 12 μm emission from that of the galaxy. This correlation is due to the fact that the AGN intrinsic emission is absorbed and re-emitted by the obscuring torus in the mid-IR bands.

We compared these luminosities with those from the X-ray spectral analysis. For the X-ray detected sources we obtain a mean of $L^{x}_{2-10keV} = (5.4 \pm 1.0) \cdot 10^{43}$ erg/s.

For the X-ray undetected sources we did not have their intrinsic luminosity from X-ray spectral analysis. However, we used the rest-frame 2 − 10 keV flux upper limits. To compute those we assumed a power-law model with $\Gamma = 0.4$ and derived upper limits on the 2 − 10 keV intrinsic luminosity, as showed in section 3.3.2. In figure 3.6 we show the comparison between the AGN rest-frame 2 − 100 keV intrinsic luminosities obtained from the AGN 12 μm luminosities with those from the X-ray spectral analysis (including also the intrinsic luminosity upper limits, computed as stated above). For the data derived

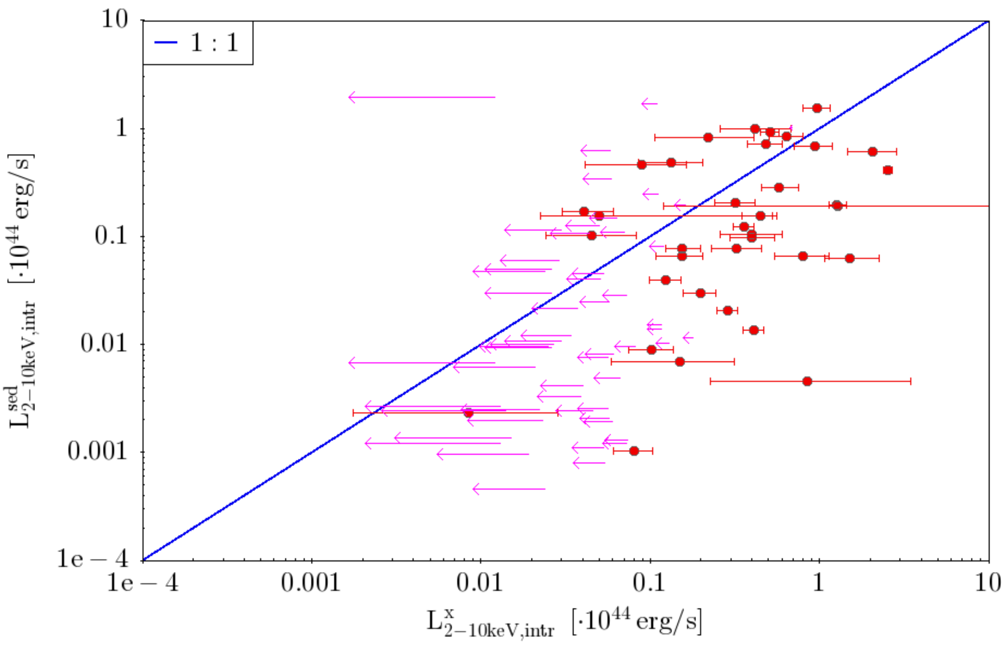

Figure 3.6: Comparison of the rest-frame 2 − 10 keV intrinsic luminosities between those obtained from the AGN 12 μm luminosity ($L^{sed}_{2-10keV}$) and those from the X-ray spectral analysis ($L^{x}_{2-10keV}$). The $L^{x}_{2-10keV}$ upper limits are those of the X-ray undetected sources and were obtained from the count-rate upper limits assuming a power-law model with $\Gamma = 0.4$.

from X-ray analysis there is a clear separation between the detections and the upper limits, with the former populating an area with $L^{x}_{2-10keV} > 10^{43}$ erg/s and the latter restricted to $L^{x}_{2-10keV} < 10^{43}$ erg/s. A similar separation is not present in the intrinsic luminosities derived from the SED-fitting.

The mean ratio between the $L^{sed}_{2-10keV}$ and the $L^{x}_{2-10keV}$, without considering the upper limits, is $\approx 1.0 \pm 0.2$ and its median is ≈ 0.3. Similar to the correlation between the bolometric luminosities, this one allows to estimate the AGN power, using only optical-to-FIR photometric data. The use of both the correlations can be an extremely valuable aid in computing AGN luminosities when X-ray observations are not available.

AGN significance

The SED-fitting procedure allowed us to separate the contribution of the AGN from that of the galaxy. However, this process is subject to a certain intrinsic degeneracy: an over-estimation of the AGN fraction will result in an under-estimation of the IR emission from the galaxy and thus of the SFR and vice versa. To further assess the reliability of the chosen torus models, as well as to estimate the importance of the AGN component on the total emission on a *per-source* basis, we estimated the AGN significance using an F-test between the best-fit χ^2 with and without AGN component, similarly to what was done in Delvecchio et al. (2014). We carried out a second run of *sed3fit*, using the same optical and IR models, without any torus model. We compared the obtained $\chi^2_{\mathrm{NO-AGN}}$ with the χ^2_{AGN} obtained using the torus models. The F-value was computed as

$$F_{\mathrm{test}} = \frac{\chi^2_{\mathrm{NO-AGN}} - \chi^2_{\mathrm{AGN}}}{\bar{\chi}^2_{\mathrm{AGN}}} \tag{3.5}$$

where $\bar{\chi}^2 = \chi^2/dof$ and dof is the number of degrees of freedom. As dof we used the number of photometric points, in case of the model without AGN, and the number of photometric points minus one, to take into account the additional parameter (the torus), for the model with AGN.

We obtained seventy-four sources (79%) with an AGN significance $\geq 1\,\sigma$, forty-eight sources (51%) with $\geq 2\,\sigma$, thirty-three (35%) with $\geq 3\,\sigma$ and twenty-four (26%) with $\geq 4\,\sigma$. Considering only the best value of the chi squared, thirteen sources (14%) were better fitted with a model without AGN component; seven of them were sources for which we had only upper limits on their AGN IR luminosity, due to the low contribution of the AGN to the total SED. The fact that nearly half of the sample has an AGN significance $< 2\,\sigma$ can be attributed to intrinsic low torus luminosities, with the torus emission largely diluted in the host-galaxy emission. In this regard, all the thirteen sources which were better fitted without the AGN component were not detected in the X-ray. Moreover, all the objects with AGN significance $\leq 1\,\sigma$ are not X-ray detected, except for three sources. Two of these, however, have luminosity in the lower end of our distribution ($L_{2-10\mathrm{keV,intr}} < 10^{43}$ erg/s).

We studied the distributions of the [Ne v] luminosities for the sources with AGN significance $< 2\,\sigma$ and those with AGN significance $\geq 2\,\sigma$. We found a segregation of the sources with low AGN significance at low [Ne v] luminosities. In fact, while 44% of the sources with $\geq 2\,\sigma$ AGN significance has low [Ne v] luminosity ($L_{[\mathrm{Ne\,v}]} < 10^{41}$ erg/s), this percentage goes up to 85% for the sources with AGN significance $< 2\,\sigma$. This segregation is in support to the fact that low AGN luminosities (the [Ne v] emission is a proxy of the nuclear intrinsic emission) may be challenging in separating the AGN component from the galaxy emission.

Stellar mass

The median (and 16 – 84th percentiles) stellar mass is $\log(M_* / M_\odot) = 10.91^{+0.28}_{-0.46}$. We noted that the stellar masses are well constrained by the SED-fitting procedure ($\Delta \log M_* / \log M_* \approx 2\%$). We did not find any significant trend of the stellar mass with the X-ray detection.

Star formation rate

The SFR of our sample was obtained in two different ways: using the best-fit model (SFR$^{\rm sed}$) and using the 8 – 1000 μm SF luminosity (SFR$^{8-1000\mu m}$). The SFR$^{\rm sed}$ is the mean

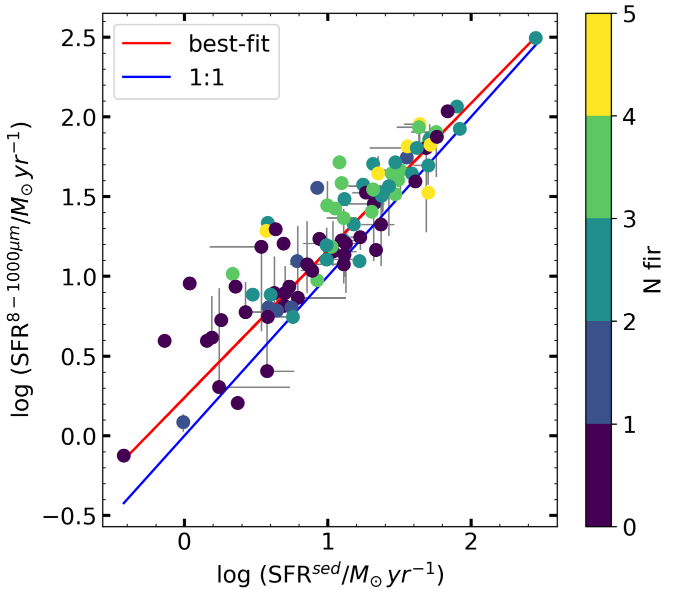

Figure 3.7: Comparison between the SFR obtained from the optical-NIR bands with those from FIR. SFR$^{\rm sed}$ are derived through the modeling of the stellar emission in the UV-to-NIR. The color code indicates the number of photometric detections for each source in the FIR band. A low number of FIR detections may influence the goodness of the FIR SED-fitting, hence the SFR$^{8-1000\mu m}$, but we did not find any significant trend between the SFR$^{8-1000\mu m}$–SFR$^{\rm sed}$ offset and the number of IR photometric detections (the apparent larger differences for the sources with few IR photometric points is a consequence of the logarithmic scale of the plots and of the segregation of low $N_{\rm FIR}$ sources to low SFR). The blue line is a 1:1 line; the red line is the best-fit line with a slope of m = 0.92 and c = 0.24, in the $\log(\mathrm{SFR}^{8-1000\mu m}/M_\odot\,\mathrm{yr}^{-1}) = c + m \log(\mathrm{SFR}^{\rm sed}/M_\odot\,\mathrm{yr}^{-1})$ notation.

SFR of the last 0.01 - 0.1 Gyr as obtained from the modeling of the stellar component in the UV-to-NIR regime with *sed3fit*. The code uses the UV-optical-NIR library of Bruzual (2007), that produced the optical-to-NIR spectra, by considering the spectral evolution of stellar populations for different metallicities and star formation histories and assuming a Chabrier IMF (Chabrier 2003). We obtained a median $\mathrm{SFR}^{\mathrm{sed}} = 12.9^{+30.4}_{-9.1}\, M_\odot\, yr^{-1}$ ($\Delta\mathrm{SFR}/\mathrm{SFR} \approx 0.6\%$). The $\mathrm{SFR}^{8-1000\mu\mathrm{m}}$ is the SFR averaged over the last 100 Myr computed by the emission of dust heated by young stars as well as of evolved stellar populations. It is derived from the IR luminosity, once the AGN contribution is removed, assuming the Kennicutt (1998) relation $\mathrm{SFR}^{8-1000\mu\mathrm{m}}\,(\mathrm{M}_\odot/\mathrm{yr}) = 4.5 \cdot 10^{-44}\, \mathrm{L}_{8-1000\mu\mathrm{m}}\,(\mathrm{erg/s})$ and using $\mathrm{SFR}_{\mathrm{Chabrier}} = 0.67 \cdot \mathrm{SFR}_{\mathrm{Salpeter}}$ to convert it to a Chabrier IMF. The median SFR is $\mathrm{SFR}^{8-1000\mu\mathrm{m}} = 20.4^{+43.4}_{-9.2}\, M_\odot\, yr^{-1}$. Because the $\mathrm{SFR}^{8-1000\mu\mathrm{m}}$ is heavily dependent on the fitting of the FIR band, the reliability of this value is linked to the accuracy at which the far-IR is measured. As we can see from Fig. 3.7, in which we compared the two SFR values for each source, we have a systematic ~ 0.15 dex offset, but we did not find a clear correlation between this offset and the number of IR photometric detections, nor with the SFR uncertainties.

SFR-M$_*$ relation

We investigated whether the hosts of the [Ne v]-selected AGN lie within the SFR-M$_*$ "main sequence" (Noeske et al. 2007), see Fig 3.8 (*red dots*). We used the Schreiber et al. (2015) "main sequence" (MS), in which the SFR is a function of both the stellar mass and the redshift:

$$\log\,(\mathrm{SFR}_{\mathrm{MS}}/\mathrm{M}_\odot\,\mathrm{yr}^{-1}]) = m - m_0 + a_0 r - a_1[\max(0, m - m_1 - a_2 r)]^2 \tag{3.6}$$

where $r \equiv \log(1+z)$, $m \equiv \log(\mathrm{M}_*/10^9\mathrm{M}_\odot)$, $m_0 = 0.5\pm0.07$, $a_0 = 1.5\pm0.15$, $a_1 = 0.3\pm0.08$, $m_1 = 0.36 \pm 0.3$ and $a_2 = 2.5 \pm 0.6$. Using the masses obtained from the SED-fitting and the Schreiber et al. (2015) SFR-M$_*$ relation (eq 3.6), we obtained a median $\mathrm{SFR}_{\mathrm{MS}} = (\,33.2^{+21.0}_{-14.8}\,)\,\mathrm{M}_\odot/\mathrm{yr}$. We used the $\mathrm{SFR}_{\mathrm{norm}} = \mathrm{SFR}^{\mathrm{sed}}/\mathrm{SFR}_{\mathrm{MS}}$ to trace how much a source deviates from the MS. We found a median $\mathrm{SFR}_{\mathrm{norm}} = 0.48^{+0.76}_{-0.37}$, which indicates that a significant fraction our sample has a SFR lower than what is expected for SF galaxies.

We investigated if the SFR and the position in the SFR-M$_*$ plane depend on the number of IR photometric detections (which are fundamental in constraining the overall SED in *sed3fit*). For the sources with at least one IR detection, we performed another run of SED-fitting without the IR photometric points. The comparison of the SED-derived properties with those obtained using all the photometric points shows a larger uncertainties in the parameters, but no systematic effect is present.

Mignoli et al. (2013) performed a morphological classification of the [Ne v] sample and a comparison with the morphologies of a control sample. The galaxies were classified following Nair & Abraham (2010) using the F814W-band images. The position of the [Ne v] in our SFR-M$_*$ plane, with the majority of the galaxies within the MS, is in agreement with the Mignoli et al. (2013) morphological classification, in which early-spirals are the most commonly found type for the [Ne v] sample (rather than irregular or late-spirals). Moreover, Mignoli et al. (2013) found a lower fraction of elliptical galaxies with respect to their control sample; this is in agreement with the paucity (with respect to our control sample) of low-sSFR ($\log \mathrm{sSFR/yr^{-1}} \leq -11$) galaxies we found in the [Ne v] sample (see § 3.5).

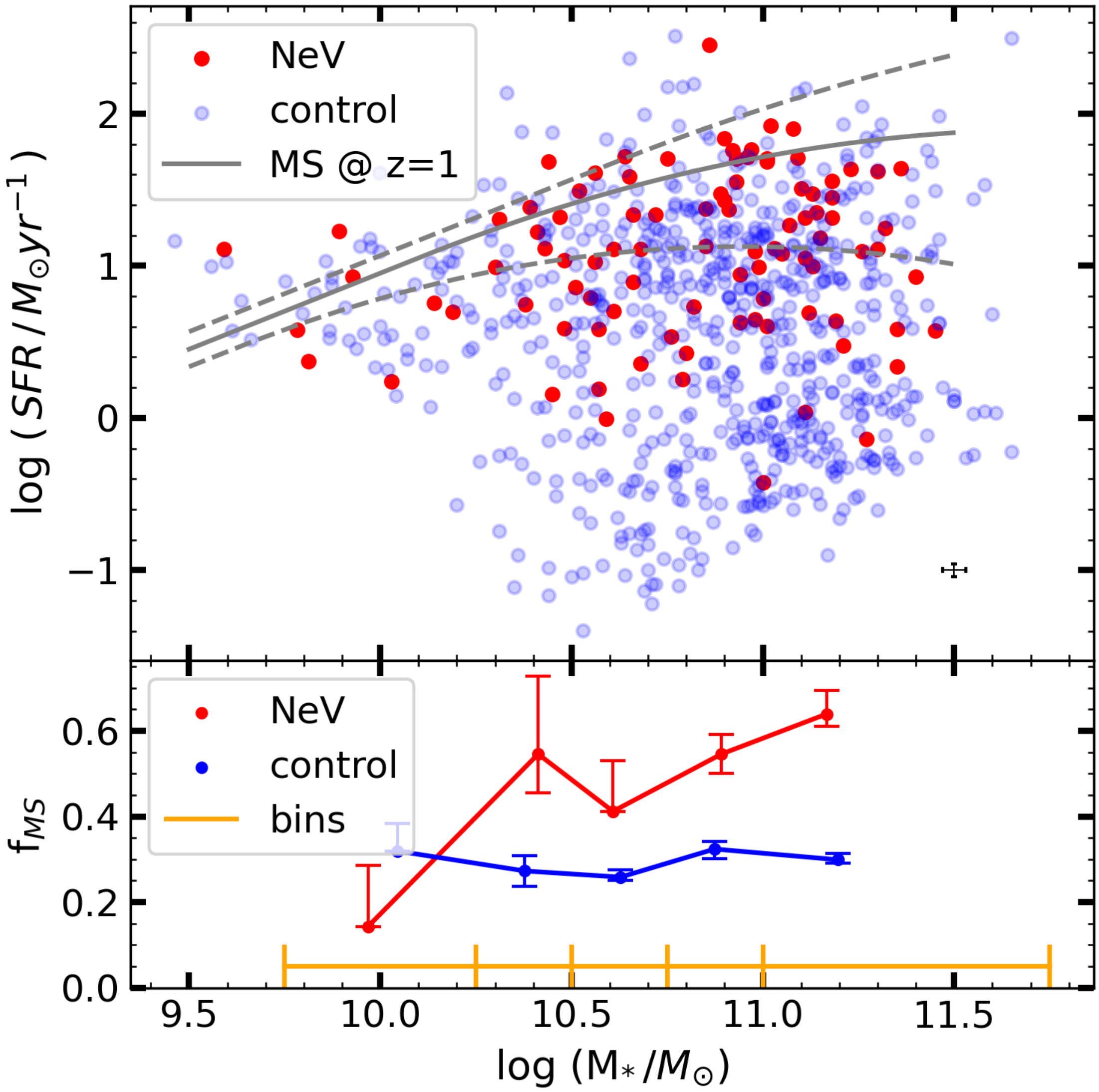

Figure 3.8: Comparison of the positions of the [Ne v] (*red*) and stellar mass- and redshift-matched control (*blue*) samples in the SFR−M_* plane. *Upper panel*: The grey solid line is the Schreiber et al. (2015) MS, the grey dashed lines its 1 σ dispersion. The black error bars in the bottom right are the mean uncertainties for the [Ne v] sample sources. *Bottom panel*: Fraction of sources within the 1 σ dispersion of the MS. In orange the adopted stellar mass binning.

Eddington Ratio

An effective way of estimating the SMBH growth is via the Eddington ratio $\lambda_{\rm Edd} = L_{\rm bol}/L_{\rm Edd}$: SMBHs already grown that are accreting slowly should have $\lambda_{\rm Edd} \ll 1$, while those in main episode of growth $0.1 \leq \lambda_{\rm Edd} \leq 1$.

We used the Suh et al. (2020) $M_{\rm BH} - \mathrm{M}_*$ relation to estimate the SMBH masses using the stellar masses obtained from the SED-fitting. This relation was obtained from a sample of 100 X-ray selected AGN in the COSMOS field with host galaxy masses from SED-fitting decomposition and BH masses computed considering single epoch Hα, Hβ, and Mg II broad line widths and line/continuum luminosity as proxy for the size and velocity of the Broad Line Region (BLR). We obtained a median $\log{(\mathrm{M_{BH}/M_\odot})} = 7.5^{+0.4}_{-0.7}$.

Due to the high number of different $M_{\rm BH} - M_*$ relations in literature, their large uncertainties and the different ways used to estimate the $M_{\rm BH}$, we tested three different $\mathrm{M_{BH}} - \mathrm{M}_*$ relations and compared them in Fig.3.9.

The Reines & Volonteri (2015) relation (black line) comes from a sample of 262 broad-line AGN and 79 galaxies. For the AGN, their $M_{\rm BH}$ estimations are derived from single epoch spectra of sources with broad Hα, using the line FWHM as well as its luminosity, under the virial assumption. For 15 AGN they used reverberation-mapped $M_{\rm BH}$ from literature. Finally, the $M_{\rm BH}$ for the 79 galaxies were obtained from measurements based on stellar dynamics, gas dynamics, and maser disk dynamics. The blue line is the Shankar et al. (2016) relation, obtained with sources from five different literature samples of galaxies with BH dynamical mass measurements.

We chose the Suh et al. (2020),$M_{\rm BH} - \mathrm{M}_*$ relation to estimating the $M_{\rm BH}$ mass for the following considerations: as our sample is composed of AGN, we preferred to avoid relations derived only with non-active galaxies; in addition, the Suh et al. (2020) relations is a "middle ground" compromise between the three relations and, considering the uncertainties, is compatible with the other relations. Finally, we performed KS-tests to check if the $M_{\rm BH}$ values of the [Ne v] sample obtained using the other two relations would differ significantly from those obtained with the Suh et al. (2020). We found that the $M_{\rm BH}$ obtained with Shankar et al. (2016), Reines & Volonteri (2015) and Suh et al. (2020) relations are not significantly different.

From the $M_{\rm BH}$ we computed the Eddington luminosity and used it with the AGN bolometric luminosity to estimate the Eddington ratio (Fig. 3.10). We found a median $\lambda_{Edd} = 0.12^{+0.31}_{-0.10}$. For 25 sources ($\sim$ 25%) we obtained only an upper limit of their $\lambda_{\rm Edd}$. Forty percent of the sources have $\lambda_{\rm Edd} \geq 0.1$, and 5% are accreting near or above the Eddington limit. Although the Eddington ratio distribution is strongly dependent on the choice of the $\mathrm{M_{BH}} - \mathrm{M}_*$ relation, all three investigated functions suggest that a significant fraction of the [Ne v] sample are AGN in a high accretion phase.

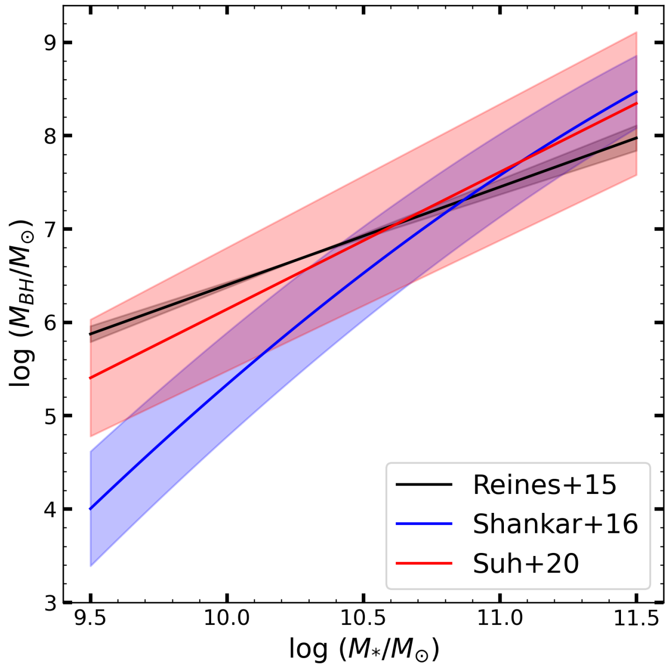

Figure 3.9: Comparison of different $M_{\rm BH} - M_*$ relations and relative uncertainties. The red line refers to the Suh et al. (2020) relation used in this work, while the black, and blue to the Reines & Volonteri (2015), and Shankar et al. (2016) respectively. KS-test on the obtained [Ne v] $M_{\rm BH}$ values assured us that using the Reines & Volonteri (2015), or Shankar et al. (2016) would not provide significantly different $M_{\rm BH}$.

Molecular gas fraction

We used the Kaasinen et al. (2019) relation (eq 3.7) to estimate the molecular gas mass of the sample. They obtained this relation linking the $L_{850\mu m}$ to the CO luminosity, which, in turn, is a proxy of the molecular gas mass.

$$\mathrm{M}^{850\mu\mathrm{m}}_{\mathrm{molgas}}\ (\mathrm{M}_\odot) = \left(\frac{\mathrm{L}_{850\mu m}}{\mathrm{erg\ s^{-1}\ Hz^{-1}}}\right)\left(\frac{1}{6.2\times10^{-19}\left(\mathrm{L}_{850\mu\mathrm{m}}/10^{31}\right)^{0.07}}\right) \tag{3.7}$$

We computed $L_{850\mu m}$ from the host-galaxy component, derived via the SED-fitting. The median molecular gas mass of the sample was $\log(M^{850\mu m}_{molgas}/M_\odot) = 10.4^{+0.4}_{-0.5}$. As sanity check, we obtain comparable M_{molgas} from the dust mass using a gas-to-dust mass ratio (G/D) in the in the 100 − 200 range (best fit G/D ∼ 130), in agreement with e.g., Tacconi et al. (2020). We used the $\mathrm{M}^{850\mu m}_{\mathrm{molgas}}$ to obtain the molecular gas fraction of the galaxies, defined as $f_{\rm mol} = \mathrm{M}^{850\mu m}_{\mathrm{molgas}}/(\mathrm{M}_{\mathrm{molgas}} + \mathrm{M}_*)$. The sample had a median $f_{\rm mol} = 0.24^{+0.30}_{-0.14}$. We found that our sample has $f_{\rm mol}$ in agreement with those in Dessauges-Zavadsky et al. (2020a), who collected from literature the $f_{\rm mol}$ for CO detected main sequence galaxies at $0 \le z \le 6$. The $f_{\rm mol}$ distribution indicates that the majority of the [Ne v] sources have cold

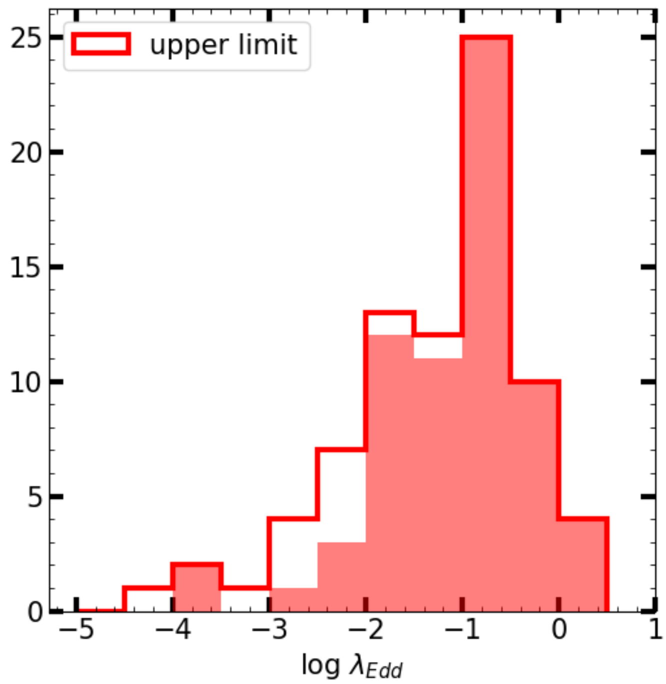

Figure 3.10: Eddington ratio distribution of the [Ne v]-selected sample of type 2 AGN. The $\lambda_{\rm Edd}$ was computed with the M_{BH} obtained from the stellar mass (using the Suh et al. 2020 $M_{\rm BH} - M_*$ relation) and with the AGN bolometric luminosity from the SED fitting. The sources for which we had only an upper limit on their bolometric luminosity are reported as a white histogram. 40% of the sources have $\lambda_{\rm Edd} \geq 0.1$, suggesting that a significant fraction of the [Ne v] sample are in a highly accreting phase.

gas available to fuel the SF, and, on the contrary of "read and dead elliptical galaxies", will likely continue to forming stars, not having depleted their reservoir yet.

3.5 Comparison with the control sample

To be able to make comparison between our sources and non-active galaxies in the COSMOS field, we built a control sample. For each source in the [Ne v] sample we selected 8 sources in the zCOSMOS20k catalog, matched in redshift (with a 0.01 – 0.05 accuracy) and stellar mass (0.1 dex accuracy). From this sample we excluded the 51 sources with the X-ray source flag in the COSMOS2020 catalog. We performed the SED-fitting in the same way as we did for the [Ne v] sample (see section 3.4), and excluded 30 sources for which the SED-fitting failed to properly reproduce the emission. The final control sample was composed of 618 sources. Fig. 3.11 shows the comparison of the z and M_* distributions between the [Ne v] and the control samples.

We report in Table 3.2 the comparison between the main properties of the [Ne v] sample and of the control sample.

The [Ne v] sample has, on average, a higher SFR with respect to the control sample,

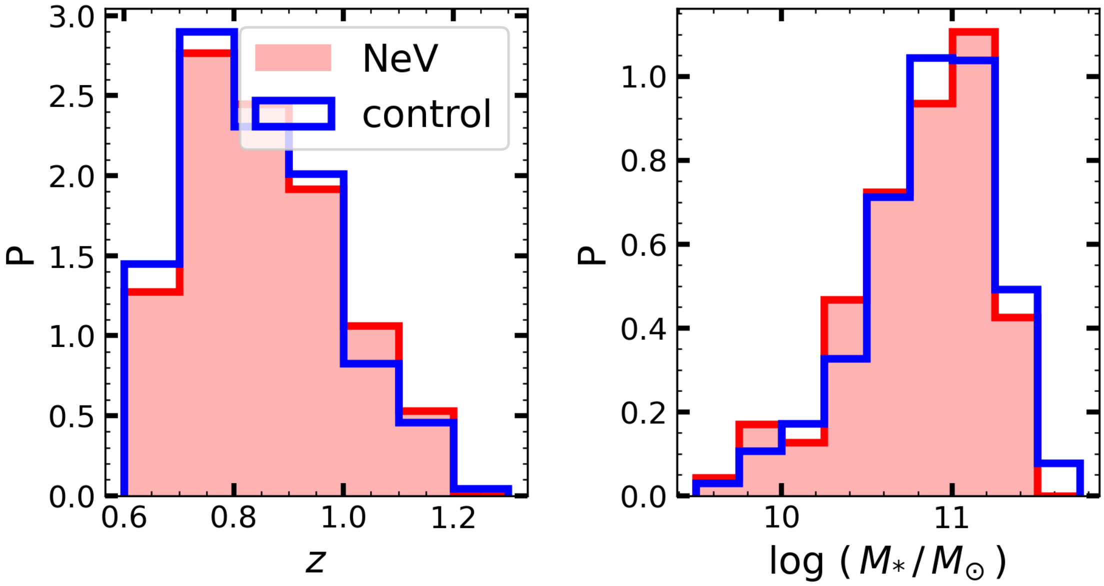

Figure 3.11: Comparison of redshift (*left*) and stellar mass (*right*) distributions of the [Ne v] (*red*) and control (*blue*) samples.

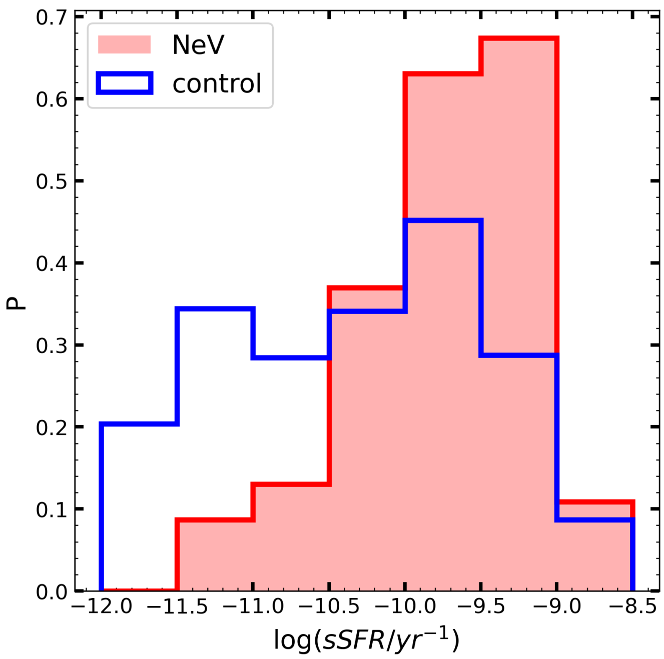

Figure 3.12: Comparison of the specific SFR (sSFR=SFR/M_*) distributions of the [Ne v] (*red*) and control (*blue*) samples. The two sSFR distributions are significantly different (KS-test P= $1.3{\times}10^{-7}$), with the [Ne v] sample lacking a population of "quiescent" galaxies (log(sSFR/yr^{-1}) ≤ -11).

although both reach values as high as $\sim 300 M_{\odot}/yr$. Considering the position in the SFR-M_* plane, we found that the [Ne v] sample has a higher fraction of sources in the MS (thus

Table 3.2: Comparison of the main properties of the [Ne v] sample and of the control sample. We report the median values (with the 16^{th} and 84^{th} percentiles) for the redshift z, stellar mass, SFR, specific SFR (sSFR=SFR/M_*), SFR_{norm}=SFR/SFR_{MS}, fraction of sources within the Schreiber et al. (2015) MS (f_{MS}), molecular gas mass (M_{molgas}), the molecular gas fraction ($f_{mol}=M_{molgas}/(M_{molgas}+M_*)$), age of the galaxy ($t_{age}$, from the SED-fitting), BH mass (M_{BH}) and Eddington ratio (λ_{Edd}).

	[Ne v]	control
z	$0.86^{+0.11}_{-0.15}$	0.85 ± 0.11
$\log(M_* / M_\odot)$	$10.91^{+0.28}_{-0.46}$	$10.91^{+0.30}_{-0.42}$
SFR^{sed} ($M_\odot\, yr^{-1}$)	$12.9^{+30.4}_{-9.1}$	$4.3^{+14.4}_{-3.7}$
$\log(sSFR/yr^{-1})$	$-9.7^{+0.5}_{-0.7}$	$-10.2^{+0.8}_{-1.1}$
SFR_{norm}	$0.48^{+0.76}_{-0.37}$	$0.16^{+0.56}_{-0.14}$
f_{MS}	$0.52^{+0.08}_{-0.02}$	0.30 ± 0.01
$\log(M_{molgas} / M_\odot)$	$10.4^{+0.4}_{-0.5}$	$10.3^{+0.5}_{-0.7}$
f_{mol}	$0.24^{+0.30}_{-0.14}$	$0.20^{+0.26}_{-0.15}$
$\log(t_{age}/yr^{-1})$	$9.48^{+0.20}_{-0.25}$	$9.63^{+0.15}_{-0.30}$
$\log(M_{BH}/M_\odot)$	$7.5^{+0.4}_{-0.7}$	
λ_{Edd}	$0.12^{+0.31}_{-0.10}$	

a lower fraction of sources below). As we can see from Fig. 3.8, for $\log(M_* / M_\odot) \geq 10.25$, there is a clear difference in the fraction of sources within the MS, with the control sample having a constant fraction $f_{MS} \sim 40\%$, while the [Ne v] $f_{MS} \sim 60\%$. This difference between the two samples appears more evident in Fig. 3.12: the control sample specific SFR (sSFR=SFR/M_*) has a broader distribution (almost bimodal depending on the chosen binning), while the [Ne v] has a higher fraction of sources at high sSFR and seems to lack a population of "quiescent" galaxies ($\log(sSFR/yr^{-1}) \leq -11$). The two sSFR distribution are significantly different (KS-test P= 1.3×10^{-7}). The comparison of the t_{age} (the age of the oldest stars in the galaxy, as provided by *sed3fit*) distributions shows a significant difference (KS-test P$\sim 6 \times 10^{-7}$) between the two samples, with the [Ne v] sources being, on average, younger.

We investigated the cold gas content of the galaxies to test whether the difference in sSFR (and the lack of "quiescent" galaxy) was originated by a different amount of gas available to fuel the SF. As in § 3.4.2, we computed M_{molgas} and f_{mol} for both samples. We did not find any significant difference in the M_{molgas} (Fig. 3.13) and f_{mol} distributions. Moreover, binning in M_* and SFR, we found that the median f_{mol} in each bin is practically the same for both samples (with the exception of the lowest M_* and highest SFR bin, which have only one [Ne v] source within).

We conclude that the highest fraction of sources within the MS for the [Ne v] sample and the lack of a population of "quiescent" galaxies is not related to a different amount of cold gas reservoir in the host galaxy, rather it is linked to a different efficiency in forming stars. We propose three explanations for this higher SF efficiency: the AGN may have the effect of enhancing the SF; the AGN is more likely to be triggered in galaxies with higher SFR; the [Ne v] selection efficiently allows us to pick up AGN in an obscured growth phase, where we expect high SFR and obscured AGN activity. This last hypothesis is consistent with the [Ne v] sample being slightly younger on average: the IR host galaxy are still growing and will reach higher masses than the control sample.

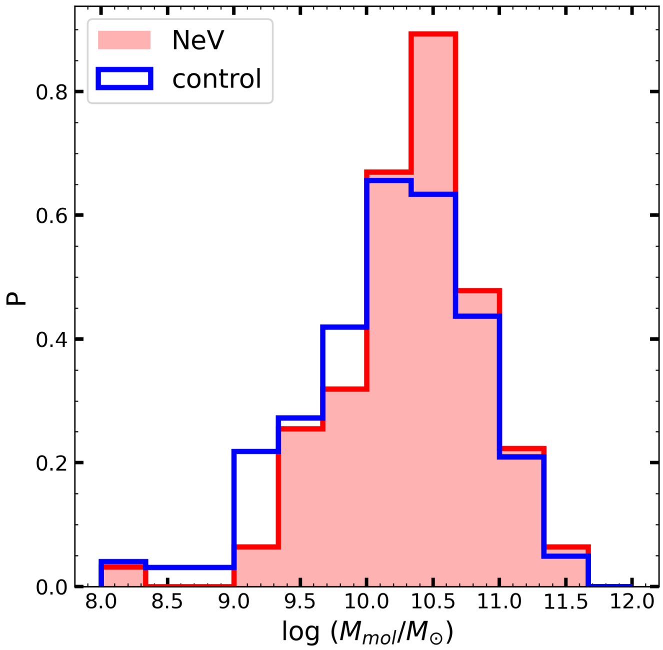

Figure 3.13: Comparison of the molecular gas mass distributions of the [Ne v] (*red*) and control (*blue*) samples. We obtained the M_{mol} from the $L_{850\mu m}$, using the Kaasinen et al. (2019) relation. The two M_{mol} distributions are not significantly different.

3.6 Interpretation within the *in-situ* co-evolution model

We place our sources in the context of the AGN-galaxy co-evolution paradigm by comparing them with the theoretical prediction of the *in-situ* BH-galaxy evolution model (Mancuso et al. 2016a,b, 2017; Lapi et al. 2018, ; see section 1.4.2).

Briefly, in the *in-situ* co-evolution model, the SF is a local process that is, at first, regulated by the feedback from SNe, then by the one from the central SMBH. In the early stages, the BH mass grows in gas-rich obscured environment, the AGN emits at mildly super Eddington ratio and its luminosity rises exponentially. In this phase the SF is regulated by its own feedback. However, when the AGN luminosity reaches values similar to those of the SF, it becomes dominant and the AGN feedback quenches the SF. Then the galaxy moves below the MS and evolves passively (i.e., its stellar mass growth becomes negligible), while the AGN emits at sub-Eddington ratios and its AGN luminosity declines.

In figure 3.14 we illustrate the position of the [Ne v] sources in the SFR$-L_{bol}$ diagram, along with the prediction of the *in-situ* model (Mancuso et al. 2016a). The colored contours show the number density of sources predicted by the model. The dashed lines represent two evolutionary track from the *in-situ* model for AGN with different peak luminosity (the maximum bolometric luminosity an AGN can reach before the feedback kicking in and lowers the AGN output), with the forward time direction indicated by the arrows. The evolution of the sources begins from the left, with the SFR roughly constant while the AGN luminosity grows exponentially. When the source crosses the blue $L_{bol} = L_{SF}$ line, the AGN luminosity becomes dominant. Then, when the feedback starts to 'kick-in' the SF is abruptly quenched and the SFR drops quickly, while the AGN emission fades with a slower trend. Binning the [Ne v] sources in three L_{bol} bins (black pentagons) shows that the mean SFR follows the same trend reported in Mancuso et al. (2016a): an approximately constant SFR until a rise near the position where the feedback kicks-in. This rise is attributed to the mean SFR being statistically dominated by objects with higher SFR, due to the fact that, for reaching this AGN luminosity, the BH must be hosted in galaxies with higher mass, thus likely higher SFR.

Color coding the [Ne v] sources on the basis of their t_{age} shows that the oldest galaxies are preferentially on the lower-left part of the diagram, in the locus where we expect the AGN to have started quenching the SF. A similar behavior is found if we consider their sSFR. The position of the [Ne v] sources in the SFR$-L_{bol}$ plane confirms that our sample is mainly composed of sources in the obscured accretion phase, with the oldest sources being exactly where we would expect them to be if they were at the beginning of the 'quenching phase'.

3.7 Conclusions

In this chapter, we investigated the AGN and host galaxies properties of a sample of 94 [Ne v]-selected type 2 AGN. We performed an X-ray spectral analysis of the 36 X-ray detected sources, to characterize their AGN intrinsic luminosity and obscuration. For the X-ray undetected sources, we used the X/[Ne v] ratio to estimate their amount of obscuration and the fraction of CT-AGN. We performed optical-to-FIR SED-fitting, using the *sed3fit* algorithm, to characterize both the AGN and the host galaxy properties. We used the stellar mass to obtain the BH mass and, therefore, the Eddington ratio of the [Ne v] sample. Finally, we compared the host galaxy properties (stellar mass, SFR, sSFR, cold gas content) of the [Ne v] sample with those of a non-active control sample and interpret our results in the light of the *in-situ* BH-galaxy evolution scenario. Our main results are the following:

1. The [Ne v] selection is an optimal tool to select very obscured ($N_H \geq 10^{23}$ cm^{-2})

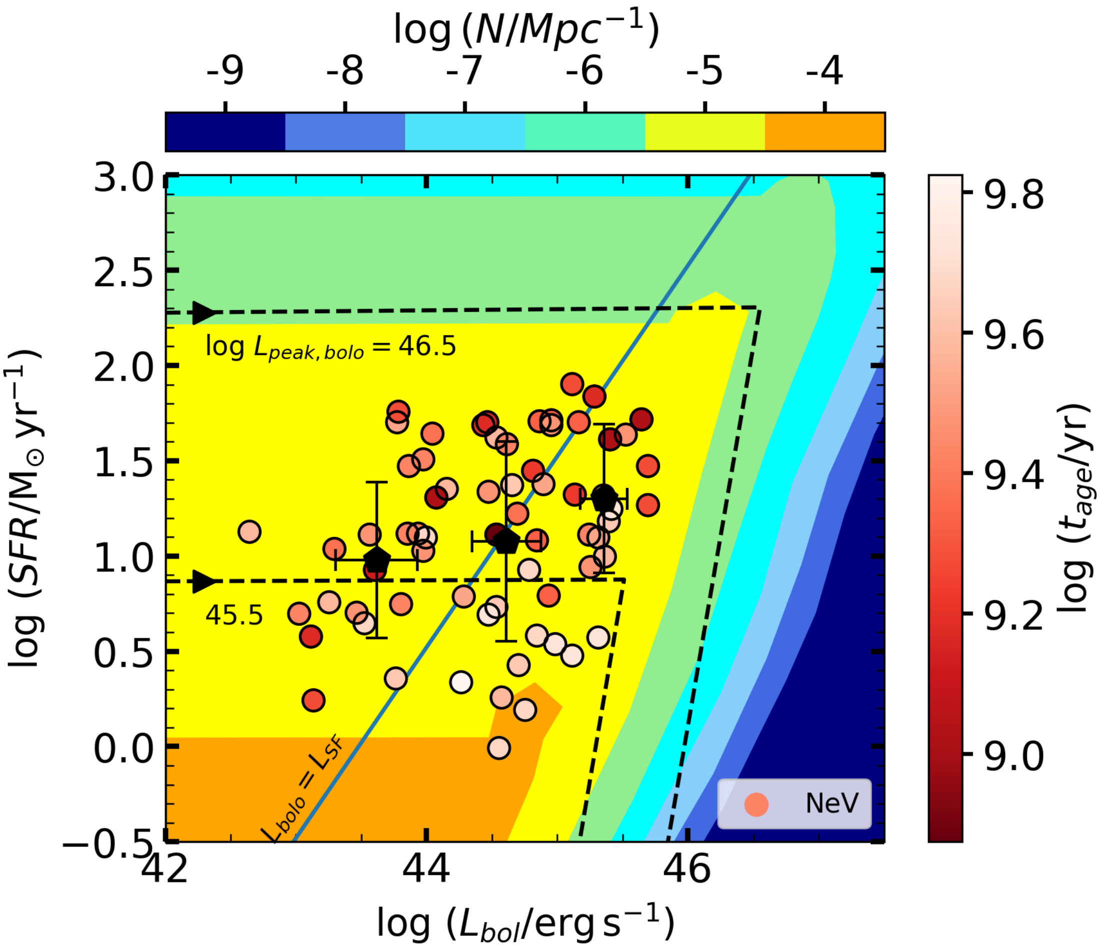

Figure 3.14: Distribution of the [Ne v] sources on the SFR−L_{bol} plane. The color code indicates the age of the oldest stars in the host galaxy, as provided by *sed3fit*. Colored contours illustrate the number density of galaxies plus AGN at $z \sim 1$ as predicted by the *in-situ* co-evolution model (Mancuso et al. 2016a); orange, yellow, green, cyan, light blue, blue, and dark blue contours refer to number density of 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}, 10^{-8}, 10^{-9} Mpc^{-1}, respectively. The black dashed lines show two evolutionary tracks (forward time direction indicated by the arrows) for AGN with peak bolometric luminosity of $10^{45.5}$, and $10^{46.5}$ erg/s. The blue continuous line indicates where the SF luminosity is equal to the AGN luminosity. The black stars with error bars refer to the mean and standard deviation of the [Ne v] sources binned in AGN bolometric luminosity.

and CT-AGN at $z \sim 1$. More than two-thirds of our sample is composed of very obscured sources, and $\sim 20\%$ of the sources are candidate CT-AGN.

2. Almost half of the sample is composed of AGN with high Eddington ratio ($\lambda_{\rm Edd} \geq 0.1$) in a strong episode of SMBH growth.

3. The [Ne v] sample has a significantly higher fraction of sources within the MS than normal galaxies, and seems to lack a population of "quiescent" (low sSFR) galaxies.

4. This difference is not due to the amount of cold gas reservoir, as both samples show similar M_{mol}, but to the higher efficiency in forming stars of the [Ne v] sample. This higher efficiency could be related to the AGN enhancing the SF, to the high SF triggering the AGN activity; alternatively, it could be due to a selection effect, with the [Ne v]-selection picking-up AGN in their obscured growing phase.

5. We interpret our results in the context of the *in-situ* co-evolution scenario and find that the [Ne v] sample is mostly composed by sources in the 'pre-quenching' phase, with only the few oldest sources showing the effect of the AGN quenching the SF. Therefore, we favor the latter hypothesis: [Ne v] sources are preferentially hosted in obscured AGN efficiently accreting and forming stars.

In chapter 4, we will exploit the C IV emission line to extend this analysis to obscured AGN at higher redshifts.

Chapter 4

C IV-selected obscured AGN at $z \sim 2-3$

In this chapter, we extend the search and the analysis of obscured AGN to higher redshift compared to the [Ne V] selection. We perform an analysis similar to the one of chapter 3 exploiting the C IV 1550Å line selection to move from $z \sim 1$ to $z \sim 2-3$. This allows us to cover the entire cosmic noon, with the [Ne V] sample targeting the late phase and the C IV selection targeting its early phase.

4.1 Introduction

In chapter 3, we exploited the [Ne V] emission line to select obscured AGN up to $z \sim 1.2$. We analyzed both the AGN and their host galaxies, and showed that these sources are exactly the obscured accreting AGN that are needed to investigate the BH-galaxy co-evolution paradigm. The evolution of the SFRD and BHAD (section 1.4.3) reveals that most of the SF and BH accretion should have taken place at the cosmic noon, i.e., around $z \sim 2$. However, the [Ne V] line is redshifted out of the VIMOS optical spectral range at $z \sim 1.2$. Therefore, the use of lower rest-frame wavelengths are needed to cover the cosmic noon. The C IV 1550Å emission line is a good candidate, given its high-ionization potential, likely a signature of the AGN activity, and since it is usually the most intense line in the UV spectra (with the exception of the Lyα). Furthermore, it falls in the VIMOS spectral coverage for sources in the $1.45-3.05$ redshift range.

4.1.1 The C IV-selected sample

The C IV sample was drawn by Mignoli et al. (2019, hereafter M19) from the zCOSMOS-deep survey, similarly to the [Ne V] sample discussed in section 3.2. We refer to M19 for an in-depth description of the sample selection. Briefly, the zCOSMOS-deep survey includes 9523 spectroscopically observed objects, 80% of them with available redshift. By limiting the sample to the $1.45-3.05$ redshift range (4391 galaxies), it is guaranteed that the C IV 1550Å line fell within the observed wavelength range. Their VIMOS spectra were visually inspected to identify C IV emitters, as the complexity of the line (the C IV profile can include an AGN emission component, absorption due to the ISM, and P-Cygni profile from stellar winds) did not allow an automatic detection procedure. Finally, all the sources with a C IV emission peak five time higher than the nearby (i.e., in a 50Å window around the C IV line) continuum rms were classified as C IV emitters . The final sample was composed of 192 AGN; adopting a 2000 km s^{-1} FWHM threshold, 90 were classified as type 2 AGN (FWHM< 2000 km s^{-1}) and 102 as type 1 AGN.

In this chapter, we focus only on the type 2 C IV-selected AGN, and we will refer to them as the C IV sample; however, in the future we will also analyze the type 1 AGNs and draw comparisons between the samples.

The median (and 16th – 84th percentile) redshift of the sample is $2.16^{+0.56}_{-0.50}$. In Fig. 4.1, we show the redshift distributions of the C IV and [Ne V] samples; the use of both samples allows us to cover the entire cosmic noon.

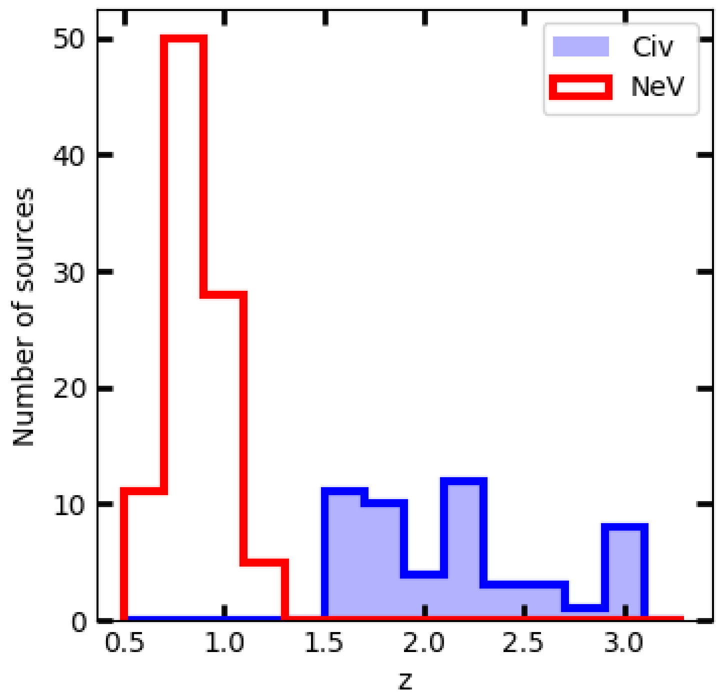

Figure 4.1: Redshift distributions of the C IV sample (*blue*) and of the [Ne V] sample (*red*). The use of both samples allows us to cover the entire cosmic noon.

4.1.2 X-ray Data

All 90 C IV AGN fall in the *Chandra* COSMOS Legacy mosaic, composed of the *Chandra* COSMOS survey (the central ~ 0.9 deg^2; Elvis et al. 2009) and the COSMOS Legacy survey (covering the external ~ 1.7 deg^2 with a similar depth of the C-COSMOS survey; Civano et al. 2016). Fifty-two C IV-selected AGN were detected by *Chandra* within 1.5" from the optical position, with a median displacement of 0.5". Similarly to the [Ne V] sample, the X-ray spectra were extracted using the *CIAO* (Fruscione et al. 2006) tool `specextract` from circular regions of radius r_{90}+. The background spectra were extracted from annuli centered on the source position with inner radius r_{90} + 2.5" and outer radius of r_{90}+20". For each source, the spectra of its multiple observations were combined in a single spectrum via the *CIAO* tool `combine_spectra`.

4.1.3 Photometric Data

For the SED-fitting analysis of the C IV AGN, we used multi-band photometric data from the COSMOS 2020 catalog (Weaver et al. 2022, improved version of the previous COSMOS 2015 catalog from Laigle et al. 2016) and from the "super-deblended" catalogue (Jin et al. 2018). The two catalogs were cross-matched based on the source position. Not all the sources in our sample had a reliable super-deblended counterpart; we limited ourselves to a conservative separation of 0.3".

The use of both catalogs allowed us to cover from the optical bands down to the mm band and also to have a radio VLA photometric point at 1.4 GHz from Schinnerer et al. (2010). The complete list of photometric bands used for the SED-fitting is reported in Table 4.1.

4.2 X-ray spectral properties

Using the *XSPEC* software (Arnaud 1996), we performed the X-ray spectral analysis of 52 type 2 AGN with X-ray detection. The median number of net-counts is 48^{+36}_{-22}, two sources have less than 20 counts, only one source has more than 200 counts. Due to the low number of net-counts, we used unbinned data and C-statistic (Cash 1979) for all the sources.

The models used in the spectral analysis were the same as the [Ne V] sample analysis (see section 3.3.1); i.e., a simple power-law corrected for the Galactic absorption to obtain the spectral slope and a fixed $\Gamma = 1.8$ power-law with an additional absorption component to estimate the amount of obscuration.

We visually inspected all the spectral fittings, and try to add more complex components if the fit was not dimmed satisfactory. For example, for source zCOSMOS 403380, we obtained a better fit with a partial covering component with a covering fraction of $f = 0.95^{+0.03}_{-0.08}$. Finally, for 22 of the sources, instead of fitting the background subtracted

Table 4.1: Photometric bands used in the SED-fitting.

Observatory /Instrument	Filter	Central λ [Å]	Ref.
X-ray	$2-10\,\mathrm{keV}$	2.1	1
	$0.5-2\,\mathrm{keV}$	9.9	1
CFHT/MegaCam	u	3823	2
Subaru/Suprime-Cam	B	4400	2
	g^+	4804	2
	V	5487	2
	r^+	6305	2
	i^+	7693	2
	z^+	8978	2
	z^{++}	9063	2
Subaru/HSC	g	4847	2
	r	6219	2
	i	7699	2
	z	8894	2
	y	9761	2
HST/ACS	F814W	8333	2
VISTA/VIRCAM	Y	10216	2
	J	12525	2
	H	16466	2
	K_s	21577	2
Spitzer/IRAC	ch1	35686	2
	ch2	45067	2
	ch3	57788	2
	ch4	79958	2
Spitzer/MIPS	24μm	238433	3
Herschel/Pacs	green	1.03×10^6	3
	red	1.67×10^6	3
Herschel/SPIRE	PSW	2.51×10^6	3
	PMW	3.52×10^6	3
	PLW	5.12×10^6	3
JCMT/SCUBA2	850	8.63×10^6	3
VLA	1.4 GHz	20 cm	3

Notes. (1) Flux computed via X-ray spectral analysis, see section 4.3.2.

References. (2) Weaver et al. (2022); (3) Jin et al. (2018).

data, we performed an additional run of X-ray spectral fitting modeling the background as well. These sources were those for which the background dominated the spectrum, thus a simple background subtraction was leaving residuals, or otherwise, making the spectrum challenging to fit. The background was modeled with a phenomenological model, following Fiore et al. (2012); briefly, it consists of two powerlaw components (to reproduce the continuum), three narrow Gaussian components (to model the emission lines at 1.48, 1.74, and 2.16 keV), one broad Gaussian component (to reproduce the broad bump between 1 and 2 keV), and a thermal component. We fitted the background of each source separately, and not all the components were present in all the spectra. Once we found a good fit of the background, we rescaled its normalization with the ratio between the extraction area of the source and the area of the background, and used it alongside the source model to fit the AGN spectrum. For all the sources, we then computed the intrinsic luminosity and amount of obscuration value assuming the same AGN model as the background-subtracted sources.

The median spectral index, intrinsic luminosity, and obscuration of the C IV sample are reported in Table 4.2. On the basis of their N_H, almost all the X-ray detected sources (48) are obscured ($\log(N_H / cm^{-2}) \geq 22$), 37 (71 %) are extremely obscured ($\log(N_H / cm^{-2}) \geq 23$), and 2 (4 %) are CT-AGN ($\log(N_H / cm^{-2}) \geq 24$).

From Fig. 4.2, it easy to see that the C IV-selected AGN with X-ray detection are more luminous and more obscured with respect to the [Ne V] sample AGN.

Table 4.2: X-ray spectral properties (median and $16^{th} - 84^{th}$ percentile) of the C IV and [Ne V] samples.

	C IV	[Ne V]
z	$2.16^{+0.56}_{-0.50}$	$0.88^{+0.10}_{-0.17}$
$N_{net-counts}$	48^{+36}_{-22}	83^{+129}_{-63}
Γ	$0.6^{+0.7}_{-0.6}$	$1.1^{+0.6}_{-1.1}$
$\log(L_{2-10,intr} / erg\,s^{-1})$	$44.1^{+0.4}_{-0.5}$	$43.6^{+0.4}_{-0.6}$
$\log(N_H / cm^{-2})$	$23.4^{+0.4}_{-0.6}$	$22.6^{+0.7}_{-0.9}$

Notes. Γ is the photon index obtained from a simple power-law (corrected for the Galactic absorption) model. $L_{2-10,intr}$ and N_H are the intrinsic (i.e., absorption corrected) 2 – 10 keV luminosity and the amount of obscuration, obtained from a power-law model (corrected for the Galactic absorption) with fixed photon index $\Gamma = 1.8$ plus an absorption component.

4.2.1 X-ray undetected sources

Similarly to the [Ne V] sources (see sec. 3.3.2), we analyzed the C IV X-ray undetected sources. We used the *CIAO* tool `dmextract` to obtained from the 0.3 – 7 keV COSMOS Legacy mosaic image the number of counts for each undetected source at the corresponding optical position, and the tool `modelflux` to calculate the flux upper limit. Using

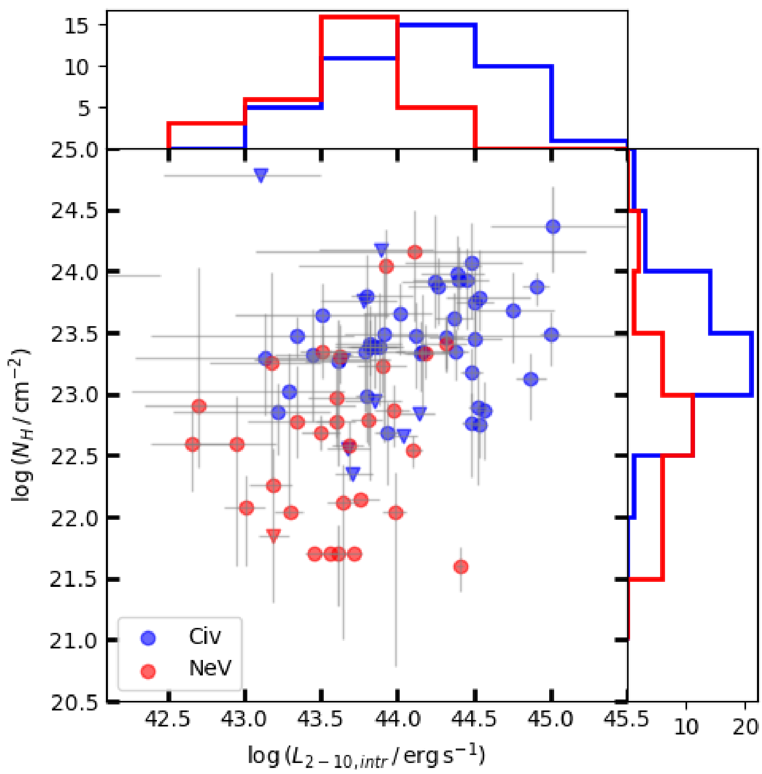

Figure 4.2: Obscuration as a function of the intrinsic $2-10\,$keV luminosity for the C IV sample (*blue*) and for the [Ne V] sample (*red*). The *grey lines* are the associated 90% uncertainties. The *triangles* denote upper limits on the obscuration (usually due to the limits of the X-ray fitting in low X-ray count statistics).

`dmextract`, we calculated the $0.5-7$ keV (observed-frame) net count-rate in a circular region centered on the source position. The radius was computed as follows: for each of our undetected AGN, we found the nearest detected source, we circularized all the extraction regions that were used for this source (each of these corresponds to the PSF size and shape at the source position in each of the available observations of the mosaic), and took the 80 percentile radius. By using the extraction regions of the nearest detected source, we are assuming that the PSF would not differ significantly for close positions in the field of view. The chosen radii varied, depending on the source, between ~ 2 and ~ 7 arcsec. Background counts were extracted in an annular region centered on the source position, using as inner and outer radii of 2 and 5 times the radius of the source region. We visually confirmed that no other source fell within the chosen radius or in the background region. The uncertainties were computed at $1\,\sigma$ confidence level, exploiting the Gehrels (1986) approximation to confidence limits for Poisson distributions due to the low number of counts*. We divided the net-counts by the effective (i.e., vignetted-corrected) exposure time at the source position to obtain the net-count rate. Because these sources were not

*`https://cxc.cfa.harvard.edu/ciao/ahelp/dmextract.html`

detected in the X-ray band, the net count-rate + uncertainties can be considered as 1 σ upper limits. We used `modelflux` to calculate, from the net count-rates, the upper limit for the flux in the 2 – 10 keV rest-frame energy range. We used a power-law model with photon index $\Gamma = 0.6$ (the average spectral index of the X-ray detected sources), modified by Galactic absorption. This model takes into account the source obscuration via a flatter photon index than the intrinsic one $\Gamma = 1.8 - 1.9$. Finally, we converted this flux to a rest-frame 2 – 10 keV luminosity.

As sanity check, we performed the same procedure for the X-ray detected C IV AGN and compared the resulting $L^{ul}_{2-10\mathrm{keV}}$ with the luminosity obtained from the X-ray spectral analysis ($L^{\mathrm{X-ray\,spec}}_{2-10\mathrm{keV}}$). As we can see from Fig. 4.3 (*left panel*), the X-ray luminosities obtained from the mosaic using the method described above do not differ significantly from those coming from the X-ray spectral analysis. Notable exception are the few sources with high $L^{\mathrm{X-ray\,spec}}_{2-10\mathrm{keV}}$ and low $L^{ul}_{2-10\mathrm{keV}}$: these are the most obscured sources with photon index $\Gamma \sim 0$ or negative, for which the assumption of $\Gamma = 0.6$ underestimates their intrinsic luminosity. At the moment, we do not have a reliable way to estimate the obscuration of the X-ray undetected C IV sources, therefore we assume that these sources have N_H similar to the X-ray detected sample (thus $\Gamma = 0.6$, which correspond to $\sim 10^{23}\,\mathrm{cm}^{-2}$).

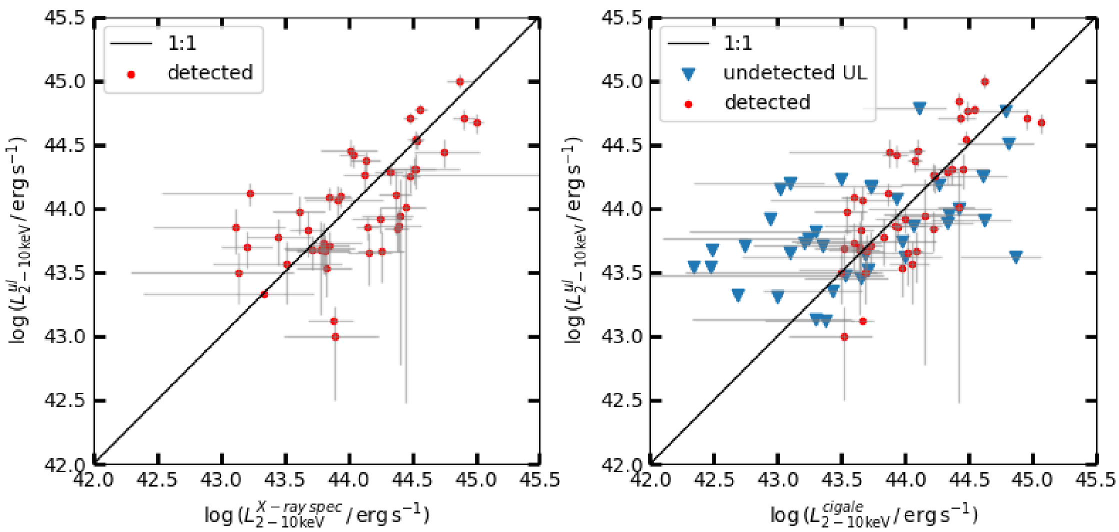

Figure 4.3: *Left panel*: Comparison between the rest-frame 2 – 10 keV luminosities obtained from the X-ray spectral analysis ($L_{2-10\mathrm{keV}^{\mathrm{X-rayspec}}}$) with the ones extracted from the mosaic ($L^{ul}_{2-10\mathrm{keV}}$) for the X-ray detected sources. The $L^{ul}_{2-10\mathrm{keV}}$ were computed following the same procedure used to obtain the luminosity upper limits for the X-ray undetected sources. *Right panel*: Comparison between the rest-frame 2 – 10 keV luminosities obtained from the SED-fitting ($L_{2-10\mathrm{keV}^{\mathrm{cigale}}}$) with the luminosities obtained from the mosaic. The *red dots* are the X-ray detected sources, while the *blue arrows* refer to the $L_{2-10\mathrm{keV}^{\mathrm{ul}}}$ upper limits for the X-ray undetected sources.

In Fig. 4.3 (*right panel*), we compare the X-ray luminosities from the mosaic with those obtained from the SED-fitting. The latter is an output of CIGALE, linked to the

AGN intrinsic luminosity, obtained from the best-fit AGN component of the SED. The median (and 16$^{\rm th}$–84$^{\rm th}$ percentile) luminosity for the X-ray detected AGN is $\log{(L^{ul}_{2-10\rm keV}/\rm erg\,s^{-1})} = 44.0^{+0.5}_{-0.3}$ and $\log{(L^{ul}_{2-10\rm keV}/\rm erg\,s^{-1})} = 43.8^{+0.4}_{-0.3}$ for the X-ray undetected. On average, the X-ray undetected sources are less luminous than the X-ray detected ones. More interestingly, for the X-ray undetected sources CIGALE does not have any prior on the AGN luminosity from the X-rays, therefore, finding agreement between $L^{ul}_{2-10\rm keV}$ and $L^{\rm cigale}_{2-10\rm keV}$ was not obvious, and it seems to confirm the reliability of our analysis.

zCOSMOS 411636 is the only source for which the luminosity from SED-fitting is significantly higher than the upper limit in the X-rays. Upon inspection, this source has ~ 20 counts in the mosaic and was probably classified as X-ray undetected due to its low significance. In the future, we plan to extract and analyze its spectrum, as it could be another extremely obscured or CT-AGN.

4.3 X-ray to IR SED-fitting

To estimate the AGN and galaxies properties, we chose to use the `CIGALE` code (Boquien et al. 2019; Yang et al. 2020, 2022). The choice of a different SED fitting code with respect to `sed3fit` used in chapter 3 is driven by the larger wavelength range of `CIGALE`. In fact, `CIGALE` allows us to fit the SED from the X-ray to the radio bands, while `sed3fit` is limited to the UV-to-mm wavelength range. As we mentioned in chapter 2, one of the challenges when performing SED-fitting of type 2 AGN with a code like `sed3fit` is to separate the AGN and host-galaxy emission. In fact, for obscured AGN, the UV-optical part of the SED is dominated by the galaxy luminosity, and the AGN emission is detectable only in the few photometric points in the IR bands. However, at these wavelengths there is also the galaxy emission, thus, the separation of the two components is not trivial. Having additional photometric data in the X-ray and radio bands helps significantly in constraining the AGN emission. In fact, for the redshift range we are studying, the AGN component is the only source of emission in the X-rays and the main source in the radio band (our SED-fitting models take into account the SF-related radio emission, but we found that is usually sub-dominant with respect to the AGN one).

4.3.1 CIGALE

The Code Investigating GALaxy Emission (`CIGALE`) has been developed to study the evolution of galaxies by comparing modeled galaxy spectral energy distributions to observed ones. The code was presented in Burgarella et al. (2005); Boquien et al. (2019) and improved with the addition of the X-ray and radio modules in Yang et al. (2020, 2022). In this thesis we use `CIGALE` Version 2022.0.

In `CIGALE`, the SED models are built through a series of "modules" defined by the user. This architecture allows to easily customize and tailor the code for our own specific use. For example, the X-ray or radio modules can be added only when there are X-ray and radio data, thus saving up computational time (or allowing more complex stellar models) when there are not. Similarly to `sed3fit`, the code makes use of the energy balance between the energy absorbed by the dust in the UV-optical bands and the dust emission in the IR.

For the SED-fitting of the C IV sample, we use a delayed star formation history (SFH) with optional exponential burst, the Bruzual & Charlot (2003) population synthesis model with a Chabrier (2003) initial mass function, nebular emission lines, a Calzetti et al. (2000) dust attenuation law, and the Draine et al. (2014) dust models. The AGN emission is implemented via the SKIRTOR models (Stalevski et al. 2012, 2016), in which the torus is modeled as a clumpy two-phases medium (we refer to Yang et al. 2020, for further details). We also makes use of the X-ray module, which takes into account the emission in the X-ray from the AGN and from the X-ray binaries of the host-galaxy. Finally, as we had at our disposal VLA photometric points at 1.4 GHz, we also used the radio module. In Table 4.3, we report the modules and parameter used for the SED-fitting of the C IV sample.

Table 4.3: Parameter space for the `CIGALE` SED fitting. For each parameter, N_{sample} values are simulated in the *values* range.

	N_{sample}	values	description
SFH		sfhdelayed	delayed SFH with optional exponential burst
τ_{main}	4	100, 500, 1000, 5000	e-folding time of the main stellar population model in Myr.
age_{main}	6	300, 500, 750, 1000, 3000, 4000	Age of the main stellar population in the galaxy in Myr.
IMF		Chabrier	Initial Mass Function
Z	2	0.004, 0.02	Metallicity
Z_{gas}	2	0.004, 0.02	Nebular component: gas metallicity
dust attenuation		Calzetti 2000	
$E(B\text{-}V)_{\text{lines}}$	4	0.05,0.1,0.5,0.9	Color excess of the nebular line light.
Dust emission		Draine+2014	
AGN		Skirtor16	
$\tau_{9.7\mu m}$	4	3, 5, , 11	Average edge-on optical depth at 9.7 μm.
i	5	10°, 30°, 50°,70°, 90°	Viewing angle (w.r.t. the AGN axis).
f_{AGN}	7	0.01, 0.05, 0.1, 0.3, 0.5, 0.7, 0.9	$8 - 1000\,\mu$m AGN fraction.
X-ray			
α_{ox}	7	−1.9, −1.7, −1.5, −1.4, −1.3, −1.2, −1.1	$\alpha_{\text{ox}} = 0.3838 \times \log(L_{\nu,2\text{keV}}/L_{\nu,2500\text{Å}}$
radio			
q_{IR}	2	2.5, 2.6	FIR/radio correlation coefficient for SF.
R_{AGN}	8	0.001, 0.01, 0.1, 1, 5, 10, 50, 100	AGN radio-loudness $R_{\text{AGN}} = L_{\nu,5\text{GHz}}/L_{\nu,2500\text{Å}}$ at $i = 30°$.

Notes. For each module, we report only the parameters that we changed with respect to the default values. We refer to Boquien et al. (2019); Yang et al. (2022) for the complete list.

We performed several run of SED-fitting, carefully inspecting the resulting SED after each run. In particular, we checked firstly if the best-fit models chosen by `CIGALE` was a good fit at all wavelengths (e.g., if the IR photometric points were fitted as well as the optical ones); secondly, if the posterior parameter space was well sampled. For the latter, we run additional SED-fitting enlarging the prior parameter space, when the the posterior parameter distribution was suggesting that we need a larger parameter space (i.e., when it was skewed towards one of the previous prior limit). The problem of having a non equally good fit depending on the wavelengths was more challenging to resolve. In fact, as we can see from Fig. 4.4 *left panel*, we encounter many difficulties in having a good fit for the IR bands. This problem was not related to the modules we used or to the investigated parameter space, as we run several SED-fitting with different parameters and the results were not improving. We concluded that this problem arises from having a significantly larger number of photometric observations in the UV-optical part of the spectrum with respect to the IR bands. This problem was also worsen by the fact that the uncertainties associated with the UV-optical fluxes were smaller than those associated with the IR fluxes, thus, when computing the chi-square, the UV-optical bands had significant higher weights. Finally, we also had three sources for which the best-fit SEDs were obliviously incorrect, with reduced χ^2 as high as 30, and only one photometric point fitted. As we will see, when we managed to have a equally good fit of the SED at all wavelengths, we also obtained decent fit for these three sources as well.

The solution was to enlarge the uncertainties of the UV-optical bands. This led to a lower weight of these bands, and it is a technique not unheard of (e.g., Gruppioni et al. 2008), as it is well known that the uncertainties of the photometric fluxes tend to be underestimated. We proceeded as follow: we run a SED-fitting with the best parameter configuration we found and checked the χ distribution of the flux for each filter, where $\chi = (F_{\rm obs} - F_{\rm model})/eF_{\rm obs}$, with $F_{\rm obs}$ the observed flux, $eF_{\rm obs}$ its uncertainty, and $F_{\rm model}$ the best-fit predicted flux. We fitted the χ distribution with a Gaussian model; in case of a perfectly good fit with just random uncertainties, we would expect the distribution to be a Gaussian centered on 0 and with a dispersion $\sigma = 1$. Instead, we found that all the filters (with the exception of the X-ray and radio ones) had larger distributions (in the $1.5 - 3\,\sigma$ range, see Fig. 4.5). We refer to Appendix A.1 for the χ flux distributions of all the filters. We proceeded by multiplying each UV-optical filter uncertainties by the σ of their flux distribution, then we run again the SED-fitting. The results justifies our procedures: overall, almost all sources showed better fitting (see Fig. 4.4 *right panel*), without, or with reduced, gap between the best-fit models and the IR observed fluxes. Moreover, we were also able to obtain good fit for the three sources for which we were previously unable.

4.3.2 X-ray fluxes

As for the current version, `CIGALE` requires that the input X-ray fluxes are intrinsic. This means that the fluxes by definition should already be corrected for the instrumental response and for any obscuration of the AGN itself. To this goal, we used the intrinsic $2-10$ keV (rest-frame) luminosity that we obtained from the X-ray spectral analysis (section 4.2) to compute the $0.5 - 2$ and $2 - 10$ keV (observed-frame) intrinsic fluxes. We used the same $\Gamma = 1.8$ assumption as in the X-ray spectral analysis. Finally, we used box-

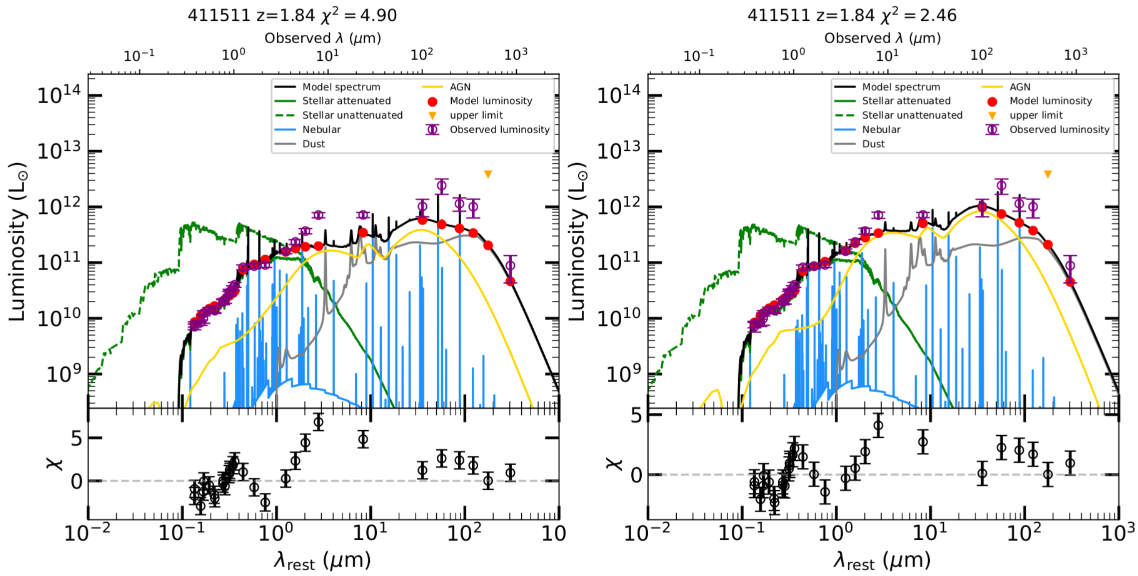

Figure 4.4: Comparison of one of the SED best-fit model obtained with (*right panel*) and without (*left panel*) enlarging the UV-optical flux uncertainties. We do not show the X-ray and radio part of the spectrum to better show the effects of the larger uncertainties on the IR part of the spectrum.

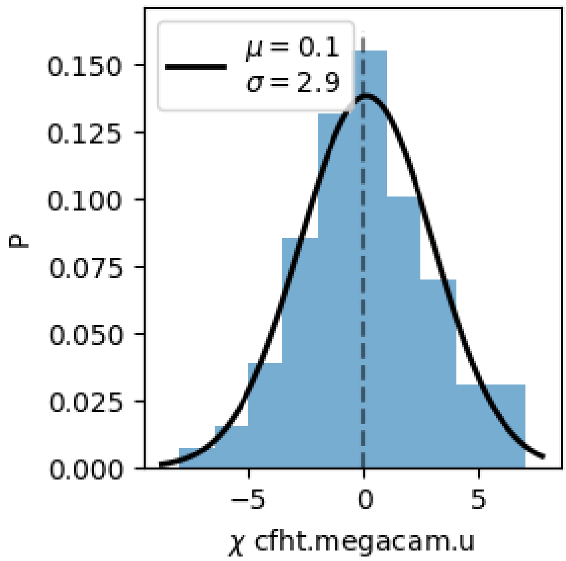

Figure 4.5: Flux χ distribution for the *MEGACAM u* filter before enlarging the flux uncertainties. $\chi = (F_{\rm obs} - F_{\rm model})/eF_{\rm obs}$, with $F_{\rm obs}$ the observed flux, $eF_{\rm obs}$ its uncertainty, and $F_{\rm model}$ the best-fit predicted flux. Fitting the χ distribution with a Gaussian model (*black line*) revealed a σ larger than expected for a random uncertainty. We ascribed this to an underestimation of the flux uncertainties. Enlarging the flux uncertainties by a factor equal to the dispersion σ of the χ distribution proved to improve the SED fitting.

shaped filters for the two filter responses (a fair assumption as we have already corrected for the instrumental response, see above).

4.3.3 The control sample

Similarly to what we did for the [Ne V] sample in section 3.5, we built a control sample of non-active galaxies. We extracted the control sample from the same zCOSMOS deep survey from which the C IV sample comes, and matched the sources in stellar mass and redshift.

In particular, we had at our disposal 9523 zCOSMOS sources, we took only the 6645 with a sky separation between the zCOSMOS catalogue and the COSMOS2020 catalog < 0.3". We also rejected from the COSMOS2020 catalogue all the sources with flags (meaning that some of the COSMOS2020 fluxes were unreliable) and that were not classified as "galaxies" by the `LePhare` code (Arnouts et al. 2002; Ilbert et al. 2006), which was used in the catalog to estimate the photometric redshift and provide a quick first estimation of the galaxy properties. We also excluded the sources with a detection in *Chandra* COSMOS Legacy, as at these redshifts, a Chandra detections implies luminosity that are mostly associated with AGN (i.e., $L_{2-10} > 10^{42}\,\mathrm{erg\,s^{-1}}$). We limited our sample to the sources in the 1.3 – 3.3 redshift range. In the end, the sample was composed of 4131 sources. We proceeded by matching this sample in redshift and stellar mass with the C IV AGN. For the redshift, we used the spectroscopic redshift of zCOSMOS, while for the stellar mass, the M_* from `LePhare` in the COSMOS2020 catalogue (both for the control sample and the C IV sample). For each of the C IV sources, we randomly selected 8 sources with $z = z_{\mathrm{C\,IV}} \pm 0.05$ and $\log(M_*/\mathrm{M}_\odot) = \log(\mathrm{M}_{*,\mathrm{C\,IV}}/\mathrm{M}_\odot) \pm 0.1$, where $z_{\mathrm{C\,IV}}$ and $M_{*,\mathrm{C\,IV}}$ are, respectively, the redshift and the stellar mass of the C IV source. If there were not 8 sources in the bins, we proceeded by iteratively enlarging them (both in z and M_*), until 8 sources were found. The matched sample was so composed of 688 non-active galaxies, with a mean (maximum) bin widths of 0.11 (0.85) for the redshift and 0.12 (0.4) dex for the stellar mass.

We performed SED-fitting for these sources, using the same filters and a similar `CIGALE` configuration of the C IV sample (section 4.3.1). As we did not have any X-ray detected sources in our sample, we did not use the `x-ray` module. Unfortunately, the M_* we obtained from the SED-fitting, were not always in agreement with the ones from `LePhare` in the COSMOS2020 catalogue (the median $\Delta M_* = \log(M_*^{\mathrm{CIGALE}}/\mathrm{M}_\odot) - \log(M_*^{\mathrm{LePhare}}/\mathrm{M}_\odot)$ was $\Delta M_* = -0.04^{+0.23}_{-0.17}$ with a maximum of $\Delta M_* = 1.16$, see Fig. 4.6). This led to the `CIGALE` M_* distributions for the C IV sample and control sample not being similar. Therefore, we matched again the two samples, using, this time, the stellar mass obtained from the SED-fitting, and choosing only 4 sources in the same z and M_* bins (following the same procedure discussed above). The final control sample was composed of 344 mass- and redshift-matched non-active galaxies. We show in Fig. 4.7 the comparison of redshift and stellar mass distributions for the C IV and the control samples.

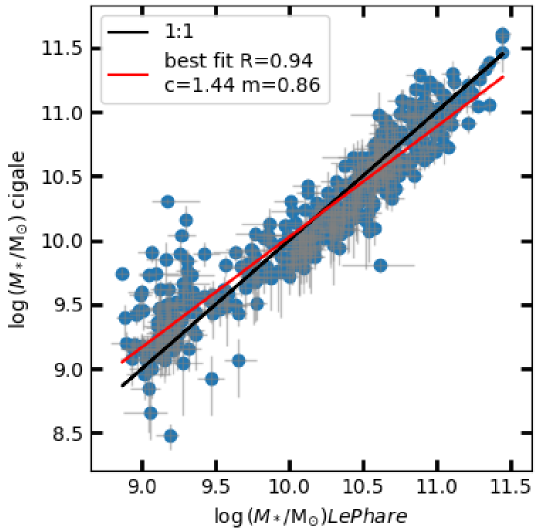

Figure 4.6: Comparison of the stellar mass of the control sample obtained from the `CIGALE` SED-fitting with the one from the COSMOS2020 catalogue computed via the `LePhare` code. The *black line* is the 1:1 relation, while the *red line* is the best fit line (via chi-square minimization) with a slope of $m = 0.86$ and an intercept of $c = 1.44$ in $\log(M_*^{\rm CIGALE}/{\rm M}_\odot)$ scale. Although the overall mass values are in agreement, the small differences lead to the stellar mass distributions of the C IV and control sample to be different.

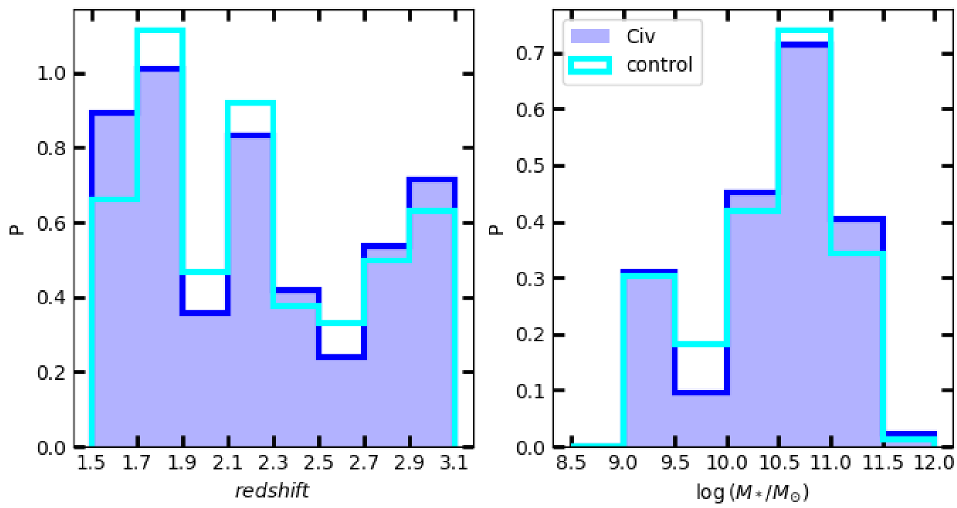

Figure 4.7: Comparison of redshift (*left*) and stellar mass (*right*) probability density distributions of the C IV (*blue*) and control (*cyan*) samples.

4.4 Comparison C IV and control samples

In this section, we discuss the similarities and differences between the C IV sample and the control sample introduced in section 4.3.3. We summarize the main properties of the two samples in Table 4.4.

Table 4.4: Summary of the properties (median and 16th – 84th percentile) for the C IV-selected AGN, for their control sample of non-active galaxies, and for the [Ne V]-selected AGN of chapter 3.

	C IV	control sample	[Ne V]
z	$2.17^{+0.60}_{-0.49}$	$2.17^{+0.65}_{-0.47}$	$0.88^{+0.10}_{-0.17}$
$\log(L_{2-10,\mathrm{intr}} / \mathrm{erg\,s^{-1}})$		$44.1^{+0.4}_{-0.5}$	$43.6^{+0.4}_{-0.6}$
$\log(N_H / \mathrm{cm^{-2}})$		$23.4^{+0.4}_{-0.6}$	$22.6^{+0.7}_{-0.9}$
$\log(L_{\mathrm{bolo}}/\mathrm{erg\,s^{-1}})$	45.2 ± 0.8		$44.6^{+0.6}_{-0.8}$
$\log(M_{\mathrm{BH}}/\mathrm{M_\odot})$	$7.2^{+0.6}_{-0.7}$		$7.5^{+0.4}_{-0.7}$
$\log \lambda_{\mathrm{Edd}}$	$-0.1^{+0.9}_{-0.8}$		$-0.9^{+0.6}_{-0.8}$
$\log(M_*/\mathrm{M_\odot})$	$10.7^{+0.4}_{-1.0}$	$10.6^{+0.5}_{-1.1}$	$10.9^{+0.3}_{-0.5}$
$SFR_{100\mathrm{Myr}}/\mathrm{M_\odot\,yr^{-1}}$	30^{+54}_{-24}	35^{+55}_{-28}	13^{+30}_{-9}
$\log(\mathrm{SSFR/yr^{-1}})$	$-9.0^{+0.6}_{-0.5}$	$-8.9^{+0.4}_{-0.7}$	$-9.7^{+0.5}_{-0.7}$
$\log(\mathrm{SFR/SFR_{MS}})$	-0.2 ± 0.5	$-0.1^{+0.4}_{-0.6}$	$-0.3^{+0.4}_{-0.7}$
f_{MS}	$0.39^{+0.11}_{-0.10}$	$0.46^{+0.07}_{-0.17}$	$0.43^{+0.15}_{-0.09}$

Notes. $L_{2-10,\mathrm{intr}}$ and N_H are the intrinsic (i.e., absorption corrected) 2 – 10 keV luminosity and the amount of obscuration, obtained from a power-law model (corrected for the Galactic absorption) with fixed photon index $\Gamma = 1.8$ plus an absorption component. L_{bolo} is the AGN bolometric luminosity (from the SED-fitting), M_{BH} is the BH mass, obtained from the host-galaxy stellar mass via the M_*-M_{BH} Suh et al. (2020) relation. λ_{Edd} is the Eddington ratio, with $\lambda_{\mathrm{Edd}} = L_{\mathrm{bolo}}/L_{\mathrm{Edd}}$, where L_{Edd} is the Eddington luminosity. M_* and SFR refer to the host-galaxy stellar mass and star formation rate, respectively. $SSFR$ is the specific SFR ($SSFR$= SFR/M_*). f_{MS} is to the fraction of sources within the 1 σ dispersion of the Schreiber et al. (2015) MS.

4.4.1 SFR-M$_*$ relation

For the analysis of the SFR, we used for both samples the 100 Myr-averaged SFR from `CIGALE`, which is obtained from the best-fit SFH. The distributions of the SFR are not significantly different, with median $SFR_{\mathrm{C\,IV}}/\mathrm{M_\odot\,yr^{-1}} = 10.6^{+0.4}_{-1.0}$ and $SFR_{contr}/\mathrm{M_\odot\,yr^{-1}} = 10.6^{+0.5}_{-1.0}$, and a KS test probability of P=0.82.

In Fig. 4.8, we show the position of the C IV AGN and of the control sample with respect to the SF Main Sequence at $z = 2.2$ (i.e., the median redshift of the C IV sample). As we can see, their position with respect to the MS seem to show a small trend with the stellar mass. High-mass ($M_* > 10^{11}$ M$_\odot$) sources are preferentially below the MS, while those with $\log(M_*/\mathrm{M_\odot}) < 9.5$ tend to be above the MS. While we are planning to conduct

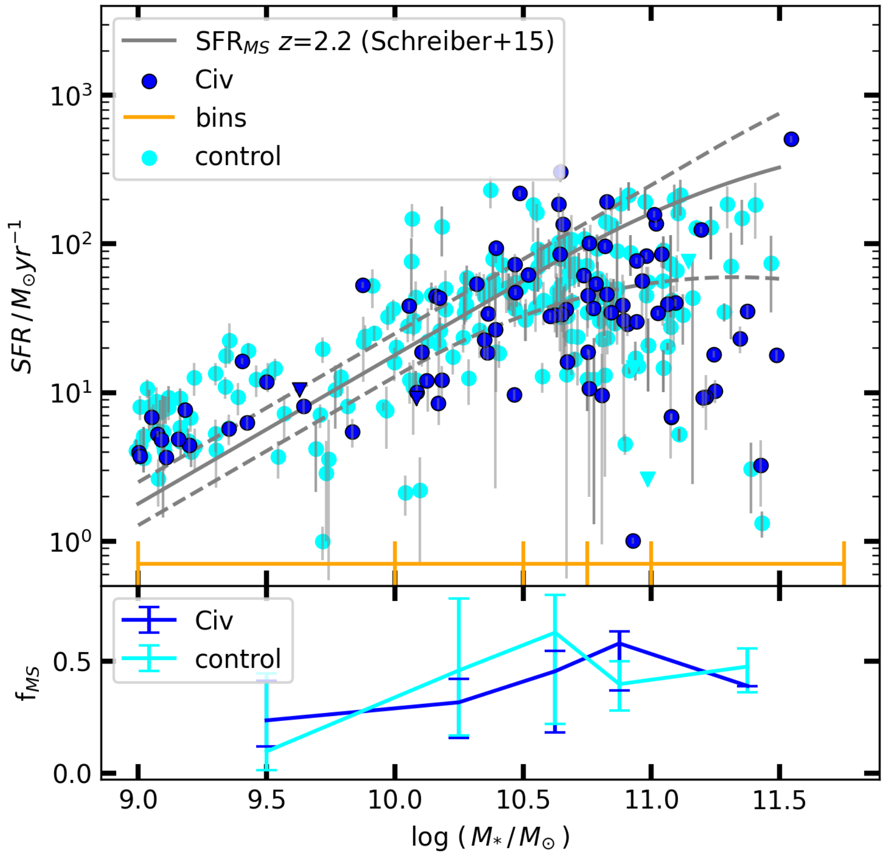

Figure 4.8: Comparison of the positions of the C IV (*blue*) and stellar mass- and redshift-matched control (*red*) samples in the SFR−M$_*$ plane. *Upper panel*: The grey solid line is the Schreiber et al. (2015) MS, the grey dashed lines its 1 σ dispersion. The *triangles* indicate SFR upper limits. In *orange* the adopted stellar mass binning. *Bottom panel*: Fraction of sources within the 1 σ dispersion of the MS.

further investigation on the sources with $\log(M_*/\mathrm{M}_\odot) < 9.5$, we are confident about the results of the SED-fitting for the high-mass sources; we note that a higher number of high-mass sources below the MS is expected by the downsizing paradigm (i.e., high-mass galaxies are formed earlier and in shorter times; e.g., Cowie et al. 1996; Thomas et al. 2005; Cimatti et al. 2006). Comparing the position of the C IV and control samples, we did not find any significant difference. The fraction of sources within the MS (f_{MS}) is the same for the two samples in all the stellar mass bin (Fig. 4.8 *bottom panel*).

4.4.2 SSFR and MS ratio

We further investigate the similarities between the two samples by analyzing the SSFR and the MS ratio. The latter is defined as SFR/SFR$_{\mathrm{MS}}$ (where SFR$_{\mathrm{MS}}$ is the expected SFR for a MS object given its stellar mass) and is a useful tool to indicate the position on the SFR-M$_*$ with respect to the MS. As we can see from Fig. 4.9, the C IV

AGN have, on average, slightly lower SSFR, although the distributions are not significantly different, with median $\log(SSFR_{\rm C\,IV}/{\rm yr}^{-1}) = -9.0^{+0.6}_{-0.5}$ for the C IV sample and $\log(SSFR_{\rm contr}/{\rm yr}^{-1}) = -8.9^{+0.4}_{-0.7}$ for the control sample. The MS ratio shows the same trend, with the control sample having more sources above the MS than the C IV sample. In both cases, the differences between the two distributions are not significant according to the KS test.

Another feature of the SSFR distribution is the absence of a bimodality. In fact, in both samples the number of objects at $\log(SSFR_{\rm C\,IV}/{\rm yr}^{-1}) < -9.5$ rapidly decreases, with no indication of the population of quiescent galaxies that we found at lower redshift (section 3.5 and Fig. 3.12).

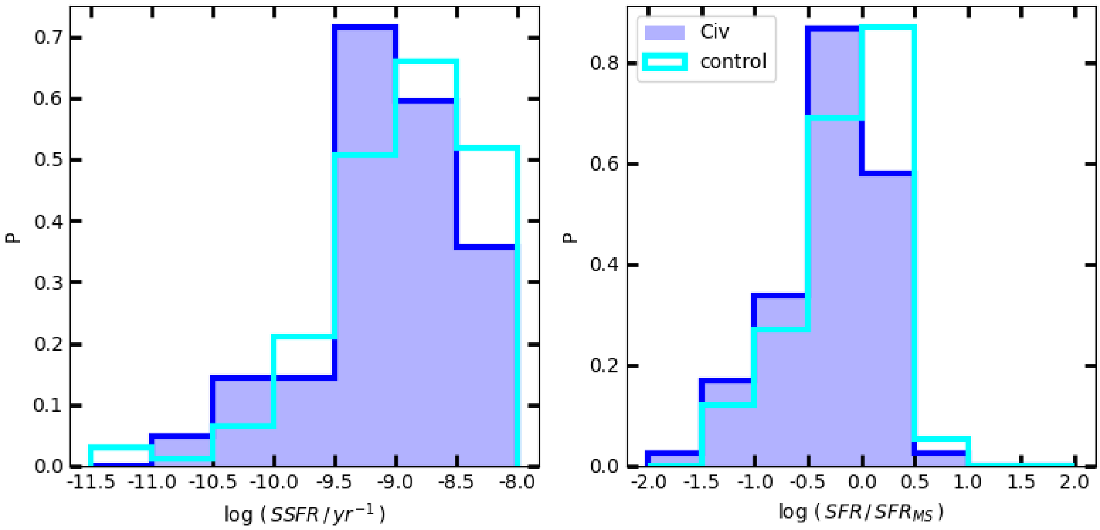

Figure 4.9: Comparison of the specific SFR (*left*) and MS ratio (*right*) probability density distributions of the C IV (*blue*) and control (*cyan*) samples. The specific SFR is defined as SSFR= SFR/M$_*$, the MS ratio is SFR/SFR$_{\rm MS}$, where SFR$_{\rm MS}$ is the SFR predicted by the Schreiber et al. (2015) MS given the galaxy stellar mass.

4.5 Comparison C IV and [Ne V] samples

In this section, we compare the results we obtained for the C IV- and the [Ne V]-selected AGN. We summarize the properties of the two samples in Table 4.4.

As the SEDs of the C IV and [Ne V] samples were fitted with different SED-fitting code (`CIGALE` and `sed3fit`, respectively), we performed additional tests to check whether the use of different routines were affecting the results. We re-analyzed the [Ne V] sample using `CIGALE`, and compared the results with those of sec. 3.4.2. In particular, we run two different configurations for `CIGALE`, one with the exact same parameters used for the fitting of the C IV sources, and one in which we used the same AGN torus models available to sed3fit (i.e., the Feltre et al. 2012, torus models instead of the newest SKIRTOR models, see sec. 4.3.1).

We did not found any significant difference in the SED-fitting results between the two torus models. This is likely due to the limited amount of mid- and far-IR photometric points at our disposal and to the use of broad-band filters, while the main differences between the models may possibly emerge when higher spectral resolution data are used. Regarding the comparison between CIGALE and sed3fit, as we can see from fig. 4.10, we did not find any significant bias either in stellar mass or AGN bolometric luminosity. The stellar mass from sed3fit seems to prefer slightly higher values for $\log(M_*/M_\odot) > 10.5$, but this is an inheritance of the magphys code well attested in the literature. Although with larger dispersion with respect to the stellar mass, we found that for most of the sources, the bolometric luminosity from CIGALE and sed3fit are compatible within their uncertainties, with no systematic bias. The few sources that differ more than 1 σ are planned to be accurately analyzed. Finally, the comparison of the SFR between the two codes is still under investigation.

Overall, we concluded that, despite the use of different SED-fitting code, we can compare the properties of the [Ne v] and C IV samples, as no systematic effect has been found so far, with in-depth comparisons between the SED-fitting codes planned for the future.

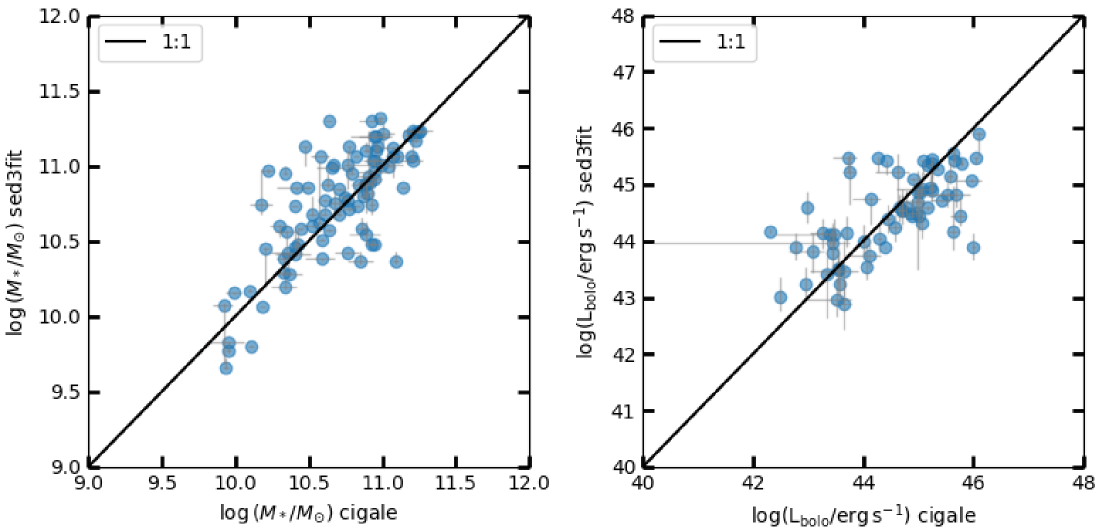

Figure 4.10: Comparison of the stellar mass (*left*) and AGN bolometric luminosity (*right*) between the CIGALE and sed3fit SED-fitting code, for the [Ne v] sample.

4.5.1 Stellar mass

In Fig. 4.11, we show the comparison of the stellar mass and SFR for the C IV and [Ne v] samples. As we can see, the majority of the sources in both samples covers the $10^{10-11.5}$ $M_\odot$ stellar mass range. However, ~ 10 % of the AGN in the C IV samples has

$\log(M_*/M_\odot) < 9.5$. At the moment, we are not able to exclude that these values could be related to the use of `CIGALE` instead of `sed3fit`. Therefore, we refrain from drawing any definitive conclusion on these sources. The median stellar mass of the whole C IV sample is $\log(M_{*,\rm C\,IV}/M_\odot) = 10.7^{+0.4}_{-1.0}$, while excluding the "uncertain" sources with $\log(M_*/M_\odot) < 9.5$ is $\log(M_{*,\rm C\,IV}/M_\odot) = 10.8^{+0.3}_{-0.4}$. The median M_* for the [Ne v] sample is $\log(M_{*,\rm C\,IV}/M_\odot) = 10.9^{+0.3}_{-0.5}$.

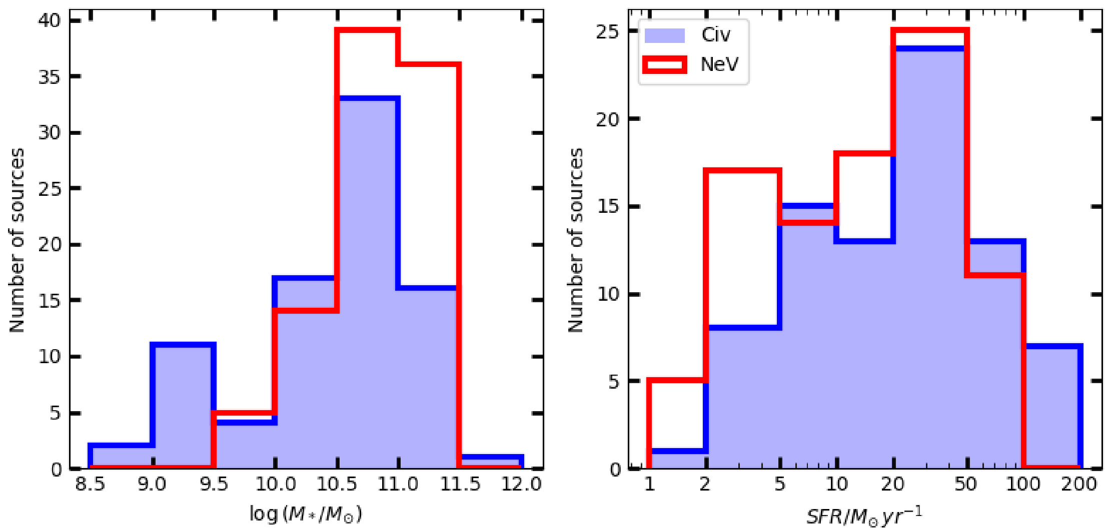

Figure 4.11: Comparison of the stellar mass (*left*) and SFR (*right*) distributions of the C IV (*blue*) and [Ne v] (*red*) samples.

4.5.2 SFR

We compared in Fig. 4.11 the 100 Myr-averaged SFR. Although the two samples spawn almost the same SFR range (with the exception of few C IV sources with SFR> 100 M$_\odot$ yr^{-1}), the C IV AGN show, on average, a higher SFR with respect the [Ne v] ones. This is reflected in the median (and 16$^{\rm th}$ – 84$^{\rm th}$ percentile) SFR: $SFR_{\rm C\,IV}/M_\odot\,{\rm yr}^{-1} = 30^{+54}_{-24}$ for the C IV sample and $SFR_{[Ne\,v]}/M_\odot\,{\rm yr}^{-1} = 13^{+30}_{-9}$ for the [Ne v] one.

4.5.3 SFR-M$_*$ relation

In Fig. 4.12, we show the position of the C IV-selected AGN and the [Ne v]-selected ones in the M$_*$-SFR plane. If we exclude the C IV sources with $\log(M_*/M_\odot) < 9.5$ (see section 4.5.1), the main difference between the two samples is in the highest stellar mass bin. While the [Ne v] sources have $f_{\rm MS} \sim 0.65$ at $\log(M_*/M_\odot) > 11$ (and a comparable fraction in the $10.75 < \log(M_*/M_\odot) < 11$ stellar mass bin), the fraction drops to $f_{\rm MS} \sim 0.35$ for the C IV sample.

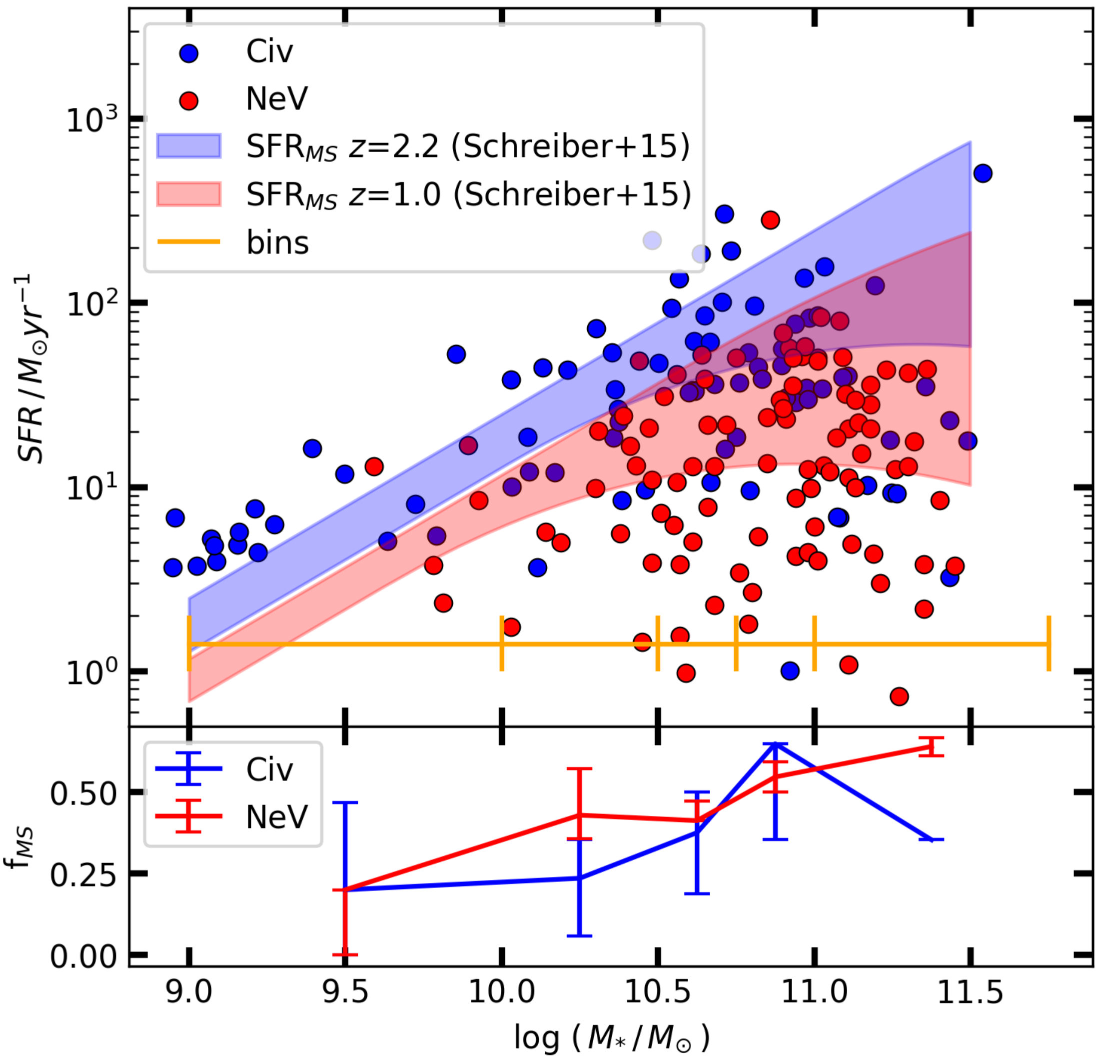

Figure 4.12: Comparison of the positions of the C IV-selected AGN (*blue*) and [Ne V]-selected ones (*cyan*) in the SFR–M$_*$ plane. *Upper panel*: The *blue area* is the 1σ dispersion of the Schreiber et al. (2015) MS at z = 2.2 (the median z of the C IV sample), while the *red area* is the one for the MS at z = 1 (the median z of the [Ne V] sample). In *orange* the adopted stellar mass binning. *Bottom panel*: Fraction of sources within the 1σ dispersion of the MS.

4.5.4 SSFR and MS ratio

Similarly to the comparison between the C IV and the control sample, we further investigate the position with respect to the MS of the C IV- and [Ne V]- selected AGN by analyzing the SSFR and the MS ratio. As we can see from Fig. 4.13, while their M_* and SFR are somewhat similar, their SSFR distributions are extremely different. The median SSFR are $\log(SSFR_{\mathrm{C\,IV}}/\mathrm{yr}^{-1}) = -9.0^{+0.6}_{-0.5}$ for the C IV sample and $\log(SSFR_{\mathrm{contr}}/\mathrm{yr}^{-1}) = -9.7^{+0.5}_{-0.7}$ for the [Ne V] sample. The obvious different distributions are confirmed by the $P < 10^{-11}$ of the KS-test.

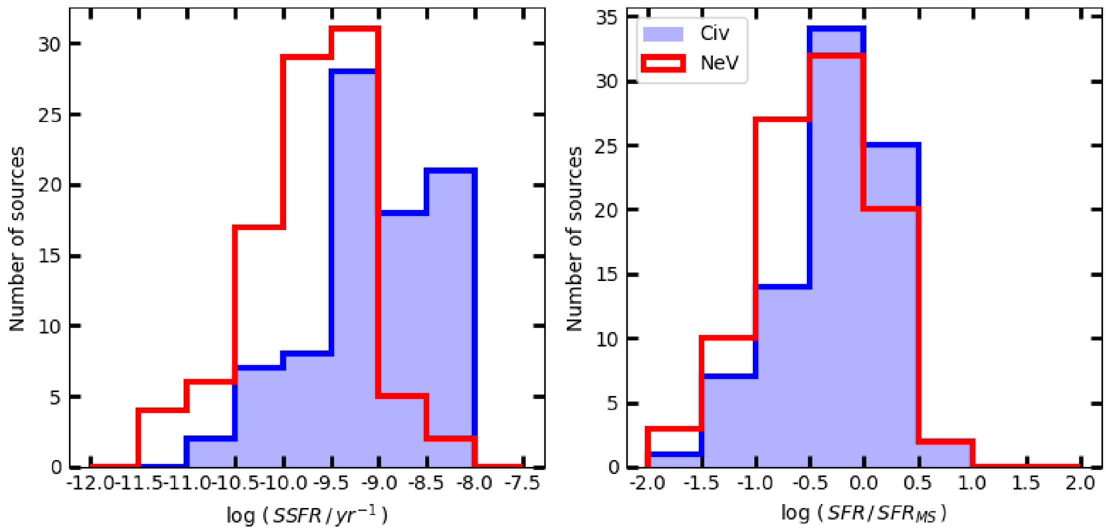

Figure 4.13: Comparison of the specific SFR (*left*) and MS ratio (*right*) distributions of the C IV (*blue*) and [Ne V] (*red*) samples.

The difference in SSFR, however, is expected due to the redshift evolution. This is easy to see by comparing the $z = 1$ and $z = 2.2$ Schreiber et al. (2015) MS: the *red* $z = 2.2$ MS is higher than the *blue* one, thus, for a given stellar mass, the expected SFR is higher and so the SSFR must also be higher. In this case, the effect of the redshift evolution can be avoided by considering the MS ratio. Similarly to the M_* and SFR, the MS ratio distributions of the two sample cover the same range. However, it emerges that C IV AGN occupy, on average, a locus higher than the [Ne V] AGN with respect to the MS. The median MS ratios are -0.2 ± 0.5 for the C IV-selected AGN and $-0.3^{+0.4}_{-0.7}$ for the [Ne V]-selected ones. This seems to indicate that, although the galaxies of both samples are, on average, slightly below the MS, the [Ne V] sample has more sources well below the MS sequence (i.e., with a SFR < 0.1 SFR$_{\rm MS}$).

4.5.5 AGN bolometric luminosity

The use of SED-fitting allows us to separate the host-galaxy and AGN emission and to obtain the AGN properties alongside those of the galaxy. As such, we are able to obtain the AGN luminosity also for the X-ray undetected sources. From this point onward, we excluded from the plots and from our consideration the C IV-selected AGN with $\log(M_*/\mathrm{M}_\odot) < 9.5$, as we are not entirely confident on the reliability of their fits (see sections 4.5.1 and 4.5.3) and as their values would affect significantly our results. In Fig. 4.14, we show the distributions of the AGN bolometric luminosity for the C IV-selected AGN and for the [Ne V]-selected ones. Similarly to what we found when analyzing the $2-10$ keV AGN luminosity from the X-ray spectra (Fig. 4.2), the C IV sample is composed of intrinsically more luminous AGN, reaching $L_{\rm bolo} > 10^{46}\,{\rm erg\,s^{-1}}$†. Taken together, the two samples uniformly cover the $43 \leq \log{(L_{\rm bolo}/{\rm erg\,s^{-1}})} < 46.5$ luminosity range.

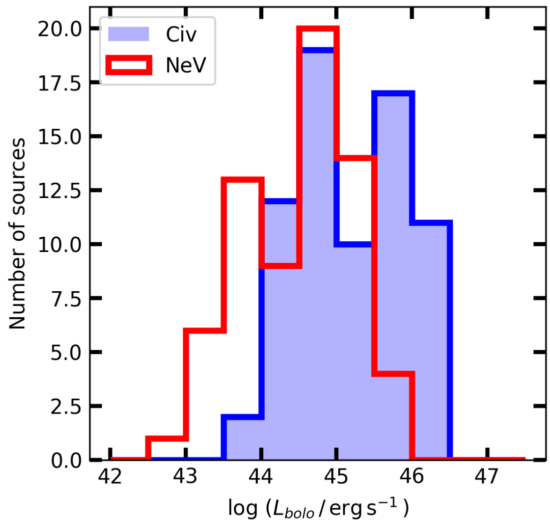

Figure 4.14: Comparison of the AGN bolometric luminosity distribution of the C IV (*blue*) and [Ne V] (*red*) samples.

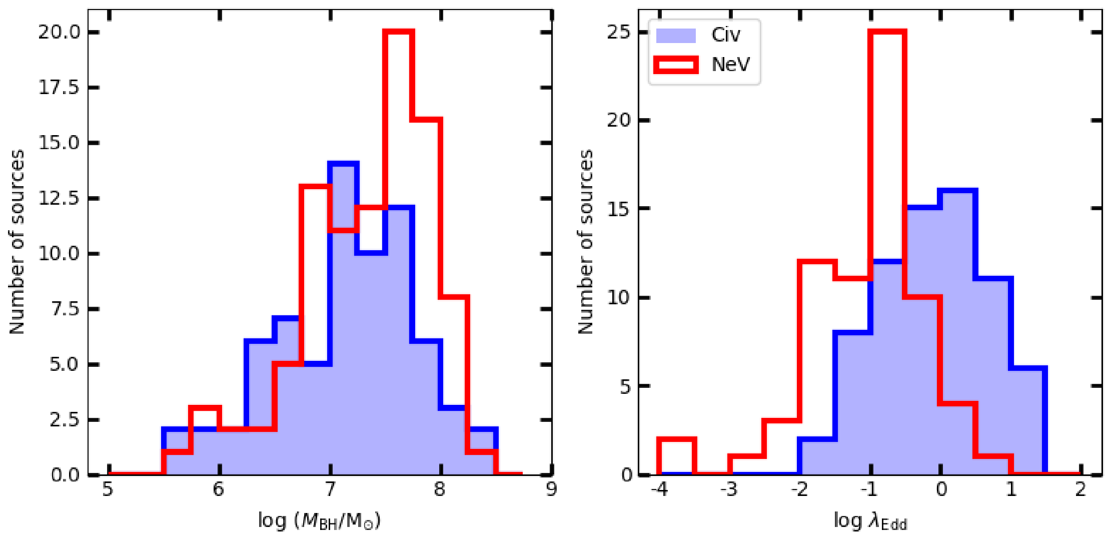

Figure 4.15: Comparison of the BH mass (*left*) and Eddington ratio (*right*) distributions of the C IV (*blue*) and [Ne V] (*red*) samples. The BH mass is computed from the stellar mass assuming the Suh et al. (2020) relation (see section 3.4.2). The Eddington ratio is defined as $\lambda_{\rm Edd} = L_{\rm bolo}/L_{\rm Edd}$, where $L_{\rm Edd}$ is the Eddington luminosity (see section 1.1).

†We plan to soon estimate the contribution of selection effects in the AGN luminosity and obscuration.

4.5.6 M_{BH} and Eddington ratio

Similarly to what we did in section 3.4.2, we used the Suh et al. (2020) relation to estimate the mass of the SMBH from the host-galaxy stellar mass. Therefore, the distributions of the $M_{\rm BH}$ follow those of the M_*, with the two samples covering the same $5.5-8.5$ (in $\log(M_{\rm BH}/{\rm M}_\odot)$) BH mass range (Fig. 4.15).

More interesting are the distributions of the Eddington ratio. Due to the C IV-selected AGN having similar stellar masses, but higher AGN bolometric luminosities than the [Ne V]-selected AGN, their $\lambda_{\rm Edd}$ are significantly higher. The median Eddington ratios are $\lambda_{\rm Edd} = -0.1^{+0.9}_{-0.8}$ for the C IV sample and $\lambda_{\rm Edd} = -0.9^{+0.6}_{-0.8}$ for the [Ne V] one. This means that 73 % of the C IV-selected AGN are accreting more than the usual 10 % of the $L_{\rm Edd}$, in fact, 40 % of them are accreting at a super-Eddington level ($\lambda_{\rm Edd} > 1$).

We investigated the position of our sources in the column density - Eddington ratio diagram (Fig. 4.16) of Ricci et al. (2017)(previously proposed by Fabian et al. 2006, 2008, 2009), who studied a sample of ~ 800 local (median $z = 0.037$) AGN selected with *Swift-BAT* in the $14-195$ keV range. They identify a locus in the $N_H - \lambda_{\rm Edd}$ plane called the *blownout region* (*green area*), where the AGN radiation pressure would be high-enough to push away the obscuring material. This region has been computed assuming that the material pushed by the AGN radiation is a combination of dust and gas (while the classical Eddington limit assumes pure ionized hydrogen). Briefly, the presence of dust makes easier for the radiation to sweep away the material. As only $\sim 1.4\%$ of the Ricci et al. (2017) sample resides in this region, they state that the radiative feedback from the AGN can efficiently clear out its surroundings in very short timescales. Remarkably, we found that $\sim 70\%$ of our X-ray detected C IV AGN lies in the *blownout region*. Although the uncertainties on the Eddington ratio are in the order of 1 dex, we found the same trend either using the bolometric luminosity derived from the SED-fitting or the one computed from the fitting of the X-ray spectra. Regarding the [Ne V]-selected sample, only $\sim 17\%$ of the AGN are in the *blownout region*.

The position of the [Ne V] sources in the $N_H - \lambda_{\rm Edd}$ plane seems in agreement with what we expect from the *in-situ* scenario: most of the [Ne V] AGN are obscured and accreting, but they have not reached their peak luminosity, and quenched the SF and their own accretion. Reading our results along with the Ricci et al. (2017), the [Ne V] sources in the *blownout region* should be expelling their surrounding material and soon move to lower N_H and $\lambda_{\rm Edd}$. The interpretation of the position of the C IV AGN is more complex, due to most of them being in the *blownout region*. One possible explanations for this high fraction are that our selection is extremely sensitive to AGN in this phase, *i.e.* at $z \sim 2-3$ these sources are more common than in the local universe. Alternatively, as suggested by Ballo et al. (2014), if the dust-to-gas ratio is lower than the typical of the ISM assumed by Fabian et al. (2009), the radiation pressure would have a harder time pushing away the material, and the *blownout region* would move to higher $\lambda_{\rm Edd}$. In any case, we must be cautious in the interpretation of these results, as the $\lambda_{\rm Edd}$ is heavily dependent on the BH

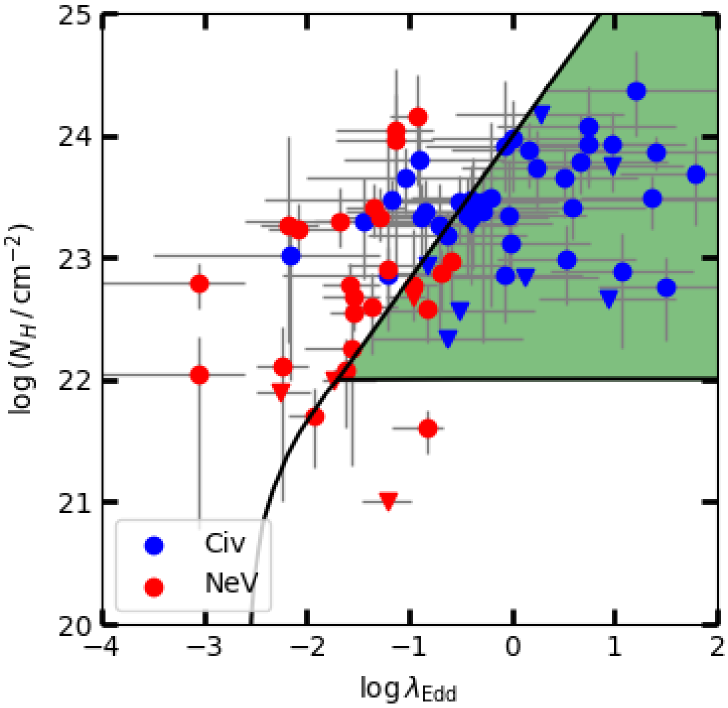

Figure 4.16: Comparison of the position of the C IV (*blue*) and [Ne V] (*red*) sources in the column density - Eddington ratio diagram of Ricci et al. (2017). The continuous line is the effective Eddington limit $\lambda_{\rm Edd}^{\rm eff}$ for different values of column density (see Ricci et al. 2017). The horizontal line at $N_H = 10^{22}\,{\rm cm}^{-2}$ delimits the region where the host-galaxy could contribute to the AGN observed obscuration. The *green area* represents the blowout region, *i.e.*, the locus where the AGN radiation pressure should push away the obscuring material.

mass, and those were computed assuming the Suh et al. (2020) relation. Unfortunately, our knowledge of $M_{\rm BH} - M_*$ relation is mostly local. It could be possible that the BH and the stellar mass grow in synchronous at high-redshift; alternatively, it could be that one of the two grows faster and the other one catches up at later times. For example, if the SMBH grows faster than its host-galaxy (as it seems the case for GS-14, the source we investigated in chapter 5, Übler et al. 2023), assuming the local relation leads to underestimate the BH mass and to overestimate the Eddington ratio.

4.6 Conclusion

In this chapter, we extended the study of obscured AGN to higher redshift. In particular, we exploited the C IV emission line to select 90 obscured AGN in the 1.5 − 3.1 redshift range. The combined analysis of the C IV-selected AGN and of the [Ne V]-selected ones allowed us to completely cover the cosmic noon, from $z \sim 0.5$ up to $z \sim 3$.

We performed an X-ray spectral analysis of the *Chandra* 52 X-ray detected sources to characterize the AGN intrinsic luminosity and amount of obscuration. We also exploited the UV-to-radio photometric coverage of the COSMOS field to obtain the SED of the entire C IV sample, thus further characterizing both the AGN and the host-galaxies. We compared our results with a mass- and redshift-matched control sample of non-active galaxies, as well as with the results we obtain from the analysis of the [Ne V]-selected AGN in chapter 3. Our main results are the following:

- The X-ray analysis confirmed that 92 % of the AGN in the C IV type 2 sample are obscured. In fact, the majority of them are extremely obscured ($N_{\rm H} > 10^{23}\,{\rm cm}^{-2}$). We also found 2 CT-AGN. The C IV sample results, on-average, almost 1 dex more obscured than the [Ne V] sample. Therefore, the C IV line resulted even more powerful than the [Ne V] line to select extremely obscured AGN, although, we can not differentiate yet if it is due to selection effects or to the redshift evolution.
- The majority of the C IV-selected AGN are in a strong episode of SMBH growth ($\lambda_{\rm Edd} \geq 0.1$), and 40 % of them are accreting at super-Eddington ratios.
- While the [Ne V] sample was mostly composed of AGN in the obscured accretion phase, with some galaxies showing the first effects of quenching, the C IV-selected sources are the most obscured and highly accreting AGN in the obscured accretion phase of the BH-galaxy co-evolution paradigm.
- The main difference between the C IV- and the [Ne V]-selected AGN emerged from the comparison with their control sample of non-active galaxies. The [Ne V]-control sample showed a population of quiescent galaxies that the [Ne V]-AGN seems to lack. However, both the C IV-selected AGN and their control sample do not have any significant quiescent populations, meaning that they are all still forming stars and that the quenching is most likely yet to take place.
- The lack of differences between the C IV-selected AGN and the non-active galaxies seems to suggest not only that these AGN are in the early stage of the obscured accretion phase, when the effects of the AGN feedback have not significantly influenced the host-galaxy yet, but also that in the stellar mass range we investigated, the feedback effects will come later in the galaxy evolution.

Chapter 5

Double stellar population and AGN activity in a galaxy at $z \sim 5.5$

5.1 Introduction

Galaxies are found to follow the star-forming main-sequence (MS), a tight correlation between the stellar mass and the SFR up to very high-z ($z \approx 5-6$, e.g. Speagle et al. 2014; Noeske et al. 2007). What are the physical processes that give rise to this tight sequence? How do physical process as feedback (from SNe, AGN) and feeding of pristine gas influence and self-regulate each other? One of the key components is the cold gas. Cold molecular gas is first ingredient for the SF, as its low temperature allows the gas to reach high densities and to form protostellar clouds, kickstarting the star formation chain. Moreover, the cold gas is also extremely important for the AGN activity. Cold streams of gas are able to funnel huge quantity of gas toward the galactic nucleus, where they will accrete into the SMBH and feed its growth. Finally, the cold gas is also tracer of the quenching phase of the AGN-galaxy co-evolution with molecular outflows, extending on kpc scales from the nucleus, being detected in some AGN hosts and ULIRGs (e.g., Feruglio et al. 2010; Cicone et al. 2014). In these systems the mass loss ranges from one to several times the SFR. Similarly, using various emission line diagnostics, neutral (H I , [C II]) and ionised ([O III], Hα) outflows were also detected (Nesvadba et al. 2008; Alexander et al. 2010; Harrison et al. 2018).

This cold gas is mostly composed of molecular hydrogen H_2; however, H_2 is very difficult to detect, as it has no emission line at temperature below 100 K. Luckily, along with the H_2, other elements are present in the cold molecular gas: Oxygen, Carbon, Silicates, that lead to a variety of different proxies, that we can exploit to trace the cold gas content. The CO rotational transition (Bolatto et al. 2013), the dust mass from IR SED-fitting (Leroy et al. 2011; Kaasinen et al. 2019) and the dust continuum emission in the sub-mm (Scoville et al. 2014) are the most commonly used methods to estimate the molecular gas

mass.
ALMA represents the most powerful telescope for observing the cool Universe, i.e. molecular gas and dust. Its extraordinary angular resolution down to sub-arcsec scales allows to spatially resolve the CO emission within the galaxies, while its sensitivity is crucial to investigate the gas content even in faint and distant sources.
In this chapter, we investigate a sub-sample of SF galaxies from the ALPINE survey with extreme values of molecular gas fraction ($f_{mol} < 0.1$ or $f_{mol} > 0.9$), with the goal of providing a characterization of the properties of these sources. Then, we focus on the source with the lowest f_{mol} of the entire sample, as we found it harbors a double stellar population and sign of hidden AGN activity.

5.1.1 ALPINE survey

The ALMA Large Program to INvestigate C$^+$ at Early times (ALPINE; Le Fèvre et al. 2020; Béthermin et al. 2020; Faisst et al. 2020) features 118 galaxies observed in the [C II] 158μm line and far infrared (FIR) continuum emission during the period of rapid mass assembly, right after the end of the HI reionization, at redshifts of $4 < z < 6$.

The ALPINE sample contains sources from both the cosmic evolution survey (COSMOS; Scoville et al. 2007) field and the Chandra deep field south (CDFS; Giacconi et al. 2002). The galaxies were selected on the basis of having a reliable spectroscopic redshift at $4.4 < z < 5.9$ (excluding 4.65 < z < 5.05, where [C II] falls in a low transmission atmospheric window). Galaxies are UV-selected with $L_{\rm UV} > 0.6\,L_*$ to include most of the star formation traced by the UV, and excluding type 1 AGN identified from broad spectral lines. The absolute UV luminosity cut ($M_{\rm UV} < -20.2$) is equivalent to SFR > 10 $\rm M_\odot\ yr^{-1}$. The redshift are obtained from extensive spectroscopic campaigns at the VLT (VUDS; Le Fèvre et al. 2015) and Keck (DEIMOS; Hasinger et al. 2018). The ALMA observations comprise 69 hr taken between 2018 May (Cycle 5) and 2019 February (Cycle 6). The galaxies have been observed in Band 7 at a spatial resolution of ~ 1.0“. The detection rates is 64% for the [C II] line and 19% for the continuum (at the 3.5 σ detection limit; Béthermin et al. 2020). We refer to Le Fèvre et al. (2020) for further details about the sample selection and the ALMA observing strategy, to Béthermin et al. (2020) for the ALMA data reduction and analysis, and to Faisst et al. (2020) for the ancillary data and multi wavelength studies.

5.2 The ALPINE f_{mol} outliers

As introduced in sec 5.1, we would like to investigate the huge spread in molecular gas fraction f_{mol} in the ALPINE sample. As we can see from Fig. 5.1, the f_{mol} values range from 0.1 up to almost 1. We focused on the ALPINE sources with [C II] 158μm detection (to have a reliable value of the molecular gas content) in the $5 < z < 6$ redshift range,

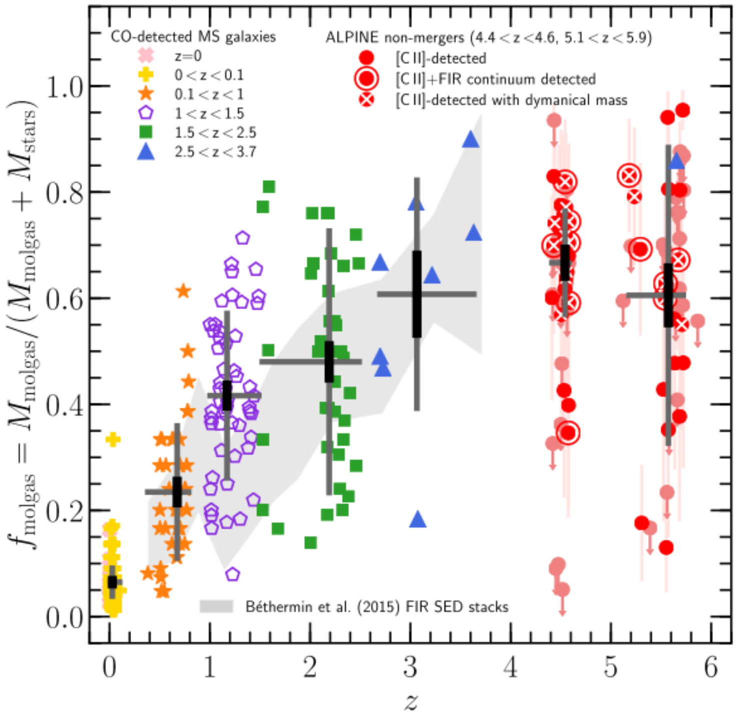

Figure 5.1: Molecular gas fraction as function of redshift from Dessauges-Zavadsky et al. (2020b). The ALPINE sources detected in [C II] are indicated with a *dark red circle*, while the *light red circles* are the [C II] undetected.

as these sources showed the most extreme molecular gas fractions. Our sample then comprises 32 sources. We excluded 10 sources which showed sign of on-going merger (Faisst et al. 2020). All the galaxy properties in this section were computed by Faisst et al. (2020), while the molecular gas mass and the molecular gas fraction by Dessauges-Zavadsky et al. (2020b). In particular, the molecular gas mass was computed exploiting the eq. 5.1 calibration of Zanella et al. (2018) between [C II] luminosity and M_{mol}, which seems to be valid regardless of metallicity, redshift (in the $0 \leq z \leq 6$ range), and MS or starburst nature of the galaxies. In Dessauges-Zavadsky et al. (2020b), they also computed the M_{mol} adopting different molecular gas tracer: they found compatible molecular gas masses, either using the relation between the IR luminosity and the CO(1 – 0) luminosity, or the $L_{850\,\mu\mathrm{m}} - M_{mol}$ relation of Kaasinen et al. (2019)(we refer to Dessauges-Zavadsky et al. 2020b, for further details).

$$\log \frac{L_{[C\,\textsc{ii}]}}{\mathrm{L}_\odot} = (-1.28 \pm 0.21) + (0.98 \pm 0.02) \log \frac{M_{mol}}{\mathrm{M}_\odot} \tag{5.1}$$

We identified 4 sources, two with $f_{mol} < 0.2$ (hereafter low f_{mol} outliers) and two with $f_{mol} > 0.9$ (hi f_{mol} outliers), and we focused our analysis on them. Fig. 5.2 shows the f_{mol} and redshift distribution for our ALPINE subsample and for the 4 outliers. Putting the sources in the $\mathrm{M}_* - \mathrm{SFR}$ plane (Fig. 5.3), it is easy to see that the outliers show extreme

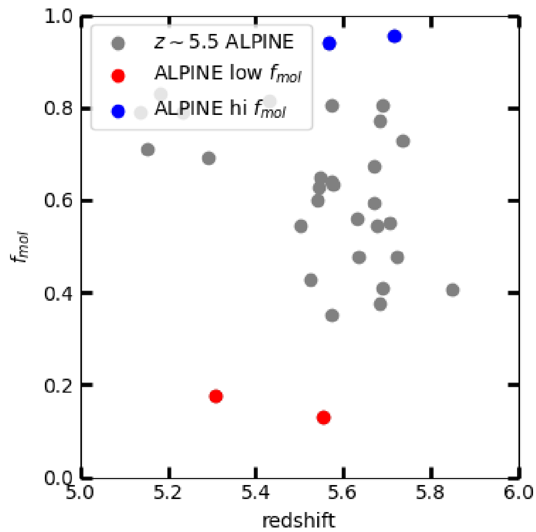

Figure 5.2: Molecular gas fraction as function of redshift for the ALPINE subsample we focused on.

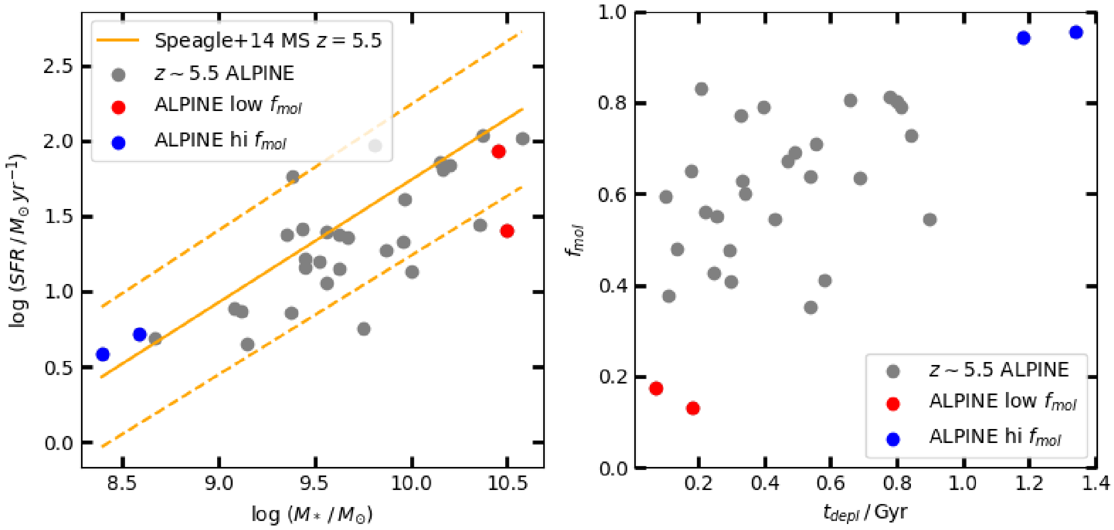

Figure 5.3: *Left:* $M_* - SFR$ plane for our ALPINE subsample. The *yellow solid line* is the Speagle et al. (2014) MS and its $1\,\sigma$ dispersion. GS-14 is the *red* source with the lowest SFR. *Right:* $f_{\rm mol}$ as function of the depletion time ($t_{\rm depl} = M_{\rm mol}/SFR$) for the same sample.

values of stellar mass as well. Considering the depletion time ($t_{\rm depl} = M_{\rm mol}/SFR$), the low f_{mol} outliers have short $t_{\rm depl}$; however, other sources share similar values. On the other hand, the hi f_{mol} outliers are well isolated also in $t_{\rm depl}$ (Fig. 5.3). We can see in Fig. 5.4 that, considering the molecular gas mass, the four outliers are remarkably similar to the other sources in the sample, being almost at the center of the $M_{\rm mol}$ distribution. Finally, the two hi f_{mol} outliers and one of the low f_{mol} outlier are young galaxies (with respect

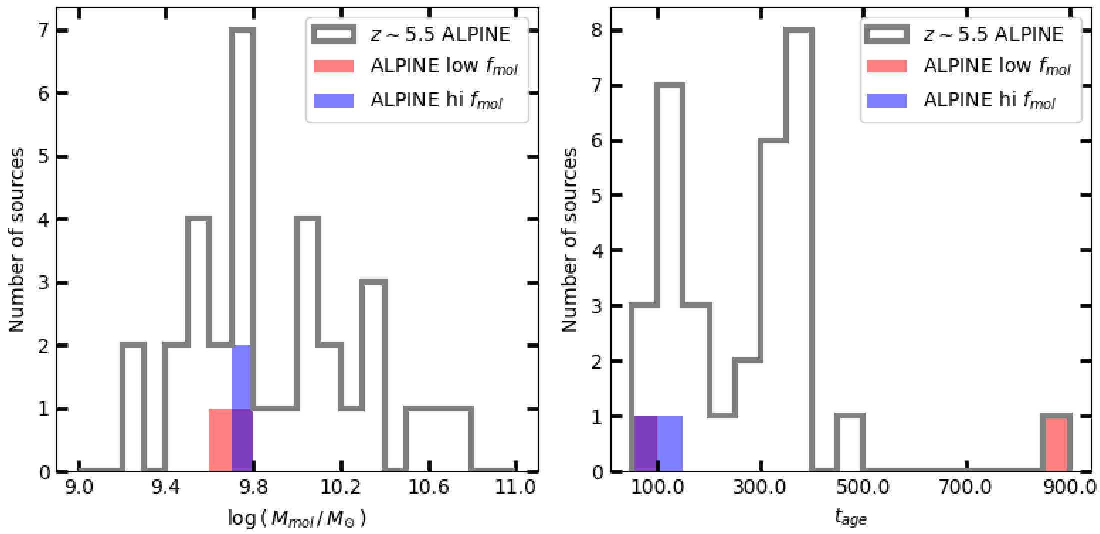

Figure 5.4: *Left:* Distribution of the molecular gas mass for our ALPINE subsample. *Right:* Distribution of the age of the galaxy for the same sample. GS-14 (*red* source on the right) has an age of almost ~ 900 Myr.

to the age distribution of our sample) with t_{age} ~ 100 Myr. However, the other low f_{mol} outlier (that is GS-14) is three times older than the median age of the sample.

We also performed a multivariate correlation analysis among all the galaxy properties and the [C II]-derived ones to search for further correlations in our sample. As we can see from Fig. 5.5, we did not find any significant correlation, with the exception of already known (e.g., $M_* - SFR$) relations or trivial ones (e.g., $L_{CII} - M_{mol}$).

In conclusion, the analysis of the outliers properties reveals that the extreme values of the f_{mol} are not driven just by the amount of molecular gas, or by the extreme values of the stellar mass, but, is a combination of the two. As we can see from Fig. 5.6, for the same stellar mass, the outliers show lower (or higher) content of molecular gas. Considering the other properties we investigated, the outliers do not appear to be significantly different from the rest of the sample. With one exception: the extreme age of GS-14. As the age of the universe at $z = 5.5$ is ~ 1 Gyr, a massive galaxy with an age of almost 900 Myr is quite remarkable and worth of further investigations.

5.3 GS-14

GS-14 is a $z = 5.56$ (Raiter et al. 2010; Vanzella et al. 2010) MS galaxy with unusual spectral features that have been interpreted as signatures of a double stellar population or linked to AGN activity. The nature of GS-14 emission has been extensively debated since its discovery in the southern field of the Great Observatories Origins Deep Survey (GOODS) during the ESO/FORS2 survey (Vanzella et al. 2006). Fontanot et al.

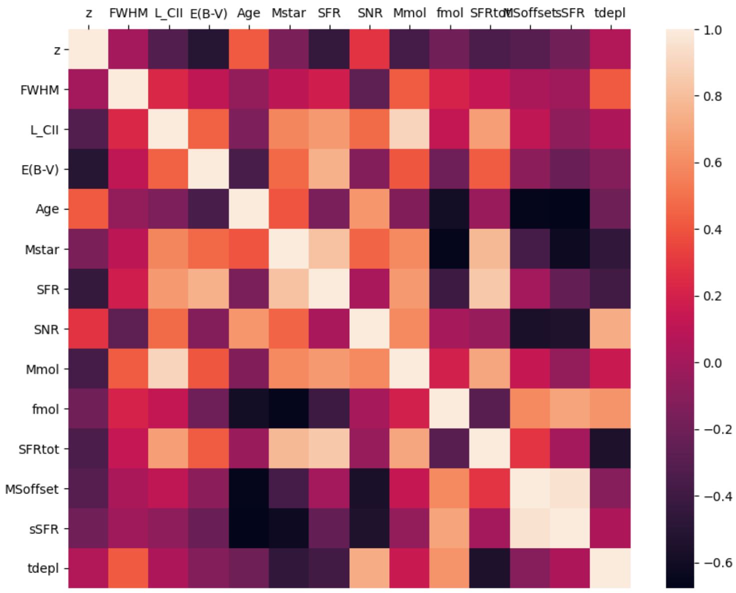

Figure 5.5: Multivariate correlation analysis between the [C II]-derived properties and the galaxy properties for our ALPINE subsample. The color code indicates the sample correlation coefficient (r). An r value of 1 (*white*) indicates perfect correlations, a value of −1 (*black*) a perfect anti-correlation, a value between −0.4 and 0.4 (∼*red*) no correlation.

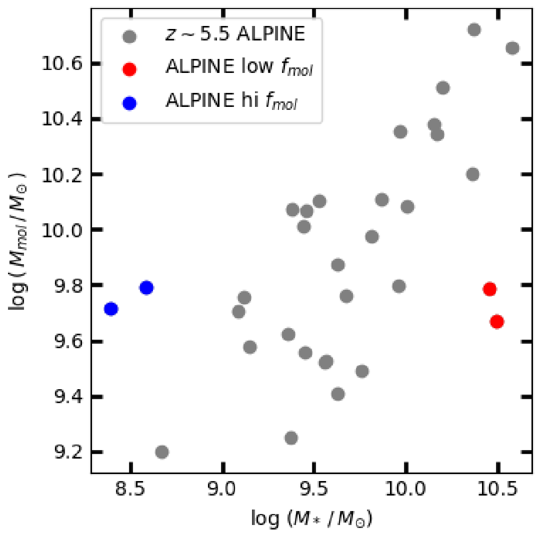

Figure 5.6: Molecular gas mass as function of the stellar mass for our ALPINE subsample.

(2007) selected this source as a QSO candidate on the basis of its z_{850} magnitude and color selection, but reclassified it as an H II star-forming galaxy due to the presence of the

N IV] 1483, 1486Å and the lack of the N V 1240Å lines in the FORS2 spectrum (Vanzella et al. 2006). GS-14 appears as a compact-source at all wavelengths (Fig. 5.7), although it is marginally resolved both in the i_{775} and z_{850} bands (Vanzella et al. 2010). Wiklind et al. (2008) classified the source as a Balmer galaxy, with the discontinuity between the Ks and 3.6 μm filter indicating an evolved stellar population. Vanzella et al. (2010) interpreted the bright Lyα line and the detection of N IV] emission line as signature of a young population of massive stars or, alternatively, of an AGN. They also performed spectral energy distribution (SED) fitting using 16 photometric bands from UV to near-IR (NIR), and found that the source may host a double stellar population, composed of an evolved/aged population and of a young population of massive stars. Grazian et al. (2020) analyzed the same FORS2 spectrum as Vanzella et al. (2010) in depth together with new VIMOS and X-SHOOTER spectra of this source. In addition to the two lines detected by Vanzella et al. (2010), they also found O VI 1032Å, and N V 1240Å emission lines, which led them to regard GS-14 as an AGN. Regarding the X-ray bands, GS-14 is undetected in the ultradeep 7 Ms X-ray image by Chandra with a flux limit of 10^{-17} erg cm^{-2} s^{-1} in the observer frame 0.5 – 2.0 keV band (Giallongo et al. 2019).

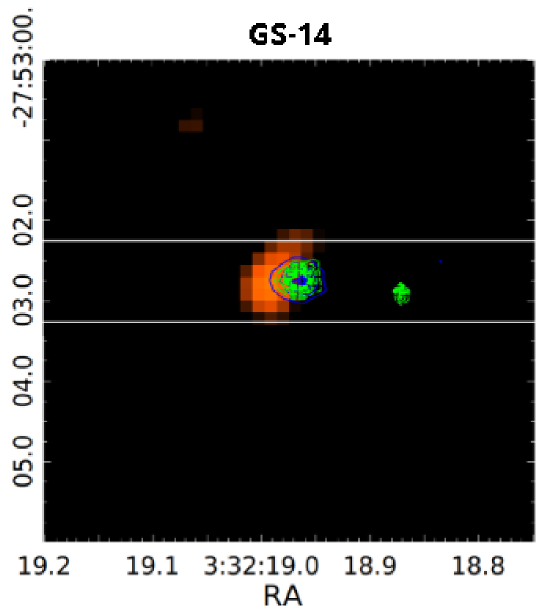

Figure 5.7: GS-14 [C II] 158μm flux map. The *green contours* represent the emission in the $r+$ SUBARU filter in which the Lyα line falls, and the *blue contours* show the emission in the ULTRAVISTA Ks filter(Cassata et al. 2020). The source is not resolved in any of the three wavelength ranges. Vanzella et al. (2010) associated an effective radius to the Ks emission of r_e[3300Å] < 0.9 kpc.

GS-14 has also been selected as part of the ALPINE survey, as a normal star-forming galaxy (SFG). It has been detected in the [C II] 158μm emission line with a signal-to-noise ratio (S/N) of 4.6, but it is not spatially resolved (with a beam of 0.7″, ~ 4.2 kpc, Fig. 5.7). There is no detection of continuum near [C II] with an upper limit of $L_{\rm IR} < 2.0 \times 10^{10}\, L_\odot$

in the far-infrared dust emission (Béthermin et al. 2020). With a molecular gas fraction of $f_{mol} = M_{mol}/(M_* + M_{mol}) = 0.10^{+0.13}_{-0.06}$ (derived from the [C II] luminosity; Dessauges-Zavadsky et al. 2020b), GS-14 is the [C II]-detected source with the lowest f_{mol} of the entire ALPINE sample.

5.4 Spectral analysis

5.4.1 Spectroscopic data

Different rest-UV spectroscopic observations are available for GS-14 in terms of depth and resolution: a 4-hour FORS2 spectrum (Vanzella et al. 2010), a 20-hour VIMOS spectrum (ID 194.A-2003, McLure et al. 2018), and a 49-hour X-SHOOTER spectrum obtained under two observing programs (384.A-0886 and 089.A-0679). We focus on the VIMOS spectrum, as it has the best S/N of the three available spectra, and it allows us for the first time to detect the rest-frame UV continuum of this galaxy at $z \sim 5.5$ at a $S/N = 3.7$ (as computed in the 8200 – 9200 Å observed-frame wavelength range).

Figure 5.8 shows four windows of the one-dimensional spectrum of GS-14. The main features of the spectrum are the Lyα line and the almost total intergalactic medium (IGM) absorption blueward of it, at a wavelength of $\lambda_{\rm obs} < 8000$ Å. N V 1240Å and N IV] 1483, 1486Å are detected at a S/N of 4.6 and 10.1, respectively*. O VI 1032Å and C II* 1335Å are detected at an S/N of 4.4 and 4.0, respectively. C IV 1550Å is not detected because this line falls at the edge of the VIMOS spectral range and over a sky emission line.

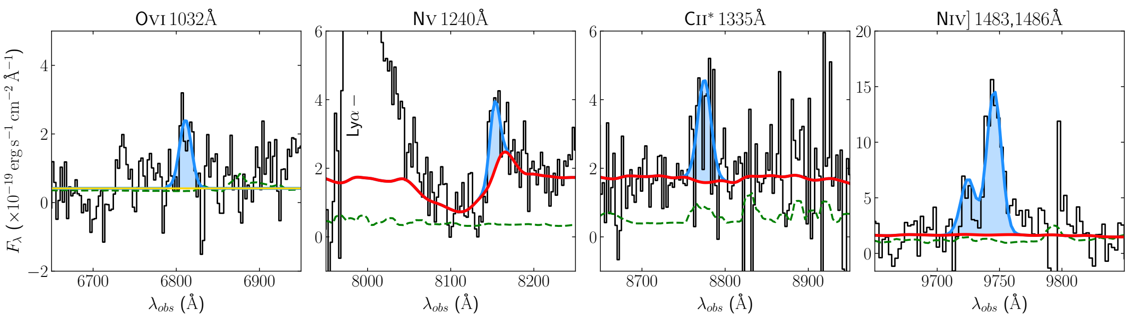

Figure 5.8: Spectra and best-fit models for the GS-14 main spectral features. From left to right: O VI 1032Å, N V 1240Å, C II* 1335Å, and N IV] 1483, 1486Å. The *red line* is the SS99 best-fit model, the Gaussian best fit for the emission feature is plotted in *blue*, the *yellow line* represents the continuum in the Lyman forest region, and the *green line* shows the spectrum noise. In the second panel, the strong emission line at $\lambda_{\rm obs} \sim 8000$ Å is the Lyα line. For the complete VIMOS spectrum, see Figure 2 in Grazian et al. (2020).

* $S/N = F_{\rm line} / (\int_{\lambda_0-3\sigma}^{\lambda_0-1.5\sigma} N_\lambda \, d\lambda + \int_{\lambda_0+1.5\sigma}^{\lambda_0+3\sigma} N_\lambda \, d\lambda)$, where $F_{\rm line}$ is the line-integrated flux, N_λ is the spectrum noise, λ_0 and σ are the line centroid and width.

5.4.2 Spectral fitting

The latest release of the VIMOS spectrum reaches a continuum S/N of 3.7, which is high enough to show a clear P-Cygni N v line profile for the first time. P-Cygni profiles in N v are not unheard of (Jaskot et al. 2017; Vanzella et al. 2018; Matthee et al. 2022) and can be produced by stellar winds of very young and massive stars (e.g., Prinja & Howarth 1986) or be mimicked by broad or narrow absorption line quasar (BAL/NAL QSO; Bentz et al. 2004; Appenzeller et al. 2005). As our spectrum does not show any absorption line and because there are no signs that GS-14 is a type 1 QSO (e.g., $> 1000\,\mathrm{km\,s^{-1}}$ broad emission lines or high X-ray luminosity), the origins of the P-Cygni profile are likely associated with a young population of massive stars. We fit the spectrum with the models from the STARBURST99 (S99) spectral synthesis code (Leitherer et al. 1999), adopting the same method as Marques-Chaves et al. (2021). S99 models allow the fitting of the stellar continuum and of the lines produced in the stellar atmosphere, but not lines that are produced in the ionized gas in the interstellar medium (ISM) of the galaxy. High-resolution UV model spectra with burst ages in the $0.01 - 20\,\mathrm{Myr}$ range and metallicities between 0.03 and $0.5\,\mathrm{Z_\odot}$ were rebinned and smoothed to the VIMOS spectrum resolution. To take dust attenuation into account, the Calzetti et al. (2000) extinction curve was used to match the E(B − V) of the models with the one measured on the spectrum of GS-14 (E(B − V)= 0.05, from the 9200 − 12000Å observed-frame wavelength range). The free-from-lines spectral window 8400 − 8600Å was used to normalize the flux of the S99 models to the GS-14 flux. Finally, we performed a χ^2 minimization on the 8050 − 8400Å range to find the best fit for the N v line profile. We found that the continuum and the N v line profile of GS-14 are best fit with a $2.7 \pm 0.1\,\mathrm{Myr}$ old stellar population, a mass of $(5 \pm 1) \times 10^7\,\mathrm{M_\odot}$, and a metallicity of $0.5\,\mathrm{Z_\odot}$. We note that the P-Cygni fit exhibits a slight degeneracy with metallicity: the N v P-Cygni profile varies little with different values of metallicity, while it is extremely sensitive to the age of the stellar population (see Marques-Chaves et al. 2021). Our best-fit model is able to reproduce the absorption part of the N v line profile as well as the continuum redward of the Lyα emission line. However, the best-fit S99 model is not able to fully reproduce the observed emission in N v, and we need an additional emission component. The Gaussian fit of this component provides an emission line centered at $8152.6^{+0.6}_{-0.1}$ Å, with a full width half maximum of $FWHM$= $13.9^{+2.4}_{-1.9}$ Å and a flux of $2.9^{+0.9}_{-0.1} \times 10^{-18}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$.

We also fit the O vi, C ii*, and N iv] emission lines using Gaussian components. Their best-fit values are reported in Table 5.1. Due to the IGM absorption, the measured O vi flux should be considered as a lower limit.

We compared the EW of the N v line in absorption and emission with the binary population and spectral synthesis code (BPASS; Stanway & Eldridge 2018) and with the S99 models. We find that some models are able to reproduce the observed EW in absorption and have a very young stellar population, in agreement with what we find from the P-Cygni N v line profile. However, regarding the emission component, all the models have EW> −3 Å, which is far from the measured $EW_{\mathrm{N\,V,em}} = -5.2$ Å. This result suggests that two different mechanisms are likely at the origin of the N v emission: Stellar winds from young and massive stars create the absorption component and contribute to the emission, while an additional emission that is not linked to stellar wind phenomena provides the rest of the observed N v flux. This additional contribution is likely due to the ionized gas in the ISM, which is not modeled by stellar synthesis codes such as BPASS or S99.

Table 5.1: Properties of the emission lines of GS-14, derived from the VIMOS spectrum and fitting the lines with a Gaussian profile.

Line	λ_0	*FWHM*	*F*
	Å	Å	erg s^{-1} cm^{-2}
O vi 1032Å	$6811.0^{+0.1}_{-1.0}$	$17.3^{+4.1}_{-2.2}$	$> 3.6^{+0.9}_{-1.2} \times 10^{-18}$
N v 1240Å	$8152.6^{+0.6}_{-0.1}$	$13.9^{+2.4}_{-1.9}$	$2.9^{+0.9}_{-0.1} \times 10^{-18}$
C ii* 1335Å	$8775.0^{+1.0}_{-1.0}$	19^{+14}_{-5}	$6.0^{+4.5}_{-3.0} \times 10^{-18}$
N iv] 1483Å	$9724.8^{+0.4}_{-0.1}$	$13.9^{+2.4}_{-1.9}$	$7.3 \pm 1.3 \times 10^{-18}$
N iv] 1486Å	$9745.6^{+0.4}_{-0.1}$	$13.9^{+2.4}_{-1.9}$	$1.9 \pm 0.3 \times 10^{-17}$

Notes. λ_0 refers to the observed-frame centroid, *FWHM* is the full width at half maximum, and *F* is the line flux. For the N v emission line, the values refer to the fit obtained after the stellar component of the P-Cygni profile was subtracted. The O vi line lies in the Lyman forest and its underlying continuum is not detected, therefore its flux should be considered a lower limit.

5.4.3 Comparison with literature

We performed a deep search for sources in the literature with O vi 1032Å, N v 1240Å, or N iv] 1483, 1486Å emission lines; our goal was to discern the nature (SF or AGN) of GS-14. We found eight sources with all the three lines, seven from Dietrich et al. (2003) at redshift $3.9 < z < 5$, and one from Baldwin et al. (2003) at $z = 1.96$. These are all spectroscopically confirmed AGN. The three lines can also be found in the Hainline et al. (2011) composite spectrum of 33 narrow-line AGN at $z \sim 2 - 3$. In the collection of sources, we also verified whether the flux of N iv] was higher than that of N v, as is the case for GS-14. None of the above-mentioned AGN shows this characteristic. We note that the $z = 3.36$ lensed galaxy of Fosbury et al. (2003) and the stacked spectrum of $z \sim 2 - 3.8$, $\mathrm{EW_{C\,III]}} \geq 20$Å galaxies of Le Fèvre et al. (2019) show a N iv] emission line that is more luminous than N v, but both lack an O vi detection. With the exception of the young SFG of Marques-Chaves et al. (2021) (which shows P-Cygni profiles in both O vi and N v, but no N iv] emission line detection), all the sources with N v and O vi detections are AGN. Because it is a high-ionization emission line, the O vi line is usually associated with AGN activity, although it has also been detected in emission in a small number of extreme SFG (Otte et al. 2003; Grimes et al. 2007; Hayes et al. 2016). We report in Table 5.2 the complete list of the sources we use for our comparison, as well as the detected lines.

Table 5.2: Compilation of literature sources with O VI, N V, or N IV] emission lines.

	ID	z	O VI	N V	N IV]	Ref
	GS+14	5.55	✓	✓p	✓	This work
AGN	Q0353-383	1.96	✓	✓	✓	Baldwin et al. (2003)
	DLS1053-0528	4.02		×	✓	Glikman et al. (2007)
	NDWFS1433	3.88		✓	✓	Glikman et al. (2007)
	UDS24561	3.21	✓	×	×	Tang et al. (2021)
	33 NL AGN	2 – 3	✓	✓	✓	Hainline et al. (2011)
	12 Seyfert 2	0 – 4		✓60%		Dors et al. (2014)
	59 radio-galaxies	0 – 4		✓25%		Dors et al. (2014)
	10 QSO2	0 – 4		✓70%		Dors et al. (2014)
	S82-20	3.08	✓	✓		Lin et al. (2022)
	J1254+0241	1.8		✓	✓	Dhanda et al. (2007)
	J1546-5253	2.0		✓	✓	Dhanda et al. (2007)
	J1553+0056	2.63		✓p	×	Appenzeller et al. (2005)
	16 QSO	2.4 – 3.8	✓26%	✓		Dietrich & Wilhelm-Erkens (2000)
	11 QSO	3.9 – 5.0	✓73%	✓	✓73%	Dietrich et al. (2003)
	PSO J006+39	6.61		✓	×	Koptelova et al. (2019)
	J1512+119	2.11	✓	×	×	Borguet et al. (2012)
	J000239+255034	5.80		✓		Fan et al. (2006)
	124 QSO stack	5.6 – 6.7		✓		Bañados et al. (2016)
	13 AGN stack	2 – 3.8		✓	✓	Le Fèvre et al. (2019)
galaxy	Lynx Arc	3.36	×	×	✓	Fosbury et al. (2003)
	31 stack $\mathrm{EW}_{[\mathrm{C\,II}]} > 20$	2 – 3.8		✓	✓	Le Fèvre et al. (2019)
	120 stack $10 < \mathrm{EW}_{[\mathrm{C\,II}]} < 20$	2 – 3.8		✓	✓	Le Fèvre et al. (2019)
	J141445+544631	5.42	×	×	✓	McGreer et al. (2018)
	J160810+352809	0.03		✓p		Jaskot et al. (2017)
	Ion3	4.0	×	✓p	×	Vanzella et al. (2018)
	17 low mass galaxies stack	1.4 – 2.9		✓	✓6%	Stark et al. (2014)
	L^* lensed galaxy	3.5		✓	✓	Patrício et al. (2016)
	J0121+0025	3.24	✓p	✓p	×	Marques-Chaves et al. (2021)

Notes. Each row refers to a single source, a collection of sources, or a stacked spectrum. In the first case, the *ID* columns refer to the source name, in the second case, to the number and type of the sources in the collection, and in the third case, to the number of sources included in the stack. O VI, N V, and N IV] report the detection (✓) or the nondetection (×) of the respective emission line. No symbol indicates that the line falls out of the spectral range. In case of a collection of sources, the percentage refers to the fraction of sources with a detection in the reported emission lines. Finally, *p* indicates that the line exhibits a P-Cygni profile.

5.4.4 Comparison with theoretical predictions

We exploited the N IV]/N V and O VI/N V flux ratios to further investigate the origin of the additional N V emission component, as well as N IV] and O VI emission lines. Figure 5.9 shows the comparison of the line ratios in GS-14 (*red star*) with those from the AGN of §5.4.3 with all the three lines detected (*cyan diamonds*). GS-14 N IV]/N V line ratio uncertainty is shown as a *black error bar* and the O VI/N V lower limit is indicated with

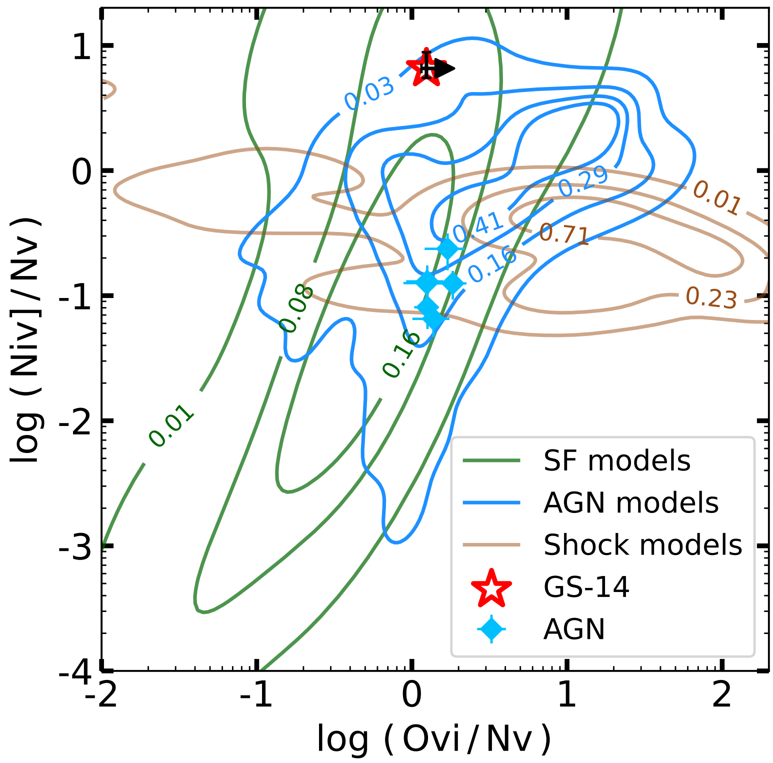

Figure 5.9: Comparison of GS-14 (*red star*), literature AGN (*cyan diamonds*), and nebular theoretical predictions of the N IV]/N V and O VI/N V flux ratios. The contours represent isoproportions of the density (i.e., 8% of the probability mass lies outside of the contour drawn for 0.08). *Brown contours* refer to shock predictions from the Morisset et al. (2015) 3MDB^{S} database, *green contours* to SF models from Gutkin et al. (2016), and *blue contours* to AGN emission as computed by Feltre et al. (2016). Literature AGN are the sources described in §5.4.3 with all the three lines detected. No shock model is able to reproduce the GS-14 line ratios; SF and AGN can both be at the origin of the observed nebular emission. As the GS-14 O VI line is not corrected for the IGM attenuation, we have only a lower limit on its flux, thus a lower limit on the O VI/N V ratio (*black arrow*).

the *black arrow*. The contours refer to the theoretical predictions for flux ratios driven by shocks, AGN, and SF. The shock predictions (*brown contours*) are from the Mexican million models database (3MDB^{S}; Morisset et al. 2015), a compilation of shock models calculated with the code MAPPINGS and evaluated with the CLOUDY (Ferland et al. 2013) photoionization code. The SF models (*green contours*) are from the Gutkin et al. (2016) models and refer to a cloud illuminated by a 10 Myr old population; the nebular emission was computed with CLOUDY. Finally, the AGN predictions (*blue contours*) are from Feltre et al. (2016) and rely on CLOUDY to simulate the emission from a NLR cloud illuminated by the central AGN. We find that no shock model is able to reproduce the observed flux ratios of GS-14, while both SF and AGN models are compatible with the observed values. We note that because we must consider the O VI flux as a lower limit, the intrinsic O vi/N v flux ratio in GS-14 should be higher; in fact, the IGM attenuation at O vi wavelength for a $z \sim 5.5$ source could be as high as a factor 4 (e.g., Inoue et al. 2014) and GS-14 should move toward the region in which fewer SF models reside. Table 5.3 reports the parameter space explored by the shock, AGN, and SF models we used. While we cannot constrain the ionization parameter using the GS-14 line flux ratios, we find that for the AGN models, only a hydrogen density of the cloud of $n_H \sim 10^4\,\mathrm{cm}^{-3}$ and a metallicity $Z \leq 0.5\,Z_\odot$ fit our data. Regarding the SF models, the observed line ratio can be reproduced only by $Z \leq 0.13\,Z_\odot$ models. With this low metallicity, stellar models predict a much weaker P-Cygni profile in the N v wind line than is observed. This suggests that the nebular emission in GS-14 is likely powered by an AGN. In the future, a definitive answer on the origin of the emission might be obtained by observing more emission lines, such as C iv 1550Å and He ii 1640Å.

Table 5.3: Parameter space for the AGN, SF, and shock models. For each parameter, N_{sample} values are simulated in the *values* range.

		N_{sample}	values
AGN	$\log U_S$	5	$[-5, -1]$
	$Z/Z_\odot$	16	$[0.06, 4.6]$
	ξ_d	3	$0.1, 0.3, 0.5$
	$\log(n_H/\mathrm{cm}^{-3})$	3	$2, 3, 4$
	α	4	$[-2, -1.2]$
SF	$\log U_S$	7	$[-1.0, -4.0]$
	$Z/Z_\odot$	14	$[0.006, 2.6]$
	ξ_d	1	0.3
	$\log(n_H/\mathrm{cm}^{-3})$	1	2
Shock	$v_{\mathrm{shock}}/\mathrm{km\,s}^{-1}$	37	$[100, 1000]$
	$\log(n_{\mathrm{H,pre-shock}}/\mathrm{cm}^{-3})$	7	$[0, 4]$
	$B_{\mathrm{transv}}\,/\,\mu\mathrm{G}$	8	$[10^{-4}, 10]$
	$Z/Z_\odot$		$[0.006, 2.6]$

Notes. U_S refers to the ionization parameter at the Strömgren radius (see Feltre et al. (2016)), Z to the metallicity, ξ_d to the dust-to-metal mass ratio, n_H to the cloud gas density, and α to the power-law index at UV and optical wavelengths of the AGN continuum ($S_\nu \propto \nu^\alpha$ for $0.001 \leq \lambda/\mu\mathrm{m} \leq 0.25$). For the shock models, v_{shock} refers to the shock velocity, $n_{\mathrm{H,pre-shock}}$ to the pre-shock density, and B_{transv} to the transverse magnetic field; the investigated metallicity parameter space is the same as in the SF models.

5.5 Photometric analysis

5.5.1 Photometric data

GS-14 has been observed in several photometric bands. It has been detected with the *MPG-ESO Wide Field Imager* (WFI; Hildebrandt et al. 2006; Erben et al. 2005), with the *Infrared Spectrometer and Array Camera* (ISAAC) instrument at the *Very Large Telescope* (VLT; Nonino et al. 2009; Wuyts et al. 2008; Retzlaff et al. 2010), with the *Canada-France-Hawaii Telescope* (CFHT) *Wide-field InfraRed Camera* (WIRCam; Hsieh et al. 2012), and with the *SUBARU Suprime-Cam* (Cardamone et al. 2010). Regarding observations from space telescopes, GS-14 has detections with the *Advanced Camera for Survey* (ACS) and *Wide Field Camera 3* (WFC3) on board the *Hubble Space Telescope* (HST; Giavalisco et al. 2004; Koekemoer et al. 2011; Grogin et al. 2011; Brammer et al. 2012; van Dokkum et al. 2013), as well as with the *Spitzer InfraRed Array Camera* (IRAC; Dickinson et al. 2003; Ashby et al. 2013; Guo et al. 2013). The complete list of photometric filters we adopted is reported in Table 5.4. The photometry of all ALPINE sources was collected and calibrated by Faisst et al. (2020). It comes primarily from the 3D-*HST* catalog and was corrected for Galactic extinction, point spread function size, and other biases. Some additional data that are not present in the 3D-*HST* catalog come from various observation programs in the Extended Chandra Deep Field South Giacconi et al. (2002, ECDFS,) and were measured and calibrated by Faisst et al. (2020). GS-14 has not been detected by MPG-ESO/WFI in band $U38$, b, and v, in the *HST*/ACS $F435W$ band, by the Subaru/Suprime-Cam $IA445$, $IA505$, $IA527$, $IA550$, $IA574$, $IA598$, $IA624$, and $IA738$ filters, or by *Spitzer/MIPS* at 24 μm. Regarding the X-ray bands, GS-14 is undetected in the ultradeep 7 Ms X-ray image by *Chandra* with a flux limit of 10^{-17} erg cm^{-2} s^{-1} in the observed frame $0.5-2.0$ keV band (Giallongo et al. 2019). We refer to Faisst et al. (2020) for further details.

The SED of GS-14 is characterized by the evident discontinuity between the *VLT/ISAAC* and *CFHT/WIRCAM* bands and the *Spitzer/IRAC* bands, which is consistent with originating from the Balmer break (see Fig.5.10). Another peculiarity is that the flux in the *IRAC 1* and *2* bands is higher than the fluxes of bands *3* and *4*. This is probably due to a significant contribution of the Hβ+[O III] and Hα lines to the *IRAC 1* and *2* fluxes.

5.5.2 SED fitting

To estimate the galaxy properties, we chose to use the `CIGALE` code (**?**Boquien et al. 2019; Yang et al. 2020, 2022) (already presented in section 4.3.1), as it is quite flexible and allows us to fit the photometric data with and without the AGN component, as well as with a double stellar population. We used the Bruzual & Charlot (2003) population synthesis model with a Chabrier (2003) initial mass function, nebular emission lines, a Calzetti et al. (2000) dust attenuation law, and the Draine et al. (2014) dust models. We tested different star formation histories (SFHs): double exponential, delayed SFH, delayed SFH plus burst or quench, and constant. The AGN emission was added via the

Table 5.4: Photometric bands. For details about the data, extraction, and calibration, we refer to Faisst et al. (2020).

Observatory /Instrument	Filter	Central λ [Å]	Ref.
MPG-ESO/WFI	I	3633.3	1
VLT/ISAAC	J^v	12492.2	2
	H^v	16519.9	2
	K_s^v	21638.3	2
CFHT/WIRCam	J^w	12544.6	3
	K_s^w	21590.4	3
Subaru/Suprime-Cam	$IA856$	8566.0	4
HST/ACS	$F814W$	8058.2	5
	$F850LP$	9181.2	6
HST/WFC3	$F125W$	12516.3	7
	$F140W$	13969.4	8
	$F160W$	15391.1	7
Spitzer/IRAC	ch1	35634.3	9
	ch2	45110.1	9
	ch3	57593.4	10
	ch4	79594.9	10

References. (1) Hildebrandt et al. (2006); Erben et al. (2005); (2) Wuyts et al. (2008); Retzlaff et al. (2010); (3) Hsieh et al. (2012); (4) Cardamone et al. (2010); (5) Giavalisco et al. (2004); (6) Koekemoer et al. (2011); (7) Grogin et al. (2011); Koekemoer et al. (2011); (8) Brammer et al. (2012); van Dokkum et al. (2013); (9) Ashby et al. (2013); Guo et al. (2013); (10) Dickinson et al. (2003).

SKIRTOR models (Stalevski et al. 2012, 2016), in which the torus is modeled as a clumpy two-phases medium (we refer to Yang et al. 2020, for further details). We report in Table 5.5 the complete parameter space investigated via the SED fitting.

Figure 5.10 shows the best-fit model for GS-14, with a delayed SFH, a stellar mass of $M_* = (4 \pm 1) \times 10^{10}\,\mathrm{M}_\odot$, a stellar age of $680 \pm 170\,\mathrm{Myr}$, and a burst of SF of SFR = $90 \pm 30\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ in the last $8 \pm 6\,\mathrm{Myr}$. We find that the stellar mass, age of the galaxy, and bulk of the SFH are well constrained and do not depend heavily on the presence of AGN or of a double population. Although the fitting is acceptable with a single population ($\tilde{\chi}^2 \sim 2.7$), we obtain a better fit when a recent burst in the SFH ($\tilde{\chi}^2 = 1.6$) is added. In particular, the presence of a younger population allows us to better fit the higher fluxes of the *IRAC 1* and *2* bands (with respect to *IRAC 3* and *4*) because the Hβ and Hα lines of the young stars contribute significantly to the *IRAC 1* and *2* fluxes.

The fitting does not reveal any significant contribution from an AGN. When AGN models are used, the SED fitting constantly prefers low-luminosity type 2 AGN models,

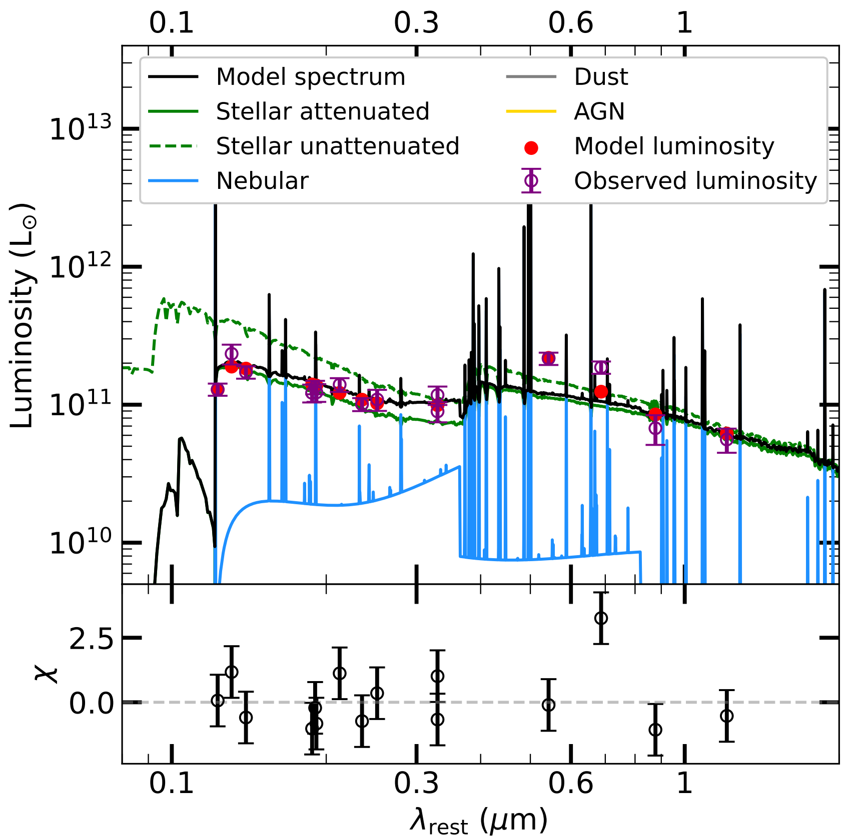

Figure 5.10: `X-CIGALE` best-fit SED of GS-14. The fitting prefers a delayed exponential (with an optional exponential burst) SFH, with an old ($t_{\rm age}$ = 680 ± 170 Myr) stellar population of $M_* = (4 \pm 1) \times 10^{10}\,{\rm M_\odot}$ and a young population of 8 ± 6 Myr, which is now experiencing a burst of SF of 90 ± 30 $\rm M_\odot\,yr^{-1}$. The dust and AGN component are not visible in this plot because they contribute most at $\lambda_{\rm rest} > 3\,\mu$m. The investigated parameter space is reported in Table 5.5.

with an AGN contribution to the optical-UV six orders of magnitude lower than the stellar emission. We note, however, that we do not have any coverage in the mid-IR, where the warm dust heated by the AGN should contribute most. The SED fitting still allows us to exclude that a type 1 AGN is present, as it would contribute significantly in the optical-UV where we have an optimal photometric coverage, but leaves open the possibility of a moderate- or low-luminosity obscured AGN. This is consistent with the narrow-line profiles observed in GS-14 (Lyα and N v).

As sanity check, we also performed an SED fitting with the SED3FIT code (Berta et al. 2013; da Cunha et al. 2008), but without the option of a double population. We find that the stellar mass and SFR are compatible with those from `X-CIGALE` within the uncertainties. Similarly, no significant AGN contribution is detected.

Table 5.5: Parameter space for the `X-CIGALE` SED fitting. For each parameter, N_{sample} values are simulated in the *values* range.

	N_{sample}	values	description
SFH		sfh2exp; sfhdelayed; sfhdelayedbq; sfhperiodic	double exponential; delayed SFH with optional exponential burst; delayed SFH with optional constant burst/quench; periodic SFH.
τ_{main}	3	50, 500, 1000	e-folding time of the main stellar population model in Myr.
$\mathrm{age}_{\mathrm{main}}$	6	100, 250, 500, 600, 700, 1000	Age of the main stellar population in the galaxy in Myr.
τ_{burst}	3	50, 100, 500	e-folding time of the late starburst population model in Myr.
$\mathrm{age}_{\mathrm{burst}}$	4	5,10,20,50	Age of the late burst in Myr.
$\mathrm{f}_{\mathrm{burst}}$	4	0, 0.001, 0.01, 0.1	Mass fraction of the late burst population.
IMF		Chabrier	Initial Mass Function
$Z/Z_{\odot}$	4	0.03, 0.3, 1.3, 3.2	Metallicity
$\log \mathrm{U_S}$	2	$-2, -1$	Nebular component: ionization parameter
$\mathrm{Z}_{\mathrm{gas}}$	2	0.0004, 0.004	Nebular component: gas metallicity
dust attenuation		Charlot&Fall 2000	
$\mathrm{A_{V,ISM}}$	3	0.3, 1.7, 3.3	V-band attenuation in the interstellar medium.
μ	3	0.3, 0.5, 1.0	$\mathrm{A_{V,ISM}}$ / ($\mathrm{A_{V,BC}}$+$\mathrm{A_{V,ISM}}$)
Dust emission		Draine+2014	
$\mathrm{q_{PAH}}$	3	0.47, 2.5, 3.9	Mass fraction of PAH.
$\mathrm{U}_{\mathrm{min}}$	4	5, 10, 25, 40	Minimum radiation field.
α	1	2.0	Power-law slope dU/dM $\propto \mathrm{U}^{\alpha}$.
γ	2	0.02, 0.1	Fraction illuminated from $\mathrm{U}_{\mathrm{min}}$ to $\mathrm{U}_{\mathrm{max}}$.
AGN		Skirtor16	
$\tau_{9.7\mu\mathrm{m}}$	5	3, 5, 7, 9, 11	Average edge-on optical depth at 9.7 μm.
p	1	1.0	Power-law index of radial gradient of dust density.
q	1	1.0	Power-law index of angular gradient of dust density.
oa	1	40°	Torus half-opening angle.
$\mathrm{R}_{\mathrm{ratio}}$	1	20	Ratio of the maximum to minimum radii of the dust torus.
M_{cl}	1	0.97	Mass fraction of dust inside clumps.
i	6	0°, 20°, 40°, 60°, 70°, 90°	Viewing angle (w.r.t. the AGN axis).
disk type	1	Schartmann et al. (2005)	Disk spectrum.
δ	1	-0.36	Power-law index modifying the optical slope of the disk.
$\mathrm{f}_{\mathrm{AGN}}$	4	0.001, 0.01, 0.05, 0.1	$8-1000\,\mu$m AGN fraction.
law	1	SMC	Extinction law of polar dust.
E(B-V)	1	0.03	E(B-V) for the extinction in the polar direction in magnitudes.
$\mathrm{T}_{\mathrm{dust}}^{\mathrm{polar}}$	1	100 K	Temperature of the polar dust.
ϵ	1	1.6	Emissivity index of the polar dust.

Notes. As the best fits were obtained with a double exponential SFH (*sfh2exp*), we report only the parameter space for this SFH, while the other investigated SFHs are indicated in *SFH*: the delayed SFH (*sfhdelayed*), the delayed SFH with a burst (*sfhdelayedbq*), and the periodic SFH (*sfhperiodic*).

5.6 Conclusion

We find several clues indicating that GS-14 has a double stellar population. The P-Cygni profile in the N v line suggests the presence of a young population of massive stars. The fitting of the line profile provides an age of 2.7 ± 0.1 Myr and a mass of $(5 \pm 1) \times 10^7\,\mathrm{M}_{\odot}$. Moreover, the best SED fitting is obtained for a double population, with a 680 ± 170 Myr old population of $(4 \pm 1) \times 10^{10}\,\mathrm{M}_{\odot}$ and a young (8 ± 6 Myr old) population of $(5.6 \pm 0.1) \times 10^8\,\mathrm{M}_{\odot}$ (see Table 5.6). In this scenario, the old population dominates and accounts for most of the stellar mass of the galaxy (similar to what was found by Laporte et al. 2021; Harikane et al. 2022; Matthee et al. 2022). It is responsible for the Balmer break and for the continuum at $\lambda_{\mathrm{rest}} \gtrsim 3000$ Å. The young population is linked to the ongoing

Table 5.6: GS-14 properties from photometric and spectral analyses.

	Photometric	Spectral
$\log(M_*^{tot}/\mathrm{M}_\odot)$	$10.62^{+0.10}_{-0.13}$	
$\log(M_*^{old}/\mathrm{M}_\odot)$	$10.61^{+0.10}_{-0.14}$	
$\log(M_*^{young}/\mathrm{M}_\odot)$	8.7 ± 0.1	$7.7^{+0.1}_{-0.2}$
age_*^{old} (Myr)	680 ± 170	
age_*^{young} (Myr)	8 ± 6	2.7 ± 0.1
SFR ($\mathrm{M}_\odot$/yr)	90 ± 30	
SFR^{MS} ($\mathrm{M}_\odot$/yr)	~ 200	
$SFR^{[\mathrm{C\,II}]}$ ($\mathrm{M}_\odot$/yr)		16^{+14}_{-7}
SFR^{UV} ($\mathrm{M}_\odot$/yr)	26 ± 5	
$Z/\mathrm{Z}_\odot$	< 0.2	0.5
f_{mol}	$0.10^{+0.13}_{-0.06}$	
E(B − V)	0.05	0.05
$\log(L^{\mathrm{SED}}_{\mathrm{bol,AGN}}/\mathrm{erg\,s^{-1}})$	<44.5	
$\log(L^{\mathrm{X-ray}}_{\mathrm{bol,AGN}}/\mathrm{erg\,s^{-1}})$	<43.5	

Notes. M_*^{tot}, M_*^{old}, and M_*^{young} refer to the total, old, and young population, respectively. age_*^{old} and age_*^{young} refer to the age of the old and young stellar population. *SFR* refers to the instantaneous SFR from the SED fitting, SFR^{MS} to the expected SFR for an MS galaxy of similar mass and redshift, obtained using the Speagle et al. (2014) MS, $SFR^{[\mathrm{C\,II}]}$ refers to the SFR estimated from the [C II] luminosity and using the Schaerer et al. (2020) relations of Figure 3, and SFR^{UV} to the SFR obtained from the UV luminosity. Z refers to the metallicity. The molecular gas fraction f_{mol} was computed with the M_{molgas} of Dessauges-Zavadsky et al. (2020b). The E(B − V) in the *Photometric column* was derived from the SED fitting by Faisst et al. (2020), while the one in the *Spectral column* from the VIMOS spectrum as described in §5.4.2. $L^{\mathrm{SED}}_{\mathrm{bol,AGN}}$ and $L^{\mathrm{X-ray}}_{\mathrm{bol,AGN}}$ are the AGN bolometric luminosity derived from the SED fitting and from the *Chandra* X-ray flux upper limit, assuming the bolometric correction from Lusso et al. (2012), respectively.

SF and produces the P-Cygni profile and most of the continuum at ~ 2000 Å. The low cold-gas fraction of GS-14 of $f_{\mathrm{mol}} \sim 0.10^{+0.13}_{-0.06}$, derived from the [C II] 158μm emission (see Dessauges-Zavadsky et al. 2020b) and the SED-fitting stellar mass favors the scenario that GS-14 is composed mostly of an old and evolved stellar population that formed ~ 600 Myr after the Big Bang, which has already consumed or expelled most of its original gas reservoir. We exclude the alternative origin of the N V P-Cygni profile, that is, a BAL-QSO absorption feature mimicking a P-Cygni profile, as there is no observational evidence that GS-14 could be a type 1 QSO and the spectra show no signs of absorption features (e.g., Lyα, C IV 1550Å, O VI 1032Å, or Si IV 1394, 1403Å; Vito et al. 2022; Vietri et al. 2022).

The specific SFR (sSFR = SFR/M$_*$) of GS-14 is $\log(\mathrm{sSFR/yr^{-1}}) = 8.7^{+0.1}_{-0.2}$. At this redshift, and considering the mass from SED-fitting, the SFR for an MS galaxy is

$SFR^{MS} \sim 200\,M_\odot\,yr^{-1}$ (Speagle et al. 2014), with the lower boundary of its $1\,\sigma$ dispersion at $\sim 60\,M_\odot\,yr^{-1}$; hence GS-14 with its SFR = $90 \pm 30\,M_\odot\,yr^{-1}$, while below the average MS, is still within its dispersion. From the SFH derived with the SED fitting without considering the recent episode of SF, the SFR should be $\sim 1\,M_\odot\,yr^{-1}$, thus GS-14 would be $> 4\,\sigma$ below the MS, in the locus of quiescent galaxies. The new episode of SF has moved GS-14 up toward the MS, although it should last only a few tens of million years due to its short depletion time ($t_{depl} = M_{molgas}/SFR = 80^{+130}_{-60}$ Myr). This kind of up-and-down movement in the SFR-M_* plane is expected: Tacchella et al. (2016) and Orr et al. (2019) suggested base on simulations that the MS dispersion can be originated by similar oscillations in sSFR on timescales $\sim 0.4\,t_{Hubble}$.

The GS-14 N v line profile shows an emission component in addition to the P-Cygni profile we fit. The comparison of the observed N v EW with N v equivalent widths from BPASS and S99 models allows us to exclude that the origin of the whole N v emission lies in stellar winds. This additional emission component therefore comes from nebular emission, but it can have various origins, depending on which process causes the radiation field that illuminates the gas. On the one hand, a shock origin can be excluded as the GS-14 flux ratios of N iv]/N v and O vi/N v are not compatible with shock-related models (Figure 5.9). On the other hand, models of AGN or SF are both able to reproduce the observed flux ratios, as shown in Figure 5.9. However, we note that an AGN origin is more likely because the high-ionization potential of the O vi requires extreme conditions for it to be of stellar origin. Moreover, we found just one star-forming galaxy in the literature (Marques-Chaves et al. 2021) with both O vi and N v, all the remaining sources with these two lines are AGN. Finally, all the sources with all the three O vi, N v, and N iv] lines are AGN (Table 5.2). While none of this is sufficient proof, the list of reasons strongly indicates an AGN origin for the GS-14 nebular emission. The lack of an X-ray detection for GS-14 in deep *Chandra* data and the fact that the SED fitting does not show a significant contribution from the AGN reveals that this AGN probably is of type 2. Obscured (type 2) AGN are more difficult to detect in the X-rays, but can be identified by their high ionization lines and/or by their mid-IR emission. Exploiting the X-ray flux upper limit and assuming a power-law spectrum (without AGN obscuration) for the X-ray emission with a photon index of $\Gamma = 1.8$ ($F(E) \propto E^{1-\Gamma}$) , we derive a 1σ upper limit on the rest frame 2 – 10 keV luminosity of the AGN: $\log(L_{2-10\,keV}/erg\,s^{-1}) < 42.5$. Assuming the bolometric correction of Lusso et al. (2012), we have an upper limit on the AGN bolometric luminosity of $\log(L^{X-ray}_{bol,AGN}/erg\,s^{-1}) < 43.5$. The upper limit on the AGN bolometric luminosity derived from the SED fitting is $\log(L^{SED}_{bol,AGN}/erg\,s^{-1}) < 44.5$. AGN activity might also be a possible explanation for GS-14 extreme low content of molecular gas. A strong past phase of nuclear activity, possibly triggered by the first episode of SF, may have expelled or heated up the molecular gas, hence lowered the f_{mol} significantly and quenched the SF. The low-power AGN signature we witness now could be the last remnant of this past, more powerful, AGN activity.

Our interpretation paints an intriguing picture of GS-14: this $z \sim 5.5$ source is an already evolved galaxy that may have formed 600 Myr after the Big Bang and that now experiences a second burst of SF. It also carries signatures of obscured AGN activity.

Future observations could shed more light on the AGN contribution in GS-14, especially if those observations were to target the mid- and far-IR part of the SED (60 μm – 3 mm observer-frame), where its emission should peak, or where it might at least be disentangled from that of SF. GS-14 also deserves additional high-resolution observations with ALMA: the [C II] 158μm is detected but not resolved, and the far-IR continuum detection (which is missing for now) would place more stringent constraints on the SF emission and the AGN contribution. Furthermore, GS-14 is the only ALPINE source with indications of obscured AGN activity, which means that follow-up with deeper and higher-resolution ALMA observations is even more important and rewarding. Finally, deep rest-UV spectroscopy, targeting medium- and high-ionization lines, such as C IV 1550Å, He II 1640Å, and C III] 1908Å, and exploiting their line ratio diagram, could provide the final evidence of whether the radiation field in GS-14 is dominated by the AGN activity or by a young stellar population.

5.7 GS-14 as archetype of high-z JWST sources

GS-14 was at first glace a unremarkable SF galaxy, with peculiar features emerging when better spectroscopic and photometric became available. The presence of a double stellar population was proposed to explain some of the features, but no definitive proof was available. Similarly, no obvious sign of AGN activity has been ever detected. However, a joint analysis of the spectra and of the photometric data confirmed the existence of the double stellar population. Moreover, when all the peculiarities and the subtle hints has been interpreted under the light of a possible hidden AGN, its presence did not seem so uncertain anymore. Having both a double stellar population and an hidden AGN is able to explain all the peculiar features of this source. GS-14 was recently observed with the *NIRSpec* on board the *James Webb Space Telescope* as part of the GTO program n.1216, and the detection of a broad ($\sim$ 3300 km s^{-1}) component in the Hα, Hβ, He I, and He II line profiles undoubtedly confirmed the AGN presence (Übler et al. 2023). In the same work, they also estimated the BH mass from the Hα broad component (exploiting the Reines et al. 2013, relation) and found a $\log (M_{\rm BH} / {\rm M}_\odot) = 8.20^{+0.11}_{-0.16}$. This indicates a BH mass $\sim$ 2 dex higher than what is expected from local $M_{\rm BH} - M_*$ relations.

Now, this question arises: how common is GS-14? How many seemingly unremarkable inactive galaxies are instead harboring hidden AGN? Furthermore, GS-14 is an already evolved galaxy at $z \sim 5.5$ and should have formed around $8 < z < 9$; could it be an archetype of early galaxies that can now be detected and studied with the *JWST*?

Chapter 6

Conclusions

Since the discovery, in the mid-90's, of the first correlations between the mass of the SMBHs and the properties of the galaxies in which they are hosted, the AGN-galaxy co-evolution paradigm has gain increasing support. It is now widely accepted that the growth of SMBH and of the stellar mass of the galaxy are strictly connected and that the AGN and SF influence each other via feedback mechanisms. However, the details of this co-evolution have not been totally unveiled yet. Despite the differences between the various co-evolution models, for example regarding the co-evolution triggering mechanism and the involved timescales, all scenario agrees that a key phase is the obscured accretion phase. In this phase, most of the BH accretion is expected to take place, however the high quantity of material fueling the SMBH growth has also the effect of hiding the core of the galaxy. Thus, this first phase of growth takes place in very obscured conditions, making its study extremely challenging. Very obscured AGN can appear as normal non-active galaxies in the UV-optical band (because the accretion disk emission is totally extincted), as a dusty SF galaxy in the IR, and/or not being detected at all in the X-rays. However, these "missing" AGN are extremely important for explaining the X-ray background, as the current spectrum of the XRB can only be explained by assuming also the presence of a population of very obscured AGN that we are not able to detect efficiently for now. Finally, a complete census of these obscured accreting AGN is also needed for reconstructing the evolution of the BHAD at high redshift. In fact, the discrepancy between the model prediction and the measurements of the BHAD could be explained by assuming that current high-redshift surveys are missing a significant population of very obscured sources. We are seeing hints of this population with IR surveys, but those are not, at present, able to cover effectively (at the proper depth) the high-z Universe, where we are instead relying mostly on the UV and X-ray bands, that are prone to miss the highly obscured AGN. Thanks to the latest ALMA (and mm/submm) surveys, a new population of extremely dusty optically-dark galaxies is now being detected; however, while they contribute significantly to the SFRD, the presence of SMBH accretion in their nuclei is still un-probed.

This thesis was devoted to the investigation of the elusive obscured AGN population, with a special consideration on the mutual effects of the AGN activity and of the SF. We tackled this challenge in three ways. Firstly, we investigated the best way to select obscured AGN, by exploiting the synergies between future X-ray and IR telescopes. Secondly, we focused on studying two populations of obscured AGN at $z \sim 1$ and $z \sim 2-3$, using rest-frame UV emission line to select them and exploiting the multi-λ coverage in the COSMOS field to characterize them. This allowed us to study the effect of the AGN activity on the galaxies at the cosmic noon, where most of the BHAD and SFRD is expected to occur. Thirdly, we moved to $z \sim 5$ to investigate the cold gas, fuel for both the AGN activity and SF, and exploited a multi-wavelength approach to reveal the AGN presence on a seemingly non-active galaxy, thus anticipating studies that are currently being carried out by the *James Webb Space Telescope* (*JWST*).

Exploiting the X-ray−IR synergies to select obscured AGN

In chapter 2 we present our investigation of the synergistic use of IR and X-ray surveys to detect and characterize AGN in deep fields. In particular, we took into consideration the capabilities of the future ESA X-ray observatory *Athena* and of template proposed IR missions, such as a ESA *SPICA*-like observatory, the NASA *OST*, and the NASA (with international collaborations) *PRIMA* telescope. We used XRB synthesis models to predict the total number of AGN as a function of intrinsic X-ray luminosity, amount of obscuration and redshift. The prediction of the AGN emission are based on the SEDs of more than 500 AGN from the COSMOS field with both X-ray spectra and optical-to-FIR SED-fitting. We differentiated if we were primary detecting either the AGN or the host-galaxy emission in the mid-IR bands and investigated if we were able to spectroscopically identify and characterize the AGN emission with *SMI-LR*-like instruments or exploiting the full R=300 spectral resolution of the *OSS*. In particular, we focused on the elusive and most-obscured CT-AGN.

We found that we need both next-generation IR and X-ray surveys to detect and characterize the obscured AGN population at high redshift. In particular, we showed that using cryogenic mid-IR spectro-photometers we will be able to detect the majority of the AGN up to $z \sim 10$. The X-ray capabilities of *Athena* will be extremely helpful in recognizing these sources as AGN (for $\log{(L_x/\mathrm{erg\,s^{-1}})} \geq 43.5$, just the *Athena* detection will suffice to prove the presence of AGN activity) and in characterizing the AGN luminosity and obscuration via spectral analysis. Moreover, the use of IR low-resolution spectra and/or SED-fitting exploiting multiple filter observations of instruments such as *OST* and *PRIMAGER*, will allow us to disentangle the AGN and host-galaxy contribution and to characterize them both, even in case of no X-ray detection. Similarly, in the case of LLAGN for which we will not have enough X-ray counts to properly constrain the AGN properties from the spectra, the AGN properties obtained by the SED-fitting could be used as priors to help the X-ray fitting. We found that new-generation IR instruments will excel in detecting the most-obscured AGN, that are missed by the current surveys. We showed that the parameter space that these new surveys will open up will be a changing factor to fill the gap between the current measurements and the prediction of the BHAD up to high-z. Concluding, we developed a tool to quickly estimate the capabilities of IR and X-ray surveys to detect and characterize AGN as a function of different instrument parameters (e.g., mirror dimension, photometric and spectroscopic capabilities). We found that only the synergies between *Athena* and IR cryogenic telescopes will allow an unbiased study of the full BHAD history, back to an age of the Universe of ~ 1 Gyr. This, along with the SF history, will give us a much better understanding of galaxy evolution as compared to what we know today.

This work has been published in Barchiesi et al. (2021).

Obscured AGN at $z \sim 1 - 3$

A large part of this thesis is devoted to the analysis of two samples of obscured AGN selected exploiting the [Ne v] and C iv emission lines, observed with the VIMOS spectrometer in the COSMOS field. The use of these rest-frame UV lines has several advantages: it allows us to study specific redshift ranges given the spectral coverage of the VIMOS instrument; the high-ionization potential of the two lines effectively selected only AGN as excitation sources; and focusing only on lines with narrow spectral profiles allowed us to exclude the type 1 AGN from our samples. We also took advantage of the deep multi-wavelength coverage of the COSMOS field, that provided us with spectro-photometric data from the X-ray to the radio band. By exploiting the *Chandra* observations, we performed the X-ray spectral analysis of the detected sources (between one third and half of the samples), thus characterizing the AGN intrinsic luminosity and amount of obscuration. Moreover, we fit the spectral energy distributions to obtain the host-galaxy properties (e.g., M_*, SFR) as well as the AGN ones (e.g., $L_{\rm bolo}$, $M_{\rm BH}$, $\lambda_{\rm Edd}$). To be able to investigate the effects of the AGN activity on the host-galaxy, for each sample we build a control sample of non-active galaxies, matching them in redshift and stellar mass.

In chapter 3, we report the results of the analysis of a sample of 94 [Ne v]-selected obscured AGN in the $0.6 - 1.2$ redshift range. We confirmed that the X/[Ne v] ratio (where X is the observed, i.e. non-corrected for the source obscuration, $2 - 10$ keV flux and [Ne v] is the [Ne v] 3426Å line flux, which is a proxy of the source intrinsic power) can be used as an obscuration proxy and we used it to give a lower limit on the $N_{\rm H}$ even for the X-ray undetected AGN, thus being able to classify two-thirds of the [Ne v] sources as very obscured AGN, and $\sim 20\%$ as candidate CT-AGN. Our analysis showed that almost half of the sample is composed of AGN in a strong episode of SMBH growth ($\lambda_{\rm Edd} \geq 0.1$). We also found that, with respect to its control sample of non-active galaxies, the [Ne v] AGN are lacking a population of quiescent galaxies, despite having the same amount of molecular gas, as estimated from the rest-frame 850 μm luminosity measured via the

SED-fitting. We compared our results with the *in-situ* co-evolution model and found that the [Ne v] sample is most-likely composed by sources in the 'pre-quenching' phase, with only the few oldest sources showing the effect of the AGN quenching the SF. We concluded that the [Ne v] sources are preferentially hosted in obscured AGN efficiently accreting and forming stars.

In chapter 4, we present the analysis of 90 C IV-selected obscured AGN. The use of the C IV emission line allowed us to target the 1.5 − 3.1 redshift range. Thus, we were able to cover the entire cosmic noon with the combined analysis of the [Ne v] and C IV samples. Similarly to the [Ne v] sample, we assumed the $M_* - M_{\rm BH}$ relation of Suh et al. (2020) and computed the mass of the SMBH. We showed that the C IV selection targeted more obscured and more highly accreting AGN with respect to the [Ne v] selection, with 40 % of the sample accreting at super Eddington ratios and composed of extremely obscured and CT-AGN. Our analysis confirmed the capabilities of the UV line selections in identifying obscured AGN and that the [Ne v]- and C IV-selected AGN form a solid high-z sample to study the co-evolution.

The work regarding the [Ne v]-selected AGN has been submitted to A&A, the one about the C IV sample is in preparation.

Obscured AGN activity in a galaxy at $z \sim 5.5$

As the cold gas is the fuel for both the SF and AGN activity, its study can provide clues on the connection between the SMBH and its host-galaxy. In chapter 5, we present our effort to investigate the cause of the huge dispersion of the molecular gas fraction in the ALPINE sample. In particular, we focused on 32 ALPINE SF galaxies, observed via ALMA in [C II] and FIR continuum at $z \sim 5.5$, for which their molecular gas fraction vary between $f_{\rm mol} \sim 0.1$ and $f_{\rm mol} \sim 0.9$. We found that the extreme values of $f_{\rm mol}$ are driven by a combination of stellar mass and actual content of molecular gas, without further differences between the sources with high and low molecular gas fraction.

During our investigation, we stumbled across an extremely peculiar source: GS-14. It was the source with the lowest $f_{\rm mol}$ of the entire ALPINE sample and, due to peculiar features in its optical spectra, its true nature has been a matter of debate for the last 20 years. We exploited the multi-wavelength coverage of GS-14 to investigate the properties and the origin of its emission by performing UV-to-NIR SED-fitting, with single and double stellar population and/or AGN component. In addition, we analyzed the latest release of the VIMOS spectrum, which showed highly ionized emission lines (O VI 1032Å, N V 1240Å, and N IV] 1483, 1486Å). The line equivalent widths and line ratios have been compared with those observed in galaxies and AGN, as well as with the predictions from radiation transfer models for star-forming galaxies, AGN, and shocks.

Ultimately, we found compelling evidence indicating that GS-14 has a double stellar population. The SED-fitting and the P-Cygni profile in the N V line point toward a second young stellar population with an age of $5 - 10$ Myr and a mass between $5 \times 10^7\,M_\odot$ and $5 \times 10^8\,M_\odot$. The old population dominates the broad band emission and accounts for most of the stellar mass of the galaxy ($\log(M_*^{tot}/M_\odot) = 10.6 \pm 0.1$). Interestingly, our analysis shows that GS-14 should have formed when the Universe was just ~ 600 Myr old, and at $z \sim 5.5$ is an already evolved galaxy with high-stellar mass and almost depleted of cold gas. Even more interestingly, we found that the N v line profile, as well as, the N iv]/N v and O vi/N v line ratios suggest the presence of an obscured AGN. AGN activity that has been revealed with the most recent JWST observations.

This paints an intriguing picture where GS-14 could be the archetype of extreme early galaxies, born at $z > 8$, the redshift range in which JWST is now detecting new galaxies, and that will quickly evolve until they are completely formed and without cold-gas at $z \sim 5$. Moreover, GS-14 raises the question of how many seemingly un-remarkable galaxies at high-z are instead harboring hidden AGN, that could be revealed through deep spectra and a multi-wavelength approach.

This work has been accepted in A&A, Barchiesi et al. (2022).

Future perspectives

We developed a tool to study the selection of AGN by exploiting synergies between *Athena* and future IR cryogenic telescopes; our code will be easily updated as new IR instruments will be proposed. At the moment, the *PRIMA* mission seems the most likely IR telescope to be adopted, although its deployment window is still far in the future. If *PRIMA* will not be chosen, we will continue to update our code with the new telescopes that will surely come out and to advocate for new mid-IR instruments. The mid- far-IR window, that has unfortunately been left without observatories for the last decade (and that will be without for at least the next 10 years) offers extraordinary opportunities to study the evolution of the SF and the AGN accretion at high-redshift. If the *PRIMA* mission will be adopted, future plans may involve the improvement of our code to investigate which photometric band combinations are the most useful to identify AGN at various redshift, by exploiting color-color diagram, decision tree, and support vector machine.

We plan to improve our work on the C iv-selected AGN as well. Firstly, by correctly estimating the contribution of the selection effects on the sample properties (both in the X-ray detected AGN and in the entire sample), we will be able to study the intrinsic redshift evolution of AGN between $z \sim 0.6$ and $z \sim 3$ (i.e., the redshift range where the bulk of the SFRD and BHAD is). We also plan to perform an in-depth investigation of the C iv-sources with $\log(M_*/M_\odot) < 9.5$ to study, once we excluded any spurious effect from the SED-fitting procedure, these AGN that, due to the low stellar masses, seem to accrete at high super-Eddington ratios. We have plans to extend the same type of analysis to AGN for which the C iv and [Ne v] lines should have been present but were not detected, with the aim of pinpointing the cause of the lines nondetection. For example, 180 sources in

C-COSMOS have X-ray luminosities typical of AGN, only narrow lines in their spectra, but no [Ne v] detection even in their stacked spectrum. This could be due to the host-galaxy medium hiding the [Ne v] emission or these sources could intrinsically lack these lines (due to combinations of metallicity and density of the NLRs and AGN radiation field strength).

In the upcoming years, the X-ray *eROSITA* mission (launched on July 2019) will scan all the sky, with deeper and deeper sensitivity. While, on average, the survey will not reach the sensitivity of the *Chandra* Deep Field South, or the C-COSMOS central region, the ecliptic poles will benefit from the deepest observations and an unmatched large survey area ($\sim 40\,\mathrm{deg}^2$ with $> 10\,\mathrm{ks}$ vignetting-corrected exposure in the $0.6 - 2.3\,\mathrm{keV}$ band). The now available *eROSITA* Final Equatorial-Depth Survey already comprises a sample of $\sim$ 22000 X-ray detected AGN (Liu et al. 2022) and $\sim$ 300 of these sources have $N_{\mathrm{H}} > 10^{23}\,\mathrm{cm}^{-2}$. Thus, we can expect to have in the near future an unprecedented number of obscured AGN to try to bridge the gap between the BHAD observations and simulations. Exploiting a multi-wavelength approach, similar to the one adopted in this thesis, can significantly help in the characterization of these AGN and in placing constraints on their properties.

Finally, the advent of *JWST* has opened up new possibilities in the search for AGN at high-redshifts. In this thesis, we have shown that high-ionization emission lines (such as the [Ne v] and the C IV) are extremely powerful tool to select obscured AGN in specific redshift ranges. Even mid-ionization UV emission lines (e.g., O III] 1663Å, N III] 1750Å, S III] 1880Å, C III] 1908Å) could be used to infer the presence of an AGN, using simulations from radiation transfer codes and line ratios (similarly to what we did in chapter 5, but see also Feltre et al. 2016). These lines can be searched for in the *JWST NIRSPEC* spectra (starting from $z \sim 0.8$ for the [Ne v], from $z \sim 2 - 3$ for the other lines, and up to $z > 10$) of already targeted sources, and the AGN presence could be supported by SED-fitting and multi-wavelength observations. In addition to *JWST*, the *MOONS* multi-objects spectrograph at the *VLT* has a planned first light for the 2024 and, with its $\sim$ 1000 fibers and $0.6 - 1.8\,\mu$m wavelength coverage, will have the capability of detecting the [Ne v] line in the $0.8 - 4.3$ redshift range and the others line in the $2 - 8$ redshift range. Exploiting the *MOONS* instrument, the MOONRISE survey will target thousands of galaxies up to $z \sim 5$, and a significant fraction of them will likely show some of these lines, thus revealing the presence of hidden AGN. *JWST NIRSPEC* and *MOONS* spectra will allow us to search for obscured AGN accretion at higher redshift, even before the cosmic noon. Furthermore, as there will be overlaps between the redshift ranges of the various lines (e.g., with *JWST NIRSPEC* and *MOONS* we will potentially detect the [Ne v] line in the C IV-selected AGN), we will be able to investigate the relations between these lines and further refine their use to select and characterize the physical properties of obscured AGN.

Chapter A

Appendices

A.1 CIGALE flux χ distributions

As we discussed in section 4.3.1, to obtain good SED-fitting, we enlarged the uncertainties of the UV-optical bands by a factor equal to the dispersion σ of the flux χ distributions. In this appendix, we report the χ distributions (where $\chi = (F_{\mathrm{obs}} - F_{\mathrm{model}})/eF_{\mathrm{obs}}$, with F_{obs} the observed flux, eF_{obs} its uncertainty, and F_{model} the best-fit predicted flux) for all the filters.

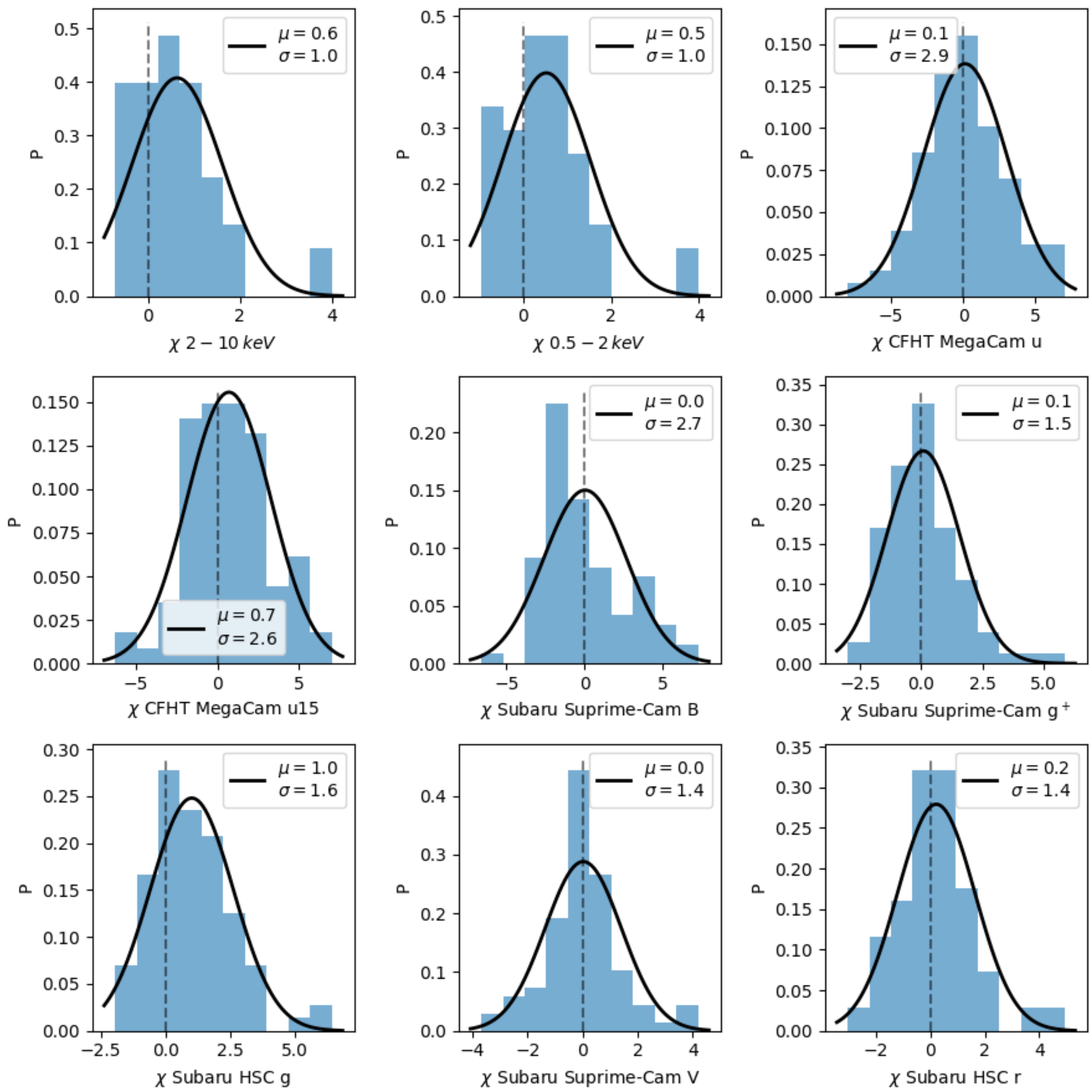

Figure A.1: Flux χ distributions before enlarging the flux uncertainties. $\chi = (F_{\text{obs}} - F_{\text{model}})/eF_{\text{obs}}$, with F_{obs} the observed flux, eF_{obs} its uncertainty, and F_{model} the best-fit predicted flux. Fitting the χ distribution with a Gaussian model (*black line*) revealed a σ larger than expected for a random uncertainty. We ascribed this to an underestimation of the flux uncertainties. Enlarging the flux uncertainties for the UV-optical filters by a factor equal to the dispersion σ of the χ distribution proved to improve the SED fitting.

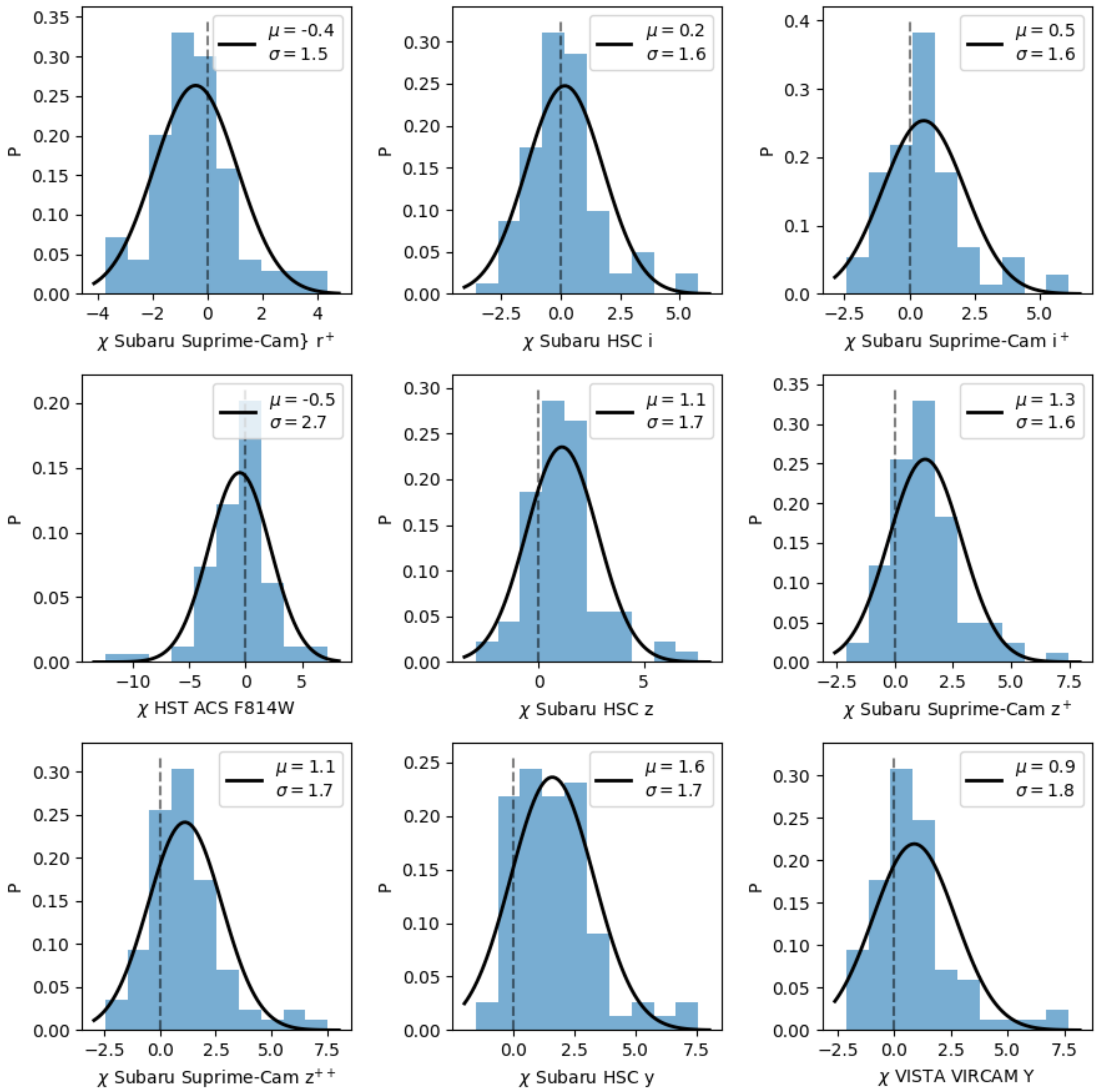

Figure A.2: Same as Fig. A.1.

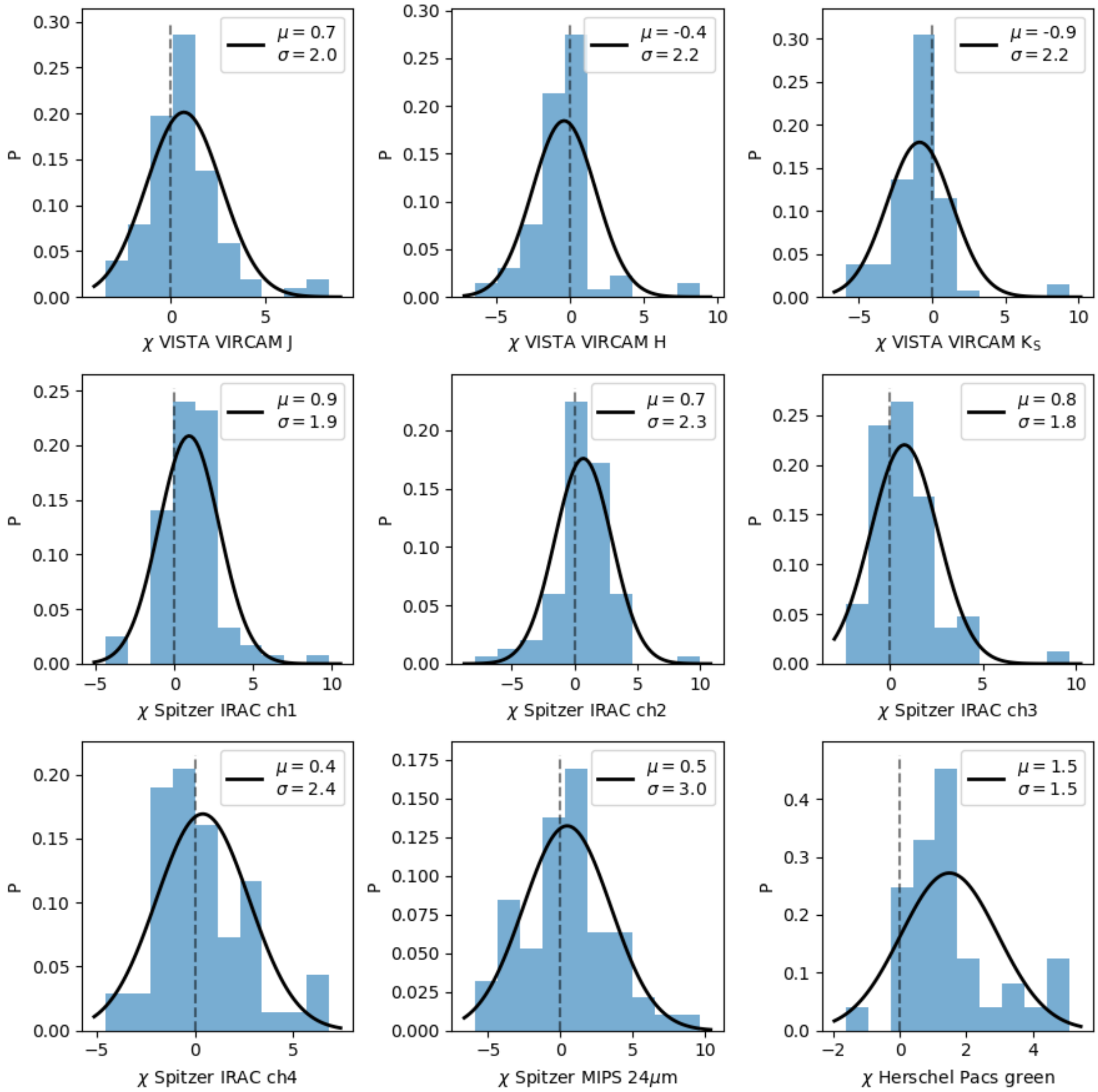

Figure A.3: Same as Fig. A.1.

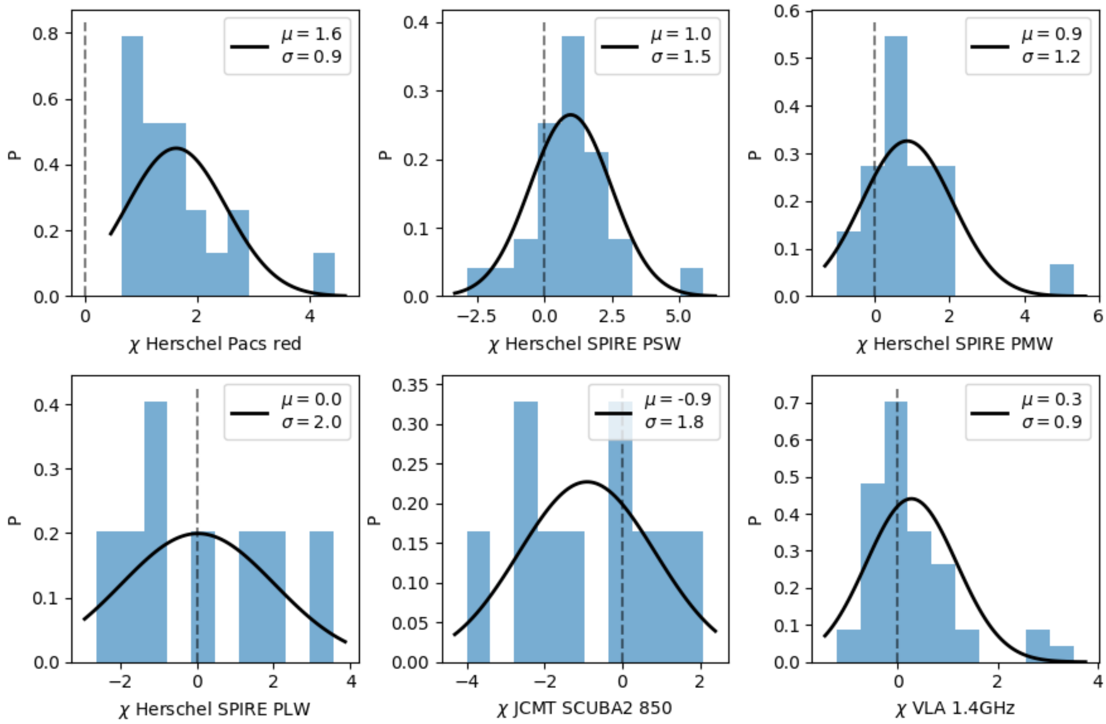

Figure A.4: Same as Fig. A.1.

Bibliography

Aird, J., Alexander, D. M., Ballantyne, D. R., et al. 2015, ApJ, 815, 66

Aird, J., Nandra, K., Laird, E. S., et al. 2010, MNRAS, 401, 2531

Akylas, A., Georgakakis, A., Georgantopoulos, I., Brightman, M., & Nandra, K. 2012, A&A, 546, A98

Alexander, D. M., Bauer, F. E., Chapman, S. C., et al. 2005, ApJ, 632, 736

Alexander, D. M., Swinbank, A. M., Smail, I., McDermid, R., & Nesvadba, N. P. H. 2010, MNRAS, 402, 2211

Almaini, O. 2003, Astronomische Nachrichten, 324, 109

Alonso-Herrero, A., Ramos Almeida, C., Mason, R., et al. 2011, ApJ, 736, 82

Ananna, T. T., Treister, E., Urry, C. M., et al. 2019, ApJ, 871, 240

André, P., Hughes, A., Guillet, V., et al. 2019, Probing the cold magnetized Universe with SPICA-POL (B-BOP)

Antonucci, R. R. J. & Miller, J. S. 1985, ApJ, 297, 621

Appenzeller, I., Stahl, O., Tapken, C., Mehlert, D., & Noll, S. 2005, A&A, 435, 465

Archibald, E. N., Dunlop, J. S., Jimenez, R., et al. 2002, MNRAS, 336, 353

Armus, L., Bernard-Salas, J., Spoon, H. W. W., et al. 2006, ApJ, 640, 204

Armus, L., Charmandaris, V., Bernard-Salas, J., et al. 2007, ApJ, 656, 148

Arnaud, K. A. 1996, in Astronomical Society of the Pacific Conference Series, Vol. 101, Astronomical Data Analysis Software and Systems V, ed. G. H. Jacoby & J. Barnes, 17

Arnouts, S., Moscardini, L., Vanzella, E., et al. 2002, MNRAS, 329, 355

Ashby, M. L. N., Willner, S. P., Fazio, G. G., et al. 2013, ApJ, 769, 80

Assef, R. J., Stern, D., Kochanek, C. S., et al. 2013, ApJ, 772, 26

Bañados, E., Venemans, B. P., Decarli, R., et al. 2016, ApJS, 227, 11

Baldwin, J. A., Hamann, F., Korista, K. T., et al. 2003, ApJ, 583, 649

Ballantyne, D. R., Draper, A. R., Madsen, K. K., Rigby, J. R., & Treister, E. 2011, ApJ, 736, 56

Ballo, L., Severgnini, P., Della Ceca, R., et al. 2014, MNRAS, 444, 2580

Barbosa, F. K. B., Storchi-Bergmann, T., Cid Fernandes, R., Winge, C., & Schmitt, H. 2009, MNRAS, 396, 2

Barchiesi, L., Dessauges-Zavadsky, M., Vignali, C., et al. 2022, arXiv e-prints, arXiv:2212.00038

Barchiesi, L., Pozzi, F., Vignali, C., et al. 2021, Publ. Astron. Soc. Australia, 38, e033

Barcons, X., Barret, D., Decourchelle, A., et al. 2017, Astronomische Nachrichten, 338, 153

Barret, D., Lam Trong, T., den Herder, J.-W., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10699, Space Telescopes and Instrumentation 2018: Ultraviolet to Gamma Ray, 106991G

Battersby, C., Armus, L., Bergin, E., et al. 2018, Nature Astronomy, 2, 596

Bavdaz, M., Wille, E., Ayre, M., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10699, Space Telescopes and Instrumentation 2018: Ultraviolet to Gamma Ray, 106990X

Behar, E., Baldi, R. D., Laor, A., et al. 2015, MNRAS, 451, 517

Bentz, M. C., Osmer, P. S., & Weinberg, D. H. 2004, in The Interplay Among Black Holes, Stars and ISM in Galactic Nuclei, ed. T. Storchi-Bergmann, L. C. Ho, & H. R. Schmitt, Vol. 222, 515–516

Berta, S., Lutz, D., Santini, P., et al. 2013, A&A, 551, A100

Béthermin, M., Fudamoto, Y., Ginolfi, M., et al. 2020, A&A, 643, A2

Bisigello, L., Gruppioni, C., Feltre, A., et al. 2020, arXiv e-prints, arXiv:2011.07074

Blandford, R. D. 1990, 161

Blustin, A. J., Page, M. J., Fuerst, S. V., Branduardi-Raymont, G., & Ashton, C. E. 2005, A&A, 431, 111

Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, ARA&A, 51, 207

Bonato, M., De Zotti, G., Leisawitz, D., et al. 2019, Publ. Astron. Soc. Australia, 36, e017

Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103

Borguet, B. C. J., Edmonds, D., Arav, N., Benn, C., & Chamberlain, C. 2012, ApJ, 758, 69

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34

Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, ApJS, 200, 13

Brightman, M. & Nandra, K. 2011, MNRAS, 413, 1206

Brightman, M. & Ueda, Y. 2012, MNRAS, 423, 702

Brusa, M., Bongiorno, A., Cresci, G., et al. 2015, MNRAS, 446, 2394

Brusa, M., Perna, M., Cresci, G., et al. 2016, A&A, 588, A58

Bruzual, G. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 374, From Stars to Galaxies: Building the Pieces to Build Up the Universe, ed. A. Vallenari, R. Tantalo, L. Portinari, & A. Moretti, 303

Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000

Burgarella, D., Buat, V., & Iglesias-Páramo, J. 2005, MNRAS, 360, 1413

Burlon, D., Ajello, M., Greiner, J., et al. 2011, ApJ, 728, 58

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682

Capetti, A., Axon, D. J., Macchetto, F., Sparks, W. B., & Boksenberg, A. 1996, ApJ, 469, 554

Cappi, M., Panessa, F., Bassani, L., et al. 2006, A&A, 446, 459

Cardamone, C. N., van Dokkum, P. G., Urry, C. M., et al. 2010, ApJS, 189, 270

Cash, W. 1979, ApJ, 228, 939

Cassata, P., Morselli, L., Faisst, A., et al. 2020, A&A, 643, A6

Castelló-Mor, N., Carrera, F. J., Alonso-Herrero, A., et al. 2013, A&A, 556, A114

Cattaneo, A., Faber, S. M., Binney, J., et al. 2009, Nature, 460, 213

Chabrier, G. 2003, Publications of the Astronomical Society of the Pacific, 115, 763

Charlot, S. & Fall, S. M. 2000, ApJ, 539, 718

Chartas, G., Rhea, C., Kochanek, C., et al. 2016, Astronomische Nachrichten, 337, 356

Cicone, C., Maiolino, R., Sturm, E., et al. 2014, A&A, 562, A21

Cimatti, A., Daddi, E., & Renzini, A. 2006, A&A, 453, L29

Civano, F., Marchesi, S., Comastri, A., et al. 2016, ApJ, 819, 62

Cleri, N. J., Yang, G., Papovich, C., et al. 2022, arXiv e-prints, arXiv:2209.06247

Collinson, J. S., Ward, M. J., Landt, H., et al. 2017, MNRAS, 465, 358

Comastri, A., Gilli, R., Marconi, A., Risaliti, G., & Salvati, M. 2015, A&A, 574, L10

Comastri, A., Ranalli, P., Iwasawa, K., et al. 2011, A&A, 526, L9

Comastri, A., Setti, G., Zamorani, G., & Hasinger, G. 1995, A&A, 296, 1

Combes, F., García-Burillo, S., Audibert, A., et al. 2019, Astronomy and Astrophysics, 623, A79

Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839

Crummy, J., Fabian, A. C., Gallo, L., & Ross, R. R. 2006, MNRAS, 365, 1067

Cucciati, O., Tresse, L., Ilbert, O., et al. 2012, A&A, 539, A31

da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595

Daddi, E., Alexander, D. M., Dickinson, M., et al. 2007, ApJ, 670, 173

De Marco, B., Ponti, G., Cappi, M., et al. 2013, MNRAS, 431, 2441

Del Moro, A., Alexander, D. M., Bauer, F. E., et al. 2017, Frontiers in Astronomy and Space Sciences, 4, 67

Del Moro, A., Mullaney, J. R., Alexander, D. M., et al. 2014, ApJ, 786, 16

Delvecchio, I., Gruppioni, C., Pozzi, F., et al. 2014, MNRAS, 439, 2736

Delvecchio, I., Lutz, D., Berta, S., et al. 2015, MNRAS, 449, 373

Dessauges-Zavadsky, M., Ginolfi, M., Pozzi, F., et al. 2020a, A&A, 643, A5

Dessauges-Zavadsky, M., Ginolfi, M., Pozzi, F., et al. 2020b, A&A, 643, A5

Dhanda, N., Baldwin, J. A., Bentz, M. C., & Osmer, P. S. 2007, ApJ, 658, 804

Di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604

Dickinson, M., Giavalisco, M., & GOODS Team. 2003, in The Mass of Galaxies at Low and High Redshift, ed. R. Bender & A. Renzini, 324

Dietrich, M., Appenzeller, I., Hamann, F., et al. 2003, A&A, 398, 891

Dietrich, M. & Wilhelm-Erkens, U. 2000, A&A, 354, 17

DiPompeo, M. A., Hickox, R. C., Carroll, C. M., et al. 2018, ApJ, 856, 76

Done, C. 2010, arXiv e-prints, arXiv:1008.2287

Done, C., Gierliński, M., Sobolewska, M., & Schurch, N. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 373, The Central Engine of Active Galactic Nuclei, ed. L. C. Ho & J. W. Wang, 121

Dors, O. L., Cardaci, M. V., Hägele, G. F., & Krabbe, Â. C. 2014, MNRAS, 443, 1291

Dowell, C. D., Hildebrand, R. H., Schleuning, D. A., et al. 1998, The Astrophysical Journal, 504, 588

Draine, B. T., Aniano, G., Krause, O., et al. 2014, ApJ, 780, 172

Duras, F., Bongiorno, A., Ricci, F., et al. 2020, A&A, 636, A73

Efstathiou, A., Hough, J. H., & Young, S. 1995, MNRAS, 277, 1134

Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, A&A, 533, A119

Elitzur, M. 2008, New Astronomy Reviews, 52, 274

Elvis, M., Civano, F., Vignali, C., et al. 2009, ApJS, 184, 158

Erben, T., Schirmer, M., Dietrich, J. P., et al. 2005, Astronomische Nachrichten, 326, 432

Evans, I. N., Ford, H. C., Kinney, A. L., et al. 1991, ApJ, 369, L27

Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2022, ApJ, 930, L12

Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019, ApJ, 875, L1

Faber, S. M. & Jackson, R. E. 1976, ApJ, 204, 668

Fabian, A. C. 1999, MNRAS, 308, L39

Fabian, A. C., Celotti, A., & Erlund, M. C. 2006, MNRAS, 373, L16

Fabian, A. C., Vasudevan, R. V., & Gandhi, P. 2008, MNRAS, 385, L43

Fabian, A. C., Vasudevan, R. V., Mushotzky, R. F., Winter, L. M., & Reynolds, C. S. 2009, MNRAS, 394, L89

Fadda, D., Yan, L., Lagache, G., et al. 2010, ApJ, 719, 425

Faisst, A. L., Schaerer, D., Lemaux, B. C., et al. 2020, ApJS, 247, 61

Fan, X., Strauss, M. A., Richards, G. T., et al. 2006, The Astronomical Journal, 131, 1203

Fanaroff, B. L. & Riley, J. M. 1974, MNRAS, 167, 31P

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

Feltre, A., Hatziminaoglou, E., Fritz, J., & Franceschini, A. 2012, MNRAS, 426, 120

Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, Rev. Mex. Astron. Astrofis., 49, 137

Fernández-Ontiveros, J. A., Spinoglio, L., Pereira-Santaella, M., et al. 2016, ApJS, 226, 19

Ferrarese, L. & Merritt, D. 2000, ApJ, 539, L9

Feruglio, C., Maiolino, R., Piconcelli, E., et al. 2010, A&A, 518, L155

Fiore, F., Puccetti, S., Grazian, A., et al. 2012, A&A, 537, A16

Fontanot, F., Cristiani, S., Monaco, P., et al. 2007, A&A, 461, 39

Fosbury, R. A. E., Humphrey, A., Villar-Martín, M., et al. 2003, in The Mass of Galaxies at Low and High Redshift, ed. R. Bender & A. Renzini, 308

Fritz, J., Franceschini, A., & Hatziminaoglou, E. 2006, MNRAS, 366, 767

Fruscione, A., McDowell, J. C., Allen, G. E., et al. 2006, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 6270, Proc. SPIE, 62701V

Galitzki, N., Ade, P. A. R., Angilè, F. E., et al. 2014, Journal of Astronomical Instrumentation, 03, 1440001

Gandhi, P., Horst, H., Smette, A., et al. 2009, A&A, 502, 457

García-Bernete, I., Ramos Almeida, C., Acosta-Pulido, J. A., et al. 2016, MNRAS, 463, 3531

Gebhardt, K., Bender, R., Bower, G., et al. 2000, ApJ, 539, L13

Gehrels, N. 1986, ApJ, 303, 336

Genzel, R., Eisenhauer, F., & Gillessen, S. 2010, Reviews of Modern Physics, 82, 3121

Genzel, R., Lutz, D., Sturm, E., et al. 1998, ApJ, 498, 579

Georgantopoulos, I. & Akylas, A. 2019, A&A, 621, A28

Georgantopoulos, I., Akylas, A., Georgakakis, A., & Rowan-Robinson, M. 2009, A&A, 507, 747

Georgantopoulos, I., Comastri, A., Vignali, C., et al. 2013, A&A, 555, A43

Georgantopoulos, I., Georgakakis, A., Rowan-Robinson, M., & Rovilos, E. 2008, A&A, 484, 671

Giacconi, R., Zirm, A., Wang, J., et al. 2002, ApJS, 139, 369

Giallongo, E., Grazian, A., Fiore, F., et al. 2019, ApJ, 884, 19

Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, ApJ, 600, L93

Gierliński, M. & Done, C. 2004, MNRAS, 349, L7

Gilli, R. 2013, Mem. Soc. Astron. Italiana, 84, 647

Gilli, R., Comastri, A., & Hasinger, G. 2007, A&A, 463, 79

Gilli, R., Vignali, C., Mignoli, M., et al. 2010, A&A, 519, A92

Glikman, E., Djorgovski, S. G., Stern, D., Bogosavljević, M., & Mahabal, A. 2007, ApJ, 663, L73

Glikman, E., Urrutia, T., Lacy, M., et al. 2012, ApJ, 757, 51

González, V., Labbé, I., Bouwens, R. J., et al. 2011, ApJ, 735, L34

Goulding, A. D., Alexander, D. M., Bauer, F. E., et al. 2012, ApJ, 755, 5

Grazian, A., Giallongo, E., Fiore, F., et al. 2020, ApJ, 897, 94

Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, A&A, 518, L3

Grimes, J. P., Heckman, T., Strickland, D., et al. 2007, ApJ, 668, 891

Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35

Gruppioni, C., Berta, S., Spinoglio, L., et al. 2016, MNRAS, 458, 4297

Gruppioni, C., Béthermin, M., Loiacono, F., et al. 2020, A&A, 643, A8

Gruppioni, C., Pozzi, F., Polletta, M., et al. 2008, ApJ, 684, 136

Gruppioni, C., Pozzi, F., Rodighiero, G., et al. 2013, MNRAS, 432, 23

Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, ApJS, 207, 24

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

Hainline, K. N., Shapley, A. E., Greene, J. E., & Steidel, C. C. 2011, ApJ, 733, 31

Harikane, Y., Inoue, A. K., Mawatari, K., et al. 2022, ApJ, 929, 1

Harrison, C. M., Alexander, D. M., Swinbank, A. M., et al. 2012, MNRAS, 426, 1073

Harrison, C. M., Costa, T., Tadhunter, C. N., et al. 2018, Nature Astronomy, 2, 198

Harrison, F. A., Aird, J., Civano, F., et al. 2016, ApJ, 831, 185

Hasinger, G., Capak, P., Salvato, M., et al. 2018, ApJ, 858, 77

Hayes, M., Melinder, J., Östlin, G., et al. 2016, ApJ, 828, 49

Herrnstein, J. R., Moran, J. M., Greenhill, L. J., & Trotter, A. S. 2005, ApJ, 629, 719

Hildebrandt, H., Erben, T., Dietrich, J. P., et al. 2006, A&A, 452, 1121

Hollenbach, D. J. & Tielens, A. G. G. M. 1999, Reviews of Modern Physics, 71, 173

Hopkins, P. F., Lidz, A., Hernquist, L., et al. 2007a, ApJ, 662, 110

Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007b, ApJ, 654, 731

Houck, J. R., Soifer, B. T., Weedman, D., et al. 2005, ApJ, 622, L105

Hsieh, B.-C., Wang, W.-H., Hsieh, C.-C., et al. 2012, ApJS, 203, 23

Hughes, D. H., Serjeant, S., Dunlop, J., et al. 1998, Nature, 394, 241

Ichikawa, K. & Tazaki, R. 2017, ApJ, 844, 21

Ichimaru, S. 1977, in Plasma physics: Nonlinear Theory and Experiments, 262

Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841

Inoue, A. K., Shimizu, I., Iwata, I., & Tanaka, M. 2014, MNRAS, 442, 1805

Ishigaki, M., Kawamata, R., Ouchi, M., et al. 2018, ApJ, 854, 73

Iwasawa, K., Gilli, R., Vignali, C., et al. 2012, A&A, 546, A84

Jaacks, J., Choi, J.-H., Nagamine, K., Thompson, R., & Varghese, S. 2012, MNRAS, 420, 1606

Jaffe, W., Meisenheimer, K., Röttgering, H. J. A., et al. 2004, Nature, 429, 47

Jaskot, A. E., Oey, M. S., Scarlata, C., & Dowd, T. 2017, ApJ, 851, L9

Jin, S., Daddi, E., Liu, D., et al. 2018, ApJ, 864, 56

Kaasinen, M., Scoville, N., Walter, F., et al. 2019, ApJ, 880, 15

Kaastra, J. S., Mewe, R., Liedahl, D. A., Komossa, S., & Brinkman, A. C. 2000, A&A, 354, L83

Kakkad, D., Mainieri, V., Padovani, P., et al. 2016, A&A, 592, A148

Kalberla, P. M. W., Burton, W. B., Hartmann, D., et al. 2005, A&A, 440, 775

Kaneda, H., Ishihara, D., Oyabu, S., et al. 2017, Publ. Astron. Soc. Australia, 34, e059

Kaspi, S., Maoz, D., Netzer, H., et al. 2005, ApJ, 629, 61

Kaspi, S., Smith, P. S., Netzer, H., et al. 2000, ApJ, 533, 631

Kawamuro, T., Ricci, C., Imanishi, M., et al. 2022, arXiv e-prints, arXiv:2208.03880

Kellermann, K. I. 1989, The Observatory, 109, 163

Kennicutt, Robert C., J. 1998, ARA&A, 36, 189

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36

Koptelova, E., Hwang, C.-Y., Malkan, M. A., & Yu, P.-C. 2019, ApJ, 882, 144

Kormendy, J. 1977, ApJ, 218, 333

Kormendy, J. & Ho, L. C. 2013, ARA&A, 51, 511

Kormendy, J. & Richstone, D. 1995, ARA&A, 33, 581

Krolik, J. H. & Begelman, M. C. 1988, ApJ, 329, 702

La Bella, N., Issaoun, S., Roelofs, F., Fromm, C., & Falcke, H. 2023, Expanding Sgr A* dynamical imaging capabilities with an African extension to the Event Horizon Telescope

La Caria, M. M., Vignali, C., Lanzuisi, G., Gruppioni, C., & Pozzi, F. 2019, MNRAS, 487, 1662

Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS, 224, 24

LaMassa, S. M., Yaqoob, T., & Kilgard, R. 2017, ApJ, 840, 11

Lamastra, A., Menci, N., Fiore, F., et al. 2013, A&A, 559, A56

Lansbury, G. B., Alexander, D. M., Aird, J., et al. 2017a, ApJ, 846, 20

Lansbury, G. B., Stern, D., Aird, J., et al. 2017b, ApJ, 836, 99

Lanzuisi, G., Civano, F., Elvis, M., et al. 2013, MNRAS, 431, 978

Lanzuisi, G., Civano, F., Marchesi, S., et al. 2018, MNRAS, 480, 2578

Lanzuisi, G., Delvecchio, I., Berta, S., et al. 2017, A&A, 602, A123

Lanzuisi, G., Gilli, R., Cappi, M., et al. 2019, ApJ, 875, L20

Lanzuisi, G., Ranalli, P., Georgantopoulos, I., et al. 2015, A&A, 573, A137

Lapi, A., Pantoni, L., Zanisi, L., et al. 2018, ApJ, 857, 22

Lapi, A., Raimundo, S., Aversa, R., et al. 2014, ApJ, 782, 69

Laporte, N., Meyer, R. A., Ellis, R. S., et al. 2021, MNRAS, 505, 3336

Le Fèvre, O., Béthermin, M., Faisst, A., et al. 2020, A&A, 643, A1

Le Fèvre, O., Lemaux, B. C., Nakajima, K., et al. 2019, A&A, 625, A51

Le Fèvre, O., Tasca, L. A. M., Cassata, P., et al. 2015, A&A, 576, A79

Leger, A., D'Hendecourt, L., & Defourneau, D. 1989, A&A, 216, 148

Lehmer, B. D., Xue, Y. Q., Brandt, W. N., et al. 2012, ApJ, 752, 46

Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, ApJS, 123, 3

Leroy, A. K., Bolatto, A., Gordon, K., et al. 2011, ApJ, 737, 12

Li, H., Dowell, C. D., Kirby, L., Novak, G., & Vaillancourt, J. E. 2008, Appl. Opt., 47, 422

Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, ApJS, 184, 218

Lilly, S. J., Le Fevre, O., Hammer, F., & Crampton, D. 1996, ApJ, 460, L1

Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70

Lin, Y.-H., Scarlata, C., Hayes, M., et al. 2022, MNRAS, 509, 489

Liu, T., Buchner, J., Nandra, K., et al. 2022, A&A, 661, A5

Luo, B., Bauer, F. E., Brandt, W. N., et al. 2008, ApJS, 179, 19

Lusso, E., Comastri, A., Simmons, B. D., et al. 2012, Monthly Notices of the Royal Astronomical Society, 425, 623

Lusso, E., Comastri, A., Vignali, C., et al. 2010, A&A, 512, A34

Lutz, D., Spoon, H. W. W., Rigopoulou, D., Moorwood, A. F. M., & Genzel, R. 1998, ApJ, 505, L103

Lutz, D., Sturm, E., Genzel, R., et al. 2003, A&A, 409, 867

Madau, P., Ferguson, H. C., Dickinson, M. E., et al. 1996, MNRAS, 283, 1388

Magnelli, B., Popesso, P., Berta, S., et al. 2013, A&A, 553, A132

Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, AJ, 115, 2285

Maiolino, R., Salvati, M., Bassani, L., et al. 1998, A&A, 338, 781

Malefahlo, E. D., Jarvis, M. J., Santos, M. G., et al. 2022, MNRAS, 509, 4291

Mancuso, C., Lapi, A., Prandoni, I., et al. 2017, ApJ, 842, 95

Mancuso, C., Lapi, A., Shi, J., et al. 2016a, ApJ, 833, 152

Mancuso, C., Lapi, A., Shi, J., et al. 2016b, ApJ, 823, 128

Manske, V., Henning, T., & Men'shchikov, A. B. 1998, A&A, 331, 52

Marchesi, S., Civano, F., Elvis, M., et al. 2016a, ApJ, 817, 34

Marchesi, S., Lanzuisi, G., Civano, F., et al. 2016b, ApJ, 830, 100

Marchesi, S., Zhao, X., Torres-Albà, N., et al. 2022, ApJ, 935, 114

Marques-Chaves, R., Schaerer, D., Álvarez-Márquez, J., et al. 2021, MNRAS, 507, 524

Mateos, S., Alonso-Herrero, A., Carrera, F. J., et al. 2012, MNRAS, 426, 3271

Mathews, W. G. & Capriotti, E. R. 1985, in Astrophysics of Active Galaxies and Quasi-Stellar Objects, ed. J. S. Miller, 185–233

Matthee, J., Feltre, A., Maseda, M., et al. 2022, A&A, 660, A10

McCammon, D., Almy, R., Apodaca, E., et al. 2002, ApJ, 576, 188

McGreer, I. D., Clément, B., Mainali, R., et al. 2018, MNRAS, 479, 435

McHardy, I. 1990, The Observatory, 110, 156

McLeod, D. J., McLure, R. J., & Dunlop, J. S. 2016, MNRAS, 459, 3812

McLure, R. J., Pentericci, L., Cimatti, A., et al. 2018, MNRAS, 479, 25

Meidinger, N., Nandra, K., & Plattner, M. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10699, Space Telescopes and Instrumentation 2018: Ultraviolet to Gamma Ray, 106991F

Meixner, M., Cooray, A., Leisawitz, D., et al. 2019, arXiv e-prints, arXiv:1912.06213

Mignoli, M., Feltre, A., Bongiorno, A., et al. 2019, Astronomy and Astrophysics, 626, A9

Mignoli, M., Vignali, C., Gilli, R., et al. 2013, A&A, 556, A29

Mor, R., Netzer, H., & Elitzur, M. 2009, ApJ, 705, 298

Mordini, S., Spinoglio, L., & Fernández-Ontiveros, J. A. 2021, arXiv e-prints, arXiv:2105.04584

Moretti, A., Vattakunnel, S., Tozzi, P., et al. 2012, A&A, 548, A87

Morisset, C., Delgado-Inglada, G., & Flores-Fajardo, N. 2015, Rev. Mex. Astron. Astrofis., 51, 103

Mushotzky, R. F., Done, C., & Pounds, K. A. 1993, ARA&A, 31, 717

Nair, P. B. & Abraham, R. G. 2010, ApJS, 186, 427

Nakagawa, T., Hayashi, M., Kawada, M., et al. 1998, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 3356, HII/L2 mission: future Japanese infrared astronomical mission, ed. P. Y. Bely & J. B. Breckinridge, 462–470

Nakagawa, T., Shibai, H., Onaka, T., et al. 2014, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9143, The next-generation infrared astronomy mission SPICA under the new framework, 91431I

Nandra, K., Barret, D., Barcons, X., et al. 2013, arXiv e-prints, arXiv:1306.2307

Nenkova, M., Ivezić, Ž., & Elitzur, M. 2002, ApJ, 570, L9

Nesvadba, N. P. H., Lehnert, M. D., De Breuck, C., et al. 2008, in Gas and Stars in Galaxies - A Multi-Wavelength 3D Perspective, 37

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, The Astrophysical Journal, 660, L43

Nonino, M., Dickinson, M., Rosati, P., et al. 2009, ApJS, 183, 244

Novak, M., Smolčić, V., Delhaize, J., et al. 2017, A&A, 602, A5

O'Dowd, M. J., Schiminovich, D., Johnson, B. D., et al. 2009, ApJ, 705, 885

Orr, M. E., Hayward, C. C., & Hopkins, P. F. 2019, MNRAS, 486, 4724

Otte, B., Murphy, E. M., Howk, J. C., et al. 2003, ApJ, 591, 821

Padovani, P., Alexander, D. M., Assef, R. J., et al. 2017, A&ARv, 25, 2

Page, M. J., Stevens, J. A., Ivison, R. J., & Carrera, F. J. 2004, ApJ, 611, L85

Paolillo, M., Schreier, E. J., Giacconi, R., Koekemoer, A. M., & Grogin, N. A. 2004, ApJ, 611, 93

Patrício, V., Richard, J., Verhamme, A., et al. 2016, MNRAS, 456, 4191

Peca, A., Vignali, C., Gilli, R., et al. 2021, ApJ, 906, 90

Perna, M., Brusa, M., Cresci, G., et al. 2015a, A&A, 574, A82

Perna, M., Brusa, M., Salvato, M., et al. 2015b, A&A, 583, A72

Peterson, B. 2006, The Broad-Line Region in Active Galactic Nuclei, ed. D. Alloin, R. Johnson, & P. Lira (Berlin, Heidelberg: Springer Berlin Heidelberg), 77–100

Pier, E. A. & Krolik, J. H. 1992, ApJ, 401, 99

Pogge, R. W. 1988, ApJ, 332, 702

Poglitsch, A., Waelkens, C., Geis, N., et al. 2010, A&A, 518, L2

Polletta, M., Weedman, D., Hönig, S., et al. 2008, ApJ, 675, 960

Pozzi, F., Vignali, C., Comastri, A., et al. 2010, A&A, 517, A11

Pozzi, F., Vignali, C., Comastri, A., et al. 2007, A&A, 468, 603

Primini, F. A. & Kashyap, V. L. 2014, ApJ, 796, 24

Prinja, R. K. & Howarth, I. D. 1986, ApJS, 61, 357

Raiter, A., Fosbury, R. A. E., & Teimoorinia, H. 2010, A&A, 510, A109

Reines, A. E., Greene, J. E., & Geha, M. 2013, ApJ, 775, 116

Reines, A. E. & Volonteri, M. 2015, ApJ, 813, 82

Reis, R. C. & Miller, J. M. 2013, ApJ, 769, L7

Retzlaff, J., Rosati, P., Dickinson, M., et al. 2010, A&A, 511, A50

Ricci, C., Ho, L. C., Fabian, A. C., et al. 2018, MNRAS, 480, 1819

Ricci, C., Trakhtenbrot, B., Koss, M. J., et al. 2017, Nature, 549, 488

Ricci, C., Ueda, Y., Koss, M. J., et al. 2015, ApJ, 815, L13

Riechers, D. A., Bradford, C. M., Clements, D. L., et al. 2013, Nature, 496, 329

Rodighiero, G., Daddi, E., Baronchelli, I., et al. 2011, ApJ, 739, L40

Rodriguez, L., Poglitsch, A., Aliane, A., et al. 2018, Journal of Low Temperature Physics, 193, 449

Rowan-Robinson, M., Mann, R. G., Oliver, S. J., et al. 1997, MNRAS, 289, 490

Rowan-Robinson, M., Oliver, S., Wang, L., et al. 2016a, MNRAS, 461, 1100

Rowan-Robinson, M., Oliver, S., Wang, L., et al. 2016b, MNRAS, 461, 1100

Saade, M. L., Brightman, M., Stern, D., Malkan, M. A., & García, J. A. 2022, ApJ, 936, 162

Sandage, A. 1965, ApJ, 141, 1560

Sanders, D. B., Mazzarella, J. M., Kim, D. C., Surace, J. A., & Soifer, B. T. 2003, AJ, 126, 1607

Sanders, D. B., Phinney, E. S., Neugebauer, G., Soifer, B. T., & Matthews, K. 1989, ApJ, 347, 29

Schaerer, D., Ginolfi, M., Béthermin, M., et al. 2020, A&A, 643, A3

Schartmann, M., Meisenheimer, K., Camenzind, M., Wolf, S., & Henning, T. 2005, A&A, 437, 861

Schinnerer, E., Sargent, M. T., Bondi, M., et al. 2010, ApJS, 188, 384

Schmidt, M., Hasinger, G., Gunn, J., et al. 1998, A&A, 329, 495

Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, Astronomy and Astrophysics, 575, A74

Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1

Scoville, N., Aussel, H., Sheth, K., et al. 2014, ApJ, 783, 84

Seyfert, C. K. 1943, ApJ, 97, 28

Shakura, N. I. & Sunyaev, R. A. 1973, A&A, 500, 33

Shankar, F., Bernardi, M., Sheth, R. K., et al. 2016, MNRAS, 460, 3119

Shankar, F., Weinberg, D. H., & Miralda-Escudé, J. 2013, MNRAS, 428, 421

Shi, Y., Rieke, G. H., Hines, D. C., et al. 2006, ApJ, 653, 127

Sijacki, D., Vogelsberger, M., Genel, S., et al. 2015, MNRAS, 452, 575

Silk, J. & Rees, M. J. 1998, A&A, 331, L1

Simmonds, C., Buchner, J., Salvato, M., Hsu, L. T., & Bauer, F. E. 2018, A&A, 618, A66

Smith, J. D. T., Draine, B. T., Dale, D. A., et al. 2007, ApJ, 656, 770

Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15

Spinoglio, L., Alonso-Herrero, A., Armus, L., et al. 2017, Publ. Astron. Soc. Australia, 34, e057

Spinoglio, L. & Malkan, M. A. 1992, ApJ, 399, 504

Spinoglio, L., Mordini, S., Fernández-Ontiveros, J. A., et al. 2021, Publ. Astron. Soc. Australia, 38, e021

Stalevski, M., Fritz, J., Baes, M., Nakos, T., & Popovic, L. C. 2012, Publications de l'Observatoire Astronomique de Beograd, 91, 235

Stalevski, M., Ricci, C., Ueda, Y., et al. 2016, MNRAS, 458, 2288

Stanway, E. R. & Eldridge, J. J. 2018, MNRAS, 479, 75

Stark, D. P., Richard, J., Siana, B., et al. 2014, MNRAS, 445, 3200

Stenholm, L. 1994, A&A, 290, 393

Sturm, E., Hasinger, G., Lehmann, I., et al. 2006, ApJ, 642, 81

Suh, H., Civano, F., Trakhtenbrot, B., et al. 2020, ApJ, 889, 32

Swinyard, B., Nakagawa, T., Merken, P., et al. 2009, Experimental Astronomy, 23, 193

Tacchella, S., Dekel, A., Carollo, C. M., et al. 2016, MNRAS, 458, 242

Tacconi, L. J., Genzel, R., & Sternberg, A. 2020, ARA&A, 58, 157

Talia, M., Cimatti, A., Giulietti, M., et al. 2021, ApJ, 909, 23

Tang, M., Stark, D. P., Chevallard, J., et al. 2021, MNRAS, 501, 3238

Teplitz, H. I., Armus, L., Soifer, B. T., et al. 2006, ApJ, 638, L1

Terashima, Y. 2005, The Nature of Far Infrared Selected Type 2 QSOs, XMM-Newton Proposal

Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, ApJ, 621, 673

Thomas, N., Davé, R., Anglés-Alcázar, D., & Jarvis, M. 2019, Monthly Notices of the Royal Astronomical Society, 487, 5764–5780

Toba, Y., Bae, H.-J., Nagao, T., et al. 2017, ApJ, 850, 140

Tombesi, F., Cappi, M., Reeves, J. N., et al. 2013, MNRAS, 430, 1102

Tombesi, F., Cappi, M., Reeves, J. N., et al. 2010, A&A, 521, A57

Tommasin, S., Spinoglio, L., Malkan, M. A., & Fazio, G. 2010, ApJ, 709, 1257

Tommasin, S., Spinoglio, L., Malkan, M. A., et al. 2008, ApJ, 676, 836

Tozzi, P., Gilli, R., Mainieri, V., et al. 2006, A&A, 451, 457

Treister, E., Schawinski, K., Urry, C. M., & Simmons, B. D. 2012, ApJ, 758, L39

Treister, E., Urry, C. M., & Virani, S. 2009, ApJ, 696, 110

Übler, H., Maiolino, R., Curtis-Lake, E., et al. 2023, arXiv e-prints, arXiv:2302.06647

Ueda, Y., Akiyama, M., Hasinger, G., Miyaji, T., & Watson, M. G. 2014, ApJ, 786, 104

Urry, C. M. & Padovani, P. 1995, PASP, 107, 803

van Dokkum, P., Brammer, G., Momcheva, I., Skelton, R. E., & Whitaker, K. E. 2013, arXiv e-prints, arXiv:1305.2140

Vanzella, E., Cristiani, S., Dickinson, M., et al. 2006, A&A, 454, 423

Vanzella, E., Grazian, A., Hayes, M., et al. 2010, A&A, 513, A20

Vanzella, E., Nonino, M., Cupani, G., et al. 2018, MNRAS, 476, L15

Vietri, G., Misawa, T., Piconcelli, E., et al. 2022, arXiv e-prints, arXiv:2205.06832

Vignali, C. 2014, in IAU Symposium, Vol. 304, Multiwavelength AGN Surveys and Studies, ed. A. M. Mickaelian & D. B. Sanders, 132–138

Vignali, C., Alexander, D. M., & Comastri, A. 2006, MNRAS, 373, 321

Vignali, C., Alexander, D. M., Gilli, R., & Pozzi, F. 2010, MNRAS, 404, 48

Vignali, C., Mignoli, M., Gilli, R., et al. 2014, A&A, 571, A34

Vignali, C., Piconcelli, E., Lanzuisi, G., et al. 2011, MNRAS, 416, 2068

Vignali, C., Pozzi, F., Fritz, J., et al. 2009, MNRAS, 395, 2189

Vito, F., Brandt, W. N., Yang, G., et al. 2018, MNRAS, 473, 2378

Vito, F., Gilli, R., Vignali, C., et al. 2014, MNRAS, 445, 3557

Vito, F., Mignoli, M., Gilli, R., et al. 2022, A&A, 663, A159

Voit, G. M. 1992, MNRAS, 258, 841

Volonteri, M., Dubois, Y., Pichon, C., & Devriendt, J. 2016, MNRAS, 460, 2979

Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, ApJS, 258, 11

Weedman, D., Polletta, M., Lonsdale, C. J., et al. 2006, ApJ, 653, 101

Wiklind, T., Dickinson, M., Ferguson, H. C., et al. 2008, ApJ, 676, 781

Wilson, A. S. & Tsvetanov, Z. I. 1994, AJ, 107, 1227

Wu, Y., Charmandaris, V., Huang, J., Spinoglio, L., & Tommasin, S. 2009, ApJ, 701, 658

Wuyts, S., Labbé, I., Förster Schreiber, N. M., et al. 2008, ApJ, 682, 985

Wyder, T. K., Treyer, M. A., Milliard, B., et al. 2005, ApJ, 619, L15

Xue, Y. Q. 2017, New Astronomy Review, 79, 59

Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2011, ApJS, 195, 10

Yan, L., Sajina, A., Fadda, D., et al. 2007, ApJ, 658, 778

Yang, G., Boquien, M., Brandt, W. N., et al. 2022, ApJ, 927, 192

Yang, G., Boquien, M., Buat, V., et al. 2020, MNRAS, 491, 740

Zakamska, N. L., Hamann, F., Pâris, I., et al. 2016, MNRAS, 459, 3144

Zanella, A., Daddi, E., Magdis, G., et al. 2018, MNRAS, 481, 1976

Zappacosta, L., Piconcelli, E., Duras, F., et al. 2018, A&A, 618, A28