UNIVERSITÀ
DEGLI STUDI
FIRENZE

PhD in
Physics and Astronomy

CYCLE XXXVI

COORDINATOR Prof. Giovanni Modugno

AGN as cosmological probes

Academic Discipline (SSD) FIS/05

Doctoral Candidate	**Supervisor**
Dr. Matilde Signorini	Prof. Guido Risaliti

Coordinator
Prof. Giovanni Modugno

The consultation of the thesis is free. Unless a specific authorization is obtained from the author, the thesis can be, however, downloaded and printed only for strictly personal purposes related to study, research and teaching, with the explicit exclusion of any use that has – even indirectly – a commercial nature.

[...] invece quest'immagine, per chissà quale disguido,
arriva a me che temo che sia troppo bella per essere vera,
troppo accetta al mio universo immaginario per appartenere
al mondo reale. Ma forse è proprio questa diffidenza verso
i nostri sensi che ci impedisce di sentirci a nostro agio
nell'universo. Forse la prima regola che devo pormi è questa:
attenermi a ciò che vedo.
Palomar guarda il cielo, Italo Calvino, 1983

Contents

Outline

Active Galactic Nuclei (AGN) are exceptionally powerful sources, characterised by the radiation coming from the accretion onto a Supermassive Black Hole (SMBH) at the centre of a galaxy. Their luminosity ranges from 10^{40} to 10^{48} erg s^{-1}, and can exceed that of the hole stellar emission from the galaxy they are found in. They are observed at all cosmic epochs, from the local Universe to up to when it was less than a billion years old, and they are very numerous. These properties makes them very interesting objects for cosmology.
AGN implementation in cosmology has two sides: first of all,, their emission can be used as a tool to measure cosmological distances and to determine the expansion rate of the Universe. At the same time, AGN emission is intricately linked to the SMBH at their core. Gaining a comprehensive understanding of SMBH formation and accretion history is pivotal to grasping the history of structure formation of the Universe, and the co-evolution of SMBH with galaxies.
In this thesis, we investigate both of these two ways in which AGN can be used to enhance our comprehension of the Universe and its history. Additionally, we probe various aspects of the physical and emission properties of AGN. Despite our current knowledge, there is still much to be understood about these objects. A deeper insight into them not only broadens our understanding but also equips us to harness them more effectively in cosmological studies.
In Chapter 1, we introduce AGN properties and the “Unified model” that described them. Then, and we present current open issues in cosmology, from the understanding of SMBH accretion to the testing of the flat ΛCDM model, where the implementation and/or the study of AGN can provide a significant help.
In Chapter 2, we discuss the use of quasars as standard candles, which is, to derive cosmic distances and from them estimate the expansion rate of the Universe at different epochs. This can be done thanks to the presence of the $L_X - L_{UV}$ relation in quasars. We sum up the most recent results in the literature, describing how removing biased objects from quasars sample has made it possible to derive (more) precise distance measurements. Moreover, we show that the fit of the Hubble diagram of quasars and Supernovae Ia shows a significant (4σ) tension with the predictions of the standard flat ΛCDM model. The significance of the tension shows the relevance fo quasars implementation for testing cosmological models..
In Chapter 3, we delve into the new results presented in this thesis. We investigate the $L_X - L_{UV}$ relation for the first time with a sample whose UV observational properties have been derived spectroscopically, and not photometrically. We aim at (i) obtaining more precise distance measurements with quasars, and (ii) understanding more of the physics behind the relation, which is relevant both for our understanding of AGN physics in general

and for the cosmological implementation of the relation.
In Chapter 4, we go on with our analysis of the $L_X - L_{UV}$ relation, investigating those factors that contribute to its observed dispersion but that cannot be removed with the sample selection. We study the contributions of quasars emission variability, inclination with respect to the line of sight, and the use of photometric X-ray data instead of spectroscopic data. We compare our estimates of these contributions with the most recent results in the literature, strengthening the belief that, although we still do not know the exact mechanism behind the $L_X - L_{UV}$ relation, it must be very tight, with an intrinsic dispersion close to zero, and be the same on a wide range of luminosities.
In Chapter 5, we implement quasars for cosmology and study the Hubble Diagram of quasars and Supernovae Ia, presenting an analysis derived using Neural Networks. The goal is to derive the shape of the Hubble Diagram in a complete non-parametric way. We compare the results obtained with thi method with the standard flat ΛCDM cosmology, confirming the presence of a strong tension, that hints towards Interacting Dark Sector models.
In Chapter 6, we discuss reverberation mapping, a technique which allows us to estimate SMBH masses and to characterize physical and geometrical properties of AGN emission. we discuss the H_0 tension and show how reverberation mapping and interferometry can allow us to use AGN at "standard rulers", providing distance estimates and an independent measurement of the H_0 constant. Moreover, we investigate how more precise mass measurements could be obtained at high redshift, with the "case study" of a lensed quasar at redshift z=2.8.
In Chapter 7, we discuss a study of the obscured fraction of AGN in the J1030 field, an X-ray deep field obtained with the *Chandra* telescope. The characterization of AGN obscuration properties is relevant for our understanding of galaxy evolution, of SMBH formation and accretion, and for the understanding of the X-ray Cosmic Background. We find evidence for an increased obscured fraction at high redshift compared to local measurements, and we compare out results with models that try to understand whether the increasing obscuration is due to changing properties of the AGN torus and/or to changing properties in the host galaxies.
In Chapter 8 we summarize our results and outline possible future work.

Chapter 1

Introduction

Active Galactic Nuclei (AGN) represent some of the most dynamic and energetic phenomena in the universe. Situated at the cores of galaxies, these luminous regions are powered by supermassive black holes (SMBHs) accreting material at prodigious rates. Doing so, they release vast amounts of energy, which makes them observable across a wide range of wavelengths and even at the earliest cosmological epochs. The interplay between AGN and their host galaxies is intricate, with feedback mechanisms that can influence both the growth of the black hole and the evolutionary path of the galaxy itself. Understanding AGN is fundamental for shedding light on the formation and accretion history of SMBH and on galaxy evolution. Furthermore, they can be used for the measurement of cosmological distances and therefore to test different cosmological models. In this Chapter, we introduce the main properties of AGN and of the standard cosmological model.

1.1 Active Galactic Nuclei

Among galaxies, 1-10% are thought to be “Active”, that is, they contain an Active Galactic Nucleus (AGN): a central region characterised by emission with intensity, spectral, and timing properties that cannot be accounted for by the standard components of a galaxy (stars, interstellar gas, dust). Among AGN, those whose emission surpasses that of the host galaxy are commonly labelled as “Quasars” (Quasi Stellar Objects, or QSOs). Quasars are the most luminous persistent sources in the Universe, observed up to a redshift of z~7, a time when the Universe was less than 1 billion years old.

AGN exhibit the following general properties:

(i) Extremely high luminosities, ranging from 10^{11} to a few $10^{14} L_{\odot}$, originating from spatially unresolved regions with a physical extension of less than one parsec.

(ii) A spectral energy distribution (SED) that spans the entire electromagnetic spectrum, from radio to γ wavelengths, which cannot be explained by a mere superposition of stellar spectra. A typical quasar SED is depicted in Figure 1.1.

(iii) Around 10% of exhibit relativistic jets, which are emitted perpendicular to the galactic plane, can extend up to $\sim$1 Mpc, and are potent radio emitters ($10^8 - 10^{10} L_{\odot}$).

(iv) Spectral emission lines are present across all wavelengths, with widths surpassing those typical of standard galaxies. Permitted lines[1]exhibit Full Width at Half Maximum

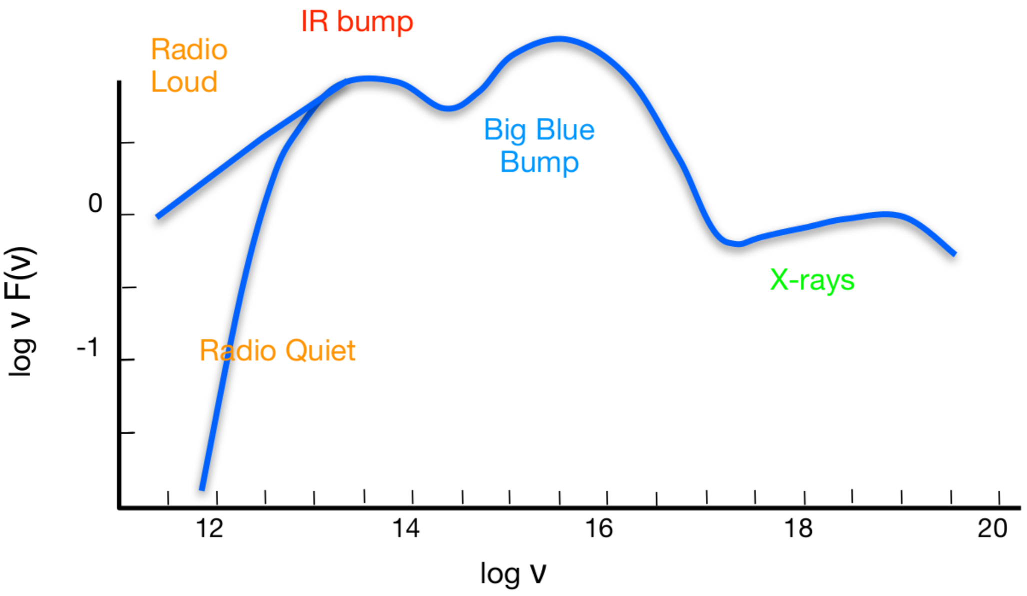

Figure 1.1: Typical Spectral Energy Distribution (SED) for an AGN. Key features include (i) the emission peak in the UV band, termed the “Big Blue Bump” , (ii) the secondary peak in the IR band, the “IR-bump”, resulting from the reprocessing of primary emission by dust, (iii) notable X-ray band emission, significantly greater than what’s found in non-active galaxies, and (iv) radio frequency emissions, differentiating between “Radio Loud” and “Radio Quiet” sources based on the radio emission intensity.

(FWHM) ranging from 10^3 to 10^5 km/s, while forbidden lines display FWHM up to 1000 km/s.
(v) Variability is observed in both the continuum and line emissions, with time scales ranging from hours to decades.

Stellar emission alone cannot account for these features. The high observed luminosities, combined with the brief emission variability time scales, suggest that if these were produced by stellar emission, a stellar mass of $\sim 10^{10} M_\odot$ would be required within a region smaller than 1 parsec. This scenario is physically implausible as such a system would collapse into a Black Hole (BH).
The sole known physical process capable of producing such potent emission is matter accretion onto a Black Hole. Here, the intense gravitational potential enables a high efficiency in mass-to-energy conversion, reaching up to $\epsilon \sim 28\%$ of the rest energy, mc^2, in contrast to $\epsilon \sim 0.7\%$ for nuclear fusion processes.
Thus, AGN emission can be explained with a Supermassive Black Hole ($10^6 - 10^9 M_\odot$)

---

[1] The terms “permitted” and “forbidden” lines denote emission lines related to electronic transitions either allowed or disallowed by the electric dipole rule, respectively. Consequently, permitted lines result from transitions from states with short lifetimes, whereas forbidden transitions relate to metastable states with lengthy lifetimes, meaning they are produced at a much reduced rate. This means that forbidden lines aren’t usually observed in high-density settings since, in such environments, collisional recombinations are more probable than radiative ones. A “critical density” can typically be defined, above which forbidden lines are not observed.

encircled by an accretion disc, where gravitational energy is the primary emission source. The observed quasar SED is intricate, considering both the emitted continuum shape and the presence of emission lines. Its characteristics are explained not solely through the accretion process but also via the interactions between the primary emission, emanating from the accretion disc, and the circumnuclear medium.

1.1.1 Classification and the Unified Model

AGN are typically classified based on their observational properties. However, most variations in these properties are attributed to the inclination angle of the source relative to the line of sight, rather than genuine physical differences. Here, we outline a concise taxonomy of the primary types of AGN, largely derived from Peterson 1997:
- Seyfert Galaxies: First classified by Carl Seyfert in 1943 (Seyfert, 1943), these objects are moderately luminous with bolometric luminosities $L_{bol} \lesssim 10^{41}$erg/s. Typically, the host galaxy is discernible and is often a spiral galaxy. Their spectra feature prominent high-ionization emission lines. Based on their emission line characteristics, they are further divided into:

  - Type-I Seyfert galaxies: Both broad and narrow lines are present.

  - Type-II Seyfert galaxies: Only narrow lines are evident.

- Quasars (or QSOs): These represent the most luminous AGN, with luminosities $L_{bol} \sim 10^{44} - 10^{48}$ erg/s. Their spectra resemble those of Seyfert galaxies, but the main distinction lies in the host galaxy, which is scarcely resolved in quasars.
- Blazars: They exhibit highly variable emissions across the electromagnetic spectrum and have jets that are oriented close to the line of sight towards Earth.
- Radio Loud Galaxies: AGN can be categorised as either radio-loud or radio-quiet based on the presence or absence of substantial radio emission. The “radio loudness” parameter is defined as $R = L_\nu(5GHz)/L_\nu(4400\text{Å})$, where L_ν represents the monochromatic luminosity at the specified frequency. Quasars with $R > 10$ are considered radio loud.
- Similarly, we can distinguish between γ-loud and γ-quiet quasars. Typically, γ-loud sources (comprising about 10% of the total AGN population) are also radio-loud.

Originally, this observational diversity resulted in an expansive taxonomy. However, from the 1980s onwards, the notion of a “unified model” began to gain traction. Observations of the Seyfert-II galaxy NGC1068 in polarised light unveiled broad line features akin to those in the Seyfert-I category. This discovery implied that Type-I and Type-II galaxies might essentially be identical, with obscuration by dust in Type-II galaxies allowing only narrow lines to be visible. This revelation paved the way for the “unified model”.
According to this model, all AGN consist of a gaseous disc accreting onto a central Supermassive Black Hole. During the accretion process, viscous forces transform potential energy into radiation, causing the disc to emit a superposition of black body spectra, peaking in the UV band. Surrounding the central region, gas and dust clouds move in Keplerian orbits around the SMBH. These clouds constitute the “Broad Line Region” (BLR)

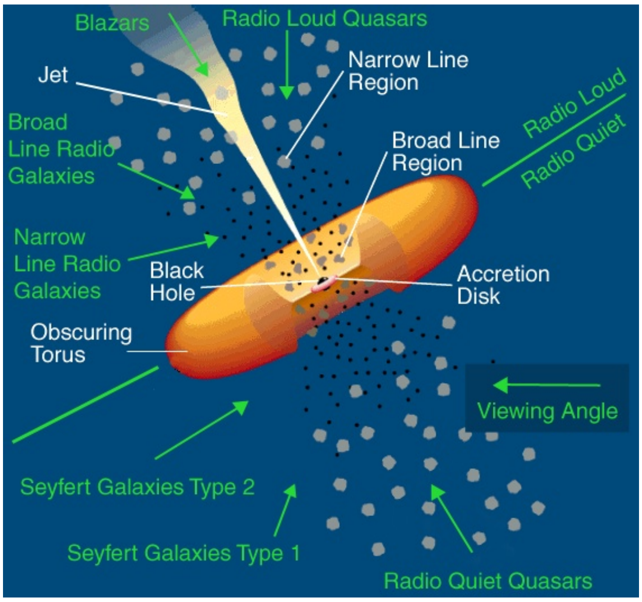

Figure 1.2: Unified model for Active Galactic Nuclei. The varied observational characteristics among AGN classes arise from the different inclination angles of the source relative to the observer. The presence or absence of the "torus" along the line of sight determines if the Broad Line Region is visible. The Narrow Line Region, situated further from the centre, always produces observable narrow lines. The presence or absence of a jet distinguishes radio loud from radio quiet objects.

and "Narrow Line Region" (NLR), where the primary emission undergoes reprocessing, resulting in pronounced emission lines. Cloud closer to the central SMBH move faster, resulting in a significant broadening of the lines (BLR), while further away could show less-broadened lines (NLR).

Encircling the central region is a dusty "torus" composed of molecular gas and warm dust (T$\sim 10^2 - 10^3$K). Though often depicted as a uniformly dense entity for simplification (as illustrated in Figure 1.2), observational data suggest it likely comprises clumps. Depending on the sightline angle, the torus can partially or entirely obscure the central AGN emission, leading to the observational distinctions mentioned previously. In particular, the BLR is thought to lie inside of the torus. Therefore, the torus can obscure the emission coming from this region and from the central disc. The NLR clouds are instead thought to be found further away from the BH. Therefore, the NLR emission is always observed. Objects were the central disc and BLR emission is blocked by the torus (partially or totally) are referred to as "obscured".

AGN show a strong X-ray emission, up to $\sim$10% of their bolometric luminosity. The accretion disc itself does not reach temperatures high enough to explain the X-ray intensity and spectral shape. Therefore, it is thought that X-ray emission comes from a region called the "Corona", which is made of a heated electron plasma positioned above the disc. UV photons coming from the disc interact with the hot electrons with Inverse Compton scattering, producing X-ray emission.

In radio-loud AGN, relativistic jets are propelled along the disc axis, emitting at radio wavelengths through synchrotron radiation and X-rays via Inverse Compton scattering.

1.2 AGN and the host galaxy

There is now abundant observational evidence of the fact that a connection exists between the AGN and its host galaxy. First of all, whilst AGNs are only found in approximately 10% of galaxies, SMBH are found to be present at the centre of all galaxies, quiescent ones included. This suggests that AGN activity might be a universal phase for all galaxies. In this context, SMBHs in non-active galaxies can be viewed as remnants of previous activity (Marconi et al., 2004). Secondly, both the space density of luminous quasars and the star formation rate of galaxies increase with the redshift, peak around redshift 2 and then decline. Observational studies of distant galaxies have pieced together the cosmic history of star formation. This history suggests a “cosmic noon” around redshift 2 when the universe was about 3 billion years old. During this epoch, the rate of star formation in the universe was at its peak. Similarly, observations have shown that AGN activity, as traced by quasar luminosity functions and other AGN diagnostics, also peaks around the same redshift. This implies that the processes feeding SMBHs and triggering star formation might be closely linked. The concurrent rise and subsequent fall of both SFR and SMBH accretion rates strengthen the co-evolution hypothesis, which suggests that galaxies and their central black holes evolve together, influencing each other’s growth. For example, gas-rich mergers or interactions can both stimulate star formation and funnel gas towards the central black hole, leading to AGN activity.

In addition to this, significant correlations are evident between the properties of the AGN (in particular, the SMBH mass) and those of the host galaxy: correlations exist between the SMBH mass and the bulge mass, the SMBH mass and the stellar velocity dispersion, and the SMBH mass and the galaxy luminosity. This may suggest a link between the evolution of the AGN and its host galaxy.

However, the understanding of the origins of this connection remains limited. The proposed mechanism underpinning this relationship is the AGN “feedback”. High-energy winds and jets from the vicinity of the SMBH can potentially heat or expel the cold gas in the host galaxy, suppressing star formation. This process can help explain why massive galaxies with powerful AGNs do not show high star formation rates, and it is called “negative feedback” (Silk & Rees, 1998; Fabian, 2012; Harrison, 2017). At the same time, AGN activity can also potentially trigger star formation (”positive feedback”). The outflows or jets from AGN, when interacting with the interstellar medium, can compress gas, leading to the formation of stars (Silk, 2013; Zubovas et al., 2013). Recently, star formation has also been detected inside gas outflows, implying that the outflows can trigger star formation (Maiolino et al., 2017; Gallagher et al., 2019).

Whilst the precise reasons for the robust association of AGN with host galaxies remain somewhat mysterious, their existence is undeniable and holds significant implications for the context of SMBH mass estimation. As will be explored in Chapter 4, the direct mass determination of is challenging and time-intensive. Hence, the relationships between SMBH masses and host galaxy attributes are invaluable, providing calibration for mass measurements across large object samples. Concurrently, it is essential to investigate these relationships across varying cosmological epochs and diverse object types to determine their consistency or evolution.

1.3 SMBH formation and growth

Observational evidence of quasars at high-redshift suggest that SMBH with masses equal or higher than 10^9 $M_\odot$ formed in the first billion years after the Big Bang. These object are very rare, with a number density around $\sim$1 Gpc^{-3}; nevertheless, it is challenging to understand how these SMBH could form and grow to such immense sizes in such a short cosmological timescale. The question had already been posed when quasars were discovered at redshift $z > 4$ (Turner, 1991), but it gets more and more intriguing as objects at higher and higher redshift are observes.

Several potential formation mechanism have been proposed and they differ in both the origin of the SMBH “seeds” and in the most relevant accretion mechanism.

Regarding the SMBH seeds, the debate has been around whether SMBH start with “light seeds”, generally though to derive from the death of the first, very massive, stars, or with “heavy seeds”, usually thought to be formed with the direct collapse of massive gas clouds. The first generation of stars, known as Pop III stars, are thought to be massive — between tens and thousands of solar masses — and, therefore, lived short lives. They are expected to form in $\sim$ $10^{5-6}\mathrm{M}_\odot$ dark matter minihalos, thanks to the H_2 cooling of primordial gas, and that the inefficiency of the cooling leads to inefficient fragmentation, making these stars more massive than the later populations. Their deaths could create black hole seeds, which could then merge and accrete material to become SMBHs. However, the initial seed masses from these remnants would be quite small, and it would need to constantly grow with an efficiency near the Eddington limit for $\sim$0.8 Gyr to reach $10^9\mathrm{M}_\odot$.

Several scenarios have been proposed to try and address this timescale issue. Usually, the options are to increase the seed BH mass, or to increase the accretion rate. One proposition is that some early Pop III seeds maintain Eddington accretion throughout cosmic history, though this is deemed unlikely for stellar-mass seeds in minihalos because of feedback processes. Nevertheless, it may be feasible in rare massive halos forming at high redshifts (Tanaka, 2014). Another angle is the acceleration of black hole assembly via mergers, mostly with other black holes. Yet, this idea encounters challenges because BH mergers might cause gravitational wave-induced kicks, expelling BHs from their dense gas reservoirs (Haiman, 2004).

A popular hypothesis postulates the formation of $\sim$ $10^5\mathrm{M}_\odot$) “massive seed” BHs through the rapid collapse of pristine primordial gas in atomic-cooling halos (ACHs). These processes could lead to the birth of a supermassive star, which subsequently collapses to form a BH of similar mass. Crucial factors enabling this include the gas staying warm, thus avoiding efficient metal or H2 cooling and fragmentation. Several conditions, such as exposure to intense Lyman–Werner (LW) radiation which suppresses H2 cooling (e.g., Omukai, 2001; Fernandez et al., 2014), or heating from rapid halo mergers (Yoshida et al., 2003; Wise et al., 2019) have been posited to sustain this thermodynamical state.

Another thing to keep in mind is that current surveys of distant quasars predominantly detect exceptionally bright and massive black holes, accreting near the Eddington limit. SMBH at $z \sim 7$ with 10^9 $M_\odot$ are rare, atypical of the general massive BH population, and they reside in large, evolved galaxies that originated in particularly dense universe regions. While comprehending these extreme objects is essential, it is crucial to understand the broader, yet obscured, massive BH population, which remains concealed due to their smaller sizes or accretion rates. Obscuration, from gas and dust near the BH or in the host galaxy, hides many early BH (Buchner & Bauer, 2017). Understanding this hidden BH population will enhance our grasp on BH and galaxy formation.

To sum up, our comprehension of SMBH formation and growth is still far form complete. The brightness of AGN allows us to detect and characterise these objects at higher and higher redshift, helping us understand the structure formation in the Universe history. Together with more observational evidence for high redshift objects, two things are needed for a better understanding of this subject: first, we need reliable, precise SMBH mass measurements for AGN at high redshift. This will be discussed in Chapter 6. Secondly, we need to understand better the obscured population of AGN and how it evolves with the redshift, if we want to understand SMBH evolution and go beyond the “tip of the iceberg”. This will be discussed in Chapter 7.

1.4 Cosmological framework

In this Section, we introduce the cosmological tools that will we used later in this thesis. We provide a brief description of the standard flat ΛCDM model, and on distance measurements. We then describe how quasars can be used as standard candles to extend the Hubble Diagram beyond what is possible with Supenovae Ia (SNIa).

1.4.1 The ΛCDM model

The flat ΛCDM cosmological model is currently the most compelling and comprehensive paradigm describing the Universe large-scale structure and its evolution over time. One of its strengths lies in its ability to encapsulate a wide array of astronomical observations with just a few core principles, showcasing elegance and explanatory power. Together with being geometrically flat, at the heart of the model are two primary constituents:

- Cold Dark Matter (CDM): This form of matter, undetectable by its electromagnetic radiation, is believed to play a pivotal role in the gravitational framework of the cosmos. Unlike “hot” dark matter, CDM moves slowly compared to the speed of light, allowing it to clump and form the basic structures of the Universe. These structures later serve as the gravitational wells where galaxies and galaxy clusters form. It is termed “cold” because of the non-relativistic speeds of the particles.
- Cosmological Constant, Λ: It represents the energy density of the vacuum of space, commonly associated with dark energy. Observational evidence, particularly from Supernovae Ia, has pointed towards an accelerated expansion of the Universe, a phenomenon this constant helps explain. This acceleration suggests a repulsive form of energy, counteracting the attractive nature of gravity at cosmological distances.

The ΛCDM model has become the leading cosmological paradigm thanks to diverse observational evidences. Cosmic Microwave Background (CMB) radiation, exhibits minute temperature fluctuations that have been mapped by missions like WMAP and Planck. These fluctuations fit remarkably well with predictions arising from a ΛCDM universe, particularly the model power spectrum. In addition, large-scale structure surveys that map the distribution of galaxies across vast cosmic volumes highlight the characteristic patterns expected from cold dark matter gravitational influences. Supernovae Ia observations have unveiled an accelerated cosmic expansion, pointing towards the repulsive influence of dark energy, consistent with the cosmological constant Λ. Moreover, the observed abundance of light elements, a product of Big Bang nucleosynthesis, concurs with predictions made under this model. Lastly, gravitational lensing studies further corroborate the existence of unseen mass, giving credence to the dark matter component. The collective weight of these observations cements the ΛCDM model position as the "standard model" for cosmology.

1.4.2 Distance measurements

We define "proper distance" $d_p(t)$ the length of the geodetic that connects wo distinct points when the scale factor of the Universe has value $a(t)$. For an object that emits at time t_e and is observed at time t_0, we have

$$D_p(t_0) = c\int_{t_e}^{t_0} dt \tag{1.1}$$

The *comoving distance*, instead, is the distance between two objects that remains constant in time when they are moving with the Hubble flow. It is the proper distance rescaled by the scale factor of the Universe at the emission time or, in other terms, $D_c = D_p(t_0)\cdot(1+z)$. Proper and comoving distances are tightly related to cosmological parameters, as we can also write D_c as

$$D_c = \frac{c}{H_0}\int_0^z \frac{dz'}{E(z')} \tag{1.2}$$

where $E(z)$ is the expression for the first Freedman equation:

$$E(z) = \frac{H(z)}{H_0} = [\Omega_{m,0}(1+z)^3 + \Omega_{r,0}(1+z)^4 + \Omega_{k,0}(1+z)^2 + \Omega_{\Lambda,0}]^{-1/2} \tag{1.3}$$

with $\Omega_{k,0} = -\frac{c^2}{R^2H_0^2} = 1 - \Omega_{m,0} - \Omega_{r,0} - \Omega_{\Lambda,0}$. $\Omega_{k,0}$ is the curvature parameter: in a flat geometry, $\Omega_{k,0} = 0$, in an open geometry $\Omega_{k,0} > 0$ and in a closed geometry $\Omega_{k,0} < 0$. So the way in which $D_c(z)$ varies with the redshift of a given source is linked to the values of the cosmological parameters found in the expression for $E(z)$. In the flat-ΛCDM frame, $\Omega_{k,0} = 0$ and the Universe is flat.

We can also define the transverse comoving distance D_M, which is used to determine the comoving distance $D_{1,2}$ between two events at the same redshift but separated by an angle $\delta\theta$. We have that

$$D_{12} = D_M\delta\theta \tag{1.4}$$

$$D_M = \begin{cases} \frac{1}{\sqrt{\Omega_k}} \sinh(c\sqrt{\Omega_k} D_c/H_0) & \text{if } \Omega_k > 0 \\ D_c & \text{if } \Omega_k = 0 \\ \frac{1}{\sqrt{|\Omega_k|}} \sin(c\sqrt{|\Omega_k|} D_c/H_0) & \text{if } \Omega_k < 0 \end{cases}$$

so that transverse comoving distance changes with the space-time geometry.
We see that the defined distances are related only to the cosmological parameters and to the source redshift. This means that for a given object, once its redshift is known, we can derive its proper or comoving distance by assuming a certain cosmological model. However, in a cosmological context, we are interested in the reverse mechanism: we want a way to use observations to *test* a given cosmological model and to understand which model better reproduces the observational data. Unfortunately, all the described quantities can not be directly measured. We can go around this issue by using “standard rulers” or “standard candles”, which are sources that allow us to determine the angular diameter distance D_A and the luminosity distance D_L, respectively:

Angular diameter distance It is the ratio of an object physical transverse size x to its angular size viewed from earth, θ:

$$D_A = \frac{x}{\theta} \tag{1.5}$$

It is related to the transverse comoving distance via the relation

$$D_A = \frac{D_M}{1+z} \tag{1.6}$$

The angular diameter distance is an observable quantity for astronomical sources that act as “standard rulers”. These are objects whose true physical dimension is known; by observing its apparent dimension, we can derive D_A. If we know the redshift of the object, we can find the $D_M - z$ relation and test it in the frame of a certain model. In Figure 1.3 we can see how D_A as a function of redshift changes when the cosmological parameters vary; the three curves represents the three scenarios $(\Omega_{m,0}, \Omega_\Lambda)$=(1,0) solid,(0.05,0) dotted,(0.2,08) dashed. Therefore, if we observe a standard ruler at different redshifts we can test how well a given cosmological model fits the data.

Luminosity distance The luminosity distance is the defined by the relation between the observed flux F and the luminosity L of a given source:

$$D_L = \sqrt{\frac{L}{4\pi F}} \tag{1.7}$$

This quantity is linked to the transverse comoving distance by

$$D_L = (1+z)D_M \tag{1.8}$$

The luminosity distance can be derived for sources that act like “standard candles”. As we measure the flux, we can use relation (1.7) to obtain an estimate for the distance without assuming any cosmological model. In Figure 1.3 we can see the change of D_L as a function

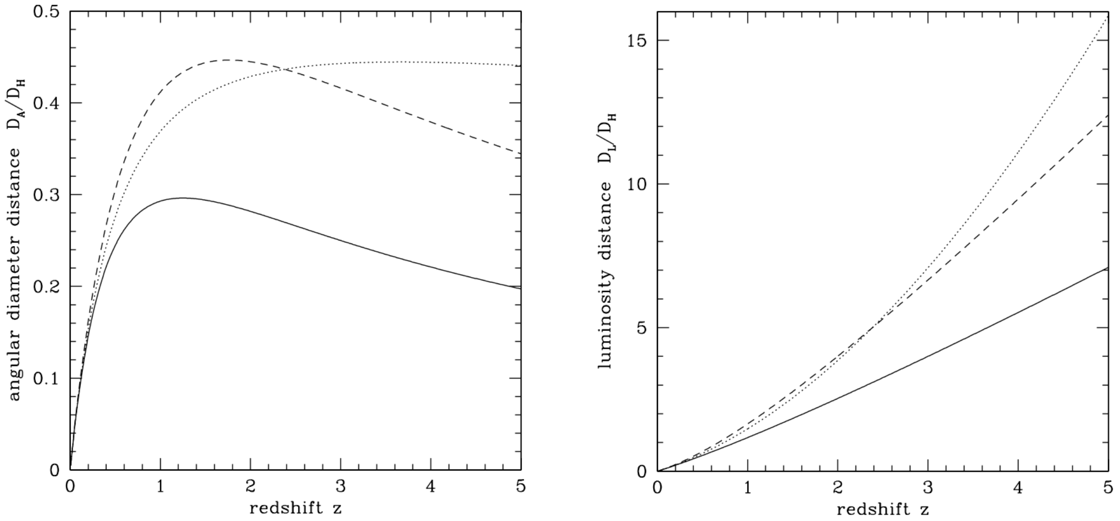

Figure 1.3: Left panel: Angular diameter distance D_A normalized for $D_H = c/H_0$ as a function of redshift z. Three different curves are shown for three different combinations for the values of cosmological parameters, which are $(\Omega_{m,0}, \Omega_\Lambda)$=(1,0) solid, $(\Omega_{m,0}, \Omega_\Lambda)$=(0.05,0) dotted, $(\Omega_{m,0}, \Omega_\Lambda)$=(0.2,0.8) dashed. Right panel: Luminosity distance D_L normalized for $D_H = c/H_0$ as a function of redshift z. Three different curves are shown for three different combinations for the values of cosmological parameters, which are $(\Omega_{m,0}, \Omega_\Lambda)$=(1,0) solid, $(\Omega_{m,0}, \Omega_\Lambda)$=(0.05,0) dotted, $(\Omega_{m,0}, \Omega_\Lambda)$=(0.2,0.8) dashed.

of redshift z as the values of cosmological parameters vary.
For historical reasons, the "distance modulus" DM is often used instead of the distance luminosity in distance-redshift diagrams. The distance modulus is simply the magnitude difference between the observed object and what it would be if it were at a distance of 10 pc:

$$DM = 5\log\left(\frac{D_L}{10pc}\right) \tag{1.9}$$

The most important examples of standard candle are Cepheid stars, at distances up to ~20 Mpc, and supernovae Ia (SNIa), at redshifts up to z~2. We will now discuss the implementation of supernovae as standard candles.

1.4.3 Supernovae as standard candles

A type Ia Supernova is a type of Supernova originating in a binary system where one of the two object is a white dwarf that accretes mass from the other object (that can be any other star) or when two white dwarfs merge with one another. White dwarfs have a mass limit, the "Chandrasekhar limit", $M_{CH} \sim 1.44 M_\odot$; as the white dwarf increases in mass, it will eventually reach $1.44 M_\odot$. Above this value, the electrons degeneracy pressure of the nucleus is not enough to withstand the gravitational collapse. This ignites Carbon fusion at the center that starts a runaway process, which brings to the explosion of the white dwarf in a Supernova, releasing in a short time an extremely high luminosity, which can be close

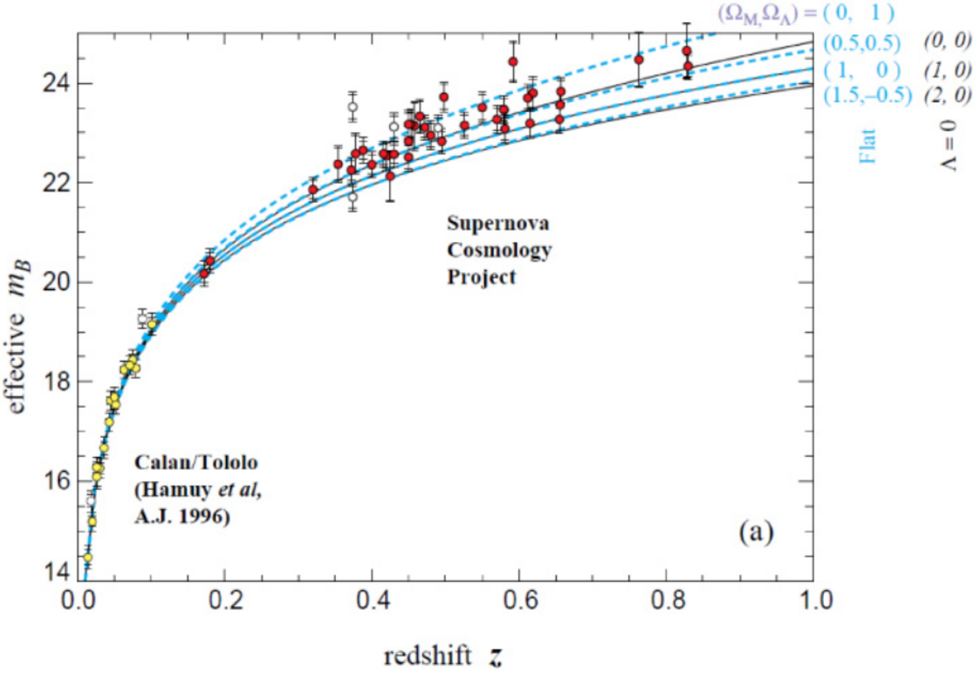

Figure 1.4: Hubble diagram for supernovae with data from Perlmutter et al. (1999); the dashed curves show different predictions based on different cosmological models.

to the one of the entire host galaxy ($10^{10} - 10^{11} L_\odot$). The emission reaches a peak and then declines in ~tens of days.
The physical process that ignites the explosion of supernovae is the same for each event, but the produced luminosity is not exactly the same. There exists a spread of ~ half a magnitude between luminosity peaks; therefore supernovae are not proper standard candles as we do not now their luminosity *a priori* . However, there is a way to "standardize" them: there exists a relation, called the Phillips relation (Phi; Hamuy et al., 1996), between the peak luminosity and the decline rate of the Supernova light curve. Observing this decline, we can correct the light curve for this factor so that all the light curves of different supernovae events match together. This way, we can derive the absolute luminosity of the objects and their luminosity distance. In order to obtain as absolute distance, supernovae need to be calibrated; this can be done using Cepheid stars, a type of variable stars that can be used as standard candles, in the redshift range in which both Cepheids and supernovae are observed ($z \sim 0.05$, corresponding to ~20 Mpc).

From the observation of different supernovae at different redshifts we can build the so-called "Hubble diagram", i.e. the relation between the distance modulus DM and the observed redshift. As described in the previous section, the shape of the $DM - z$ diagram depends on the values of the cosmological parameters, so we can compare observational data with what the prediction of different models. In the '90s, the implementation of supernovae as standard candles allowed to extend the Hubble diagram up to redshift z~1, as can be seen in Figure 1.4. The extension of the Hubble diagram allowed to discover the acceleration of the expansion of the Universe, leading to what is now defined the "cosmological standard model", consisting of a flat-ΛCDM model with $\Omega_m \simeq 0.3$ and

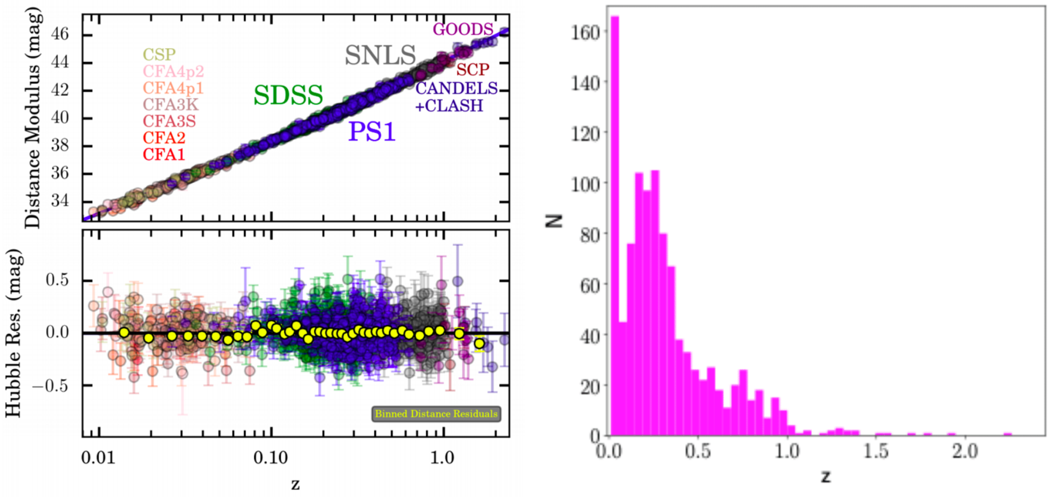

Figure 1.5: Left panel: state of the art Hubble diagram of Supernovae from Scolnic et al. (2018). Different colors and symbols refer to different supernovae subsamples. The total number of supernovae in this sample is 1048. Right panel: redshift distribution of the SN sample. While a few objects have been observed at z>1, they represent only a small tail in the distribution.

$\Omega_\Lambda \simeq 0.7$. In the last 20 years, remarkable progresses have been made in the use of supernovae as standard candles. In Figure 1.5, we can see, on the left panel, the "state of the art" of the Hubble diagram for supernovae, from Scolnic et al. (2018). We note that statistics increased significantly with respect to Perlmutter et al. (1999), with now more than 1000 observed sources. On the right panel, the redshift distribution of known supernovae is shown. We can see that supernovae are currently observed up to $z \sim 2$. However, very few objects fall at $z > 1$. This underlines the need for a cosmological probe that can not only extend the Hubble Diagram at high redshifts, but that can also provide more statistic for $z > 1$.

Another important implementation of standard candles is that Supernovae and Cepheids allow to constrain the Hubble constant H_0, which is, the expansion rate of the Universe at the present time. The most recent analysis with Supernovae and Cepheids (Riess et al., 2019) gives $H_0 = 74.03 \pm 1.42\ km\ s^{-1}\ Mpc^{-1}$. Another way to determine the H_0 parameter is based on information from the early Universe instead of the local one. This is possible using data from the Cosmic Microwave Background (CMB); from the spectrum of the anisotropies of the CMB at different angular scales, we can constrain the sound horizon at the last scattering surface and, assuming a cosmological model, extrapolate the value of H_0. Using the flat ΛCDM model, the best estimate of the Hubble constant (Planck Collaboration et al., 2020) is $H_0 = 67.36 \pm 0.54\ km\ s^{-1}\ Mpc^{-1}$. We see how the two values (the "local" one obtained with supernovae, and the one obtained with the CMB) are far from being consistent with one another, giving rise to the so-called "H_0 tension". If we assume that both results are free from hidden systematic errors, a possible way to solve this tension is to state that the flat-ΛCDM model is not the *true* cosmological model. This topic will be addressed in Chapter 6.

1.4.4 Quasars as standard candles

Nowadays, supernovae Ia are observed up to redshift z~2 (with very few objects at $z > 1$), which implies that they allow us to investigate the expansion of the Universe from when its age was ~3-4 billion years up to the present time. This means that, using only supernovae, in the redshift range between the furthest supernovae (z~2) and the CMB (z~1100) we do not have observational data to constrain cosmological models. This issue can be addressed with the implementation of quasars as standard candles. As described in Chapter 1, in fact, quasars are extremely luminous object, and are now observed up to redshift ~7, corresponding to an age of the Universe of only 0.7 Gyr. They are also extremely numerous, as there are hundreds of thousands of known quasars with spectroscopic confirmation. This provides a high statistics in all the redshift range z~1-7. Unfortunately, quasars are not intrinsically standard candles: their bolometric luminosity can vary up to five orders of magnitudes, so there is no way to directly determine the luminosity distance of a quasar. However, an observational result allow us to overcome this issue and use quasars as standard candles: the evidence of a non linear relation between the UV and X-ray luminosity that can be parametrised as:

$$\log(L_X) = \alpha \log(L_{UV}) + \beta \tag{1.10}$$

where L_X is the 2 keV monochromatic luminosity, L_{UV} is the 2500Å monochromatic luminosity and α and β are two adimensional parameters. If we substitute $L = 4\pi F D_L^2$ for both the UV and the X-ray luminosities, we see that we can derive a relation for D_L:

$$\log(D_L) = \frac{1}{2-2\alpha} \cdot (\log(f_X) - \alpha \cdot \log(f_{UV})) + \beta' \tag{1.11}$$

where $\beta' = \beta + (\alpha - 1)\log(4\pi)$. The relation holds if α is different from one. This relation allows to estimate D_L using only the UV and the X-ray fluxes, without assuming any cosmological model. Therefore, quasars can be *used* as standard candles, although they are not standard candles in the "classical" sense. This is somewhat similar to what has been described for supernovae, where the observationally-derived Phillips relation is used to derive luminosity distances, although the dispersion in the $L_{\rm X}$–$L_{\rm UV}$ relation is larger than the Phillips relation.

The exact origin of relation (1.10) is still not completely understood, as will be detailed more in the next Chapters, but it thought to arise from the interaction between the UV emitting accretion disc and the X-ray Corona. The existence of relation (1.10) has been known for around 40 years, together with its obvious cosmological implementation. It has not been used for distances estimates due to the high dispersion δ in the $\log(L_X) - \log(L_{UV})$ plane, with first estimates giving values of $\delta \sim 0.35-0.40$ dex, which means an uncertainty on the luminosity distance of a factor ~ 2.5. However, in recent years it has been shown that most of this dispersion is imputable to observational issues, that can mostly be removed through an accurate selection of the quasars sample, in which objects that might be affected by some kind of *bias* are removed from the sample (we will discuss this issue in the next Chapter). Thanks to this sample selection, in more recent studies (Risaliti & Lusso, 2015; Lusso & Risaliti, 2016; Risaliti & Lusso, 2019) the dispersion is reduced to $\delta \sim 0.22-0.25$ dex. These dispersion values imply an uncertainty on the final estimate of the luminosity

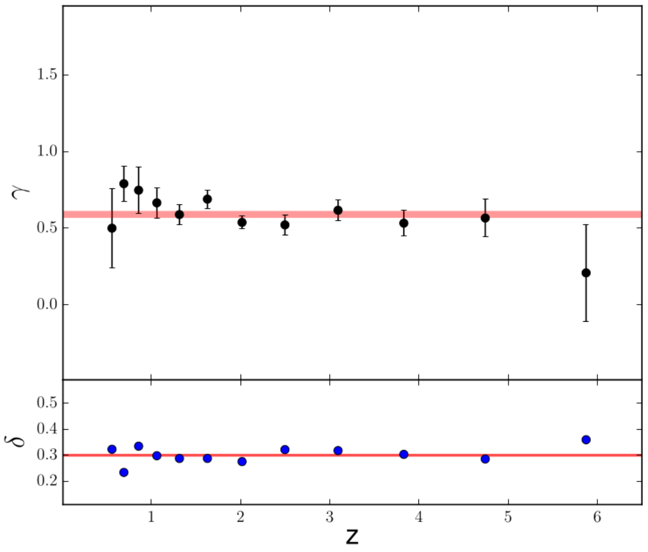

Figure 1.6: from Figure 4, Risaliti & Lusso (2015): best-fit values of the $f_X - f_{UV}$ correlation slope, γ, and the dispersion δ in narrow redshift intervals. The dispersion due to the different distances within each redshift interval is negligible with respect to the intrinsic dispersion of the $f_X - f_{UV}$ correlation (see the text for details). The horizontal lines show the average values, i.e., $\gamma = 0.60 \pm 0.02$ and $\delta = 0.3$.

distance that is still quite higher than what can be achieved with supernovae: ~20-25 quasars are needed to provide the same "cosmological power" as one single supernovae. However, at redshifts higher than 2 (where supernovae are not observed), hundreds of quasars have been observed. This means that implementing them in the Hubble diagram is equal to having found tens of supernovae at $z > 2$. The Hubble diagram for quasars has, in fact, already shown very interesting results, as will be discussed in the next Chapter.
In order to fully justify quasars as standard candles, we must demonstrate that (i) relation (1.10) does not change with redshift and (ii) our results agree with what can be found with supernovae in the common redshift range.

Redshift (in)dependence

The redshift independence is fundamental to corroborate the use of quasars as standard candles; we now briefly describe how this independence has been discussed in previous studies (Risaliti & Lusso, 2015). If we analyse relation (1.11) in a sufficiently narrow redshift range, the $\log(D_L)$ term will be nearly constant for the sources in that redshift interval. Therefore we can test relation (1.10) using fluxes instead of luminosities. If we do this in different redshift bins, we can see whether there is an evolution of the α parameter, which is, whether we see an evolution of relation (1.11) with redshift. We want the variation of the D_L parameter to be smaller than the observed dispersion of the $f_X - f_{UV}$ relation. If we consider a dispersion $\delta \sim 0.3$ dex, we must then have $\Delta \log(D_L) < 0.15$, which means $\Delta \log(z) < 0.1$ [2]. The other requirement is that in each redshift bin we have a

---

[2] the $\Delta \log(z) < 0.1$ prescription has been used in Risaliti & Lusso (2015), as the observed dispersion for that sample is $\delta \sim 0.3$. We note that in all the following works, including this thesis, the redshift bins are

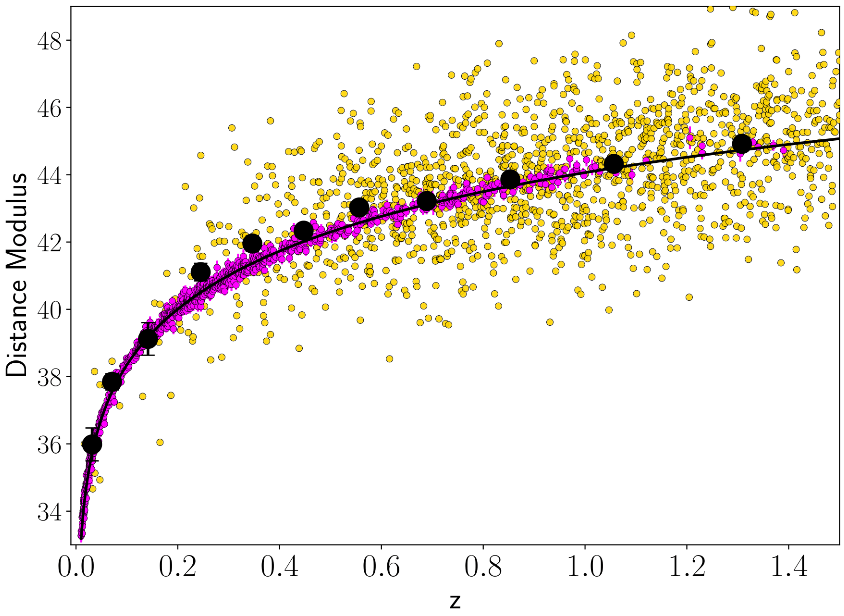

Figure 1.7: Hubble diagram for supernovae (magenta points) and quasars (yellow points) in the common redshift range. Black points represent the average for quasars in small redshift bins. The excellent match in the common redshift range demonstrates the independence of the normalization parameter β of the $L_X - L_{UV}$ relation.

number of sources sufficient for a statistical analysis. In the aforementioned paper Risaliti & Lusso (2015), the two requirements are met by splitting the quasar sample in 12 redshift bins; as can be seen in Figure 1.6, the parameter α (which is called γ in Risaliti & Lusso (2015)) does not show any trend with redshift, with a mean value among the 12 bins of $\alpha = 0.60 \pm 0.02$.

Match with supernovae

Being a normalization parameter, we can not give an estimate of β' of relation (1.11) in the absence of a theoretical model that explains the $L_X - L_{UV}$ relation. In order to evaluate β' and assure its independence from redshift, we can compare the quasar Hubble diagram with the supernovae Hubble diagram in the common redshift range, up to $z \sim 1.4$. The only difference between the two "standardized candles" is, in fact, the absolute calibration of the DM-z relation, which is measured for supernovae (based on the match with the distances derived with Cepheid stars) but is unknown for quasars. As a consequence, we can fit relation (1.11) for the joint Hubble diagram using parameter β' as a free parameter. In this way, we can cross-calibrate supernovae with quasars.

Figure 1.7 shows the Hubble diagram for quasars and supernovae in the common redshift range, and we note that there is an excellent agreement between them. This means there can not be a redshift evolution of the $L_X - L_{UV}$ relation, unless (i) there exists something that affects supernovae in the exact same way and that has not be discovered before or (ii) the normalization parameter β of the $L_X - L_{UV}$ relation starts evolving only at $z > 1.4$.

---

always chosen so that the distance spread is smaller than the dispersion we want to test

Since both cases seem very unlikely, we are confident that we can use the $L_X - L_{UV}$ relation in cosmological analysis, given the redshift independence of the parameters α and β.

Chapter 2

Quasars as standard candles: the $L_{\mathrm{X}} - L_{\mathrm{UV}}$ relation

In the previous Chapter we mentioned that we can effectively use quasars as standard candles if we accurately select the quasar sample, in order to decrease the observed dispersion of the $L_X - L_{UV}$ relation. In this Chapter we describe the details of this selection and we illustrate the "state of the art" in terms of cosmological results obtained with quasars, while the original work for this thesis on this topic will be addressed in Chapter 3 and 4. Our references are mostly the works Lusso & Risaliti (2016); Risaliti & Lusso (2019); Lusso et al. (2020). In these papers, it has been shown that (i) an accurate sample selection can decrease the observed dispersion down to $\delta \sim 0.22 - 0.25$ dex, (ii) quasars can effectively be used as standard candles, given the redshift independence of the $L_X - L_{UV}$ relation parameter α and the excellent match with supernovae in the common redshift range, and (iii) the extension of the Hubble diagram up to $z \sim 5$ with quasars show a highly significant tension with the ΛCDM cosmological model. We have addressed (ii) in the previous Chapter. In this Chapter we focus on issues (i) and (iii).

2.1 Sample selection

The starting sample in the aforementioned studies is obtained by cross-matching UV and X-ray quasars observations, as we need objects with data in both bands.
UV observations come from the Sloan Digital Sky Survey Quasar Catalogue Data Release 14 (SDSS DR14). This survey has observed around one third of the sky, providing both photometric and spectroscopic observations.
X-ray data come from the public survey catalogues of the XMM-Newton and Chandra observatories, together with smaller samples at $z \sim 3$ by Nardini et al. (2019), $4 < z < 7$ by Salvestrini et al. (2019), $z > 6$ by Vito et al. (2019) and the XMM-XXL North quasar sample published by Menzel et al. (2016).

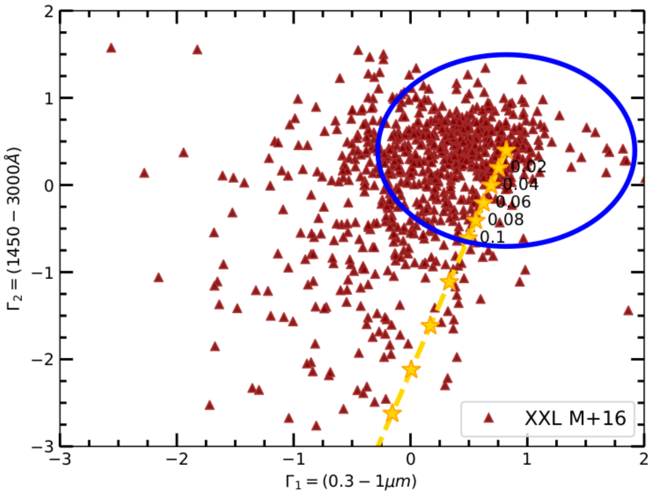

Figure 2.1: Figure 6 in Lusso et al. (2020): example of a typical quasar distribution in the $\Gamma_1 - \Gamma_2$ plane, with Γ_1 and Γ_2 the slopes of the power law in the $\log(\nu) - \log(\nu L_\nu)$ plane in the 0.3-1 μm and 1450-3000Å ranges. Yellow stars represent the values of quasars SED with increasing dust reddening, with numbers representing different $E(B_V)$ values. Quasars that fall inside the blue circle are the ones selected, as they are the ones with minimum host-galaxy and dust reddening contamination

2.1.1 Broad Absorption Line and Radio-loud quasars

Quasars classified as “broad absorption line” quasars (BAL) are removed from the sample. They are identified by the “BALnicity index of CIV absorption trough” in the SDSS catalogue. These sources are removed because broad absorption lines in the UV band are associated with absorption features in the X-ray band (Gallagher et al., 2002; Vignali et al., 2003). This means that the observed X-ray flux of these sources is not reliable, so they can not be used in the $L_X - L_{UV}$ relation analysis.

Quasars classified as “radio loud” are removed from the sample too. The radio loudness parameter is defined as $R = L_{\nu,6cm}/L_{\nu,2500}$ and sources are removed if $R > 10$. Radio-loud quasars show, in fact, an enhanced X-ray emission with respect to radio-quiet quasars (Worrall et al., 1987; Fossati et al., 1998; Shaban et al., 2022). Such emission is associated with the presence of jets and ultimately causes the photon index Γ of the X-ray spectrum to be smaller. Since the X-ray emission we are interested in is the one produced in the “Corona”, radio-loud quasars are removed from the sample.

2.1.2 Filtering in the UV: reddening and galaxy contamination

In the UV band, the most relevant phenomena that can change the observed flux with respect to the emitted one are dust reddening and host galaxy contamination. Both effects tend to “redden” the quasar spectrum: dust absorbs more efficiently high energy (blue) photons than low energy (red) ones, so that the final observed SED will results redder than the emitted one. In a similar way, the host galaxy emission can influence the observed color: quasars SED are intrinsically blue, differently from non-active galaxies. If the host galaxy emission is sufficiently strong, it can alter the observed quasar SED, resulting, again, in a "redder" observed emission. Since in the optical-UV band a quasar spectrum is generally well described by a power-law in the $\log(\nu) - \log(\nu L_\nu)$ plane, in both cases the reddening shows up as a reduced value of the intrinsic optical-UV slope.

This allows to find reddened source with a process described in the mentioned studies: from photometric SEDs, UV spectral slopes for 0.3-1μm and 1450-3000Å ranges are derived (respectively called Γ_1 and Γ_2). Typically, unobscured quasars are identified by slopes values around (Γ_1, Γ_2)=(0.82,0.40) (Richards et al., 2006). In order to minimize the reddening, only quasars with $\sqrt{(\Gamma_1 - 0.82)^2 + (\Gamma_2 - 0.4)^2} \leq 1.1$ are selected, while the others are removed from the sample, as can be seen in Figure 2.1. In this way we assure an extinction value $E(B - V) \leq 0.1$. Furthermore, this selection allows us to identify unusual SEDs or SEDs characterised by bad photometry, which are then excluded from the sample.

2.1.3 Filters in the X-rays: gas absorption

In the X-ray band, the emission of quasars can be affected by gas absorption, as described in Chapter 1. It is clear that we should remove absorbed sources from our sample, in order to avoid systematic underestimates of the X-ray flux.

Gas absorption is more efficient in the soft-X than in the hard-X band, which means that the global effect of absorption can be seen as a decrease of Γ_X, the slope of the X-ray continuum, together with the lowering of the observed X-ray flux with respect to the emitted one. Therefore we can use constrains on the Γ value in order to remove absorbed sources from our sample. We consider that the average Γ value for unabsorbed quasar is found to be $\Gamma \sim 2.0$, with a scatter of $\sim$ 0.2-0.3 (Young et al., 2009). As an upper limit, sources with $\Gamma > 3.0$ are removed, as such high Γ values may be due to observational issues such as incorrect background subtraction, or they may be associated with "exotic" physical states of the "Corona". Regarding the lower limit, we know that for values of $\Gamma < 1.5$ the source is surely absorbed; if Γ shows a value between 1.5 and 1.7, this *might* be due to absorption, although values in this range are observed in a minority of unabsorbed quasars. In order to be as conservative as possible, sources with $\Gamma < 1.7$ are removed from the sample.

2.1.4 Eddington bias

Considering the X-ray variability of quasars, sources whose X-ray flux is close to the minimum delectable flux for a given observation, are observed only in the case of a positive flux fluctuation. Therefore, faint quasars will suffer an overestimate of the average X-ray flux. This bias, which is called "Eddington bias", can significantly affect the analysis of the $f_X - f_{UV}$ relation by flattening it.

In order to remove this bias, sources are filtered from the sample if their expected flux falls below a threshold dependent on the intrinsic dispersion of the $L_{UV} - L_X$ relation and the flux limit of each observation. We require:

$$\log(f_{2keV,exp}) - \log(f_{min}) < k\delta \tag{2.1}$$

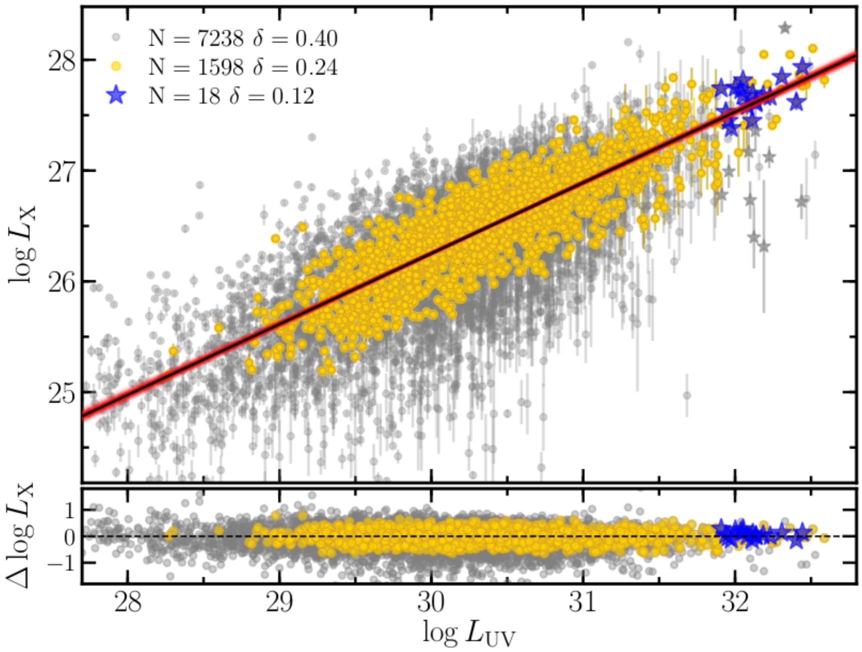

Figure 2.2: Figure 1 in Risaliti & Lusso (2019): $L_X - L_{UV}$ relation for the whole sample of quasars (gray) and for the ones that fulfill the selection criteria (yellow). Luminosities are determined assuming a ΛCDM cosmology with $\Omega_m = 0.3, \Omega_\Lambda = 0.7, H_0 = 70 km\ s^{-1}\ Mpc^{-1}$. The dispersion parameter δ goes from 0.40 dex to 0.24 dex when selection criteria are implemented. The red solid line represents the best fit of the $L_X - L_{UV}$ relation, which has a slope $\alpha = 0.633 \pm 0.002$. The lower panel shows the residuals with respect to the best-fit line. Blue points represent a subsample of $z \sim 3$ quasars with dedicated X-ray observations.

where $f_{2keV,exp}$ is the expected flux for each object given its observed UV flux at 2500 Å and the $f_X - f_{UV}$ relation, with the assumption of $\alpha = 0.60$ and a standard flat-ΛCDM cosmology:

$$\log(f_{2keV,exp}) = (\alpha - 1)\log(4\pi) + (2\alpha - 2)\log(D_L) + \alpha\log(f_{UV}) + \beta \tag{2.2}$$

f_{min} is determined for each X-ray observation taking into account the length of the observation and the position of the source in the field of view of the detector; δ is the dispersion of the $L_X - L_{UV}$ relation, and k is a number determining the strength of the filter.

In the aforementioned paper Risaliti & Lusso (2019), the value of k is found with the following procedure: a mock sample of quasars is generated, with X-ray fluxes derived from the UV fluxes assuming the $f_X - f_{UV}$ relation and an intrinsic dispersion of $\delta = 0.15$. Then luminosities are derived assuming some values of the cosmological parameters Ω_m and Ω_Λ; different mock samples are built, assuming different cosmologies. Then, for each sample, the filter in (2.1) is imposed, for various values of k. The final k is chosen to be that value above which the assumed cosmological model is met when fitting the mock sample data.
We note that this procedure is somewhat circular: we need to assume the $L_X - L_{UV}$ relation and its slope α in order to estimate the expected f_X. However, we can be confident in implementing it since it has been shown that we can retrieve the assumed cosmology for many different cosmological models (Risaliti & Lusso, 2019), therefore this "circularity" does not influence the cosmological predicting power of quasars.
After implementing these selection criteria, the observed dispersion in the $L_X - L_{UV}$ relation is reduced to $\delta \sim 0.22 - 0.25$ dex, as can be seen in Figure 2.2, where δ goes from 0.40 dex to 0.24 dex when the described selection criteria are employed.

2.2 Cosmological results

The "cleaned" quasar sample has been used in recent works to build the distance-redshift relation, i.e. the "Hubble diagram" (Risaliti & Lusso, 2015; Lusso & Risaliti, 2016; Risaliti & Lusso, 2019). This is done by employing relation (1.11) to evaluate the luminosity distance for each object, given the parameters α and β'. Regarding α, it is estimated by fitting the $f_X - f_{UV}$ relation in small redshift bins, in order to be independent from the chosen cosmology. Alternatively, it can be left as a free parameter in the fit. The parameter β' is an absolute normalization and it needs an external calibrator to be determined. It is estimated by cross-matching the Hubble diagram of quasars with the Hubble diagram of supernovae, in the common redshift range (z=0-1.4).

Fitting the Hubble diagram of quasars with the concordance flat ΛCDM model, a best fit parameter of $\Omega_m = 0.31 \pm 0.05$ is found, in agreement with current estimates.
However, the Hubble diagram shows a significant deviation from the predictions of the ΛCDM model for high redshifts ($z > 1.4$), as can be seen in Figure 2.3. In order to quantify this deviation, we need to use a fitting procedure that does not depend on a cosmological model. This can be done with a "cosmographic" approach, in which the

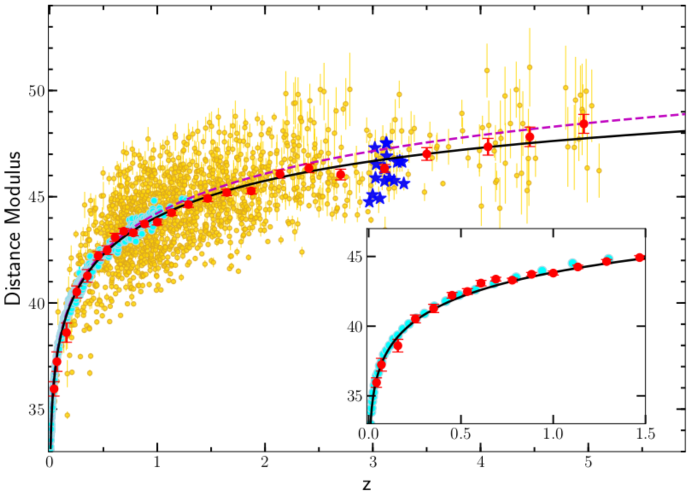

Figure 2.3: Figure 2 in Risaliti & Lusso (2019): Hubble diagram with quasars (yellow) and supernovae (cyan). Red points show the distance modulus average for quasars in small redshift bins. Blue stars represent the subsample at $z \sim 3$ from Nardini et al. (2019). The dashed magenta line shows the predicted diagram shape for a flat ΛCDM model with $\Omega_m = 0.31$; the black solid line is the best fit of the cosmographic logarithmic expansion.

Hubble diagram is fitted with a polynomial function whose coefficients can be related to the cosmological parameters of a given model. In Risaliti & Lusso (2019), this is done by fitting the polynomial function $D_L = k \sum_i a_i[\log(1+z)^i]$, where $k = \ln(10)c/H_0$ and $a_1 = 1$ in order to reproduce the Hubble diagram at low redshifts. An excellent fit is obtained with just two free parameters, a_2 and a_3. Assuming a ΛCDM model we can derive $a_2(\Omega_m)$ and $a_3(\Omega_m)$ and compare the best fit values of a_2 and a_3 with the prediction of the ΛCDM model. From Figure 2.4 we can see a 4σ discrepancy between ΛCDM predictions and the cosmographic fit. We note, again, that the data show consistency with the ΛCDM model for $z < 1.4$; the discrepancy starts to be visible from $z \sim 1.4$, i.e. in the redshift range where we only have quasars data. The excellent match between the Hubble diagram for quasars and that of supernovae in the common redshift range, together with the independence of the $L_X - L_{UV}$ relation with redshift, provide a strong validation of our method and make it extremely unlikely that the observed deviation at z~1.4 is not due to cosmological effects.

Once the tension with the flat ΛCDM is demonstrated, the next obvious step is the test of different cosmological models. If we modify the dark energy equation of state, so that w is left as a free parameter ($w = -1$ is the ΛCDM model), we obtain a best fit with a high value of the matter density parameter ($\Omega_m > 0.4$) and $w < -1$, which means a dark energy component whose density is increasing with time. As an example, in Figure 2.5 we show the results for a fit of a model with an evolving equation of state of dark energy $w(z) = w_0 + (1-a)w_a$. This is a widely used extension of the ΛCDM model for its simplicity, even if there is no particular physical reason to choose this parametrization. In order to test more complex extension of the ΛCDM model (such as allowing the dark energy equation of state to vary with time), we need to combine quasars with other cosmological probes, such as data from the CMB, in order to obtain tight constrains on cosmological parameters.

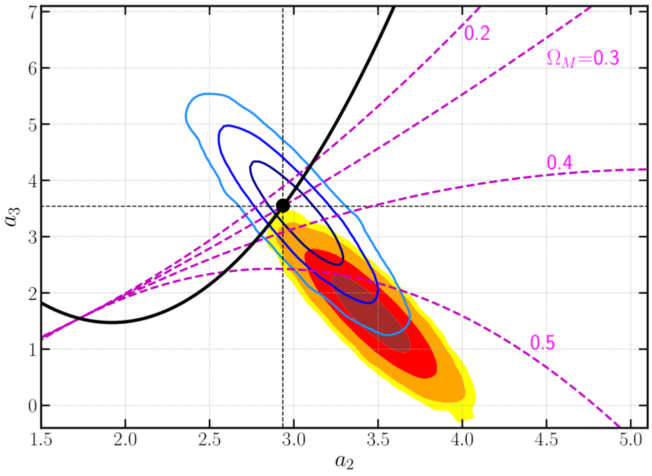

Figure 2.4: Figure 3 in Risaliti & Lusso (2019): comparison between the a_2 and a_3 parameters obtained as the best fit of the cosmographic logarithmic expansion and the ΛCDM predictions. The black curve is the relation between a_2 and a_3 for the ΛCDM model, with the black dot that remarks the value for $\Omega_m = 0.3$. Dashed magenta lines show the $a_2 - a_3$ relation for a wCDM model (the dark energy parameter is set as a free parameter) for different Ω_m values. Filled contours are obtained from the third order logarithmic polynomial fit with quasars and supernovae; empty contours are obtained with the same data for redshift $z < 1.4$. The concordance model is in agreement with observational data for $z < 1.4$ but in tension with the whole data set at a 4σ level. Implementing a wCDM model, the preferred values for the parameters are $\Omega_m > 0.3$ and $w < -1.3$, denoting a dark energy component with increasing density (the so-called "phantom" scenario).

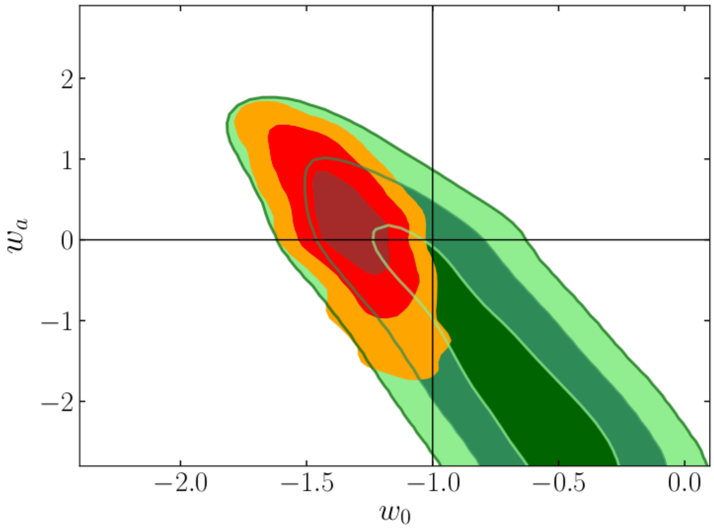

Figure 2.5: Fig. 9, Lusso et al. (2020): Results from a fit of a $w_0 w_a CDM$ model to the combined Hubble diagram of supernovae and the quasars. The green contours refer to the CMB results from Planck (Planck Collaboration et al., 2020)). The yellow-orange red contours are obtained adding the constraints from the Hubble diagram of supernovae and quasars.

2.3 Summary and open problems

In this Chapter we have shown how quasars can effectively be used as standard candles and that their implementation has already brought very interesting results. The most relevant feature of quasars as standard candles is that they are observed with sufficient statistics up to redshift $z \sim 5$, with a significant tail up to z ~7. Therefore, they allow to extend the Hubble diagram and to investigate the DM-z relation in a redshift range which was previously almost unexplored. They can be used together with other cosmological probes (supernovae Ia, CMB, BAO etc.) in order to test cosmological models as they give complementary information on the expansion rate of the Universe at high redshifts. The presence of a discrepancy with respect to the flat-ΛCDM model at high redshift is an interesting result also because it adds up to other tensions that the cosmological standard models suffers when low-z and high-z results are compared, namely the H_0 tension and the σ_8 tension (e.g., MacCrann et al., 2015). Although their precision as cosmological probes is still not as high as the one with supernovae, quasars can help to "fill the gap" between low and high redshift and give us a more complete understanding of the expansion of the Universe.

The highly significant tension with the flat-ΛCDM model is an extraordinary result in terms of implications for cosmology. Yet we acknowledge that the statement that the standard cosmological model has to be at least modified, if not completely subverted, is a very serious one. Therefore, we need to strengthen our result in order for it to be fully accepted. We point out that since a 4σ tension is already highly significant, this strengthening would not consist in increasing its value, for example by showing that its value is 5 or 6σ. What needs to be done is to unequivocally assure the reliability of our method, so that the tension with the ΛCDM model would be undeniable and not imputable to systematic effects. These are the goals of the works presented in Chapter 3 and 4.

On one hand, this could be done by deriving a fully consistent physical model that could explain the $L_X - L_{UV}$ relation. In past years, efforts have been made in terms of trying to characterize the disk-"Corona" interplay. Present models imply the existence of magnetic loops that connect the disk to the "Corona" and, in this way, do the energy transfer. However, there is still a lack of consensus on the exact physical mechanisms that come at play. Furthermore, models that have been proposed up to now strongly depend on (rather arbitrary) initial conditions, so that it is difficult to explain the $L_X - L_{UV}$ relation starting from first principles.
On the other hand, a way to strengthen the use of quasars as standard candle is to demonstrate that the intrinsic dispersion of the $L_X - L_{UV}$ relation is very small, possibly close to zero. In fact, if we have a scatter which is negligible with respect to the phenomenon we want to observe, we could rule out the possible presence of systematics, even without a complete physical model that could explain the $L_X - L_{UV}$ relation.

With this path in mind, much is still to be done regarding the $L_X - L_{UV}$ relation and its dispersion. First of all, we recall that in up-to-date studies, photometric data has been used to derive the monochromatic X-ray and UV luminosities. We argue that a complete spectroscopic analysis of quasars spectra would (i) improve our estimates of monochromatic fluxes, (ii) provide more detailed information about the quasars emission and about how spectral properties could be used to understand the $L_X - L_{UV}$ relation and, as a consequence, (iii) allow to obtain a possible further decrement of the observed dispersion. This will be investigated in Chapter 3.

In addition to this, we know there are causes of dispersion that can not be removed using the sample selection that has been described in this Chapter: we will discuss them in Chapter 4. Studying these factors we could give a reasonable estimate of the intrinsic dispersion of the $L_X - L_{UV}$ relation and therefore, hopefully, demonstrate that it is very small, possibly close to zero. This, as stated before, would corroborate the use of quasars as standard candles and validate the cosmological results.

In Chapter 5 we will discuss a fully non-parametric analysis of the Hubble Diagram of quasars and Supernovae Ia, with the aim of testing it again against the ΛCDM predictions and to gain insight on possible alternative models.

Chapter 3

The quest for indicators in the $L_X - L_{UV}$ relation

To sum up from the previous Chapters: Quasars are the most luminous persistent objects in our Universe. As described in the previous chapter, their spectral energy distribution (SED) is complex; it goes from the radio to the X-rays, with the most intense emission emerging at optical–UV wavelengths (e.g. Sanders et al., 1989; Richards et al., 2006; Elvis et al., 2012). The origin of this emission is attributed to accretion from an optically thick and geometrically thin disc around the central supermassive black hole (SMBH; Shakura & Sunyaev, 1973). For decades, the presence of a non-linear relation between the X-ray and UV luminosities of quasars has been observed (Tananbaum et al., 1979). This relation is usually parameterised as $\log(L_X) = \gamma \log(L_{UV}) + \beta$, where the slope is found to be $\gamma \simeq 0.6$ over a wide range of redshifts and luminosities (e.g. Steffen et al., 2006; Lusso et al., 2010; Young et al., 2010). This relation must be based on the interaction between the accretion disc, which emits mainly in the UV, and the so-called X-ray corona, which consists of a hot-electron plasma. UV photons coming from the disc are up-scattered in the corona, where they reach X-ray energies. It is clear that such an inverse-Compton mechanism can rapidly cool down the electron plasma, thus halting the production of X-rays. Since the X-ray emission of quasars is, instead, found to be persistent, a mechanism that refuels the corona with energy must exist. Given that the engine of a quasar is the infall of matter into the central SMBH through the accretion disc, this is also where the energy that refuels the X-ray corona most likely comes from. The existence of the $L_X - L_{UV}$ relation is therefore thought to be linked to such a mechanism, even if the exact physical process is still not completely understood. Modelling attempts have considered the reprocessing of the radiation from a non-thermal electron-positron pair cascade (Svensson, 1982), buoyancy, and the reconnection of magnetic fields as a way to dissipate the gravitational power (Haardt & Maraschi, 1991, 1993; Svensson & Zdziarski, 1994) or the presence of a viscosity-heated corona in which friction produces the heating (Meyer et al., 2000).

In addition to its relevance to quasar physics, the non-linearity of the X-ray-to-UV luminosity relation and the non-variability of its parameters make it possible to determine the luminosity distance of quasars, and therefore to use them as standard candles. This cosmological implication has been evident since the discovery of the $L_X - L_{UV}$ relation. However, its application was hindered by the very high observed dispersion (~0.4 dex), which made the luminosity distance estimates too uncertain to be useful for cosmology. In recent years, it has been shown that most of this dispersion is due to observational issues, which can largely be removed by neglecting quasars that are affected by dust reddening, gas absorption, or Eddington bias from the sample (i.e. Lusso & Risaliti, 2016; Lusso et al., 2020). This way, the observed dispersion is significantly reduced to ~0.20–0.25 dex. The resort to quasars as standard candles allows us to extend the Hubble diagram up to redshift values much higher than the ones that are achieved with supernovae Ia (Scolnic et al., 2018, 2022). This extended Hubble diagram is in agreement with the predictions of a flat ΛCDM cosmology up to a redshift of $z \sim 1.5$. At the same time, it shows a 4σ tension with the concordance model at higher redshifts (e.g. Risaliti & Lusso, 2019; Bargiacchi et al., 2021, 2022).

Given the obvious relevance of these cosmological results and the absence of a physical model explaining the X-ray-to-UV luminosity relation, it is fundamental to validate it observationally, ruling out any possible redshift dependence and systematic biases in the selection of the sample and/or in the flux measurements. In order to convince ourselves that the adoption of our method beyond $z = 1.5$ (at lower redshift it has already been validated, by the comparison with SNIa (see, e.g., Lusso et al., 2020)) is well motivated, we want to further exploit our observational data set, in particular using the spectroscopic data instead of just the photometric ones, for a better analysis of the X-ray-to-UV relation. In this Chapter we focus on the search of the optimal X-ray and UV indicators for the observed relation to do so.
In the next Chapter, we focus on the analysis of the different contributions to the observed dispersion, which is sill considerable ($\delta \sim 0.20$ dex), especially when compared to the Phillips relation for supernovae Ia. Given the absence of a comprehensive physical model for the relation, it remains unclear how much of the remaining dispersion is intrinsically tied to the relation itself (for example, due to intrinsic differences between different quasars) and how much can instead be attributed to observational factors not addressed by sample selection and flux choices.

In previous studies, the 2500-Å and 2-keV luminosities have been used as L_{UV} and L_X, respectively. The reasons behind this choice are mainly historical, and mostly motivated by the lack of emission or absorption features in this ranges. Although the exact nature of the physical interaction between the disc and the corona is not completely understood, it seems unlikely to involve two monochromatic luminosities. Conversely, it is reasonable to believe that it should depend on the UV and X-ray emission over wider bands. Indeed, by adopting monochromatic luminosities for L_X and L_{UV} to analyse the $L_X - L_{UV}$ relation, we are simply choosing two *proxies* of the overall emission. Therefore, we can ask ourselves which quantities work better to minimise the observed dispersion of the relation. The main aim of this chapter is to discuss the choice of the L_X and L_{UV} indicators to establish whether it is possible to further reduce the observed dispersion, and so gain a better understanding of the physics behind the $L_X - L_{UV}$ relation. We also note that in most of our previous studies, the UV monochromatic luminosities have always been derived from photometric data. In principle, deriving luminosities from a spectroscopic analysis is a more precise method. This is true especially in the UV, as a complete spectral fitting allows us to accurately take into account all the emission lines, while in photometric estimates some contamination from line emission that is not separated from the continuum might remain. It is then worth investigating the results of implementing UV measurements derived from a thorough spectroscopic analysis instead of the photometric ones.

3.1 Sample

Our sample consists of 1764 quasars, all of which have available UV and X-ray data in public catalogues and are included in the sample published by Lusso et al. (2020). In particular, we considered in said sample all the sources with available X-ray observations from the *XMM-Newton* satellite. More in detail, the sample is made of the following subsamples of the Lusso et al. (2020) one: SDSS-4XMM (1644 objects), XMM-Newton $z \simeq 3$ (14 objects), and XXL (106 objects). Out of 1764, 772 objects have spectra obtained with the BOSS spectrograph, while the remaining ones have spectra acquired with the original SDSS one. As discussed in the next section, our analysis was performed by dividing the sample into small redshift bins. The high-redshift ($z \geq 4$) objects discussed in Sections 2.5 and 2.6 of Lusso et al. (2020) are too sparse in redshift for the kind of analysis that we are interested in. This is why we excluded them from this analysis even though they have *XMM-Newton* observations available.

Detailed information regarding the sample can be found in Lusso et al. (2020, see their Section 2). Here, we briefly recall that the selected objects are all radio-quiet and not flagged as broad absorption line (BAL) quasars. Furthermore, they have optical colours that correspond to negligible dust absorption, and steep X-ray spectra (photon index $\Gamma > 1.7$), and they have deep enough X-ray observations to avoid biases towards brighter-than-average states. These properties make them suitable for a detailed analysis of the X-ray-to-UV relation, taking advantage of the homogeneity of both the source properties and the observational data.

3.1.1 Optical–UV spectral analysis

We performed the fitting procedure of each of the SDSS spectra using the software package `QSFit` (Calderone et al., 2017). The Signal-to-noise ratio (S/N) per pixel for our sample is typically between 8 and 20 (pixels in SDSS/BOSS spectra are logarithmically spaced, with a width of 10^{-4} dex). The software allowed us to fit the AGN continuum, the Balmer continuum, the emission-line and iron-complex properties, and the host-galaxy component. Regarding the host galaxy, a single template of an elliptical galaxy was used to determine its contribution to the total luminosity. The code is highly customizable, and our main settings are as follows: for quasars with redshifts below 0.6, given that the host-galaxy luminosity can be a relevant share of the total luminosity, the slope of the quasar continuum is degenerate with the host galaxy normalisation. Therefore, there is no way to determine it in a reliable way. We consequently fixed the value of the continuum slope for such quasars as $f_\lambda \propto \lambda^{-1.7}$. Each emission line was fitted with a Gaussian profile. `QSFit` allowed us to consider both a broad and narrow component to fit each line and to add unknown components if necessary. In case the code is not able to correctly disentangle the broad and the narrow component of the main emission lines, we used a composite profile of a broad component and an unknown one to account for the residuals. With this method, what we obtained are the *total* emission line properties, rather than the separated properties of broad and narrow components. Full width at half maximum (FWHM) upper limits were set to 2000 km/s for unknown and narrow lines and to 10,000 km/s for broad lines. To avoid fitting regions of the spectrum contaminated by intergalactic absorption, the minimum wavelength for model fitting was set to 1450 Å (rest frame) for all the objects in our sample. The strength of the `QSFit` code is that the spectral properties are all fitted simultaneously. This way, one can reasonably assume that the obtained continuum properties do not depend on local features of the spectrum itself.

At the end of the fitting procedure, for each quasar we obtained the continuum slope and the monochromatic luminosity at 2500 Å. We also recorded the luminosity values at 1350 Å, 3000 Å, 4400 Å, 5100 Å. If one or more of these wavelengths was out of the spectral range for a given spectrum, we extrapolated its value adopting the best-fit continuum slope. In terms of line properties, we obtained total flux, FWHM, offset velocity, and equivalent width. Together with the line properties and their errors, quality flags for each line were given. Regarding the quality flag, the `QSFit` code automatically raises a flag whenever one or more of the following situations occur: (i) the value of the continuum luminosity, a line's FWHM, or its offset velocity is NaN or equal to zero; (ii) any of the previous quantities hits a boundary value in the fit; (iii) the relative uncertainty on the continuum luminosity is higher than 1.5; (iv) the relative uncertainty on the FWHM of a given line is higher than 2; (v) the uncertainty on the velocity offset is higher than 500 km s^{-1}. Objects with bad-quality flags were removed from the sample.

After the fitting procedure, each spectrum was visually inspected and a second quality flag was raised if (i) the residuals of the fitting procedure showed a systematic trend as a function of wavelength or (ii) the mean reduced χ^2 value was higher than 2. Since in these cases the monochromatic luminosities estimates could not be considered reliable, such sources were removed from the sample. As a consequence, our sample size decreased from 1764 to 1705 sources. Among them, 1217 also have good Mg II emission line properties (line luminosity, FWHM, EW, velocity offset), while 403 have Hβ, 305 have [O III] λ4959Å, 291 have [O III] λ5007Å (202 have both the doublet components), 493 have C IV. In Figure 3.1 we show, as an example, the spectra of four objects at different redshifts and the best-fit results of the `QSFit` analysis.

In Figure 3.2, we compare our estimates of the total luminosity of the Mg II line with the one derived from the Wu & Shen (2022) catalogue. We can see an overall good agreement. Comparisons between other emission line properties show similar results. We also performed the same comparisons with the results of Rakshit et al. (2020), and they gave analogous results.

Wu & Shen (2022) presented the detailed measurements of the spectral properties of ~500,000 quasars from the latest release of the Sloan Digital Sky Survey (SDSS-DR16) quasar catalogue, which were used to validate the results of our spectral analysis. Regarding

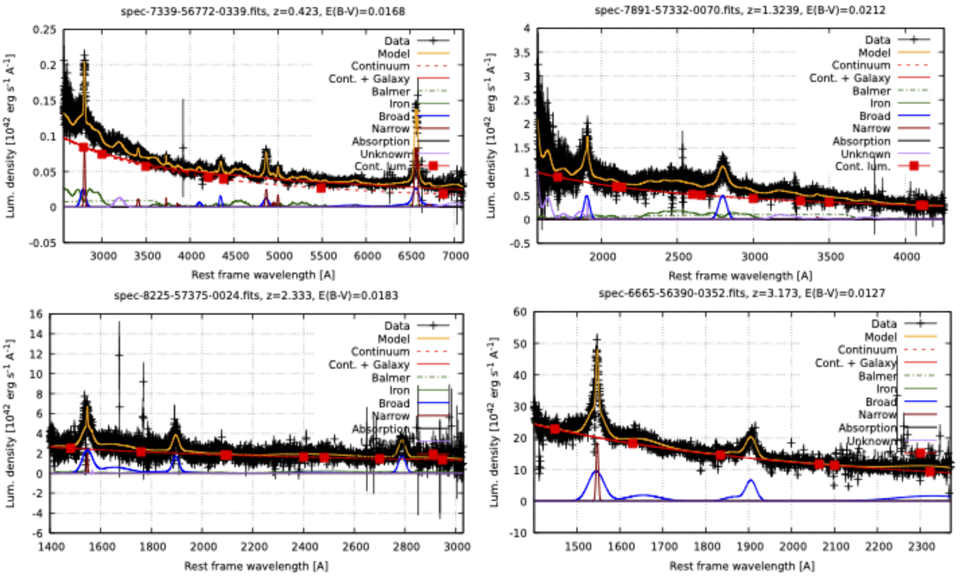

Figure 3.1: Four spectra at different redshifts and `QSFit` best-fit results, in yellow. The different components are shown in the legend on the right of each panel.

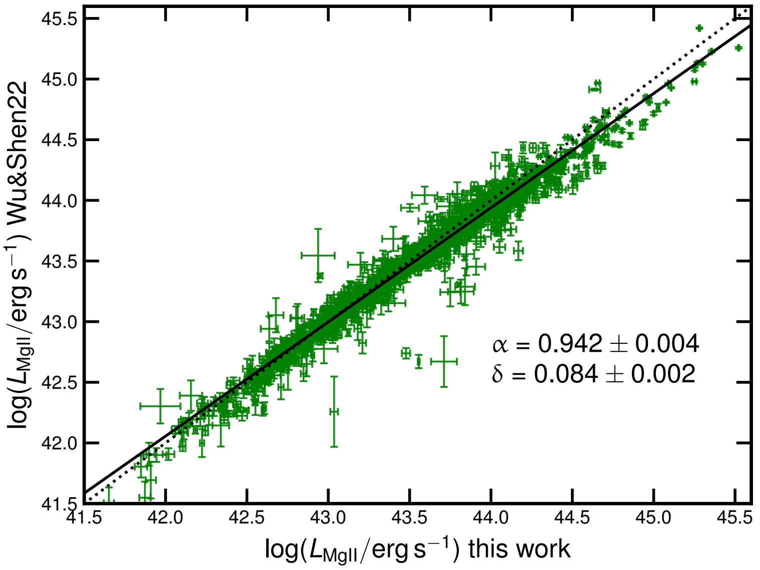

Figure 3.2: Comparison between the Mg II luminosity as derived from Wu & Shen (2022) and the Mg II luminosity derived in this thesis, in logarithmic units (erg s^{-1}). The dotted line is the one-to-one relation. We also report the best-fit slope and dispersion and the resulting best-fit-regression line as the solid black line.

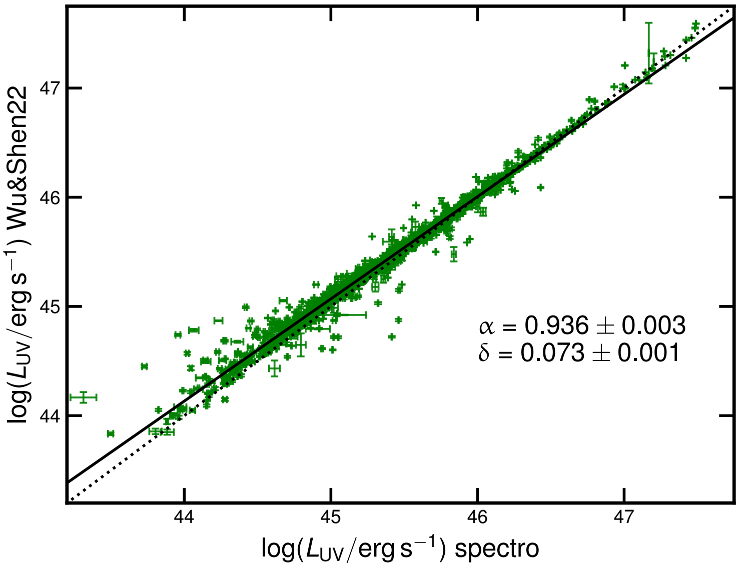

Figure 3.3: Comparison between monochromatic luminosity at 2500 Å obtained from our spectral analysis and the one from the Wu & Shen (2022), in logarithmic units (erg s^{-1}). The dotted line is the one-to-one relation. We also report the best-fit slope and dispersion and the resulting best-fit regression line with the solid black line. There is an overall good agreement between the two measurements.

the monochromatic luminosities, we find a very good match, as can be seen in Figure 3.3 for the monochromatic luminosity at 3000 Å. We can also see that the match is excellent at high luminosities, while some scatter is present for less luminous objects. This mainly depends on the different host-galaxy fitting techniques that have been employed; we used a single host-galaxy template, while in Wu & Shen (2022) the galaxy is not characterised, and can be any type. Quasars with lower luminosities are also the ones more affected by the host-galaxy contribution, so we expect higher discrepancies for such objects. Comparisons with the other monochromatic luminosities at different wavelengths (1350 Å, 4400 Å, 5100 Å) all show analogous results.
We note that the `QSFit` code returns results in luminosities, and not fluxes, assuming a standard ΛCDM cosmology with H_0= 70 km s^{-1}Mpc^{-1}, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$. Given that we are interested in flux measurements, we derived them from the obtained luminosities assuming the same cosmology. All the objects also have photometric flux estimates, derived as described in Lusso et al. (2020).

For completeness, Figure 3.4 presents the comparison between the monochromatic luminosity at 2500 Å derived from photometry as a function of the one obtained from the spectral analysis of our sample. The dotted line represents the one-to-one relation, while the solid black line is the best-fit regression between these two parameters. The resulting slope and dispersion are also reported in the figure. Interestingly, the relation between the two monochromatic flux estimates is significantly non-linear, showing a sort of Baldwin effect (Baldwin, 1977) between the two, whose origin is still not clear. The discussion of this result will address it in a forthcoming work.

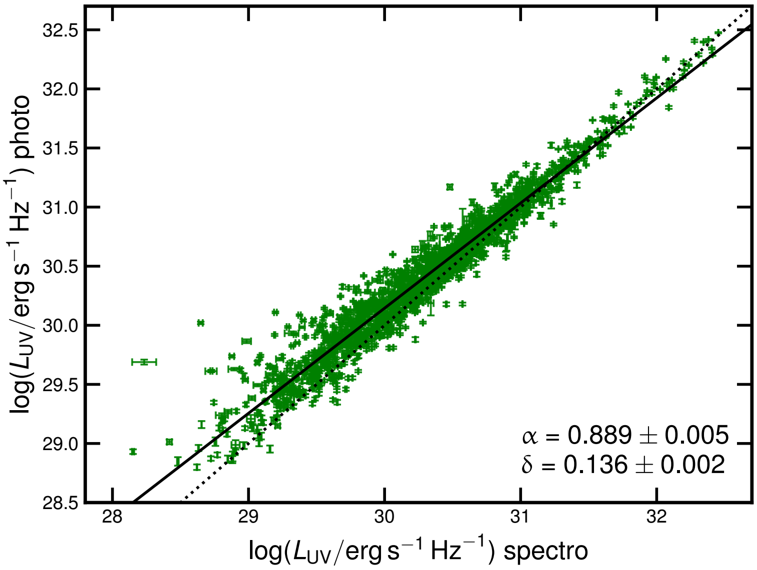

Figure 3.4: Relation between monochromatic luminosity at 2500 Å derived from photometry as a function of the one obtained from the spectral analysis of our sample. The dotted line represents the one-to-one relation. The solid black line is the best-fit regression between these two parameters. The resulting slope and dispersion are reported in the figure.

3.1.2 X-ray spectral analysis

All the objects in our sample have an estimate of the rest-frame 2-keV monochromatic flux derived from *XMM-Newton* photometric data. A detailed description of the procedure can be found in Section 4 of Lusso et al. (2020).
We have performed a full spectroscopic analysis for a subsample of 292 objects, which are all the objects in our sample at redshifts higher than 1.9, for which any discrepancy between the photometric and spectroscopic values can have major consequences in terms of cosmological applications. This comes as an extension of the X-ray spectroscopic analysis provided in Sacchi et al. (2022) for objects at redshifts higher than 2.5.
The goal of this analysis is to prove that not using fully spectroscopic X-ray fluxes does not introduce any bias in our results, while investigating if part of the residual observed dispersion of the $L_X - L_{UV}$ relation can be attributed to the lower accuracy of photometric measurements at the same time. We followed the standard procedure from the *XMM-Newton* user manual to obtain the spectra. For each object, we extracted three spectra for the three *XMM-Newton* cameras (pn, MOS1, and MOS2). We then combined the two MOS spectra into a single one.

The fit procedure was performed with the package XSPEC version 12.12 (Arnaud, 1996). We derived the source and the background spectra from the raw data, choosing an annulus around the source for the background extraction and making sure there were no additional sources in it. We then assumed a Cash statistic (Cash, 1979), and fit the data assuming a power-law model, considering Galactic photoelectric absorption: $constant * TBabs * clumin * zpowerlw$. We fitted the pn and the MOS spectra simultaneously, imposing the same spectral shape and allowing for a varying normalisation constant between the

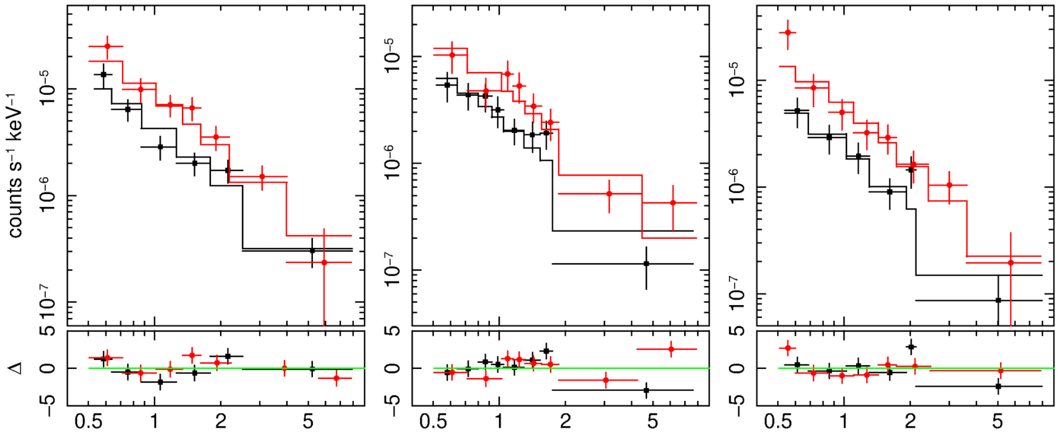

Figure 3.5: Example of X-ray spectrum and best fit of three sources at redshift 2.138, 2.144, 2.293. The pn spectrum is shown in red, and the MOS spectrum is shown in black. The relative uncertainties on the free parameters (the monochromatic flux at 2 keV and the photon index Γ) are shown as well.

two cameras. From the best fit, we estimated the monochromatic flux at 1 keV, the monochromatic flux at 2 keV, and the photon index Γ, together with their 1σ uncertainties. In Figure 3.5 we show, as an example, the spectra and the best-fit model for three objects, with data of different quality and therefore different uncertainties on the determination of the monochromatic X-ray flux at 2 keV. In Figure 3.6, we show the comparison between the monochromatic fluxes at 2 keV derived from this spectroscopic analysis with the ones obtain from photometric data, which has been the standard procedure in all previous works [1]. The relation between the two quantities was fitted with a linear relation, and the best fit is statistically consistent with a one-to-one relation, with the linear regression returning a slope of $m = 1.01 \pm 0.01$ as the best fit. In Figure 3.7, we also show the histogram of the differences between the spectroscopic and the monochromatic 2 keV fluxes, expressed in units of the standard deviation σ. We also see from this distribution that there is no systematic shift between the two quantities. Neither is there a significant skewness of the distribution, the skewness parameter being $k = 0.27 \pm 0.30$.

From this comparison, we infer that by using the spectroscopic data only for a subsample of objects we are not introducing any systematics.Analogous results in terms of the comparison between photometric and spectroscopic data are found when comparing the 1-keV monochromatic fluxes.

For the objects at redshift $z < 1.9$, the rest-frame 2-keV flux is computed starting from the observer's frame fluxes at 0.5–2 keV (F_S) and 2–12 keV (F_H) tabulated in the 4XMM-DR9 serendipitous source catalogue. An analysis of simulated power-law spectra with typical *XMM-Newton* responses and effective areas shows that the monochromatic fluxes at 1 keV ($f_{1\,\mathrm{keV}}$) and at 3.45 keV ($f_{3.45\,\mathrm{keV}}$) are 'pivot points' for the soft and hard

---

[1] the monochromatic fluxes at the pivot energies of 1 keV and 3.45 keV are derived from the photometric data, assuming the same photon index ($\Gamma_X = 1.42$) used to derive the fluxes in the 4XMM catalogue. The two pivot fluxes are then used to estimate the photometric photon index for the source, assuming a powerlaw shape of the spectrum, and then the 2 keV flux and its uncertainty is interpolated (or extrapolated) from that. For more details, see Section 4 of Lusso et al. (2020)

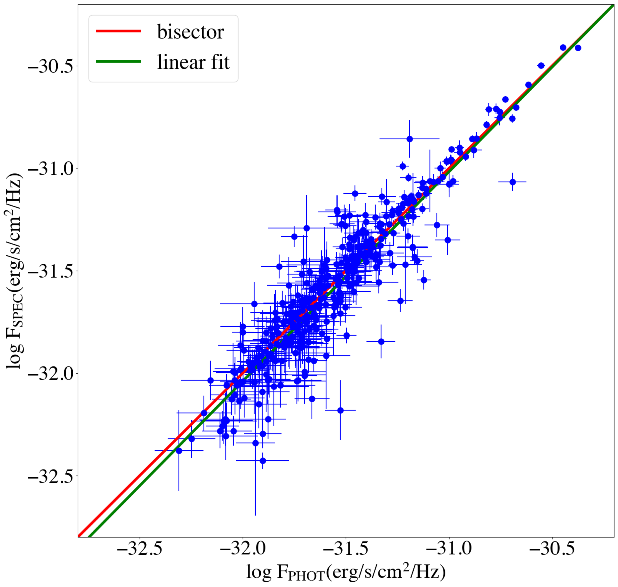

Figure 3.6: Comparison between photometrically derived 2-keV monochromatic fluxes and spectroscopically derived ones. The red line represents the one-to-one relation, while the green line is the best fit of the relation between the two quantities.

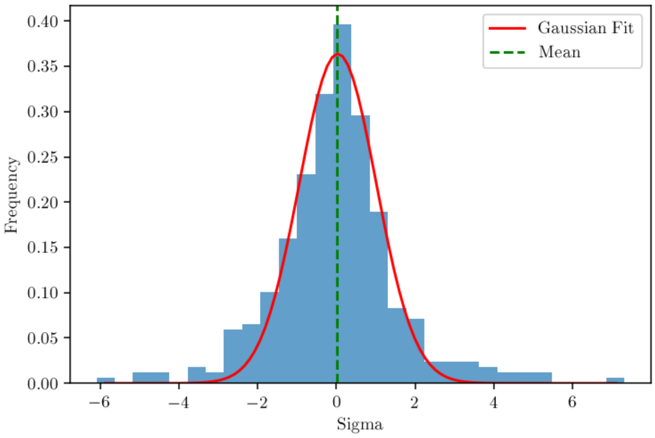

Figure 3.7: Histogram of differences between photometric and spectroscopic X-ray fluxes at 2 keV, shown in units of standard deviations. The red line shows the results of a Gaussian fit, which shows that the distribution is centred around zero. There is no significant skewness, as the skewness parameter turns out to be $k = 0.27 \pm 0.30$. This shows that there is no systematic shift between the two quantities.

bands, respectively, for example, the relation between these monochromatic fluxes and the total F_S and F_H fluxes is almost insensitive to the photon index Γ. Therefore, we estimated $f_{1\,\mathrm{keV}}$ and $f_{3.45\,\mathrm{keV}}$ from F_S and F_H assuming the same photon index used in the 4XMM catalogue ($\Gamma = 1.42$). We then used the power law defined by these two fluxes to compute both a new *photometric* photon index and the rest-frame monochromatic fluxes at different energies, as needed for our analysis (see the next sections). More details on this procedure are discussed in Risaliti & Lusso (2019) and Lusso et al. (2020).

3.2 Analysis of the X-ray-to-UV relation

The first step of our analysis of the X-ray-to-UV relation is a comparison of the results obtained adopting the new spectroscopic fluxes at the traditional reference wavelength of 2500 Å with those based on our previous UV and X-ray flux values from photometric data. Before showing the results, we outline the fitting method, which will be the same throughout the work.

As in our previous studies, we always performed our analysis by dividing the sample into a fixed number of redshift bins, such that for each bin, $\Delta \log z < 0.1$. For this sample, we chose 11 bins in the 0.38–3.48 redshift range, with a width of each bin of $\Delta \log z = 0.08$. In doing so, the differences among luminosity distances for the objects in a given bin are negligible compared to the dispersion in the relation. In particular, we checked with simulated data that the best-fit slope is always correctly recovered, provided that $\Delta \log z < 0.1$. The relation can therefore be re-written in terms of fluxes as

$$\log(f_X) = \gamma \log(f_{UV}) + \beta'' \tag{3.1}$$

where β' is related to the normalisation β of the $L_X - L_{UV}$ relation through the following expression:

$$\beta''(z) = 2(\gamma - 1) \log D_L(z) + (\gamma - 1) \log 4\pi + \beta \tag{3.2}$$

With this choice, we can perform our analysis independently of the cosmological model used to compute the luminosities. This is fundamental as we want our results on the physical relation between UV and X-ray emission not to be biased as a consequence of the adopted cosmology, which is essential to subsequently implementing quasars as cosmological probes. Furthermore, fitting the X-ray-to-UV relation in separate redshift bins allows us to investigate possible trends of its parameters with redshift, which is also crucial for cosmological applications. The parameters derived from the fitting procedure are γ, β', and the intrinsic dispersion δ. As in our previous works, we introduced the parameter δ because the observational errors on L_X (or f_X) and L_{UV} (or f_{UV}) alone cannot explain the observed scatter of the $L_X - L_{UV}$ (or $f_X - f_{UV}$) relation. For each bin, we also calculated the total dispersion δ_T, which can be considered to be the square root of the quadratic sum of the total observational error on the fluxes and the *intrinsic* dispersion δ.

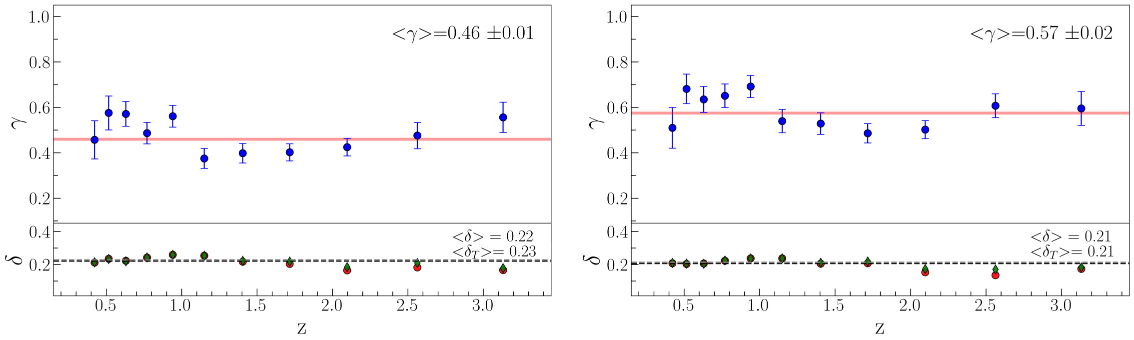

Figure 3.8: Slope γ, total dispersion δ_T (as green diamonds), and intrinsic dispersion δ (as red circles) of the X-ray (2 keV) to UV (2500 Å) flux relation in 11 narrow redshift bins. Left panel: Results with the UV flux derived from spectroscopic data. Right panel: Results with photometric UV fluxes

To perform the fit, a Bayesian approach of likelihood maximisation was used: this allowed us to estimate the parameters γ and β' while also taking the presence of the intrinsic dispersion δ into account, by modifying the likelihood function accordingly:

$$\ln \mathcal{L}(\theta|\mathrm{f_{UV}}, \mathrm{f_X}, \mathrm{ef_{UV}}, \mathrm{ef_X}) = -\frac{1}{2}\sum\left[\frac{(\mathrm{f_X} - f_{\mathrm{corr}}(\mathrm{f_{UV}}, \gamma, \beta))^2}{\mathrm{ef_X^2} + \gamma^2\mathrm{ef_{UV}^2} + e^{\ln(\delta)}} - \log(\mathrm{ef_X^2} + \gamma^2\mathrm{ef_{UV}^2} + e^{\ln(\delta)})\right] \tag{3.3}$$

where $\theta = (\gamma, \beta, \ln(\delta))$, $f_{\mathrm{corr}} = \gamma(f_{\mathrm{UV}} - \overline{f_{\mathrm{UV}}}) + \beta + \overline{f_{\mathrm{X}}}$, and $\overline{f_{\mathrm{UV}}}$ and $\overline{f_{\mathrm{X}}}$ are the mean values for the UV and the X-ray flux over the sample, respectively.
Using this approach we can obtain reliable estimates of the uncertainties on the parameters. We used the *emcee* code, which is an implementation of Goodman & Weare's Affine Invariant Markov chain Monte Carlo (MCMC) Ensemble sampler (Foreman-Mackey et al., 2013). We performed a fit for each redshift bin, adopting a sigma clipping at 3σ in order to exclude the few strong outliers that might still be present after the quality-selection procedure. The sigma-clipping procedure removes ~1% of the sample, both when photometric and when spectroscopic data are used. The parameters of the relation (γ and β') do not change with the sigma clipping; the only effect is on the dispersion parameter, which decreases by 0.02 dex for both samples.

Once we established that no significant trend of the fitted parameters in the redshift bins was to be found, we derived the average value for the parameters γ, δ, and δ_T by weighing each value for the number of objects in that bin. We remind the reader that β', instead, is expected to vary with redshift through its dependence on D_L (see Equation Eq. (3.2)).

Following this procedure, we tested the X-ray-to-UV relation using the 2500 Å monochromatic flux obtained in the spectroscopic analysis, while we keep assuming the photometric 2-keV monochromatic flux as f_{X}. The slope parameter γ of the flux-based relation is always found to be lower than in our previous analyses, where the photometric flux at 2500 Å was used (Lusso & Risaliti 2016, 2017; Risaliti & Lusso 2019; Lusso et al. 2020), with an average value of $\gamma = 0.46$, as shown in Figure 3.8, left panel. The slope does not show a systematic redshift trend. We tested this by fitting the 11 γ values as a function of redshift with a line, $\gamma = mz + q$, finding the best-fit values of the slope and intercept to be $m = -0.02 \pm 0.03$ and $q = 0.49 \pm 0.05$. The slope m is statistically consistent with zero, whilst the intercept q is statistically consistent with the average value of $\langle\gamma\rangle = 0.46$.

The redshift independence of the γ parameter in the X-ray to UV relation has already been thoroughly discussed in previous works, and it is now corroborated by employing spectroscopically derived monochromatic fluxes. Regarding the dispersion parameter δ, the fact that we obtained a value of 0.22 dex, which is similar to (although slightly larger than) the one found when using photometric fluxes, is by itself an interesting result for our understanding of the physics of the X-ray-to-UV relation. If the 'true' quantity behind the relation were the monochromatic luminosity at 2500 Å, we would expect the intrinsic dispersion to be lower, as a spectroscopic fit more accurately describes the true quasar continuum at a given wavelength when compared to the estimates we obtain from photometry. As this is not the case, it is possible that both spectroscopic and photometric fluxes are simply two different *proxies* of the quasar UV emission, which are almost equivalently effective in terms of the tightness of the resulting relation.

Starting from these considerations, we expanded our analysis in order to search for the optimal UV and X-ray indicators, under the assumption that both the UV and the X-ray spectra of quasars can be described by a power-law continuum over the range in question. We expect the 'true' physical relation to hold between the emission within a given frequency range in both the UV and the X-rays. For any frequency range, we can determine a *characteristic* frequency whose monochromatic flux would work as the best proxy according to the following argument. Let us consider a power-law spectrum between the frequencies ν_1 and ν_2. Then, for a fixed total flux in the (ν_1, ν_2) interval, there will be a frequency, ν_C, which we call 'characteristic', such that the monochromatic flux value f_{ν_C} does not depend on the slope of the power law, as this is the point that divides the spectral range into two intervals with the same weight. We argue that if we want to use a monochromatic flux as a proxy for the emission over a relatively wide wavelength range, the monochromatic flux at the 'characteristic energy' would be the best choice. Indeed, it would be (nearly) independent of the specific value of the slope of the power-law spectrum and it would be directly related to the global emission in the (ν_1, ν_2) range. This argument works for the choice of both the UV and the X-ray proxies.

In terms of frequency, if the spectral index is -1 (i.e. $f_\nu \propto \nu^{-1}$), and the instrumental response in the (ν_1, ν_2) interval is flat, then the geometric mean of the frequency interval gives us the characteristic energy $\nu_C \simeq \sqrt{\nu_1 \nu_2}$. However, this is not the case for our spectra, so we have to have another strategy. To find the characteristic frequency, we considered that any generic frequency ν can be derived from the characteristic one through the following expression:

$$f_\nu = f_{\nu_C} \left(\frac{\nu}{\nu_C} \right)^{\Upsilon}, \tag{3.4}$$

where Υ is the slope of the power law, ν_C is the characteristic frequency, and f_{ν_C} is the corresponding flux. If we assume that Equation Eq. (3.1) holds for the characteristic energy flux f_{ν_C} and we substitute Equation Eq. (3.4), we see that if we use f_ν, we also expect a correlation with the power-law slope Υ. We expect to find a correlation with Υ when a non-characteristic flux is used, while we expect no correlation with Υ when the flux at the characteristic energy is used. Furthermore, the slope of such a correlation would depend on the ratio between the frequency that we are using and the characteristic frequency. So, we can define a modified version of the flux–flux relation that also takes into account the correlation with the slope Υ:

$$\log(f_X) = \gamma \log(f_{UV}) + \zeta\Upsilon + \beta', \tag{3.5}$$

where ζ is the slope of such correlation, which we can include in our fitting procedure as an additional parameter by modifying the likelihood function accordingly. We can use this relation to search for the characteristic energy in the UV (and likewise in the X-rays) of the $f_X - f_{UV}$ relation. The value of the parameter ζ is indeed related to the characteristic frequency, as

$$\zeta = \gamma \log\left(\frac{\nu_C}{\nu}\right). \tag{3.6}$$

3.3 UV monochromatic proxy

Based on the aforementioned argument, we tested which monochromatic flux works as the best UV proxy by including the UV continuum slope as an additional parameter to the $L_X - L_{UV}$ relation. As before, the fit was performed in 11 redshift bins and by using fluxes instead of luminosities. We used the same five different monochromatic fluxes derived from the spectroscopic analysis, as discussed in the previous section. We tested relation 3.5, with the additional parameter ζ to be fitted and the UV slope Υ.

In terms of the parameter γ, we found average values in the range between 0.39 (for the 5100 Å flux) to 0.45 (for the 2500 Å and 3000 Å fluxes). The significance of the additional parameter ζ is found to be very small for all the chosen UV proxies, although a tentative increase is seen at longer wavelengths. The results are summarised in Table 3.1 and might have different explanations. It could be that a simple power law is not a sufficiently representative model for the optical–UV (1350–5100 Å) continuum, so the 'characteristic energy argument' does not completely hold, owing to, for example, some host-galaxy contribution at longer wavelengths. It could also happen that the region of the disc that is truly physically linked to the X-ray corona is emitting at wavelengths shorter than (or across) the UV peak, so the optical–UV continuum slope is not the correct quantity to be taken into account.

Given these results, each of the five proposed optical–UV proxies could, in principle, work equivalently well. The 2500 Å wavelength falls within the SDSS spectra for the widest possible redshift range in our sample. This means that there are relatively few objects for which the monochromatic flux is derived through extrapolation, which might be less reliable than the directly measured ones. Indeed, the 2500 Å flux provides the lowest dispersion, although marginally. We can thus consider this flux as the best proxy available.

We conclude this analysis with an additional fit required to compare the results obtained here to those published in previous papers on the X-ray-to-UV relation in quasars.

Table 3.1: Results of the search for the UV characteristic energy.

Wavelength	γ	ζ	δ	δ_T
1350 Å	0.43 ± 0.02	0.05 ± 0.02	0.23	0.23
2500 Å	0.45 ± 0.02	0.05 ± 0.02	0.22	0.23
3000 Å	0.45 ± 0.02	0.05 ± 0.02	0.23	0.23
4400 Å	0.41 ± 0.02	0.09 ± 0.02	0.23	0.23
5100 Å	0.39 ± 0.02	0.09 ± 0.02	0.23	0.24

Throughout this thesis, we use quantities derived from the spectroscopic analysis as f_{UV} for the first time. However, in all our previous works the UV monochromatic flux has been derived from the SDSS photometric data through an interpolation procedure (Lusso et al., 2020). We refer to this latter quantity as the ‘photometric’ UV flux. In order to facilitate comparison, we fitted the relation between fluxes with the 2-keV monochromatic flux as f_X and the photometric flux at 2500 Å as f_{UV} within the same redshift range of 0.38–3.48 and over 11 redshift bins. We obtain a mean value of the slope of $\gamma = 0.57 \pm 0.02$ and an intrinsic dispersion of $\delta = 0.21$ dex. The results are shown in Figure 3.8 in the right panel.

3.4 X-ray proxy

The X-ray spectrum of quasars can also be described by a power law. Therefore, analogous arguments regarding the choice of the ‘characteristic energy’ can be made, and we can investigate whether the 2-keV monochromatic luminosity is the best choice for L_X. The argument is the same as that described in the previous section; if we are not using the characteristic flux as f_X, we expect a dependence on the X-ray continuum slope, which in this case is described by the photon index Γ:

$$f_{2\,\mathrm{keV}} = f_{E_C} \left(\frac{E_{2\,\mathrm{keV}}}{E_C} \right)^{1-\Gamma} . \tag{3.7}$$

We note that the X-ray spectral shape is much better described by a simple power law than the UV one, as there are no other prominent spectral features. Therefore, the determination of the ‘characteristic energy’ may be more straightforward than for the UV range. The main difference with the previous section is that for the X-ray data, we still do not have a complete spectral analysis for each object, so we rely on photometrically derived estimates of both the monochromatic flux and the photon index for most of them. We tested whether any dependence on the photon index Γ is found when using the 2-keV monochromatic flux, implementing a modified version of Equation 3.5 in our fitting procedure:

$$\log(f_X) = (\Gamma - 1)\,\xi + \gamma \log(f_{UV}) + \beta', \tag{3.8}$$

where

$$\xi = \log\left(\frac{E_C}{E}\right). \tag{3.9}$$

We performed the analysis using the spectroscopically derived flux at 2500 Å as f_{UV}. As before, we used 11 redshift bins, in the redshift range of $z = 0.38 - 3.48$. We obtain a mean value of the slope of $\gamma = 0.47 \pm 0.02$, which is perfectly consistent with the previous values, and a significant $f_{\mathrm{X}} - \Gamma$ correlation parameter, with a mean value among the redshift bins of $\xi = -0.31 \pm 0.02$. The mean value of the intrinsic dispersion parameter becomes lower, with $\delta = 0.19$ dex instead of 0.22 dex.

The value of the parameter ξ allows us to determine where the X-ray characteristic energy is located according to Equation 3.9. Given the obtained value $\xi = -0.31$, we infer that the characteristic energy should be located at ~1 keV. The characteristic energy depends on the physical extent of the X-ray power law, which is not well known. Considering the minimum and the maximum energy of the (unknown) optimal X-ray range for the relation, so that the characteristic energy $E_{\mathrm{C}} \sim \sqrt{E_{\min} E_{\max}}$, our estimate of $E_{\mathrm{C}} \sim 1$ keV, might mean, for example, $E_{\min} \sim 0.01$ keV and $E_{\max} \sim 100$ keV, or $E_{\min} \sim 0.1$ keV and $E_{\max} \sim 10$ keV. Employing both the 2-keV flux and the parameter Γ in the relation mimics the employment of the 1-keV flux, which is almost insensitive to the exact value of Γ.

In order to test our assumption that the 1-keV energy is the X-ray characteristic energy, we derived, from photometric data, the monochromatic flux at 1 keV for all the sources in our sample; for objects at a redshift higher than 1.9, we derived it from the fit of the X-ray spectrum. We then performed the aforementioned analysis again. We found no significant correlation with the photon index Γ, with values of ξ consistent with zero. This confirms that we have found the true characteristic energy. Furthermore, by adopting the combination of the 1-keV flux as f_{X} and the 2500 Å spectroscopic flux as f_{UV}, we obtain a mean slope value of $\gamma = 0.46 \pm 0.01$ (consistent with the previous estimate), a value of ξ consistent with zero, and consistent values of the intrinsic and total dispersion parameters, with $\delta = 0.18$ dex and $\delta_T = 0.21$ dex.

We recall that allowing for an intrinsic dispersion parameter δ is necessary because the scatter of the observational points in the $f_{\mathrm{X}} - f_{\mathrm{UV}}$ relation cannot be entirely justified by the errors on the observed fluxes. Given that the rest-frame 1-keV emission is not between the two X-ray pivot points, and that with increasing redshift it gets progressively farther away from the observed energy range, the 1-keV fluxes are less precisely measured, with a mean observational error higher than the one obtained for the 2-keV monochromatic fluxes' estimates. However, when adopting the 1-keV flux instead of the 2-keV one, we notice a reduction of the dispersion δ and of the total dispersion δ_T. We thus conclude that, although the 1-keV flux is a less precise proxy in terms of its associated observational error, it can still be stated that, physically, it is a more accurate one, although only marginally.

To sum up, in this section we discuss the choice of the UV and X-ray proxies of the $L_{\mathrm{X}} - L_{\mathrm{UV}}$ relation. We argue that the best choice is, in both cases, in the monochromatic flux whose frequency acts as the characteristic energy of the spectral emission. Regarding the UV side of the relation, we found no clear trend as a function of the continuum slope, and therefore it can be stated that the monochromatic flux at 2500 Å is a good choice, as it provides the smallest intrinsic dispersion and it can be directly observed in optical–UV spectra for a wide redshift range. Regarding the X-ray side of the relation, we found a significant correlation with the photon index Γ when the 2-keV flux is used as f_{X}. The observed dispersion decreases when the latter dependence is taken into account. From this,

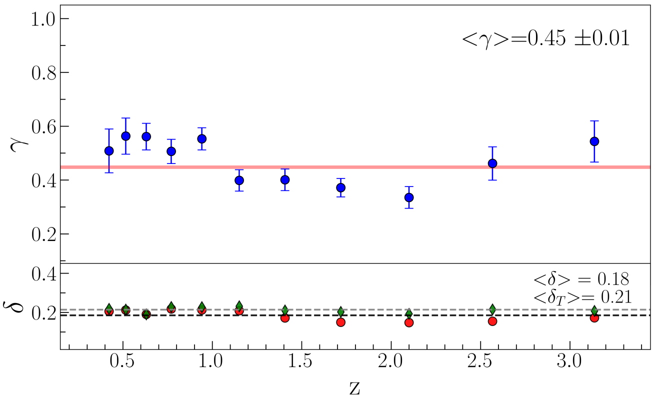

Figure 3.9: Slope parameter γ as function of redshift, as obtained by fitting Equation 3.1 when the monochromatic flux at 2500 Å derived from the spectroscopic analysis is used as f_{UV} and the monochromatic flux at 1 keV is used as f_X. The mean value of γ is 0.45±0.01, and we can see that there is no clear trend with redshift.

we can deduce that the X-ray characteristic energy should fall around 1 keV. The adoption of the 1-keV flux, indeed, results in no correlation with the photon index Γ, and a lower dispersion too. However, the photometric estimates of the 1-keV fluxes are significantly less precise than the ones at 2 keV for our sample.

In Figure 3.9, we show the results in terms of the parameter γ as a function of redshift, with the (spectroscopic) 2500 Å monochromatic flux as f_{UV} and the monochromatic flux at 1 keV as f_X. In Figure 3.10, we also show the $f_X - f_{UV}$ relation in the 11 redshift bins. With these choices, both the intrinsic and the total dispersion of the relation between fluxes are reduced. We checked again for the redshift dependence of the slope parameter γ, again by fitting the 11 values with a line as a function of redshift. We found the best-fit slope to be $m = -0.04 \pm 0.03$. Such a value is compatible with zero at 1.3σ and, overall, very small. We stress again that the non-dependence of the γ parameter with redshift is fundamental to showing that (i) the physical mechanism behind the relation is the same at all the observed redshifts, and (ii) we can safely use quasars as standard(isable) candles for cosmology.

3.5 Mg II line flux as UV proxy

The fluxes of quasar recombination lines depend on the intensity of the photoionising continuum. As such, these fluxes could constitute good indicators of the disc primary emission. We chose to test this hypothesis by analysing the $f_X - f_{UV}$ relation using the 1-keV monochromatic flux as f_X and the flux of the Mg II λ2800Å emission line as f_{UV}.

The Mg II λ2800Å resonant line (Netzer, 1980; Krolik & Kallman, 1988) requires the second ionisation of magnesium by photons with energies exceeding 15.035 eV, (i.e., with a wavelength shorter than 824.6 Å). Therefore, it is sensitive to the UV continuum blueward of the Lyman limit, which is not accessible through direct observations. Furthermore, its high equivalent width allows relatively precise flux measurements, and its rest-frame wavelength makes it observable in most of the redshift range of our spectroscopic sample. This combination of properties makes it the best candidate for our analysis, as it allows us

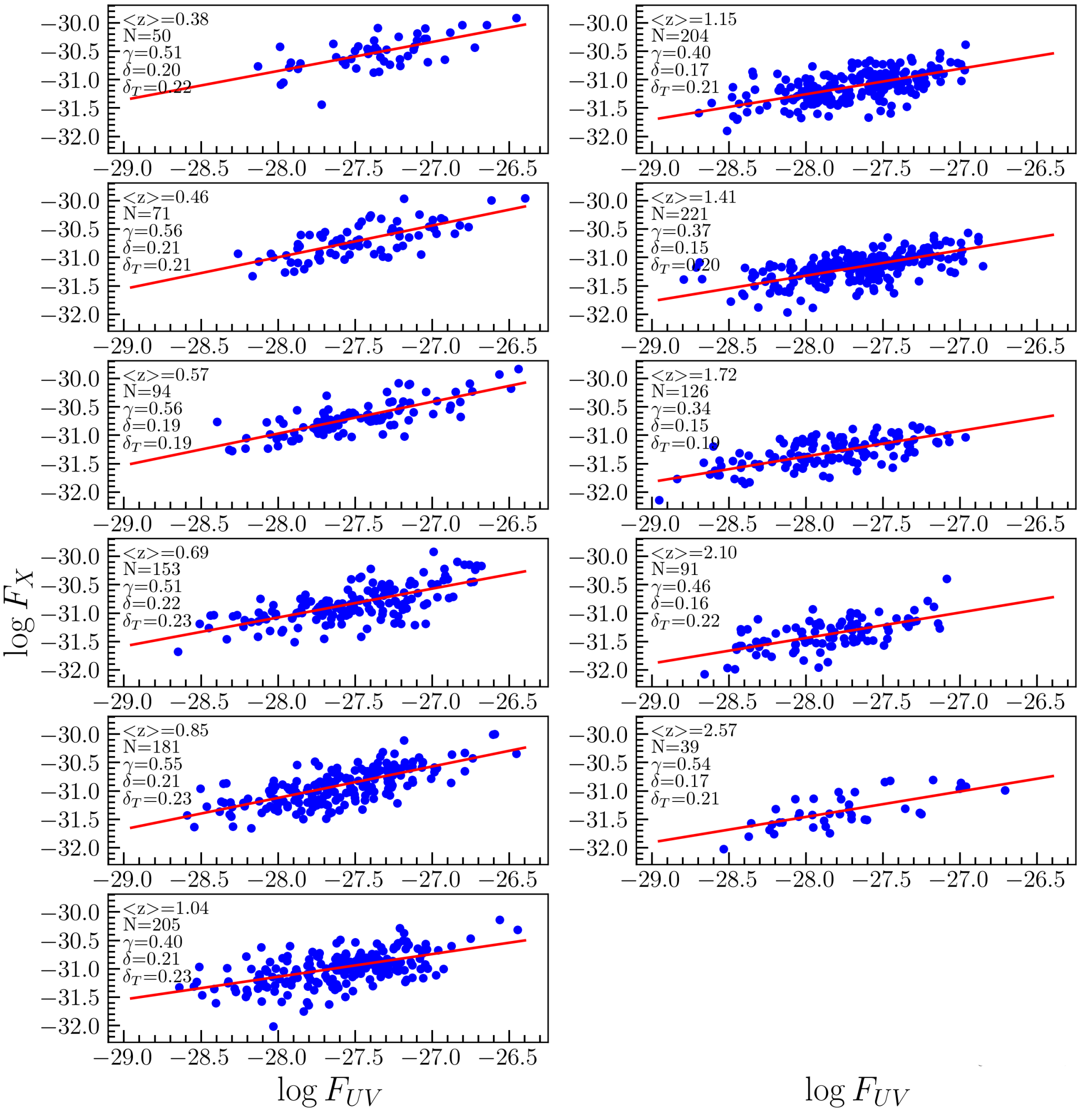

Figure 3.10: $f_{\rm X} - f_{\rm UV}$ relation in 11 redshift bins, obtained by using the 2500 Å monochromatic flux from the spectroscopic analysis as $f_{\rm UV}$ and the monochromatic flux at 1 keV as $f_{\rm X}$. The best-fit parameters (slope and dispersion), the total dispersion, the number of objects, and the average redshift are also reported.

to investigate whether the ~800 Å band is a more appropriate indicator of the disc emission than the 1350–5100 Å continuum. Alternative possibilities are the C IV λ1549Å line at higher redshifts (Lusso et al., 2021), or the Hβ λ4861Å line at lower redshift. The latter is indeed observable only up to $z \sim 0.8$.

The objects with spectra obtained with the SDSS spectrograph have Mg II emission up to a redshift of 2.28, while the ones with spectra obtained with the BOSS spectrograph have Mg II emission up to a redshift of 2.71. However, the fitting procedure of the BOSS objects in the redshift range of $2.58 < z < 2.71$ returned a poor-quality flag for the Mg II line, suggesting that the line parameters are not reliably constrained when this is located at the extremity of the spectrum. We therefore restricted the analysis to the redshift range of $0.38 < z < 2.58$ using ten redshift bins. We obtain a mean value of the slope $\gamma = 0.60 \pm 0.02$, a mean dispersion parameter $\delta = 0.16$ dex, and a mean total dispersion $\delta_T = 0.19$ dex. We notice that the Mg II line flux as a UV flux indicator gives a higher value of the slope parameter when compared with the spectroscopically derived monochromatic flux at any tested optical–UV wavelength, yet the dispersion is significantly lower. In Figure 3.11, we show the results of the fitting procedure using the Mg II line flux as f_{UV} for the ten redshift bins. We checked the dependence of the γ parameter as a function of redshift with linear regression. The fit returns a $\gamma - z$ relation with a slope of $m = -0.11 \pm 0.04$. Contrary to the previous two cases, there is evidence for a redshift trend, even if not a very strong one. We notice, however, that the analysis with the Mg II line is limited to a smaller maximum redshift, and the results in the common range are similar to those in Figs. 3 and 4 for the relations with the UV continuum. Another issue that might be affecting the Mg II results is the differences between objects with data coming from the SDSS spectrograph and the BOSS spectrograph. While we find no differences in the γ values and redshift dependence when comparing BOSS and SDSS data for the continuum, differences are found when comparing Mg II data. When fitting the $f_X - f_{UV}$ relation with only SDSS objects, we find an average slope of $\gamma = 0.64 \pm 0.03$ and an even more significant redshift trend; thus, when fitting the slope γ as a function of redshift with a line, we find a slope coefficient of $m = -0.19 \pm 0.05$. When using only the BOSS data, instead, we find $\gamma = 0.61 \pm 0.03$ and no significant redshift trend, with $m = 0.08 \pm 0.06$.

3.6 Mg II line width as a possible additional parameter

The widths of emission lines are independent observables with respect to continuum or line fluxes, and they carry complementary information, such as the gas rotational velocity or its outflow velocity. These properties are in turn related to the physical parameters of the primary source (black hole mass, accretion rate, etc.). For this reason it is interesting to investigate whether the X-ray-to-UV relation is also dependent on line widths, as already extensively discussed in Lusso & Risaliti (2017). To perform this test, we employed the following modified relation in our fit:

$$\log(f_X) = \gamma \log(f_{UV}) + \eta \log(\mathrm{FWHM}_{\mathrm{Mg\,II}}) + \beta'', \tag{3.10}$$

where f_X is the monochromatic X-ray flux at 1 keV and $\mathrm{FWHM}_{\mathrm{Mg\,II}}$ is the FWHM of the Mg II emission line that we inferred with the UV spectral fitting procedure. The goals

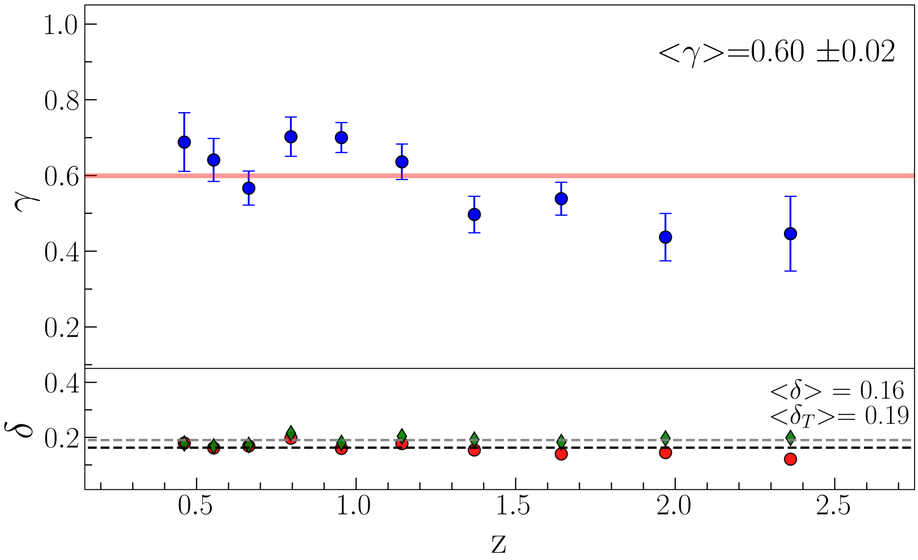

Figure 3.11: Parameter γ as function of redshift, resulting from fit of Equation 3.1 when Mg II emission line flux is used as $f_{\rm UV}$ and monochromatic flux at 1 keV is used as $f_{\rm X}$. The mean value of γ is 0.60 ± 0.02, and we can see that there is no clear trend with redshift.

of this analysis are (a) to use possible additional correlations to decrease the intrinsic dispersion of the relation, and (b) to understand whether these correlations have an effect on the use of the X-ray-to-UV relation to estimate quasar distances. The analysis is carried out in the 0.38–2.58 redshift range, where the Mg II line is available, with ten bins.

When using the spectroscopically-derived monochromatic flux at 2500 Å as $f_{\rm UV}$, we found a correlation which is highly significant, with a mean value of the parameter η equal to $\eta = 0.25 \pm 0.04$. The observed intrinsic dispersion is $\delta = 0.18$ dex and the total dispersion is $\delta_T = 0.20$ dex. We note that the δ and δ_T parameters that we obtain are comparable with the values presented in the previous section, where the dependence on the FWHM was not taken into account. It might be that the share of dispersion due to neglecting the FWHM correlation is too little for this effect to emerge and/or that the dynamical range of the FWHM values is too small for the additional correlation to impact significantly on the global observed dispersion. Regarding the parameter γ, we obtain a result consistent with the previous values, with $\gamma = 0.45 \pm 0.01$.

Using $f_{\rm Mg\,II}$ as $f_{\rm UV}$, instead, we found no correlation with the FWHM: we obtain $\eta = -0.03 \pm 0.03$, consistent with zero. This suggests that when $f_{\rm Mg\,II}$ is employed most of the significance of the correlation between $f_{\rm X}$ and line properties is already embedded in the $f_{\rm X} - f_{\rm Mg\,II}$ relation, so also considering a dependence on the FWHM does not add any new information.

Given the high significance of the dependence of the X-ray flux on the FWHM parameter, the obvious next step of the analysis is to check for a possible correlation between the UV flux indicator and the FWHM. This correlation could in principle be relevant to the inversion of the X-ray-to-UV relation to obtain quasar distances and, in the worst case scenario, introduce some redshift-dependent bias. In order to perform this check, we needed to compute the covariance matrix for the fit of Equation 3.10. However, given the small dynamical range of the quantity log(FWHM) and the relatively small number of objects in each single redshift bin, we were not able to obtain significant estimates of the correlation between UV flux and FWHM in the redshift-resolved analysis. For this reason, we carried out the fitting procedure using luminosities instead of fluxes (derived with a standard flat ΛCDM cosmology) and considering the whole sample together, using the monochromatic 1-keV luminosity as L_X, the spectroscopically derived 2500 Å monochromatic flux as L_{UV}, and the FWHM of the Mg II line. We then derived and normalised with respect to the first term the covariance matrix of γ and η, and we diagonalised it, with the goal of finding the linear combination of L_{UV} and $\mathrm{FWHM}_{\mathrm{Mg\,II}}$ required to obtain variables that are independent of one another in our fitting procedure. We found the matrix of the eigenvectors to be

$$\begin{pmatrix} -0.999 & -0.045 \\ 0.045 & -0.999 \end{pmatrix},$$

which, multiplied by the variable column vector (L_{UV}, $\mathrm{FWHM}_{\mathrm{Mg\,II}}$), gives us the following new variables:

$$\begin{aligned} X &= -0.999\, L_{UV} - 0.045\, \mathrm{FWHM}_{\mathrm{Mg\,II}} \\ Y &= 0.045\, L_{UV} - 0.999\, \mathrm{FWHM}_{\mathrm{Mg\,II}}. \end{aligned} \tag{3.11}$$

These new variables are independent of one another; if we repeat the described procedure using X, Y instead of L_{UV}, $\mathrm{FWHM}_{\mathrm{Mg\,II}}$ we find a diagonal covariance matrix.
This result also implies that we can implement a modified X-ray-to-UV flux–flux relation with the addition of the FWHM when this is available and keep using the standard relation otherwise. We performed this additional test in the whole redshift range of our sample, 0.38–3.48, again with 11 redshift bins, by modifying the likelihood accordingly. Using the photometric 2500 Å flux as f_{UV} and the 1 keV flux as f_X, we obtain $\gamma = 0.56 \pm 0.01$, $\delta = 0.16$ dex, $\delta_T = 0.19$ dex, $\eta = 0.25 \pm 0.04$. Therefore, with this mixed likelihood, we have a slight decrease of the intrinsic and the total dispersions.
Lusso & Risaliti (2017) proposed a *toy model* for the $L_X - L_{UV}$ relation, where the relation arises from the dependence of both quantities on the black hole mass M_{BH} and the black hole accretion rate $\dot{M}_{BH}$. Said model predicts the slope between the X-ray luminosity (or flux) and the FWHM of the Mg II line to be $\hat{\eta} = 0.44^{+0.21}_{-0.19}$, and they found, for a sample of 545 quasars, an observed slope between the X-ray luminosity and the Mg II FWHM equal to $\eta_{LR17} = 0.54 \pm 0.07$, consistent with the toy model. Our result, $\eta = 0.25 \pm 0.04$, is also consistent with the prediction of the toy model. It significantly differs from the Lusso & Risaliti (2017) value, but we note that they performed the analysis (i) for the sample as a whole and not in redshift bins, using ΛCDM-derived luminosities, and (ii) using the photometric 2500 Å luminosity as L_{UV} instead of the spectroscopic one. If we perform the analysis in the same way, using the photometric luminosities for our sample, we obtain $\eta_{\mathrm{phot}} = 0.45 \pm 0.04$, which is statistically consistent with the Lusso & Risaliti (2017) result. However, our results do not entirely align with the toy model presented Lusso & Risaliti (2017). The model, indeed, implied also the presence of a L_{UV}–FWHM relation, while we find the two quantities to be uncorrelated. At the same time, it might be that the small dynamic range of the Mg II FWHM does not allow us to observe the correlation.

Since the observational dependence of the X-ray flux on the Mg II FWHM line may be due to a *physical* dependence on the black hole mass, we performed an additional check by repeating our whole analysis for two subsamples with black hole mass estimates $M_{BH} < 10^{8.9} M_{\odot}$ and $M_{BH} < 10^{8.9} M_{\odot}$, respectively. We did not find any significant differences between these two subsamples. This suggests that the parameters of the relation do not directly depend on the black hole mass.

We finally have to consider that all the results that we have discussed in this subsection might be dependent on the cosmological model used to compute luminosities from fluxes, as here we are considering the whole sample and not objects in small redshift bins. To test this possible dependence, we followed a cosmographic approach to fit the quasar Hubble diagram, as described in Bargiacchi et al. (2021). We found analogous results in terms of covariance matrices, eigenvectors and $L_{\rm UV}$ – FWHM dependence.

3.7 Discussion

The main results obtained through the various fits presented in the previous sections are the following.

First of all, the slope γ and the intrinsic dispersion δ of the X-ray-to-UV relation obtained by using spectroscopically derived 2500 Å fluxes are $\gamma \sim 0.46$ and $\delta \sim 0.22$ dex, respectively. Among the different proxies of the disc emission that we tested, the monochromatic flux at 2500 Å can be considered as the best available one, given its observational coverage over a broad redshift range. For the X-ray emission in the X-ray-to-UV relation, we find instead that the best monochromatic indicator (i.e. the one providing the smallest intrinsic dispersion) is the 1-keV flux.

When adopting the Mg II line flux as UV proxy, the obtained intrinsic dispersion is smaller than that obtained with monochromatic continuum fluxes ($\delta_{\rm Mg\,II} \sim 0.17$ dex). The slope of the relation is instead significantly steeper ($\gamma_{\rm Mg\,II} \sim 0.60$) than in the spectroscopic case. We also find that when the ‘photometric estimate’ of the UV monochromatic flux is adopted, the best-fit parameters (both slope and dispersion) are consistent with those found using the Mg II line flux as UV proxy.

Finally, when the logarithm of the FWHM of the Mg II line is added to the relation, a statistically significant correlation is found, with a negligible decrease of the total intrinsic dispersion. The Mg II FWHM and UV flux parameters are not statistically correlated.

Here we discuss the interpretation and the main consequences of these results. The most remarkable result of our analysis is arguably that the spectroscopically derived 2500 Å monochromatic flux delivered a significantly lower γ value when used as $f_{\rm UV}$ compared to the Mg II line flux. This trend can be explained as a consequence of the non-linear relation between the emission-line equivalent width (EW) and the luminosity of the quasar continuum, known as the Baldwin effect (Baldwin, 1977). We analysed the relation between the Mg II luminosity and the monochromatic luminosity at 2500 Å derived from the spectroscopic analysis of our sample and we obtained a slope of ~0.8. In order to fit the whole sample simultaneously, for this analysis we used luminosities instead of fluxes, assuming a standard flat ΛCDM model with $\Omega_{\rm M}$=0.3 and H_0 = 70 km s^{-1} Mpc^{-1}. We checked that the results are not significantly dependent on the choice of the cosmological model.

If we now consider the relation between the X-ray and UV luminosities as $\log(L_{2\,\mathrm{keV}}) = \gamma \log(L_{2500}) + \beta$, with $\gamma = 0.46$ as shown in Section 2, and we use the $\log(L_{2500}) - \log(L_{\mathrm{Mg\,II}})$ mentioned above, we obtain a slope of the $\log(L_{2\,\mathrm{keV}}) - \log(L_{\mathrm{Mg\,II}})$ of $0.46/0.8 = 0.58$, which is fully consistent with the slope of the relation that we obtain when using the Mg II luminosity (or flux) as L_{UV}, as shown in Section 4. Therefore, we conclude that the reason behind different X-ray to UV relations when shifting from continuum UV proxies to line proxies is associated with the presence of the Baldwin effect itself.

Another result related to the adoption of the Mg II as UV proxy is the smaller intrinsic dispersion δ of the X-ray to UV relation with respect to the monochromatic continuum indicators. Considering that this line requires an ionising continuum at wavelengths shorter than 824 Å, a possible interpretation of this result is that an even tighter relation must exist between the X-ray flux and the UV flux blueward of the Ly limit. In this scenario, the monochromatic fluxes at optical or near-UV wavelengths are all ‘secondary indicators’ with a similar relation to the primary one.

The slope parameter that we obtain when using the ‘photometric’ UV flux has a less obvious interpretation. The photometric flux is a complex UV proxy, as it contains contributions from both the quasar continuum and line emissions. Moreover, even if it is formally a monochromatic quantity, it is derived from the wide-band photometric fluxes. We notice that its value is similar to the one obtained using the Mg II line as a UV proxy and that the dispersion of the relation is marginally better than the one obtained with truly monochromatic fluxes. This suggests that the combined information used to derive the photometric flux is similar to that contained in the Mg II flux, and is more closely related to the UV emission at $\lambda \sim 800$ Å than the monochromatic fluxes.

In Figure 3.12, we show the relation between the luminosity of the Mg II line and the photometric luminosity at 2500 Å. The slope is statistically consistent with unity, which explains the similar behaviour of these two quantities when used as f_{UV} proxies in the X-ray-to-UV relation.

3.8 Cosmological application

The most relevant result of our analysis concerning the use of quasars as standard candles through the X-ray-to-UV relation is the confirmation of such relation with spectroscopic data. Since there are no known standard candles at redshifts higher than $\sim$1.5, it is impossible to have an ‘external’ test of the validity of the relation in a cosmology-independent way. The only way to confirm or falsify the method is through an analysis of the possible redshift dependent physical and/or selection effects that may bias the distance estimate. In this sense, the complete UV (rest-frame) spectral analysis is a significant step forward; possible biases related to the use of the optical broad-band magnitudes (e.g. due to dust reddening or to the effect of strong lines moving into or out of the photometric bands depending on the redshift) are ruled out, and the flux measurements are directly done from the spectral fits.

The second fundamental outcome of our work is the first systematic search of the best X-ray and UV proxies of the relation, within the spectral range covered by the

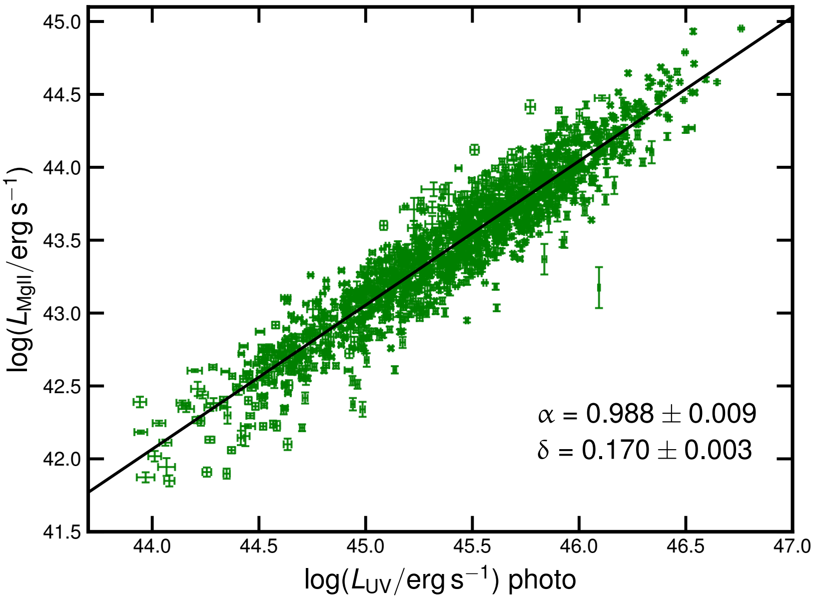

Figure 3.12: Relation between the luminosity of the Mg II line and the luminosity at 2500 Å derived from photometry.

available spectroscopic data. While the best X-ray and UV indicators turned out to be the monochromatic flux at 1 keV and the Mg II flux, respectively, we found that the intrinsic dispersion of the relation with the other proxies is only slightly larger. In particular, the 'standard' indicators, which are the monochromatic flux at 2500 Å derived from the photometric data and the monochromatic flux at 2 keV, are almost as effective as the 'best' ones.

This result suggests that none of the indicators used here is the primary driver of the relation. In order to prove this statement, we can consider our best UV indicator, that is the Mg II flux and the intrinsic dispersion of the X-ray to UV relation, $\delta_{\mathrm{Mg\,II}}$, obtained with this indicator. If the Mg II flux were the primary driver of the relation (i.e. the physical relation involves either the Mg II flux or a tightly related quantity), we would expect that the dispersion δ of the X-ray-to-UV relation using another UV proxy would be related to $\delta_{\mathrm{Mg\,II}}$ through the following relation: $\delta^2 = \delta^2_{\mathrm{Mg\,II}} + \Delta^2$, where Δ is the dispersion of the relation between the Mg II flux and the other UV indicator. However, this is not the case; for example, the dispersion of the relation between the Mg II flux and the monochromatic flux at 2500 Å is $\Delta = 0.15$ dex, which would imply $\delta \sim 0.25$ dex, while the observed value is $\delta = 0.19 \pm 0.01$ dex (see Figure 3.11). We conclude that both the UV indicators used here are proxies of a more fundamental one, probably related to the UV emission blueward of the Lyman limit.

A final result with some implications for the use of quasars as standard candles is the relation between the X-ray and UV fluxes and the FWHM of the Mg II emission line. We tested the inclusion of this parameter in the X-ray-to-UV flux relation, using the photometric 2500 Å flux as f_{UV}. Just like in the spectroscopic case, we found that there is no correlation between the UV flux and the Mg II FWHM. Therefore, we can incorporate an additional term in the relation between fluxes when the FWHM of the Mg II emission line is available, while fitting the ‘traditional’ relation if not. When doing so, we obtain $\gamma = 0.58 \pm 0.01$, $\delta = 0.16$ dex, and $\delta_T = 0.19$ dex.

We applied the results discussed above to the construction of a new Hubble diagram of quasars, based on the spectroscopic quasar sample described in this thesis. We derived the luminosity distances using Equation 3.10 for objects for which the Mg ii line is available, and the standard relation for those for which it is not. We used the 1-keV monochromatic flux as f_X. As for f_{UV}, we used both the spectroscopic and the photometric monochromatic flux at 2500 Å (we also repeated the computation using the Mg ii flux, obtaining fully consistent results). The results are shown in Figure 3.13, where we also plot the supernovae of the Pantheon sample (Scolnic et al., 2018), used for the absolute calibration. Petrosian et al. (2022) argued that the analysis in redshift intervals may bias the cosmological analysis. We stress that we never considered any binning (especially when considering redshifts) to perform the cosmological analysis in our previous works. The same consideration can be applied here as well. The distance modulus of each individual quasar is obtained from Equations 9 and 10, and the fits of the Hubble diagram are done by marginalising over the parameters γ and β of the relation. As a consequence, the red points shown in Figure 3.13 are computed only after the best fit values of these parameters are obtained, and they are shown only for illustrative purposes.

The Hubble diagram in Figure 3.13 is relevant for two main reasons. First, the two versions of the diagram, with the two different methods to derive the UV monochromatic flux, are in full agreement. This was not granted a priori; if some systematic effect related to the derivation of the photometric fluxes were present, it would have been revealed by the comparison with the spectroscopy-based values. The photometric fluxes are much easier to obtain, and indeed we always used these values in our previous works. Therefore, this result is also a validation of the Hubble diagrams in Risaliti & Lusso (2019) and Lusso et al. (2020), which revealed a strong tension with the ‘concordance’ flat ΛCDM model. Second, the dispersion in the Hubble diagram based on spectroscopic points is significantly lower than that based on photometric fluxes. This is the consequence of the different slopes of the X-ray-to-UV relation depending on which UV flux is used: $\gamma_{\rm phot} \sim 0.6$ for the photometric fluxes and $\gamma_{\rm spec} \sim 0.45$ for the spectroscopic ones. Considering the error propagation from the relation to the distance moduli plotted in Figure 3.13, the main contribution to the error is the intrinsic dispersion of the relation, divided by a factor $(1 - \gamma)$. This term implies that the uncertainty on the distances derived from spectroscopic fluxes is lower that of distances based on photometric fluxes by a factor of $(1 - \gamma_{\rm phot})/(1 - \gamma_{\rm spec}) \sim 0.75$.

The tension with the standard flat ΛCDM model of the spectroscopic sample discussed in this thesis has a statistical significance of 3σ. This is lower than previously published results based on larger samples (Lusso et al. 2020, Bargiacchi et al. 2021) but, again, it is the first result based on a sample with complete control on the derivation of the UV fluxes.

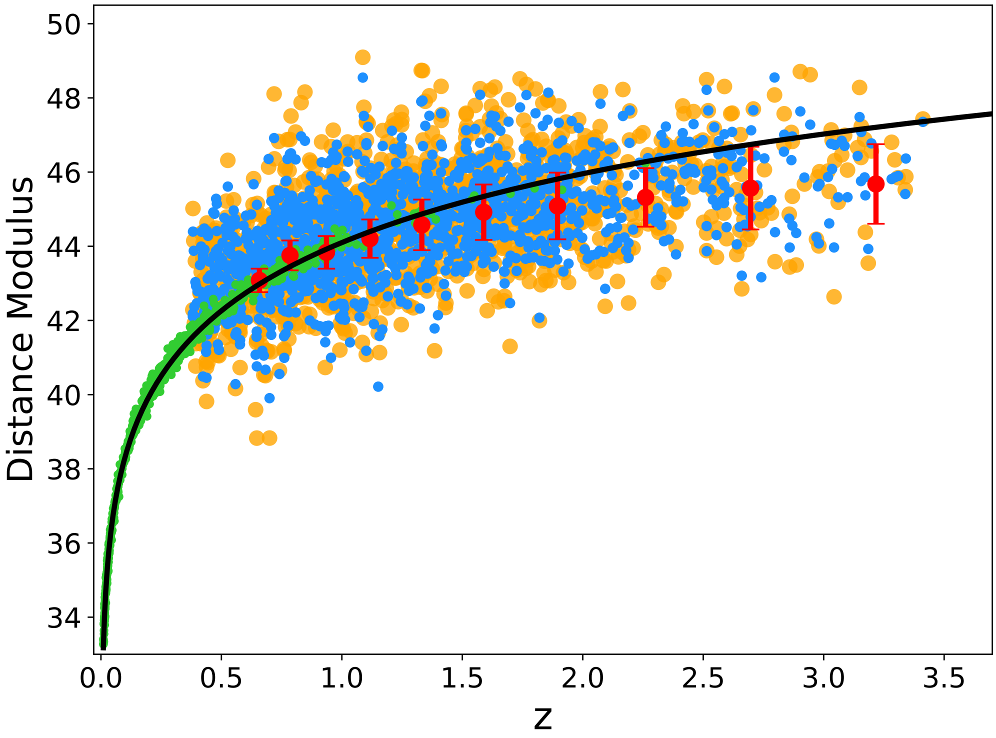

Figure 3.13: Hubble diagram of supernovae and quasars. Green points are supernovae Ia from the Pantheon sample (Scolnic et al., 2018), yellow points are quasars with distances derived using the photometric UV fluxes, blue points are quasars with distances derived using the spectroscopic UV fluxes, and red points are the average distance modulus values for spectroscopic quasars in the individual redshift bins. The normalisation parameter for quasars is chosen in order to match that of supernovae Ia. As in our previous studies (Risaliti & Lusso, 2019), we do so by cross-matching the Hubble diagram of quasars with that of supernovae in the common redshift range. The black line represents the prediction of a flat ΛCDM model with $\Omega_M = 0.3$.

3.9 Summary

In this Chapter, we present a thorough UV spectral analysis of 1764 quasars from the Lusso et al. (2020) sample with the main aims being to discuss the choice of the L_X and L_{UV} indicators and to establish whether it is possible to further reduce the observed dispersion to gain a better understanding of the physics behind the $L_X - L_{UV}$ relation. We thus derived monochromatic luminosities at five different wavelengths as well as the emission-line properties. We compared our results with the spectral analysis of Wu & Shen (2022) and we found a good agreement. We also computed the X-ray properties of our sample from a spectroscopic analysis for objects at a redshift higher than 1.9, while we used photometric data for objects at lower redshift. Our main results are summarised below.

1. We obtained the monochromatic flux at 2500 Å from a detailed spectroscopic analysis, as this is supposed to be a more accurate measurement than the one derived from the photometric data. We then analysed the $L_X - L_{UV}$ relation in narrow redshift bins so that we could (i) use fluxes instead of luminosities and therefore be independent of any chosen cosmology, and (ii) look for possible redshift trends of the slope parameter γ. While the slope is confirmed not to show any systematic trend with redshift, its value of $\gamma \sim 0.46$ is lower (flatter) than those found in our previous studies where photometric data were used. Also, the dispersion parameter δ is slightly higher when using spectroscopically derived monochromatic fluxes as f_{UV}. If the true physical quantity behind the $L_X - L_{UV}$ relation had been the monochromatic luminosity at 2500 Å (or any of the wavelengths that we tested; see below) we would have expected better results in terms of the dispersion. As this does not occur, it probably means that the spectroscopic and the photometric luminosities are simply two different *proxies* of another quantity, and thus they are similarly effective when used in the X-ray-to-UV relation.

2. We investigated what the best energies to be used as L_X and L_{UV} are, following the assumption that both the UV and the X-ray continua of quasars can be parameterised as power laws and studying the dependence of the X-ray-to-UV relation on the respective spectral slopes. We stress that we do not expect the relation to subsist between two monochromatic luminosities, but we are looking for the best possible proxies for the overall UV (disc) and X-ray (corona) emission. In the X-rays, we find that there is a preference for 1 keV as the characteristic energy, as it is the one less sensitive to the actual spectral slope. When using the 1-keV fluxes instead of the 2-keV ones, we indeed found a lower dispersion. Unfortunately, given the redshift range of our sample, the 1-keV flux is on average measured with higher uncertainties than the 2-keV ones. This partly undermines the advantage of having found the characteristic energy because the resulting total dispersion is only slightly lower. Still, the fact that the total dispersion decreases even if we are using a 'worse' proxy in terms of uncertainties means that physically the relation is indeed tighter when we are using the 1-keV flux as f_X. In the UV, we find no conclusive indications on a specific characteristic wavelength. This might be explained if the 'true' physical quantity is found at much shorter wavelengths, beyond the peak of

the disc emission, so assuming the optical–UV power law is less appropriate and/or informative. Consequently, we argue that the best choice is the 2500 Å flux simply because it is the one that results in a marginally lower dispersion. This might be because this wavelength falls in the observed spectra for the wider redshift range in our sample, implying a lower number of objects for which fluxes are determined via extrapolation and therefore with larger uncertainties.

3. When using the integrated Mg II line flux as $f_{\rm UV}$, we obtain a higher (steeper) slope value ($\gamma \sim 0.6$) and a lower dispersion ($\delta = 0.16$ dex) compared to any other UV continuum indicator derived from the spectroscopic analysis. We note that although the Mg II emission line is found at 2800 Å, its flux strongly depends on the quasar emission at much shorter wavelengths, around ~800 Å. Therefore, it might be that the physical relation behind the X-ray and UV luminosities is more strongly linked to the quasar emission at shorter wavelengths, and that as a consequence the Mg II emission-line flux works as a better proxy when compared to the fluxes in the 1350–5500 Å range. Unfortunately, the Mg II emission line is only available for the objects in our sample up to $z \sim 2.5$. Another possible explanation is that when we consider an indicator such as the Mg II emission line, which strongly depends on the extreme-UV SED shape, a tighter correlation with the soft X-ray can naturally arise due to the energy proximity of the bands involved.

4. The comparison between the values of the slopes that we found when using, respectively, the Mg II line fluxes and the spectroscopic monochromatic fluxes, is entirely consistent with the presence of the Baldwin effect.

5. We confirm a correlation between the X-ray and UV flux taking into account the FWHM of the Mg II line, while the UV flux and the FWHM turn out to be not significantly correlated. This non-correlation allows us to include the FWHM in the X-day/UV flux relation whenever available, and to keep using the standard relation otherwise. In this way, we can overcome the redshift limitations on the Mg II flux and still obtain a lower dispersion for the whole sample.

6. The Hubble diagram obtained from spectroscopic UV data is fully consistent with that obtained with photometric data. This is a validation of the previous results based on ‘photometric’ Hubble diagrams. Moreover, the ‘spectroscopic’ Hubble diagram shows a tension at statistical level of $\sim 3\sigma$ with the flat ΛCDM model. In previous works (e.g. Lusso et al. 2020, Bargiacchi et al. 2021), we obtained a higher significance of this tension, thanks to a wider redshift extent than the sample considered here. However, the Hubble diagram presented here is the first one where we have fully checked the UV spectral properties of the sources (and also the X-ray ones at $z > 1.9$).

Overall, the results presented in this chapter are another step towards the validation of the non-linear X-ray-to-UV relation of quasars as a reliable distance indicator. Since a limited of number of supernovae Ia are available at redshifts higher than ~1.5 and, by construction, a cosmology-independent validation of the method is impossible, the only way to further check our method is to search for possible evolutionary effects in the spectral emission of the quasars included in our Hubble diagram. In Sacchi et al. (2022), we demonstrated that the average continuum and line properties of quasars at $z > 2.5$ are perfectly matched to the ones of lower redshift counterparts in both the UV and X-rays. An extensive analysis of the stacked SDSS spectra in bins of BH mass and Eddington ratio for all the sources of the current sample is currently ongoing (Trefoloni et al. 2023, in preparation). We expect that future observations of supernovae at $z > 1.5$ will be able to independently probe any deviation from the concordance model found with the Hubble diagram of quasars.

Chapter 4

Observational causes of the residual dispersion in the $L_{\rm X}-L_{\rm UV}$ relation

The observed dispersion of the $L_{\rm X}$–$L_{\rm UV}$ relation is still a relevant problem for the implementation of quasars as standard candles, as is still considerable high ($\delta \sim 0.20$ dex). This poses a problem both for the implementation of quasars in the Hubble diagram, as the uncertainties on distance measurements that we obtain are much bigger than what one can achieve with SNIa, and for the reliability of the cosmological results themselves. It is indeed not fully clear how much of this dispersion is intrinsic to the physical mechanism behind the relation and how much is due to observational factors that we can not remove with the sample selection.

Recent studies have addressed this question. For example, Sacchi et al. (2022) reported a reduced dispersion of 0.09 dex for a subset of quasars at redshift $3.0 < z < 3.3$. This subset, despite its specific selection criteria, shows UV and X-ray properties consistent with the broader quasar population. Furthermore in the previous Chapter we have shown that, with the right selection of UV and X-ray proxies, it is possible to reduce the dispersion for a much larger sample from ~0.20 dex to ~0.16 dex.

These findings suggest that the intrinsic dispersion of the $L_{\rm X}$–$L_{\rm UV}$ relation is low, and that much of the observed dispersion is linked to observational, and not intrinsic, factors. In this chapter, we aim to investigate in detail these factors. We seek to quantify their contribution to the total observed dispersion and estimate the true intrinsic dispersion of the $L_{\rm X}$–$L_{\rm UV}$ relation. This exploration is key to understanding the physical connection between the accretion disc and the corona and, by extension, the viability of quasars as cosmological probes.

We address three main possible contributors to the dispersion: X-ray variability, potential biases in X-ray flux estimates via photometry, and the inclination of the accretion disc relative to our line of sight. We consider the total dispersion, δ, as a combination of the intrinsic dispersion of the $L_{\rm X}$–$L_{\rm UV}$ relation, $\delta_{\rm int}$, and the dispersion introduced by observational issues, $\delta_{\rm obs}$. With no universally accepted model explaining the $L_{\rm X}$–$L_{\rm UV}$ relation yet, our approach focuses on determining $\delta_{\rm obs}$ to better constrain and understand $\delta_{\rm int}$. Through this approach, we aim to provide valuable insights into this relationship and support the employment of quasars in cosmology. For this work we considered the objects selected in Lusso et al. (2020). This sample is made of ~2400 quasars, all of which have available UV

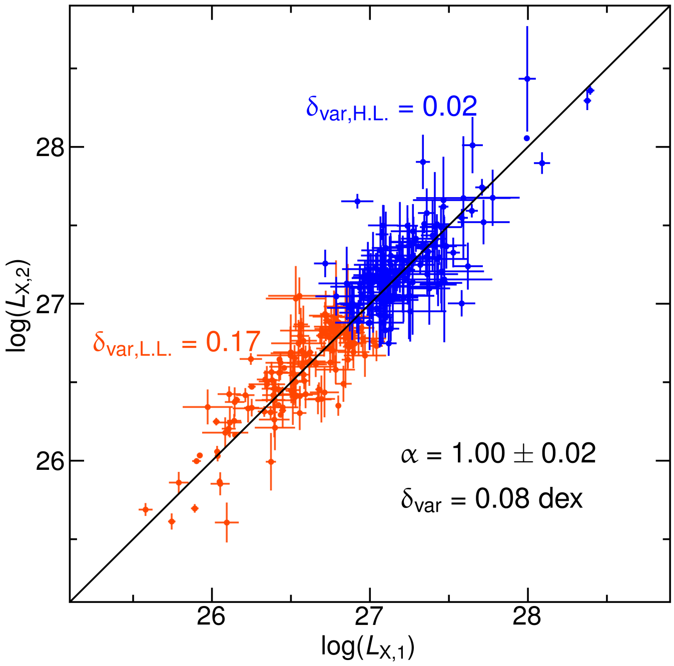

Figure 4.1: Comparison of the 2-keV luminosities for objects with multiple X-ray observations. The best fit is consistent with the bisector line. The dispersion parameter due to variability is $\delta_{\mathrm{var}} = 0.08$ dex. When distinguishing between the high- and low-luminosity subsamples, we see that the dispersion parameter is higher for lower luminosities. Luminosities are derived from photometric fluxes, assuming a standard flat ΛCDM model. We note that, as we are comparing luminosities for the same object, the results do not depend on the chosen cosmological model

and X-ray data in public catalogues. As discussed in Chapter 2 these objects have been selected in order to remove "biased" sources, which allowed us to reduce the observed dispersion from ~0.40 dex to ~0.20 dex.

4.1 Variability

Quasars exhibit non-periodic and stochastic variability across all observed wavelengths (Vanden Berk et al., 2004; Markowitz & Edelson, 2004), occurring on timescales ranging from hours (Ponti et al., 2012) to years (de Vries et al., 2005; Vagnetti et al., 2011). Despite significant efforts, the underlying mechanisms driving this variability are still not completely understood. Notably, an observed anti-correlation between luminosity and variability has been established (Hook et al., 1994; Kelly et al., 2009; Lanzuisi et al., 2014; Paolillo et al., 2017), attributed to the influence of black-hole mass and accretion rate on the variability pattern.

The timescales of quasar variability are wavelength-dependent, with X-ray emission displaying considerably faster variability compared to the optical band. In the context of the L_X–L_{UV} relation this means that, because of the nature of variability, the same UV state corresponds to a range of X-ray states. Therefore, even if the intrinsic dispersion of the L_X–

$L_{\rm UV}$ relation were null, this would introduce a scatter in the observed relation. Given that X-ray variability exhibits the largest amplitude on comparable timescales, it contributes most significantly to the observed dispersion. In this Section, we give an estimate of this contribution for the objects in the Lusso et al. (2020) sample.

To do so, we looked for the objects in the Lusso et al. (2020) sample that have more than one serendipitous observation in the XMM-*Newton* source catalogue 4XMM–DR9 (Webb et al., 2020). We found 289 objects with multiple observations; the vast majority (80%) has only two observations, so we considered the longest and the second longest observation for each object. Our goal was to compare the monochromatic 2 keV luminosities obtained from two observations that happened at random different times; we expected them to follow a one-to-one relation, with a scatter that would give us an estimate of the variability contribution to the dispersion in the $L_{\rm X}$–$L_{\rm UV}$ relation. So we fit the relation between the second longest ($L_{\rm X,2}$) and the longest ($L_{\rm X,1}$) observations with a line: $\log(L_{\rm X,2}) = \alpha \log(L_{\rm X,1}) + \zeta$, with the slope α and the normalization ζ as free parameters. The fit was performed with a Bayesian approach of likelihood maximization; we used the `emcee` code, which is an implementation of Goodman & Weare's Affine Invariant Markov Chain Monte Carlo (MCMC) Ensemble sampler (Foreman-Mackey et al., 2013). In building the likelihood, we need to consider that we have uncertainties of similar magnitude on both axes; to account for this, we adapted the BCES method (Akritas & Bershady, 1996), where the tangent ellipse is used to measure the distance of each point form the best-fit line.

The results are shown in Figure 4.1. We obtained, as expected, $\alpha = 1.00 \pm 0.05$ a slope, and an intercept value of $\zeta = -0.005 \pm 0.015$. We also derived the total dispersion of the relation as:

$$\delta_{\rm tot} = \sqrt{\sum_{\rm i}^{N} \log(L_{\rm X,2})_i^2 - (\alpha \log(L_{\rm X,1})_i + \zeta)^2 / N} \tag{4.1}$$

which turns out to be 0.17 dex. If the uncertainties on the x- and y-axis completely explained this dispersion, no intrinsic scatter due to variability would be present. However, we computed the average observational uncertainty on both $\log(L_{\rm X,1})$ and $\log(L_{\rm X,2})$ and we obtained 0.15 dex. Given this result, we can compute the intrinsic dispersion due to variability simply as $\delta_{\rm var} = \sqrt{\delta^2_{\rm tot} - \delta^2_{\rm err}}$, and get $\delta_{\rm var} = 0.08$ dex. This value can therefore be considered as the average contribution of X-ray variability to the dispersion in the $L_{\rm X}$–$L_{\rm UV}$ relation.

Comparing two fluxes for the same object, so that we believe are two random extraction from the same distribution centred around a mean flux value, might be bringing an overestimate of the variability contribution to the dispersion. However, we checked that our fit procedure, with the dispersion being derived from the distance to the best fit, is an accurate way to measure it. To do so, we created different mock samples of quasars assuming different variabilities and then we fit them with our procedure, always recovering the assumed initial value and not overestimating it.

Although the number of objects for which we have multiple observations is only ~12% of the total Lusso et al. (2020) sample, these objects are fully representative (given that the observations we are considering are serendipitous) and they span the same luminosity range as the whole sample. So we can consider the result from this analysis to be applicable

to the whole quasar catalogue.
We note that a similar analysis on objects with multiple observations was performed in Lusso & Risaliti (2016), with a smaller sample of 159 objects, and they found the variability contribution to be $\delta_{\rm var} = 0.12$ dex. Considering that the analysis described in this Section was performed with a sample which is more than doubled in statistics, we believe our estimate of $\delta_{\rm var} = 0.08$ to be more accurate.
n Chiaraluce et al. (2018), the authors derived an estimate of the variability contribution to the dispersion in the L_X–L_{UV} relation using the Structure Function (SF) of objects observed multiple times. They discovered that variability accounts for approximately 56% of the total dispersion, which is 0.31 dex. Although their selection method differs from ours, from which the higher ”starting point” of the dispersion, the SF analysis has proven to be an alternative approach to assess the variability contribution. This method warrants further investigation, especially when a larger dataset of objects with more than two observations becomes available.

We also tested the dependence of variability on luminosity for our sample, by dividing it into two subsamples, above and below the value of $\log(L_X/{\rm erg\,s^{-1}\,Hz^{-1}}) = 26.9$. The objects in the subsamples are 145 and 144, respectively, and for each of them we derived the total dispersion and the average uncertainty as described above. As can be seen in Fig. 4.1, the high-luminosity subsample shows a smaller dispersion due to variability ($\delta_{\rm var,H.L.} = 0.02$ dex) than the low-luminosity one ($\delta_{\rm var,L.L.} = 0.17$ dex). This is not only consistent with results in the literature, but it can, at least partially, also explain the results of Sacchi et al. (2022), where a dispersion as low as 0.09 dex is observed for a subsample of objects at redshift $3.0 < z < 3.3$. Together with having spectroscopic data and a subsample of pointed X-ray observations, these objects have an average luminosity of $\log(L_X/{\rm erg\,s^{-1}\,Hz^{-1}}) \sim 27.7$, which is in the “high luminosity” regime. Therefore, the contribution of variability to their total dispersion must be very little, if not even zero.

4.2 Inclination

The second factor contributing to the observed dispersion we are considering, is the inclination of the quasar accretion disc relative to the line of sight. The optical–UV intrinsic emission from quasars is typically attributed to a disc-like component. The angle at which a quasar is viewed then crucially influences its observed flux. Specifically, unless the quasar is perfectly face-on, the observed flux, $f_{\rm obs}$, is derived from the intrinsic UV flux ($f_{\rm int}$) as $f_{\rm obs} = f_{\rm int} \cos\theta$, where θ is the angle between the observer’s line of sight and the quasar disc axis. Notably, while the inclination affects the observed UV flux, it has no effect on the X-ray flux. This distinction arises because X-ray ‘coronal’ emission is believed to be isotropic. The exact location and geometry of the corona are still largely unknown, although polarization results are now providing new perspectives on the topic (e.g., Gianolli et al., 2023). We note here that we will assume the X-ray emission to be isotropic throughout this work.
This inclination effect, by changing the UV flux, changes the slope of the relation, differently for each different inclination angle. Overall this results in an increase in the observed

dispersion for the global quasar sample. Moreover, this effect is asymmetric, impacting quasars differently based on their relative brightness with respect to the detection limit. Bright quasars, surpassing by far the detection threshold, will still be detected even with a diminished observed flux $f_{\rm obs}$ at large inclination angles. However, for quasars nearing the flux limit, there is a range of $\cos\theta$ values where they become undetectable.
In an ideal scenario, if we had accurate knowledge of the inclination angle θ for every quasar observed, we could correct the observed flux to counterbalance inclination effects. Unfortunately, we do not have consistent estimators for quasars inclination. Some works suggest that the intensity of the [O III] line is an indicator of the inclination angle of a quasar (e.g. Bisogni et al., 2017). Yet, this result is valid in a statistical sense, but can not be reliably used for an object by object correction. Furthermore, this line escapes the SDSS spectral range for redshifts above ~0.7.
Given that we lack observational methods to measure inclinations directly, we will tackle this challenge using mock samples of quasars. While we might not be able to eliminate this source of dispersion, we can estimate its contribution. This is essential for accurately determining the intrinsic dispersion of the $L_{\rm X}$–$L_{\rm UV}$ relation. In this Section, we discuss how to build a mock sample of quasars to correctly represent the effect of inclination, and derive an estimate for its contribution to the total dispersion.

4.2.1 Correction to a luminosity function

The first step to build our mock sample is to determine from which luminosity distribution we should extract our objects. Numerous studies have established the observed UV luminosity function for quasars and its redshift evolution (e.g. Boyle et al., 2000; Croom et al., 2009; Ross et al., 2013). However, within a particular luminosity range, the quasars we observe are only those that, once inclined, have a flux above the survey flux limit. Consequently, the observed luminosity function for quasars is biased, and a priori we do not know how much the shape of the luminosity function changes because of that. Here we therefore attempt to deduce the intrinsic luminosity function of quasars from the observed one, and then use it as the starting point for our mock sample.
We designate θ as the inclination angle, signifying that a face-on quasar has $\theta = 0$ whilst an edge-on quasar has $\theta = \pi/2$, as shown in Fig. 4.4. We postulate that quasars are randomly inclined in the sky. For a given intrinsic luminosity $\mathcal{L}$, the observed luminosity is $L = \mathcal{L}\cos\theta$, where $L = \mathcal{L}$ for face-on quasars and $L \to 0$ for those seen at increasingly large inclinations.

Assuming the true quasar luminosity distribution in the Universe is a continuous function $m(\mathcal{L})$, and given that quasars are randomly inclined, a specific observed luminosity bin would contain objects with particular combinations of intrinsic luminosity and inclination angle. Hence, the observed distribution $n(L)$ can be expressed in terms of the intrinsic distribution $m(\mathcal{L})$:

$$n(L)dL = \int_0^{\pi/2} m\left(\frac{L}{\cos\theta}\right)\sin\theta\, d\theta d\mathcal{L} \tag{4.2}$$

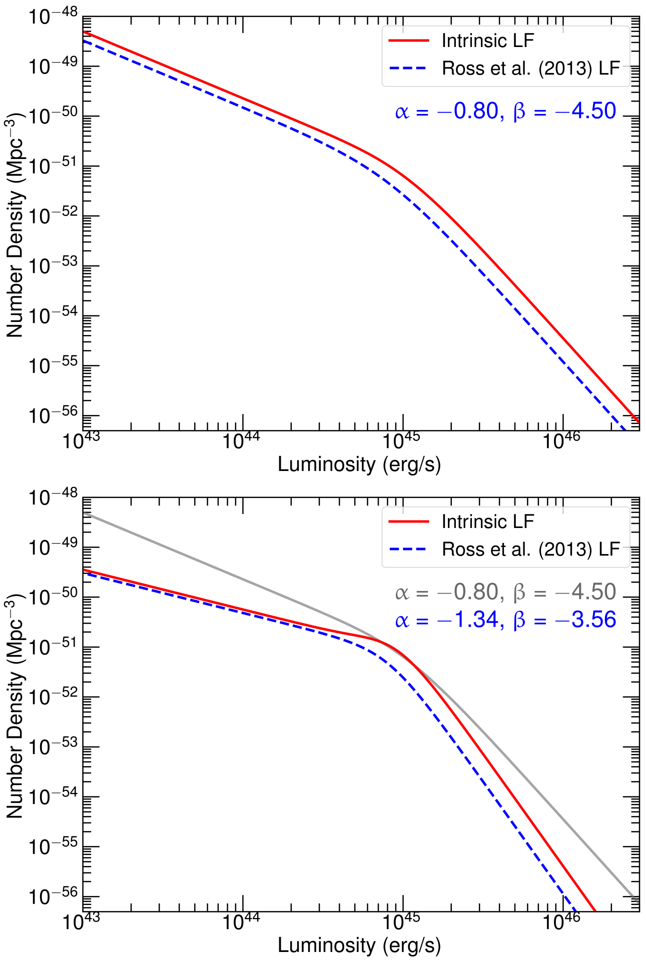

Figure 4.2: *Left*: observed luminosity function $n(L)$, in blue dashed, as per the Ross et al. (2013) parametrization: the luminosity function is a broken power law with indexes $\alpha = -1.34$ and $\beta = -3.56$, and $\log(\phi_*) = -6.15$. A Pure Luminosity Evolution is assumed. In this Figure we show, as an example, the results for $z = 1$. In solid red, the intrinsic luminosity function $m(\mathcal{L})$, which is corrected for inclination effects. *Right*: same as the left panel, but with an observed luminosity function assumed to have indexes $\alpha = -0.8$ and $\beta = -4.5$. The comparison between the two panels shows that a greater difference between α and β accentuates the 'knee' distortion and the variance between $m(\mathcal{L})$ and $n(L)$ at higher luminosities. For a better comparison, the red solid line of the left panel is also plotted in the right panel, in grey.

Detailed calculations in the Additional Material Section 4.6 allow us to derive the true luminosity function $m(\mathcal{L})$ from the observed one, $n(L)$:

$$m(\mathcal{L}) = n(L) - n'(L)L \tag{4.3}$$

The result of this correction is therefore the distribution of intrinsic quasar luminosities, which would be the observed distribution if all the objects were face-on. To properly compare with the observed $n(L)$, we multiply $m(\mathcal{L})$ by the average $\cos\theta$, which is 0.5. In Figure 4.2, we show, as an example, the juxtaposition the observed $n(L)$ with the derived intrinsic luminosity function $m(\mathcal{L})$, assuming the quasar luminosity function shape from Ross et al. (2013):

$$\frac{dn}{dL} = \frac{\phi_*}{\left(\frac{L}{L^*}\right)^{-\alpha} + \left(\frac{L}{L^*}\right)^{-\beta}} \tag{4.4}$$

with parameters α and β again following the results from Ross et al. (2013): $\alpha = -1.34$ and $\beta = -3.56$, and $\log(\phi_*) = -6.15$. As in the Ross et al. (2013) parametrization, we assume the quasar luminosity function to follow a Pure Luminosity Evolution (PLE). In Fig. 4.2, as an example, we show the results assuming a redshift value of $z = 1$. Notably, at both low and high luminosities, the shape of the intrinsic luminosity function aligns with the observed one. The main differences are found around the change in the slope, commonly called the 'knee'. Altering the α and β values impacts the relative shapes: a greater difference between α and β accentuates the 'knee' distortion and the variance between $m(\mathcal{L})$ and $n(L)$ at higher luminosities. The right panel of Fig. 4.2 showcases the results for $\alpha = -0.8$ and $\beta = -4.5$.

4.2.2 Mock sample

Now that we know how to derive an inclination-corrected luminosity function, we can use it to build our mock sample of quasars to determine the effect of inclination on the $L_{\rm X}$–$L_{\rm UV}$ observed dispersion. We do so, at first, allowing the inclination angle to vary between 0 and $\pi/2$.

We start building a sample of 100,000 objects, and we consider the quasar luminosity function obtained by Ross et al. (2013) for the SDSS as the starting point. Such a luminosity function assumes a redshift dependence. Therefore, we first assign a random redshift to the objects in the mock sample, extracting them from the redshift distribution of L20. Then, for each object, we derive the luminosity function corresponding to that redshift from Ross et al. (2013), we correct it for the inclination effect described in the previous subsection and we use it to extract a luminosity value for that object. The luminosity function described in Ross et al. (2013) is derived for the i-magnitude, but we are interested in the monochromatic luminosity at 2500 Å. Therefore, for each object, we derive it by assuming an SED with $f_\nu \sim \nu^{-\alpha}$, with $\alpha = 0.5$. We tested all of the following also assuming $\alpha = 1$ and we always obtained consistent results. Given the 2500 Å luminosities, we assume the $L_{\rm X}$–$L_{\rm UV}$ relation with γ=0.6 and derive the corresponding $\log(L_{\rm X})$ for each object. The values of $\log(L_{\rm X})$ are then shifted by a random quantity extracted from a Gaussian with mean zero and standard

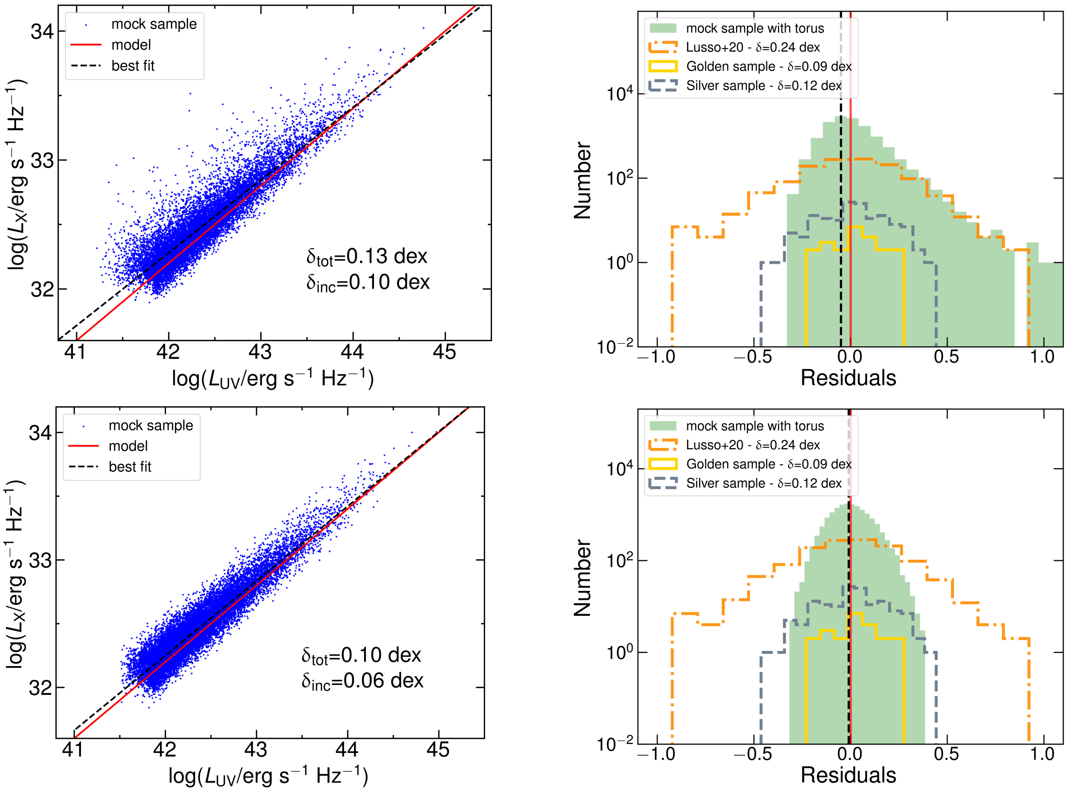

Figure 4.3: *Upper left*: mock sample of 100.000 quasars assuming an intrinsic relation with zero intrinsic dispersion, a contribution from variability to the observed dispersion of 0.08 dex, and an inclination angle between 0 and 90°. For each object, a redshift is assigned using the Lusso et al. (2020) (L20) sample redshift distribution; given the redshift, the corresponding luminosity function from Ross et al. (2013) is derived, and the luminosity of the objects is extracted. The red solid line represents the starting sample, with a slope $\gamma = 0.6$ and a zero dispersion. The blue points show the sample after the dispersion due to variability is added and the objects are assigned a random inclination. The total dispersion of the sample is $\delta_{\rm tot} = 0.13$ dex, the inclination effect accounts for $\delta_{\rm inc}$=0.10 dex. *Upper right*: in green filled, the histogram of the fit residuals. This distribution is skewed (skewness parameter $s = 1.18$). In dot-dashed orange, the histogram of the residuals for the L20 sample. The L20 distribution, which is the observed one, is much more symmetric (skewness parameter $s_{\rm L20} = 0.20$). In dashed silver and in solid gold, the residuals corresponding to the “silver sample” and the “golden sample” of Sacchi et al. (2022). The red solid vertical line corresponds to zero, while the dashed black vertical line corresponds to the peak of the mock sample distribution, equal to −0.05. All the histograms are shown in logarithmic units to enhance the readability. *Lower left*: the same as the Upper Left panel, but assuming the presence of an obscurer with an angle width of $\psi_{\rm torus} = 25$, which means that the inclination angle for the objects in the mock sample can go from 0 to 65°. The total dispersion is reduced to $\delta_{\rm tot} = 0.10$ dex, and the inclination effect accounts for $\delta_{\rm inc}$=0.06 dex. *Lower right*: the same as the Upper right panel, but assuming the presence of an obscurer with an angle width of $\psi_{\rm torus} = 25°$. Here we see that the residuals distribution is symmetric, with a skewness parameter of $s = 0.19$ which is consistent with the L20 value.

deviation equal to 0.08; this mimics the contribution to the dispersion due to the variability, which, as we found out in Section 4.1, is $\delta_{\rm var} = 0.08$. dex Then, for each object, we assign a random value of the inclination angle θ, from 0 to $\pi/2$, as we are here assuming no absorbing torus. The angle is extracted from a distribution which is uniform in $\cos\theta$. Given θ, the inclined UV luminosity is derived as $L_{\rm obs} = \mathcal{L}\cos\theta$, while the X-ray luminosity is left untouched. We now have a mock sample of inclined objects and we have to consider the presence of an observational threshold. We consider as a threshold the *i*-magnitude flux limit of the first SDSS release (Richards et al., 2002), $m_i = 19.1$, from which we derived the corresponding monochromatic flux at 2500 Å [1]. So for each object, given its redshift, we derived its 2500 Å flux by assuming a standard flat ΛCDM cosmology and then removed the object from the mock sample if it fell below the threshold. We now fit the relation between $\log(L_{\rm X})$ and $\log(L_{\rm UV})$, where $L_{\rm UV}$ has the inclination-affected values. The fit is done with a Bayesian approach of likelihood maximisation, assuming the function $\log(L_{\rm X}) = \gamma\log L_{\rm UV} + \beta$, where γ and β are two free parameters, and where we take the presence of a dispersion δ into account by modifying the likelihood function accordingly. As before, we performed the fit using the *emcee* code. The mock sample and the results of the fit are shown in the top left panel of Fig. 4.3. We retrieve a slope coefficient $\gamma_{\rm fit} = 0.58 \pm 0.02$, consistent with the assumed value of $\gamma = 0.6$. This reassures us that the inclination effect is not significantly affecting the slope of the relation. The total dispersion is $\delta_{\rm tot} = 0.13$ dex. Given that we assumed a dispersion of $\delta_{\rm var} = 0.08$ dex due to the variability, to obtain the contribution of the inclination we have to quadratically subtract the two, so that $\delta_{\rm inc} = \sqrt{\delta^2_{\rm tot} - \delta^2_{\rm var}}$. The result is $\delta_{\rm inc} = 0.10$ dex. On the top right panel of Fig. 4.3 we show, in green, the histogram of the fit residuals (which we obtain by subtracting the best-fit model from the mock data). In dot-dashed orange, we show the same histogram for the Lusso et al. (2020) sample. It is clear that the point distribution of our mock sample is not representative of the observed scenario, as it is (slightly) off-centred and significantly skewed, with a skewness parameter of $s = 1.18$, while the Lusso et al. (2020) distribution is much more symmetric, with a skewness parameter of $s_{\rm L20} = 0.2$. In the top right panel, we also show the distribution of the residuals for the "silver" and the "golden" samples of Sacchi et al. (2022) which will be further discussed in Section 4.4.

To enhance the accuracy of our mock sample, we introduced in our model an obscurer, characterized by a maximum angle, $\psi_{\rm torus}$ assumed from the accretion disc, as shown in the scheme in Fig. 4.4. We assume this torus to be a homogeneous dust distribution that extends from the plane of the accretion disc to $\psi_{\rm torus}$. Therefore, if a quasar has an inclination angle that exceeds $\theta_{\rm max} = \pi/2 - \psi_{\rm torus}$, the torus absorbs its emission, making it undetectable[2] . This results in an accessible angle range of $[0, \theta_{\rm max})$, instead of the initial $[0, \pi/2)$ range. We find that by increasing $\psi_{\rm torus}$, the residual histogram becomes more and

---

[1] The flux selection for the true observed sample is going to be much more complicated than a simple flux threshold. However, here we want to recreate a simpler scenario of a uniform sample with a given flux limit. We tested that the final results in terms of the inclination contribution to the dispersion do not depend on the exact choice of the flux limit.

[2] We basically assumed an infinite optical depth for the torus. This is not truly representative of the real scenario, but does not affect the consistency between our mock sample and the observed sample, as in the observed sample even mildly obscured objects are removed by our sample selection.

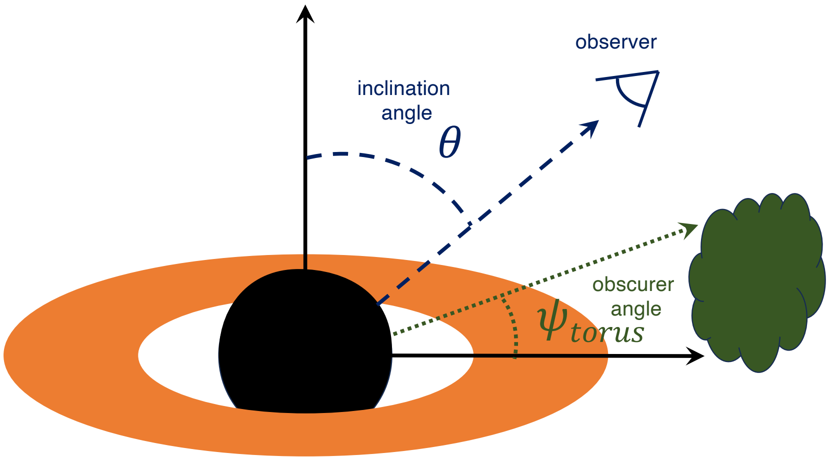

Figure 4.4: Schematic representation of a quasar observed at a certain angle θ. The angle is measured starting from the accretion disc axis. If we assume the presence of an obscurer with a certain opening angle $\psi_{\rm torus}$, measured from the disc surface, it means that the inclination angle θ can vary between zero (face-on) and $\theta_{\rm max} = \pi/2 - \psi_{\rm torus}$.

more symmetric. After some tests, we deduced that a $\theta_{\rm max}$ of ~65° (which corresponds to an obscurer angle of $\psi_{\rm torus}$=25°) is the maximum value that achieves a distribution of the residuals similar to the Lusso et al. (2020) sample, as visualized in the lower right panel of Figure 4.3. This new histogram has a skewness parameter of $s = 0.19$, consistent with what we find for the Lusso et al. (2020) sample residuals histogram, which is $s_{\rm L20} = 0.20$. We note that the Lusso et al. (2020) residuals histogram is wider because while our mock sample has an overall dispersion of $\delta = 0.10$ dex, Lusso et al. (2020) has $\delta = 0.21$ dex. The lower left panel of Fig. 4.3 displays the new fit. Due to the narrower angle range, objects disperse less around the best fit, resulting in a total dispersion of $\delta_{\rm tot} = 0.10$ dex. By considering this dispersion and subtracting the variability contribution quadratically, we obtain $\delta_{\rm inc} = 0.06$ dex. Given the match in the residual distribution shape with the actual observed sample, we believe this to be a more reliable inclination dispersion estimate.
The concept of a toroidal absorber to describe the emission of quasars is not novel and it is a fundamental part of the AGN "unified model" (see for example Bianchi et al. 2012 for a review). At the same time, it is noteworthy that our mock sample necessitated an obscuring torus based solely on the comparison with the histogram of the residuals from the Lusso et al. (2020) sample, as this comes out as a somewhat indirect way to assess the minimum required angular width of the torus for an average population of quasars. In summary, our mock quasar sample analysis allows us to estimate the inclination contribution to the observed dispersion, approximating it at $\delta_{\rm inc} = 0.06$ dex.

We note that, in order to retrieve the intrinsic luminosity distribution, we could not start from the luminosity distribution of L20, as the latter is not only affected by inclination but also by additional filtering criteria that make the selection function very complex. As the latter effects cannot be corrected, we adopted instead the luminosity function of Ross et al.

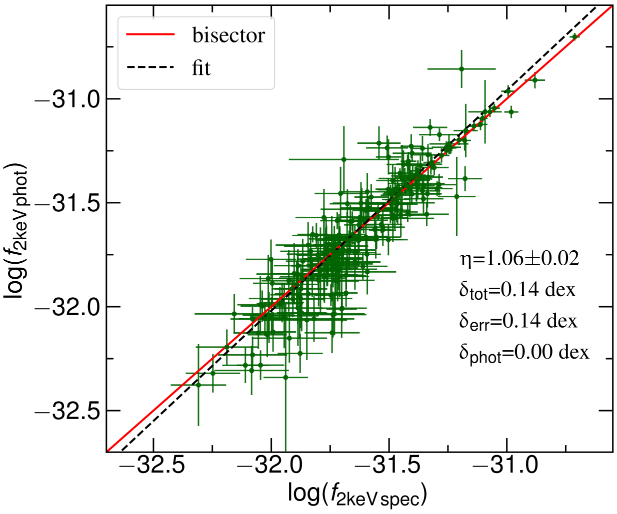

Figure 4.5: Comparison of the 2-keV monochromatic flux estimate derived from photometric data (see L20) and the one derived from spectroscopic data for the subsample of objects with only one observation in the 4XMM-DR9 catalogue. The solid red line represents the bisector, while the black dashed line is the best fit, with a slope of $\eta = 1.06 \pm 0.02$. The total scatter around the best fit is equal to the average uncertainties, $\delta_{\rm err} = \delta_{\rm tot} = 0.14$ dex. The "intrinsic" dispersion of the relation is therefore zero. Overall, we can say that there is no significant contribution to the observed dispersion that comes from using photometric X-ray data instead of spectroscopic data.

(2013) to build our mock sample. At the same time, it is important to note that the selection of Lusso et al. (2020) does not depend on the UV luminosity itself, so it should not alter the contribution of inclination to the total observed dispersion. To address this assumption, we performed two additional tests. First, we estimated the dispersion contribution in small redshift bins ($\Delta z \sim 0.2$), in the redshift range of the Lusso et al. (2020) sample (which goes from $z \sim 0$ to $z \sim 5$). At all redshifts, the resulting contribution of the inclination to the final dispersion is $\delta_{\rm inc} \sim 0.06$ dex, confirming the absence of a luminosity trend in this contribution. As an additional test, we build another mock sample starting from the Lusso et al. (2020) luminosity distribution, and correcting that distribution for the inclination factor. We also assumed the Lusso et al. (2020) redshift distribution to derive the redshift, and performed all the steps described above. The results are shown in the Additional material 4.7, and again a contribution of the inclination to the total dispersion of $\delta_{\rm inc} \sim 0.06$ dex is confirmed.

The results presented here demonstrate that the dispersion due to the inclination does not depend on luminosity or redshift and that, overall, we can safely consider its contribution to be $\delta_{\rm inc} \sim 0.06$ dex. Differences in the exact shape of the starting luminosity distribution for the mock sample do not seem to affect the estimate.

4.3 X-ray analysis

Finally, another potential source of observational dispersion might arise from the use of photometric data to calculate the 2-keV monochromatic luminosities of our sources, instead of spectroscopic data. In this Section we show that this is not the case, and we can safely use photometric X-ray data. In the previous Chapter, we tested whether there is any offset between the photometric and spectroscopic flux determination. To do that, we performed the X-ray spectral analysis on a sample of 231 objects, which are all the quasars in the Lusso et al. (2020) catalogue with XMM-*Newton* observations and a redshift higher than 1.9. The findings indicate that employing photometrically derived fluxes (or luminosities) instead of spectroscopically derived ones does not introduce any significant offset. However, it is possible that photometrically derived fluxes exhibit a greater scatter around the "true value" compared to spectroscopically derived ones. If this were the case, our prevalent use of photometric data would not introduce any bias in determining the relation parameters, but it would still increase the observed dispersion.

To investigate this further, we used the X-ray spectroscopically analysed sample from the previous Chapter, and we tested the equation:

$$\log(f_{2\,\mathrm{keV,\,spec}}) = \eta \log(f_{2\,\mathrm{keV,\,phot}}) + \epsilon. \tag{4.5}$$

We employed the same fitting procedure described in Section 4.1. From the 231 sources, we removed the 38 that have multiple observations. Indeed, these objects have a photometric flux estimate that has been obtained by averaging multiple observations (Lusso & Risaliti, 2016; Lusso et al., 2020), while the spectroscopic analysis has been performed on the longest observation. As the flux estimates have been obtained in different ways, there is no point in comparing them.

From the fit analysis, we obtained a slope close to unity ($\eta = 1.06 \pm 0.02$), as previously discussed in the previous Chapter. The results are illustrated in Figure 4.5. We found both δ_{tot} and δ_{err} to be equal to 0.14 dex, indicating that the "intrinsic" dispersion δ_{phot} is consistent with zero. The outcome of this analysis is clear: when comparing the spectroscopic and photometric 2 keV flux estimates for a subsample of objects where the same observations were used to derive both estimates, there is no additional dispersion beyond that arising from observational uncertainties. Consequently, utilising photometric data instead of spectroscopic ones does not introduce any offset or bias, nor does it contribute to the total observed dispersion. This finding supports the validity and reliability of employing photometrically derived fluxes (or luminosities) in our astrophysical analyses, which is particularly relevant considering that deriving spectroscopic X-ray monochromatic flux estimates for thousands of objects is very time-consuming.

Another potential issue with X-ray observations is their off-axis nature. The majority of X-ray data for the Lusso et al. (2020) sample is derived from serendipitous observations, exhibiting a mean off-axis angle for the XMM-*Newton* sources—which constitute most of the sample—of $\phi_{\mathrm{offax}} \sim 6.4'$. The objects not being at the detector centre might introduce additional dispersion. This was tested on a subset of 1778 objects corresponding to those with serendipitous XMM-*Newton* observations in the Lusso et al. (2020) sample. We examined the L_X–L_{UV} relation across six off-axis angle bins: 0-2', 2'-4', 4'-6', 6'-8', 8'-10', and 10'-12'. In each bin, we observed a dispersion of $\delta \sim 0.24$ dex, with no significant trend relating to the off-axis angle. This suggests that the off-axis angle does not substantially contribute to the dispersion. However, it is important to note that when the observed dispersion is large, minor differences become indiscernible. Assuming that objects at an off-axis angle of ~10' have an additional dispersion factor of $\delta_{\rm offax} \sim 0.04$ dex, this difference would be obscured in a comparison with a sample having a smaller off-axis angle, since $\sqrt{0.24^2 - 0.04^2} \sim 0.24$, and the variance is masked by uncertainties in the dispersion estimate. One approach to further investigate this might involve analysing a subsample with a range of off-axis values but a lower initial observed dispersion, such as the 'golden sample' from Sacchi et al. (2022). However, in the 'golden sample', objects with a smaller off-axis angle (the pointed objects from Nardini et al. (2019)) also exhibit significantly higher average luminosity compared to those at a greater off-axis angle. Given this, along with the small overall statistics for this sample, we can not confidently distinguish the effect of lower variability due to higher average luminosity from the potential effect of dispersion reduction due to pointed observations. Future targeted observations on select subsamples might provide clearer insights into the actual impact of off-axis observations. For now, we can conclude that if this effect exists, it is likely a minimal contribution, with $\delta_{\rm offax} \leq 0.04$ dex.

4.4 Comparison with observational results

We have determined that the contribution from X-ray variability to the dispersion is approximately $\delta_{\rm var} \sim 0.08$ dex. It is also evident that low-luminosity objects exhibit a greater contribution than their high-luminosity counterparts. Furthermore, our findings indicate that the use of X-ray photometric data introduces no additional dispersion to the relation. Through the construction of mock samples, we have determined that the inclination of the quasar accretion disc relative to our line of sight contributes to the total observed dispersion with $\delta_{\rm inc} \sim 0.06$ dex. The observational sources of dispersion we have assessed collectively contribute therefore with a residual dispersion of $\delta_{\rm res} = \sqrt{0.06^2 + 0.08^2} = 0.10$ dex. We can now compare these results with the most recent estimates of the observed dispersion.

Firstly, we consider the results of the previous Chapter. Using UV spectroscopic data and the best UV and X-ray proxies for the correlation, we found a dispersion of $\delta_{\rm obs} = 0.16$ dex. When compared to our current estimate of the residual dispersion $\delta_{\rm res} = 0.10$ dex, it is evident that the intrinsic dispersion of the correlation must be less than $\delta_{\rm int} \leq \sqrt{\delta^2_{\rm obs} - \delta^2_{\rm res}} = 0.12$ dex.

The work of Sacchi et al. (2022) offers additional clues. They presented a sample of quasars, a subsample of the Lusso et al. (2020) one, with high-quality data and a one-by-one UV and X-ray spectral analysis. This sample, which we call here the 'silver sample', showed a dispersion of $\delta_{\rm silver} = 0.12$ dex. Within this sample, they also highlighted a subsample of objects at redshift $z \sim 3$, referred to as the 'golden sample', with an even lower observed dispersion, $\delta_{\rm golden} = 0.09$ dex. We display the residuals with respect to the $L_{\rm X}$–$L_{\rm UV}$ relation for these two samples in the right panels of Fig. 4.3, alongside the residuals for the mock samples and for the Lusso et al. (2020) sample discussed previously. Considering the 'silver' sample, if we sum up quadratically the variability (0.08 dex) and the inclination (0.06 dex), up to 0.10 dex of its dispersion can be attributed to variability and inclination. Hence, the intrinsic dispersion for the relation is estimated to be $\delta_{\rm int} \leq \sqrt{0.12^2 - 0.10^2} \sim 0.07$ dex. For the golden sample, its dispersion is already smaller than the total of 0.10 dex found for our mock sample. This can be explained in terms of the sample high average luminosity, $\log(L_X/\mathrm{erg\,s^{-1}\,Hz^{-1}}) \sim 27.7$. Consequently, the overall variability contribution to the dispersion is minimal, approximately $\delta_{\rm var,H.L.} \sim 0.02$ dex, as detailed in Section 4.1. The intrinsic dispersion of the L_X–L_{UV} relation for the 'golden' sample can therefore be estimated as $\delta_{\rm int} \leq \sqrt{0.09^2 - 0.06^2 - 0.02^2} \sim 0.06$ dex. Remarkably, the estimates for both subsamples are similar.

To sum up, when examining the entire quasar sample, data-quality constraints limit us to a dispersion no lower than 0.21 dex for the Lusso et al. (2020) data set and 0.16 dex for the one utilising UV spectroscopic data along with the optimal proxies for UV and X-ray emission. Of this total dispersion, 0.10 dex is attributable to the combined impact of variability and inclination. In cases of high-quality data, the dispersion can drop to 0.12 dex when the variability contribution is still significant, and to 0.09 dex when high luminosities reduce the variability contribution, as demonstrated in Sacchi et al. (2022). From these analyses, we deduce that the intrinsic dispersion of the L_X–L_{UV} relation must be equal to or lower than $\delta_{\rm int} \sim 0.06$.

4.5 Summary

In this Chapter we investigated those factor that, although not being intrinsic to the $L_X - L_{UV}$ relation, might contribute to its observed dispersion, and that we can not remove by selecting unbiased samples. We identified three possible dispersion sources:

- Variability: quasar emission is known to be variable both in the UV and in the X-ray bands, which inevitably causes an increase in the observed dispersion. Given the shorter timescales, X-ray variability is the one that is going to predominantly affect our results. To test the contribution of variability, we selected the 289 objects in the Lusso et al. (2020) sample that have multiple X-ray observations in the XMM-*Newton* 4XMM-DR9 catalogue. We found that the average scatter between different estimates of the 2-keV monochromatic luminosity is $\delta_{\rm var} \sim 0.08$ dex, which we can therefore consider as the variability contribution. We also found, consistently with literature results, that more luminous objects show less variability than the less luminous ones, with the "high luminosity" subsample having an estimate of only $\delta_{\rm var,H.L.} = 0.02$ dex variability contribution to the total observed dispersion.

- Inclination: the inclination of the accretion disc is believed to affect the observed UV luminosity, but not the X-ray one. Therefore, the different quasar inclinations introduce a scatter in the $L_X - L_{UV}$ relation. Unfortunately, we do not have observational methods to derive inclination estimates and correct the UV luminosities accordingly. Therefore, we relied on mock-sample estimates to derive the inclination contribution to the observed dispersion. We discussed how to recover the intrinsic luminosity distribution given an observed one, correcting for the inclination effect. Starting from the Ross et al. (2013) luminosity function we then build a mock sample of quasars for which we found the inclination contribution to the dispersion to be $\delta_{\rm inc} \sim 0.06$ dex, once we account for the presence of an obscuring torus.

- X-ray photometry: in the previous Chapter we have shown that using photometric estimates of the 2-keV monochromatic luminosities instead of spectroscopic ones does not introduce any systematic offset (e.g. due to high offaxis angles). However, it might still introduce an additional (observed and intrinsic) dispersion. In this work we tested this for a sample of 193 objects at redshift $z > 1.9$ and we found that photometric estimates do no introduce additional intrinsic additional scatter, compared to spectroscopic ones. Even though the observed scatter is larger with photometric measurements, this is accounted for by the larger uncertainties on the photometric data with respect spectroscopic ones. This is reassuring because it allows us to keep using photometric data when spectroscopic ones are not available, confident that we are not introducing additional systematic offset in the dispersion.

Comparing our results with recent observational estimates of the dispersion from Sacchi et al. (2022), we conclude that the intrinsic dispersion of the L_X–L_{UV} relation is exceedingly low, $\delta_{\rm int} \leq 0.06$. This finding reinforces the hypothesis that the physical mechanism governing the L_X–L_{UV} relation is remarkably consistent across a broad range of redshifts and luminosities. In doing such, this outcome gives additional credibility to using quasars as standard candles in cosmological studies, and it particularly highlights the significant tension existing within the flat ΛCDM cosmological model. The findings on variability confirm the results presented in Sacchi et al. (2022): subsamples with higher luminosities tend to exhibit the lowest observed dispersion values.Looking ahead, new targeted X-ray observations of high-redshift and correspondingly high-luminosity quasars promise to yield subsamples with exceedingly low dispersion, thereby enhancing our understanding of the high-redshift Hubble diagram and associated cosmological tensions.

4.6 Additional material: deriving the luminosity function correction

Here we report how we derive the "intrinsic" luminosity function $m(\mathcal{L})$ from the observed one, $n(L)$. We recall that the observed luminosity function consists of objects where the inclination effect has modified each object observed luminosity, while we are interested in recovering the intrinsic luminosity distribution. We start from the relation between $n(L)$ and $m(\mathcal{L})$.

$$n(L)dL = \int_0^{\pi/2} m\left(\frac{L}{\cos\theta}\right)\sin\theta\, d\theta d\mathcal{L} \tag{4.6}$$

We substitute $\cos\theta = x$ and $dx = -\sin\theta\, d\theta$:

$$n(L) = \int_0^1 m(L/x)dx \tag{4.7}$$

We can now write that $L/x = \mathcal{L}$ and $-(L/x^2)\,\mathrm{d}x = \mathrm{d}\mathcal{L}$ so that

$$n(L) = \int_L^{\infty} m(\mathcal{L})\frac{L}{\mathcal{L}^2}d\mathcal{L} \tag{4.8}$$

We derive with respect to L:

$$\frac{\partial}{\partial L}n(L) = \frac{\partial}{\partial L}\int_L^{\infty} m(\mathcal{L})\frac{L}{\mathcal{L}^2}d\mathcal{L} = \int_L^{\infty}\frac{m(\mathcal{L})}{\mathcal{L}^2}d\mathcal{L} - \frac{m(L)}{L} \tag{4.9}$$

and then derive again:

$$\frac{\partial^2}{\partial L^2}n(L) = -\frac{m(L)}{L^2} - \frac{Lm'(L) - m(L)}{L^2} = -\frac{m'(L)}{L} \tag{4.10}$$

from which:

$$m(\mathcal{L}) = \int_L^{\infty} n''(\mathcal{L})\mathcal{L}d\mathcal{L} \tag{4.11}$$

and finally get:

$$m(\mathcal{L}) = n(L) - n'(L)L \tag{4.12}$$

4.7 Additional material: results for additional mock samples

We provide here the fit of the $L_{\rm X}$–$L_{\rm UV}$ relation and the histograms of the residuals for the mock samples derived by starting from Lusso et al. (2020) luminosity distribution. In Fig. 4.6 we show the results obtained without the torus assumption in the upper panel, and with the assumption of a torus with an angle width of $\psi_{\rm torus} = 25$ in the lower panels. The histograms scale is set to logarithmic to better visualise the different shapes. In the Upper left panel, we see that without the torus assumption, we obtain a high estimate of the dispersion due to the inclination, $\delta_{\rm inc} = 0.19$ dex, and a slope of the relation equal to $\gamma = 0.46 \pm 0.01$, not consistent with the assumed value of $\gamma = 0.6$ for the mock sample. In the Upper right panel we also see that, as in the case discussed in the text where we started from the Ross et al. (2013) luminosity function, without the torus assumption we obtain a highly skewed histograms of the residuals, with a skewness parameter of $s = 1.74$. The peak of the histogram is also shifted from zero, and is found at -0.1. In the lower panel, we show the results obtained assuming a $\psi_{\rm torus} = 25$ torus. In the lower left panel, we see that the best fit slope is now $\gamma = 0.60 \pm 0.01$, perfectly consistent with the assumed value of $\gamma = 0.6$. We note that with this assumption, the dispersion due to inclination is estimated to be $\delta_{\rm inc} = 0.06$ dex, which is the exact same value that we obtain starting from Ross et al. (2013) luminosity distribution. In the lower right panel, we see now that the histograms of the residuals is now symmetric, with a skewness parameter of $s = 0.20$ and with the peak corresponding to -0.06. To sum up, starting with the Lusso et al. (2020) luminosity distribution instead of with the Ross et al. (2013) luminosity function, we obtain the same results both in terms of the need for a $\psi_{\rm torus} \sim 25$ torus, and for the estimate of the contribution of inclination to the total dispersion, with $\delta_{\rm inc} = 0.06$. This result shows us that our estimate for the inclination contribution does not strongly depend on the exact shape of the starting luminosity distribution.

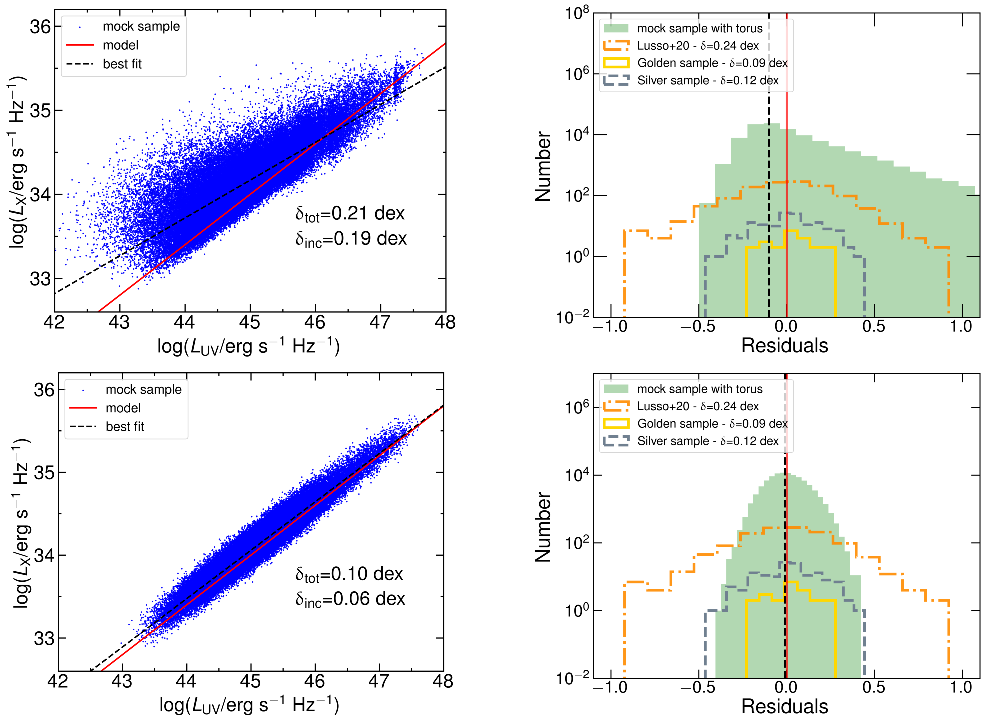

Figure 4.6: *Upper left*: mock sample of 100.000 quasars assuming an intrinsic relation with zero intrinsic dispersion, a contribution from variability to the observed dispersion of 0.08 dex, and an inclination angle between 0 and 90°. For each object, a redshift is assigned using the Lusso et al. (2020) (L20) sample redshift distribution. To derive the luminosity, we correct the L20 luminosity distribution for th einclination effect, and extract a luminosity value from there. The red solid line represents the starting sample, with a slope γ = 0.6 and a zero dispersion. The blue points show the sample after the dispersion due to variability is added and the objects are assigned a random inclination. The total dispersion of the sample is δ_{tot} = 0.21 dex, and the inclination effect accounts for δ_{inc}=0.19 dex. *Upper right*: in green filled, the histogram of the fit residuals. This distribution is skewed, with a skewness parameter of s = 1.74, and the peak is shifter from zero, at -0.16. In dot-dashed orange, the residuals histogram for the L20 sample. We see that the L20 distribution, which is the observed one, is instead much more symmetric, with a skewness parameter of s_{L20} = 0.20. In dashed silver and in solid gold, the residuals histograms corresponding to the ”silver sample” and the ”golden sample” of Sacchi et al. (2022) (details in Section 4.4). The red solid vertical line corresponds to zero, while the dashed black vertical line corresponds to the peak of the mock distribution, equal to -0.1. *Lower left*: the same as the Upper Left panel, but assuming the presence of an obscurer with an angle width of ψ_{torus} = 25, which means that the inclination angle for the objects in the mock sample can go from 0 to 65. The total dispersion is reduced to δ_{tot} = 0.10 dex, the inclination effect accounts for δ_{inc}=0.06 dex. *Lower right*: the same as the Upper right panel, but assuming the presence of an obscurer with an angle width of ψ_{torus} = 25. Here we see that the residuals distribution is symmetric, with a skewness parameter of s = 0.20, which perfectly matches the L20 value of s_{L20} = 0.20. We also note that the residuals distribution is now slightly wider than the one for the Sacchi et al. (2022) sample.

Chapter 5

Non-parametric analysis of the Hubble Diagram with Neural Networks

The Hubble diagram (i.e. the distance-redshift relation) describes the expansion of the Universe with time, and is one of the fundamental tools of observational cosmology. The "kinematic" information encoded in this diagram include the Hubble parameter H_0 (from the first-order derivative at redshift $z = 0$) and the acceleration parameter (from the second-order derivative). When a dynamical model is adopted, its physical parameters can be derived from the fit of the Hubble diagram. Typical examples are the estimate of the matter density at $z = 0$, Ω_M, within a flat ΛCDM model, or the evaluation of Ω_M and Ω_Λ within a non-flat ΛCDM model. Moreover, the physical meaning of the relevant parameters is to some extent reflecting the chosen model. Likewise, the obtained numerical estimates are also model-dependent: assume for example data to follow a ΛCDM model, with prescribed Ω_M and non-zero curvature. Then, it is easy to demonstrate through numerical simulations that, if a flat ΛCDM is adopted, the best fit value of Ω_M will be different from the correct (simulated) one.

In the past few years, possible new physics beyond the flat ΛCDM model has been suggested by several observational results, such as the mismatch between the direct measurements of H_0 in the local Universe (Riess et al. 2019; Wong et al. 2019) and the extrapolations based on the Cosmic Microwave Background (CMB), the comparison between the high- and low- multipole spectra of the CMB (Di Valentino et al. 2021), and the tension between the power spectrum of density perturbations measured on different scales (Macaulay et al. 2013; Battye et al. 2015; Lin & Ishak 2017; Heymans, C. et al. 2021; Nunes & Vagnozzi 2021). Recently, a significant deviation from the flat ΛCDM model has been observed in the Hubble diagram at high redshift, populated with quasars and gamma-ray bursts (GRB): while no significant tension is found at $z < 1.5$ with either supernovae, quasars, or GRB, the data at $z > 1.5$ suggest a slower expansion of the Universe than predicted by the flat ΛCDM model (Risaliti & Lusso 2019; Lusso et al. 2020; Demianski et al. 2017; Lusso et al. 2019).

These results make it particularly important to analyze the Hubble diagram in a way that is as model-agnostic as possible, in order to obtain an "absolute scale" for the comparison with specific models, and to infer the global, "cosmographic" properties of the expansion which, in turn, could suggest the optimal class of models to fit to the data.

Cosmographic expansions (Aviles et al. 2014; Capozziello et al. 2020; Bargiacchi et al. 2021) represent a viable approach to pursue this goal. The method is based on a standard fitting procedure and assumes that observational data can be interpolated by an appropriate series of functions, truncated to include a limited number of terms (hence of free parameters). While this is not dependent on a specific physical model, it still relies on the flexibility of the chosen functions to reproduce the shape of the observational Hubble diagram. Moreover, cosmographic techniques are rigorously valid only within a convergence radius, which is z=1 for standard methods (Cattoën & Visser 2007). At higher redshifts, no method has an absolute validity based on mathematical principles, and the effectiveness of the cosmographic analysis relies on the similarity between the chosen expansion functions and the actual shape of Hubble diagram.
An example of a robust, well checked, non-parametric approach is that based on Gaussian Process (GP) regression (Holsclaw et al. 2010; Seikel et al. 2012; Shafieloo et al. 2012), which has been used to test the hypothesis of a constant density of the dark energy term (i.e. the cosmological constant Λ). Despite their flexibility, however, GPs may under-estimate the error associated to the predictions (Col) and come with an intrinsic convergence problem for $z > 1$.

Starting from these premises, we here propose, and consequently apply, a novel analysis framework for the Hubble diagram, based on Neural Network Regression (see Dialektopoulos et al. 2022 for a similar approach). Deep feedforard fullyconnected Neural networks (FFNN) are well known universal approximators. Their ability to represent functions extends far beyond the need for this problem. However the core of their efficacy lies in the assumption that the proper regression function result from a collection of several multilevel hierarchical factors (or features) which could enable one to account in the analysis for unknown features which bear - to some extent - cosmological relevance. To sum up, we try to merge the concept of features with cascading relevance proper of a cosmographic expansion with the need of convergence and high function representation capabilities, typical of kernel methods or GP.
We will first describe the method, and check its reliability with simulated data sets. Then we will apply it to a Hubble diagram at high redshift, showing a high-redshift inconsistency with the ΛCDM model. Finally we will speculate on the class of models that could fix the discrepancy.

5.1 The cosmological background

In a Friedmann-Robertson-Walker Universe, the *luminosity distance* of an astrophysical source is related to the redshift through the equation:

$$d_L = \frac{c\,(1+z)}{H_0\,\sqrt{-\Omega_K}}\,\sin\left(\sqrt{-\Omega_K}\int_0^z dz'\,\frac{H_0}{H\,(z')}\right) \tag{5.1}$$

where $H(z)$ is the Hubble function and Ω_K stands for the curvature parameter, defined as $\Omega_K = 1 - \sum_i \Omega_i$, with Ω_i representing the density of the constituents of the Universe, normalized to the closure density. In the simplest form, assuming a flat Universe, a constant total content of matter in the Universe, a cosmological constant, and considering the redshift range where standard candles are observed (i.e. $z < 7$, where the contribution of the radiation and neutrino terms is negligible), $H(z) = H_0\sqrt{\Omega_M(1+z)^3 + 1 - \Omega_M}$. However, a wide range of different physical and cosmological models have been considered, including a non-zero curvature, an evolving dark energy density, and/or interactions between dark energy and dark matter. In this work, we want to analyze a subset of these models, represented by the equation:

$$H(z) = H_0\sqrt{\Omega_M(1+z)^3 + (1-\Omega_M)\,e^{3\int_0^z \frac{1+w(z')}{1+z'}dz'}} \tag{5.2}$$

where $w(z)$ is a generic redshift evolution of the dark energy component density. Our main goal is to test the consistency of the flat ΛCDM hypothesis (which amounts to setting $w = -1$, in the previous equation) with the present Hubble diagram of supernovae and quasars, and draw comparison with other possible functional forms for $w(z)$, as proposed in the literature. To this aim, we will carry out a non-parametric fit, via a suitably designed Neural Network. This latter enables us to reach conclusions on the predicted profile of $w(z)$ without resting on any a-priori assumption.

One key problem in any non-parametric reconstruction attempt is the so-called “inversion problem”: it is easy to demonstrate that the inversion of Equation (5.2), which involves the first and second derivatives of $H(z)$ (see e.g. Seikel et al. 2012), is inherently unstable, due to strong dependence on the Ω_M and H_0 parameters (in particular, a change of the quantity $H_0^2\Omega_M$ by as little as 0.1% can alter the predicted value of $w(z)$ by orders of magnitude, and/or flip its sign). As a consequence, constraints on $w(z)$ at very low redshift can be obtained, but the uncertainties become very large already at $z \sim 0.5$. This makes it hard to reach conclusive evidences about the supposed consistency of available data with the reference scenario with $w = -1$. In principle, better data could help to reduce the uncertainties. While we will discuss this issue in more detail in a dedicated paper, here we just mention the relevant point for the present work: it is not possible to obtain significant information on $w(z)$ from the Hubble diagram without (a) assuming some analytic form of the function and/or (b) having a combined estimate of Ω_M and H_0 with a much higher precision than available today and in the foreseeable future. There are only two possible direct ways to overcome this limitation: either we restrict our analysis to very narrow ranges of the parameters, or we constraint the shape of the function $w(z)$. Since neither of these approaches is satisfactory (and both of them have been already explored in the literature), we chose a different strategy. We do not attempt to carry out a full inversion of Eq. (5.2). On the contrary, we overcome the aforementioned numerical problems by aiming at estimating the quantity:

$$I(z) = \int_0^z \frac{w(z')+1}{1+z'}dz' \tag{5.3}$$

which can be determined from the observational data by solely invoking the first derivative of $H(z)$ (see Cárdenas 2015 for an early application of this technique). We notice that within the ΛCDM model, $w = -1$ implies $I(z) = 0$. As an obvious limitation, we will just recover the integral of the physical quantity of interest, the function $w(z)$: the degeneracy on $w(z)$ implies that different forms of $w(z)$ lead to indistinguishable shapes of $I(z)$. Nonetheless, we can achieve some remarkable results. First, we can compare the results on $I(z)$ with the prediction of the flat ΛCDM model: an inconsistency in this check would be a powerful and general proof of a tension between the model and the data (note that the opposite is not true: an agreement based on the analysis of $I(z)$ does not necessarily imply an invalidation of the ΛCDM model). More in general, we can explore the family of $w(z)$ functions leading to the observation-based reconstruction of $I(z)$, to determine which class of physical models can reproduce the observed Hubble diagram.

5.2 Regression via Deep Neural Networks (NN)

For our purposes we have chosen to deal with a fully connected feedforward architecture, as illustrated in Section 5.4. Function (5.3) is hence approximated by a suitable NN, denoted with I_{NN}, to be determined via an apposite optimization procedure, hereafter outlined. After a few manipulations, as detailed in the SI, the dataset takes the form $\mathcal{D} = \{(z^{(i)}, y^{(i)}, \Delta y^{(i)})\}$ with $i \in 1 \ldots |\mathcal{D}|$ where $y^{(i)}$ is connected to the modulus of luminosity distance $d_L^{(i)}$ and $\Delta y^{(i)}$ stands for the associated empirical error. The predictions $y^{(i)}_{\text{pred}}$ and the supplied input $y^{(i)}$ are linked via:

$$y^{(i)}_{\text{pred}} = \int_0^{z^{(i)}} dz' \left[\Omega_M (1+z')^3 + (1-\Omega_M)\, e^{3I_{NN}(z')}\right]^{-\frac{1}{2}} \tag{5.4}$$

Notice that the prediction is a functional of I_{NN}, the neural network approximation that constitutes the target of the analysis. To carry out the optimization we introduce the loss function $L(I_{NN}, \mathcal{D}) = \sum_{i=1}^{|\mathcal{D}|} \left(\frac{y^{(i)} - y^{(i)}_{\text{pred}}}{\Delta y^{(i)}}\right)^2$. The weights of the network which ultimately defines I_{NN} are tuned so as to minimize the above loss function, via conventional stochastic gradient descent methods. The *hyper-parameters* have been optimized with mock data samples, as illustrated in the SI. To quantify the statistical errors Δy_{pred} (associated to the predictions) and ΔI_{NN} (referred to the approximating neural network) we implemented a *bootstrap* procedure, further detailed in the SI. The code is freely available online [1].

The regression scheme introduced above was challenged against a selection of mock data samples. In carrying out the test we considered:

(A) A sample of 4,000 sources with no dispersion, with a flat distribution in $\log(z)$ between $z = 0.01$ and $z = 6$, and following a flat ΛCDM model with $\Omega_M = 0.3$ and $h = H_0/(100 km/s/Mpc) = 0.7$. This sample (as well as the next in the list) represents a highly idealized, hence non realistic setting. It is solely used as a reference benchmark model, for preliminary consistency checks.

(B) The same as above, but the model used is a Chevallier-Polarski-Linder (CPL) parametrization, (which assumes a Dark Energy equation of state that varies with the redshift as

---

[1] `https://github.com/Jamba15/Cosmological-Regression-with-NN.git`

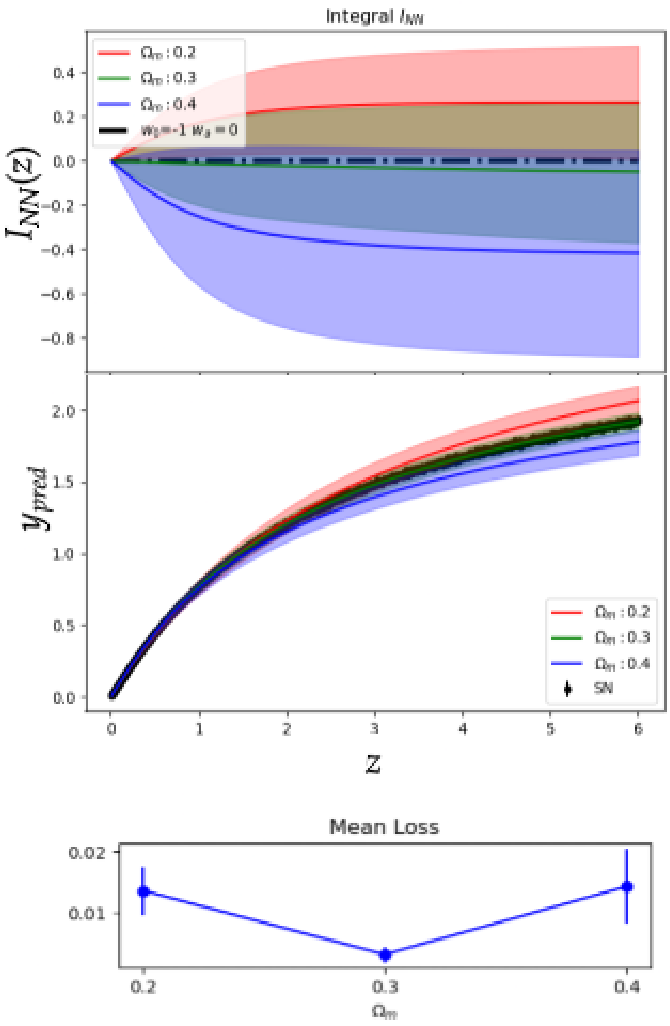

Figure 5.1: Simulations with a "perfect" sample, dataset A. Results of the NN analysis of a simulated sample of 4,000 objects with a log-flat redshift distribution and a negligible dispersion with respect to a flat ΛCDM model with Ω_M=0.3. Top panel: estimated values of $I(z)$ for different values of Ω_M (Eq. 3, the "correct" value for the simulated data is I(z)≡0). Central panel: Hubble diagram with the reconstructed best fit function obtained from the NN analysis. Bottom panel: LOSS values for different values of Ω_M. The minimum is at Ω_M=0.3, i.e. the "true" value. The corresponding $I(z)$ is consistent with zero at all redshifts. These results demonstrate that the NN analysis is able to recover the correct model and the "true" value of Ω_M.

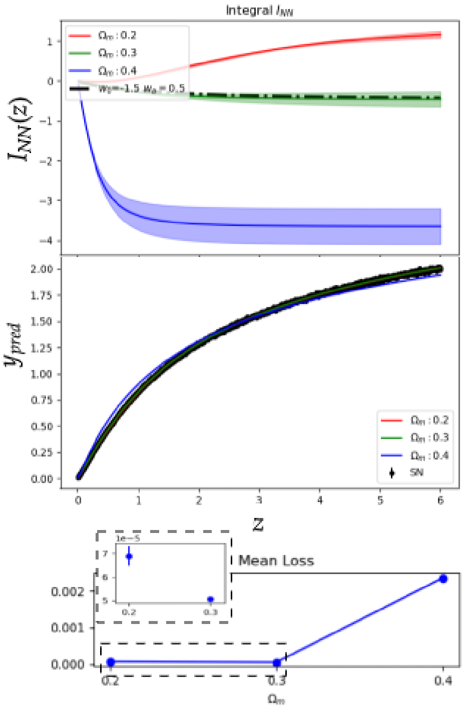

Figure 5.2: Results for dataset B. The governing model is a CPL with $w_0 = -1.5, w_a = 0.5$

$w(z) = w_0 + w_a \frac{z}{1+z}$ (Chevallier & Polarski, 2001)), with $w_0 = -1.5$ and $w_a = 0.5$. We note that this choice of the parameters would be hard to justify from a physical point of view. In particular, it implies a turning point in the Hubble parameter H(z), which violates the null-energy condition. However, even if this scenario can be considered as unphysical, it well serves our purpose of testing our regression method with extreme models.

(C) A sample with the same size, redshift distribution and dispersion as the Pantheon supernovae Ia sample (Scolnic et al., 2018), assuming a flat ΛCDM model with $\Omega_M = 0.3$.

(D) A Pantheon-like sample, as above, assuming a CPL model with $w_0 = -1.5$ and $w_a = 0.5$.

(E) A sample with the same size and redshift distribution as the combined Pantheon (Scolnic et al., 2018) and quasar (Lusso et al., 2020) samples. The quasar sample consists of 2,244 sources with redshift in the $z = 0.5 - 7.5$ range (the whole Lusso et al. 2020 sample contains 2,421 objects, but the 178 ones at redshift $z < 0.5$ are non considered in this analysis). We assume the same dispersion as in the real sample and a flat ΛCDM model with Ω_M =0.3.

(F) The same as above, assuming a CPL model with $w_0 = -1.5$ and $w_a = 0.5$.

More specifically, we generated synthetic data following the different recipes evoked above. The regression scheme, as implemented via the neural network, enables us to solve an inverse problem, from data back to the underlying physical model. The correspondence between postulated and reconstructed physical instances, readily translates in a reliable metric to gauge the performance of the proposed procedure, in a fully controllable environment and prior application to the experimental dataset.

The results for the analysis of settings A and B is shown in Figure 5.1 and 5.2. The

outcomes confirm that our NN method can consistently recover the "true" model and parameters with simulated data of (unrealistic) high quality.
The outcome of the analysis for respectively settings C (top left), D (top right), E (bottom left) and F (bottom right) is displayed in Figure 5.3. Both $I_{NN}(z)$ (the neural network approximation for $I(z)$) and $y_{pred}(z)$ are represented as function of the redshift z. For settings E and F, the associated mean loss is also plotted against the parameter Ω_M, which can be freely modulated to explore different scenarios. Working with a dataset of type C cannot yield definite conclusions: indeed the NN is unable to recover the correct value of Ω_M, as different ΛCDM models ($I_{NN}(z) \simeq 0$, within the explored range) provide an equally accurate interpolation of the (simulated) data within statistical errors. The above degeneracy is however removed when extending the examined sample so as to include quasars, see bottom-left panel of Figure 5.3 which refers to dataset E. In this case, the minimum displayed by the loss function points to $\Omega_M = 0.3$, the value assumed in the simulations, and the corresponding function $I_{NN}(z)$ is approximately equal to zero (green shadowed domain) within errors, and at variance with what it is found by employing the other chosen values of Ω_M. Datasets D and F (rightmost panels in Figure 5.3) returns similar conclusions when operating with data generated according to a CPL prescription. Working with supernovae (over a limited range in z) does not allow to distinguish between ΛCDM and CPL model, while the underlying model, assumed for data generation, is correctly singled out when quasars are accounted for (green shadowed region that encloses the dashed line, that represents the exact profile), i.e. when extending the dataset to higher redshifts.

Overall, working on synthetic data suggest that (a) the regression method is reliable, (b) with the current Hubble diagram of supernovae it is not possible to test the ΛCDM model against possible extension such as the CPL model with "phantom like" dark energy. Such a degeneracy is removed with a combined supernovae+quasar sample extending up to z~7.

Motivated by the outcome of these simulations, we applied the NN to the experimental dataset consisting of the Pantheon supernovae sample and the Lusso et al. (2020) quasars sample). The quasars luminosity distances and their errors are estimated from the observational data (X-ray and UV fluxes) following the procedure described in Lusso et al. (2020). In summary, the X-ray and UV data are first fitted in narrow redshift bins, in order to derive a cosmology-independent slope α of the X-ray to UV relation. This value and its uncertainty are used to derive luminosity distances in an arbitrary scale. The absolute calibration β is obtained from the cross-match of the quasar and supernovae sample in the common redshift interval. This calibration has a negligible uncertainty with respect to the other components on the error on the luminosity distances, i.e. the flux measurement errors, the intrinsic dispersion of the relation, and the error on the slope. In principle, rather than fitting the so-derived luminosity distances, it would be preferable to use the observables, i.e. the fluxes, and to marginalize over the parameters α and β. In practice, we have checked that the two methods provide identical results, and the use of luminosity distances and Equation (5.4) makes the analysis much faster and easier to implement within a NN method.

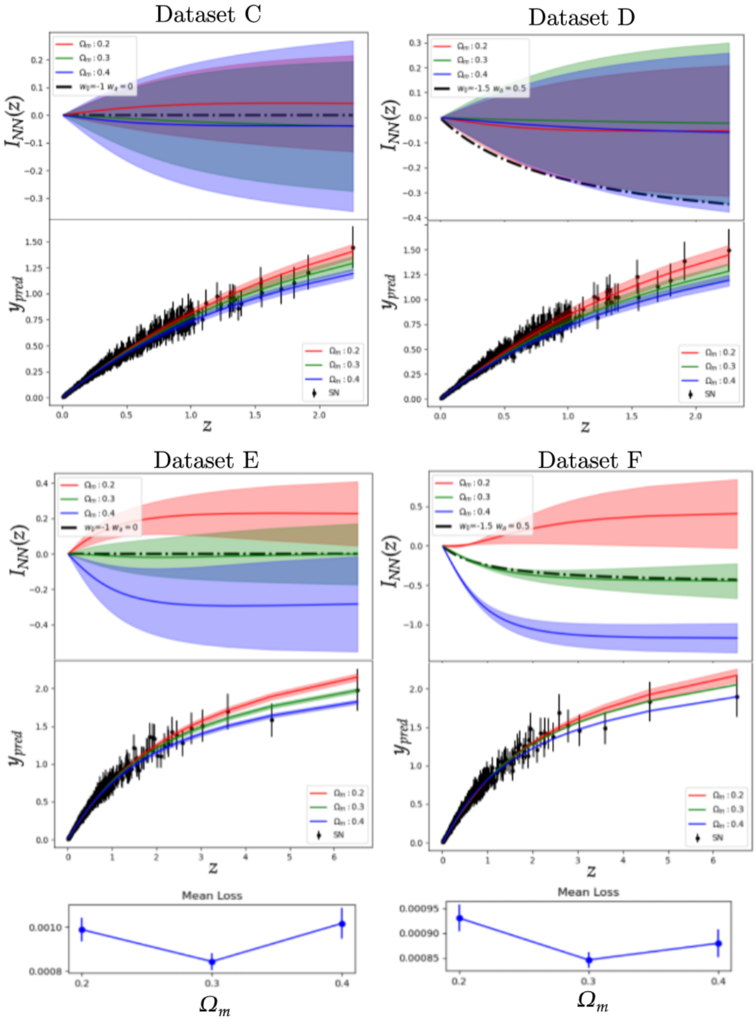

Figure 5.3: $I(z)$ Results of the NN analysis of the Hubble diagram of simulated data. Top left: Dataset C, with the same redshift distribution and dispersion as the Pantheon supernovae sample. Bottom left: Dataset E, where combined Pantheon and quasars are considered. In this case the NN is able to identify the model assumed for data generation (the green shadowed region contains the exact profile for $I_{NN}(z)$, depicted with a dashed line). The corresponding loss function is also shown and displays a minimum at the correct value of Ω_M. Top right: a Pantheon-like sample is assumed, for a CPL generative model (dataset E). The NN is unable to distinguish between different scenarios (ΛCDM vs, CPL). Bottom right: CPL model with the inclusion of quasars. The degeneracy is resolved and the NN can correctly identify the underlying model (see dashed line). The loss shows a minimum for the correct value of Ω_M, which yields the green shadowed solution for $I_{NN}(z)$ vs. z.

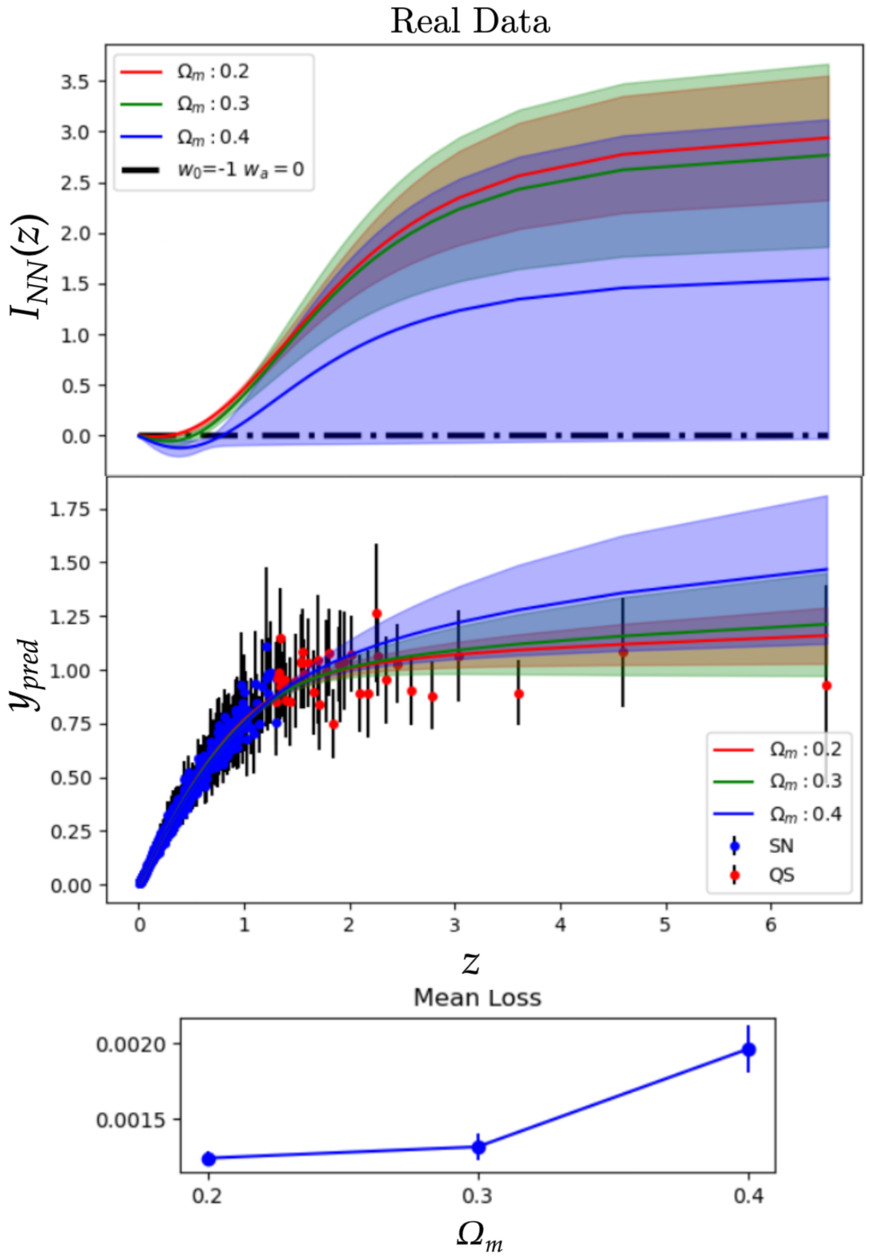

Figure 5.4: Results of the NN analysis of the Hubble diagram of supernovae (blue points in the middle panel) and quasars (red points). Top panel: estimated values of $I(z)$ for different values of Ω_M. Central panel: Hubble diagram with the reconstructed best fit function obtained from the NN analysis. Bottom panel: *Loss* values for different values of Ω_M. Notice that the solution that is closer (accounting for statistical errors) to the reference ΛCDM profile, which is the one with $\Omega_m = 0.4$, yields significantly larger value of the loss, and as such should be disregarded. The *Loss* is indeed nearly flat for $\Omega_M < 0.3$.

The results of the NN-based fits are shown in Figure 5.4. The shape of $I(z)$ is clearly non consistent with the flat ΛCDM model ($I(z) \equiv 0$). This is the main result of our work, and has been obtained without assuming any a priori knowledge on the function $I(z)$.

As a next step in the story, we introduce a dedicated indicator to quantitatively measure the compatibility of the examined data with the reference ΛCDM model. Imagine to naively access the distance of the fitted profile y_{pred} to the reference $y_{\Lambda\text{CDM}}$ ($I = 0$) curve and divide it with the error associated to the fitted function Δy_{pred}. Assume that the computed ratio (averaged over z) is smaller than unit. Then, the distance between y_{pred} and $y_{\Lambda\text{CDM}}$ is eclipsed by statistical uncertainty and thus ΛCDM cannot be ruled out as a candidate explanatory model. The above procedure can be cast on solid grounds (see SI), yielding a scalar indicator that fulfills the purpose of quantifying the sought distance, normalized to the associated error. This is denoted by $\Delta_{\Lambda\text{CDM}}$ and takes the form:

$$\Delta_{\Lambda\text{CDM}}(\mathcal{D}, I_{\text{NN}}) = \frac{1}{|\mathcal{D}|} \sum_{i \in \mathcal{D}} \frac{\delta y_{\text{pred}}^{\Lambda\text{CDM}}(I_{\text{NN}}; z^{(i)})}{\Delta y_{\text{pred}}(I_{\text{NN}}; z^{(i)})} \tag{5.5}$$

The fitted integral function I_{NN} is deemed compatible with the ΛCDM model, if $\Delta_{\Lambda\text{CDM}} < 1$. When this latter condition holds true, the predictions deviate from a ΛCDM by an amount that, on average, is smaller than the corresponding prediction error. The indicator in equation (5.5) has been computed for different mock samples, mimicking ΛCDM, with progressively increasing errors sizes Δy. The latter is assumed uniform across data points and varied from zero to 0.15, thus including the value - ~ 0.14 - that is believed to apply to real data. This information is used as a reference benchmark to interpret the results of the analysis for the Pantheon + quasar experimental dataset. To sum up our conclusions (see SI) the portion of the dataset at small redshift is compatible with a ΛCDM model with $\Omega_m = 0.3$, within statistical errors. Conversely, for $z > 2$ (notably quasars), $\Delta_{\Lambda\text{CDM}}$, as computed after available experiments, is 5σ away the expected mean value. Hence, accounting for quasars, enables us to conclude that the ΛCDM model is indeed extremely unlikely.

Finally, we comment on the results depicted in Figure 5.5 where the best fit $I(z)$ for $\Omega_M = 0.3$ (the same as in the upper panel of Figure 5.4) is plotted in logarithmic scale, and compared to $I_{\text{MATTER}}(z) = \log(z)$, the function obtained from equation (5.3) by assuming $w(z) \equiv 0$, i.e. a pure matter contribution. We recall that a cosmological constant, or equivalently a dark energy component with constant energy, implies $w(z) \equiv -1$ and $I(z) \equiv 0$. It is therefore tempting to speculate as follows, when qualitatively analyzing the profile of $I(z)$: the redshift intervals with negative derivative represent a dark energy component with density increasing in time (the "phantom" dark energy scenario); the intervals with positive derivatives, smaller than the constant derivative of $I_{\text{MATTER}}(z)$ represent a dark energy component with decreasing density; last, the intervals where the derivative is larger than that displayed by $I_{\text{MATTER}}(z)$ are matter terms, with increasing density.

The prior-free NN solution suggests therefore an "interacting dark sector" scenario, where a matter component decreases with time, and correspondingly a dark energy component rises. This interpretation is also consistent with the nearly constant $Loss$, for $\Omega_M < 0.3$: choosing values larger than 0.3 worsen the agreement, because this amounts to overestimate the total

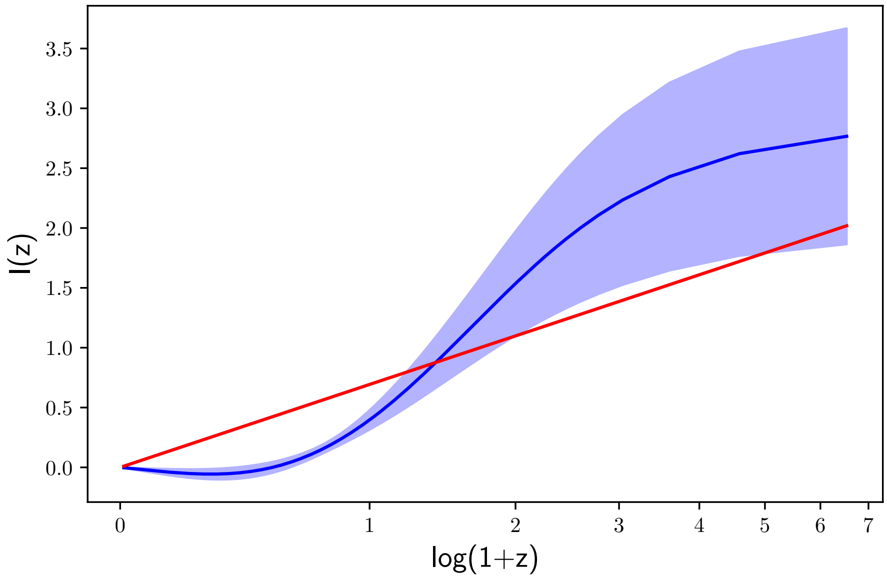

Figure 5.5: Best fit $I(z)$ from our NN regression (as in the upper panel of the previous figure) in logarithmic scale, compared with the function $I_{\mathrm{MATTER}}(z)$ obtained by assuming $w(z) \equiv 0$ in Eq. (5.3). The redshift intervals where the derivative of $I(z)$ is higher than that of $I_{\mathrm{MATTER}}(z)$ represent "matter-like" contributions, while intervals with a lower derivative refer to energy-like contributions.

matter component at $z \sim 0$. On the other hand, value smaller than 0.3 can be compensated by the matter component in $I(z)$. This interpretation is consistent with recent claims of an increasing value of Ω_M with redshift within a ΛCDM scenario (e.g. Colgáin et al. 2022).

5.3 Conclusions

Our conclusions are multifold. We have proposed and rigorously tested a Neural Network (NN) approach to analyse the Hubble diagram. Then, the NN model-agnostic regression of the combined supernovae and quasars catalogue enables us to unequivocally reveal a strong tension with the "concordance" flat ΛCDM model. Finally, the analysis carried out with the proposed NN approach suggests an "interacting dark sector" scenario, where a dark matter component flows into dark energy, at least down to redshifts $z \sim 1.5$.

5.4 Additional material

Data processing Data come as the set $\mathcal{D} = \{(z^{(i)}, y^{(i)}, \Delta y^{(i)})\}$ with $i \in 1 \dots |\mathcal{D}|$. Each component $y^{(i)}$ is linked to d_L, the physical quantity of interest, by $y^{(i)} = 5 \log(d_L^{(i)}/10\text{pc})$. The first applied transformation is defined as follows:

$$y'^{(i)} = y^{(i)}/5 + 1, \quad \Delta y'^{(i)} = \Delta y^{(i)} \tag{5.6}$$

By doing so data are traced back to the logarithm of the luminosity distance; every entry of the inspected dataset is indeed equal to $y^{(i)} = \log(d_L^{(i)})$.

Carrying out a first order expansion of equation 5.1, assuming a flat Universe ($\Omega_k \sim 0$) and inserting the expression of $H(z)$ as reported in the main text, yields:

$$d_L = \alpha(z) \int_0^z dz' \left[\Omega_M (1+z')^3 + (1-\Omega_M)\, e^{I(z')}\right]^{-\frac{1}{2}} \tag{5.7}$$

where $\alpha(z) = \frac{c(1+z)}{H_0}$. Then we proceed by setting:

$$y''^{(i)} = y'^{(i)} - \log(\alpha(z^{(i)})), \quad \Delta y''^{(i)} = \Delta y'^{(i)} \tag{5.8}$$

It is worth noticing that the relative errors associated with c, z and H_0 are negligible. The above relation transforms into:

$$y'''^{(i)} = 10^{y''^{(i)}}, \quad \Delta y'''^{(i)} = 10^{y''^{(i)}} \Delta y''^{(i)} \tag{5.9}$$

To simplify the notation we drop the apex by setting $y''' \to y$ and obtain the sought connection between every $y^{(i)}$ and the function to be fitted $I(z)$, namely:

$$y^{(i)} = \int_0^{z^{(i)}} dz' \left[\Omega_M (1+z')^3 + (1-\Omega_M)\, e^{I(z')}\right]^{-\frac{1}{2}} \tag{5.10}$$

We point out that being $\arg\min_x f(x) = \arg\min_x C \cdot f(x)$ every manipulation that results in a constant factor in front of the Loss function can be ignored. This is the case of every operation in the form of $\Delta y^{(i)}_{\text{new}} = C \cdot \Delta y^{(i)}_{\text{old}}$ which, indeed, results in a factor $\dfrac{1}{C^2}$ in front of the Loss function, see equation (5.11).

The employed Neural Network model To approximate the non linear scalar function $I(z) : z \in \mathbb{R} \mapsto I(z) \in \mathbb{R}$ we make use of a so called *feedforward* architecture. The information flow from the input neuron, associated to $z^{(i)}$ to the output neuron where the predicted value of $I_{\text{NN}}(z^{(i)})$ is displayed.
The transformation from layer k to its adjacent homologue $k+1$, following a feedfoward arrangement, is characterized by two nested operations: (i) a linear map $W^{(k)} : \mathbb{R}^{N_k} \to \mathbb{R}^{N_{k+1}}$ and (ii) a non linear filter $\sigma^{(k+1)}(\cdot)$ applied to each entry of the obtained vector. Here k ranges in the interval $1 \dots \ell$ where $N_1 = 1$ and ℓ is the number of *layers*, i.e. the depth of the NN. We have chosen $\sigma^{(k)} := \tanh$, $\forall k < \ell - 1$ whereas $\sigma^{(\ell)} = \nVdash$.

The activation of every neuron in layer k can be consequently obtained as:

$$\vec{x}^{(k)} = W^{(k-1)}(\dots \sigma(W^{(2)}(\sigma(W^{(1)}z))) \dots)$$

Furthermore, we have fixed $N_k = N_{k+1}$ $\forall k \in 2 \dots \ell - 2$, meaning that every layer (but the first and the last) has the same size as the others. The size of the so called *hidden layer* N_2 and the total amount of layers ℓ are, consequently, the only hyper-parameters to be eventually fixed.
Occasionally a neuron-specific scalar, called *bias*, can be added after application of each

linear map $W^{(k)}$. To allow for the solution $I_{\rm NN}(0) = 0$ to be possibly recovered, we have set the bias to zero.

The output $I_{\rm NN}(z)$ hence depends on $N = \sum_{k=1}^{\ell-1} N_k \times N_{k+1}$ free scalar parameters (the weights $W^{(k)}_{i,j}, i \in 1 \dots N_{k+1}\ j \in 1 \dots N_k,\ k \in 1 \dots \ell - 1$), that constitute the target of the optimization.

Model Optimization The optimization herefter described has been carried out by using parallel computing on GPU. The minimization of the Loss function as defined in the main text is performed via a variant of the stochastic gradient descent (SGC) method, recalled below.

First, the dataset $\mathcal{D}$ is shuffled and divided into smaller subsets $\mathcal{B}_i$ of size $|\mathcal{B}_i| = \beta$. These are the *batches*, and meet the following condition: $\mathcal{D} = \sqcup_i^{N_b} \mathcal{B}_i$. Obviously the number of batches N_b is equal to $\lceil \frac{|\mathcal{D}|}{\beta} \rceil$.

The gradient with respect to every weight W entering the definition of the function L is computed, within each batch, as:

$$G^{(i)} = \nabla_W L(W, \mathcal{B}_i) = \nabla_W \sum_{j:y^{(j)} \in \mathcal{B}_i} \left(\frac{y^{(j)} - y^{(j)}_{\rm pred}(z^{(j)}; W)}{\Delta y^{(j)}} \right)^2 \tag{5.11}$$

While i takes values in the range $1 \dots N_b$, the weights W are updated so as to minimize, via a stochastic procedure, the Loss function. This is achieved as follows:

$$W \leftarrow W - \alpha_{lr} G^{(i)} \tag{5.12}$$

The hyper-parameter α_{lr} is called *learning rate* and drives the amount of stochasticity in the Loss descent process. In the present work a more complex yet conceptually equivalent variant of the SGD called Adam is implemented.

A so called *epoch* is completed when all batches have been used. The number of epochs N_e is another hyper-parameter that has to be fixed a priori, as well as the batch size bs. Usually a high number of epochs (such as 400 or 600, as employed in the present application) is chosen. To avoid overfitting, the *early stop* technique is employed. Such technical aid consists in taking a small subset, $\mathcal{V}$, of the dataset ($\sim$ 15% of $\mathcal{D}$) and exclud it from the training process. During training stages, hence, the employed dataset is $\mathcal{D}' = \mathcal{D} - \mathcal{V}$. While applying SGD to the Loss so as to minimize it, Loss evaluation on dataset $\mathcal{V}$, $L(I_{\rm NN}, \mathcal{V})$ is also performed. When the latter function reaches a plateau, the optimization process is stopped. This latter procedure relies on two hyper-parameters: δ the absolute variation of L that can be considered as a real Loss change, and p, the number of consecutive epochs with no recorded variation, before the fitting algorithm can be eventually terminated.

Moreover, one additional hyper-parameter needs to be mentioned: as already explained above, the prediction $y_{\rm pred}$ involves a numerical integral of the NN approximating function, $I_{\rm NN}$. The integration step dz' is thus to be set, and was object of a meticulous optimization. An hyper-optimization process designed to find the best set of hyper-parameters has been carried out, employing several CPL and ΛCDM like models. Such process has led to a set of parameters which have been fixed and left unchanged during the trials. In Table I the chosen hyper-parameters list is provided.

N_2	ℓ	N_e	bs	α_{lr}	δ	p	dz'
20	5	600	100	10^{-6}	10^{-6}	35	$5\ 10^{-4}$

Table 5.1: Hyper-parameters employed

Estimating the errors To estimate the prediction error $\Delta y_{\text{pred}}(z)$ we have employed a *Bootstrap method*. To this end the fitting procedure is arranged so as to produce B independent estimators of the quantity y_{pred} and $I_{\text{NN}}(z)$, namely $y^{[k]}_{\text{pred}}$ and $I^{[k]}_{\text{NN}}$ with $k \in 1 \ldots B$. Each $y^{[k]}_{\text{pred}}$ is the result of an optimization process started from a subset $\mathcal{D}^{[k]} \subseteq \mathcal{D}$ obtained from $\mathcal{D}$ by uniform sampling with replacement of $|\mathcal{D}|$ elements. The prediction errors Δy_{pred} and ΔI_{NN} are then computed by extracting the standard deviation from both sets as:

$$\begin{aligned} \Delta y_{\text{pred}}(z) &= \sum_{k=1}^{B} \sqrt{\frac{\left(\bar{y}_{\text{pred}}(z) - y^{[k]}_{\text{pred}}(z)\right)^2}{B-1}} \\ \Delta I_{\text{NN}}(z) &= \sum_{k=1}^{B} \sqrt{\frac{\left(\bar{I}_{\text{NN}}(z) - I^{[k]}_{\text{NN}}(z)\right)^2}{B-1}} \end{aligned} \tag{5.13}$$

where symbols $\bar{y}_{\text{pred}}(z)$ and $\bar{I}_{\text{NN}}(z)$ represent the arithmetic mean of the estimates $y^{[k]}_{\text{pred}}$ and $I^{[k]}_{\text{NN}}$. All across this work, the errors are computed after $B = 80$ bootstrap samples.

As a next step we shall comment on the derivation of the indicator to gauge the correspondence of the fitted model with a conventional ΛCDM scheme. We begin by formally expressing δy_{pred}, the distance of the obtained prediction with respect to the reference ΛCDM model, as

$$\begin{aligned} \delta y^{\Lambda\text{CDM}}_{\text{pred}}(I_{\text{NN}}; z) &= \left.\frac{\delta y_{\text{pred}}}{\delta I}\right|_{I=\Lambda\text{CDM}} \delta I \\ &= \left.\frac{\delta y_{\text{pred}}}{\delta I}\right|_{\Lambda\text{CDM}} (I_{\text{NN}} - I_{\Lambda\text{CDM}}) \end{aligned} \tag{5.14}$$

where $\frac{\delta y_{\text{pred}}}{\delta I}$ stands for the functional derivative and $I_{\Lambda\text{CDM}} = 0$. The above equation can be further expanded so as to yield:

$$\left.\frac{\delta y_{\text{pred}}}{\delta I}\right|_{I=\Lambda\text{CDM}} = -\frac{1}{2} \int_0^z \alpha(z')^{-\frac{3}{2}} (1 - \Omega_m) e^{I(z')} \Big|_{I=0} \tag{5.15}$$

where $\alpha(z') = \Omega_M (1 + z')^3 + (1 - \Omega_M)\, e^{I(z')}$. By eventually setting $\delta I = I_{\text{NN}}$ one gets therefore:

$$\delta y^{\Lambda\text{CDM}}_{\text{pred}}(I_{\text{NN}}; z) = \frac{\Omega_m - 1}{2} \int_0^z \left(\alpha(z')\big|_{I=0}\right)^{-\frac{3}{2}} I_{\text{NN}}(z') dz' \tag{5.16}$$

We are finally in a position to introduce the scalar indicator that fulfills the purpose to quantifying the sought distance, normalize to the associated error. This is denoted by $\Delta_{\Lambda\mathrm{CDM}}$ are takes the form:

$$\Delta_{\Lambda\mathrm{CDM}}(\mathcal{D}, I_{\mathrm{NN}}) = \frac{1}{|\mathcal{D}|} \sum_{i \in \mathcal{D}} \frac{\delta y_{\mathrm{pred}}^{\Lambda\mathrm{CDM}}(I_{\mathrm{NN}}; z^{(i)})}{\Delta y_{\mathrm{pred}}(I_{\mathrm{NN}}; z^{(i)})} \tag{5.17}$$

The fitted integral function I_{NN} is deemed compatible with the ΛCDM model, if $\Delta_{\Lambda\mathrm{CDM}} < 1$. When this latter condition holds true, the predictions deviate from a ΛCDM by an amount that, on average, is smaller than the corresponding prediction error.

The indicator in (5.17) has been computed for different mock samples, mimicking ΛCDM, with progressively increasing errors sizes (assumed uniform across data points), Δy (ranging from zero to 0.15, thus including the value - ~ 0.14 - that is believed to apply to real data).

For every choice of the assigned error, 30 mock samples with $\Omega_m = 0.3$ have been generated and subsequently fitted, assuming different choices of Ω_m, namely $\{0.2, 0.3, 0.4\}$. For every selected Ω_m a bootstrap procedure is implemented (see Section 5.4) to estimate $y_{\mathrm{pred}}, \Delta y_{\mathrm{pred}}$ and $I_{\mathrm{NN}}, \Delta I_{\mathrm{NN}}$. The best fit values are selected to be those associated to the smaller mean loss functions (evaluated against the imposed Ω_m). Following this choice, the mean an the variance of $\Delta_{\Lambda\mathrm{CDM}}$ are computed, from the outcomes of the fits, performed on the corresponding (30) independent realizations.

In Figures from 5.6 to 5.9 the results of the analysis for the different datasets are displayed. The solid line stands for the average estimates, as obtained following the above procedure. The shadowed region is traced after the computed errors, namely, the variance of the indicator across the realizations.

In Figure 5.6 SNe data ($z < 2$) are solely considered for carrying out the regression. The symbol refers to the experimental dataset (Lusso et al., 2020) and is set in correspondence of the estimated error (0.14). The displayed point falls within the shadowed domain, thus implying that the examined dataset is compatible with a ΛCDM model.

In Figure 5.7 we analyze the full dataset (Pantheon + quasars). The regression is hence carried out by considering data spanning the whole range in z. After the fitting has been performed, data are split into two different regions, respectively at small ($z \leq 2$) or large ($z \geq 2$) redshift. The symbols refers to the experimental dataset and are set in correspondence of the estimated error (0.14). The portion of the dataset at small redshift (mostly populated by Supernovae) is compatible with a ΛCDM model with $\Omega_m = 0.3$), within statistical errors (the agreement is even more pronounced if the regression is carried out by solely accounting for Supernovae, see Figure 5.6). Conversely, for $z > 2$, the point computed after available experiments, notably quasars, is at a distance of about 5σ from the expected value of the indicator $\Delta_{\Lambda\mathrm{CDM}}$. Hence, accounting for quasars enables us to conclude that the ΛCDM model is indeed extremely unlikely.

In Figures 5.8 and 5.9 we repeat the analysis by employing a dataset generated from a CPL model, with an error compatible with that estimated experimentally (equivalent to datasets D and F). The results indicate that accounting for data at large redshifts is mandatory to resolve the degeneracy between distinct generative models.

Figure 5.6: $\Delta_{\Lambda\mathrm{CDM}}$ vs. the imposed error, for the Pantheon dataset (i.e. just supernovae). The symbol stands for to the experimental data, while the solid line and the shadowed regions refer to the corresponding theoretical benchmarks, obtained as described in the text.

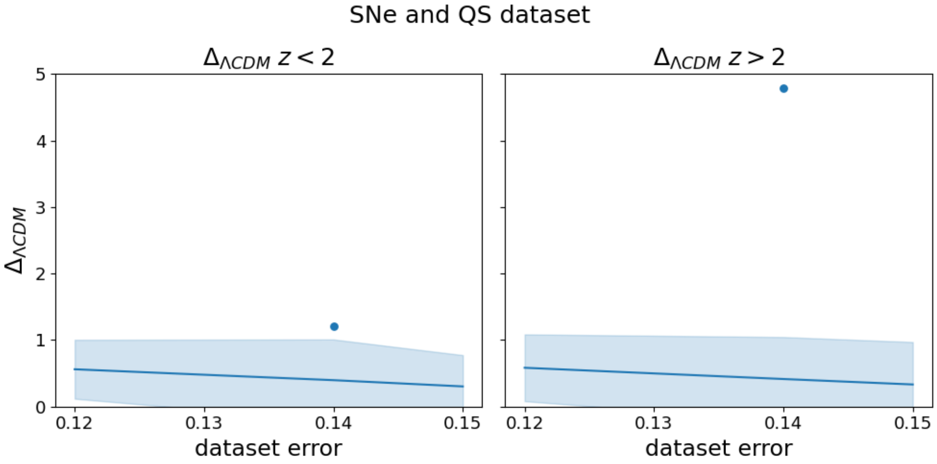

Figure 5.7: $\Delta_{\Lambda\mathrm{CDM}}$ vs. the imposed error, for the combined supernovae + quasars sample at redshifts $z < 2$ (left panel) and $z > 2$ (right panel). Symbols refer to the experimental data, while the solid line and the shadowed regions stand for the corresponding theoretical benchmarks, obtained as described in the text.

Figure 5.8: $\Delta_{\Lambda\text{CDM}}$ vs. the imposed error, for the Pantheon dataset (i.e. just supernovae). In blue the reference mean and variance (represented as a shaded region). The symbol is obtained upon processing the synthetic example generated via the CPL model.

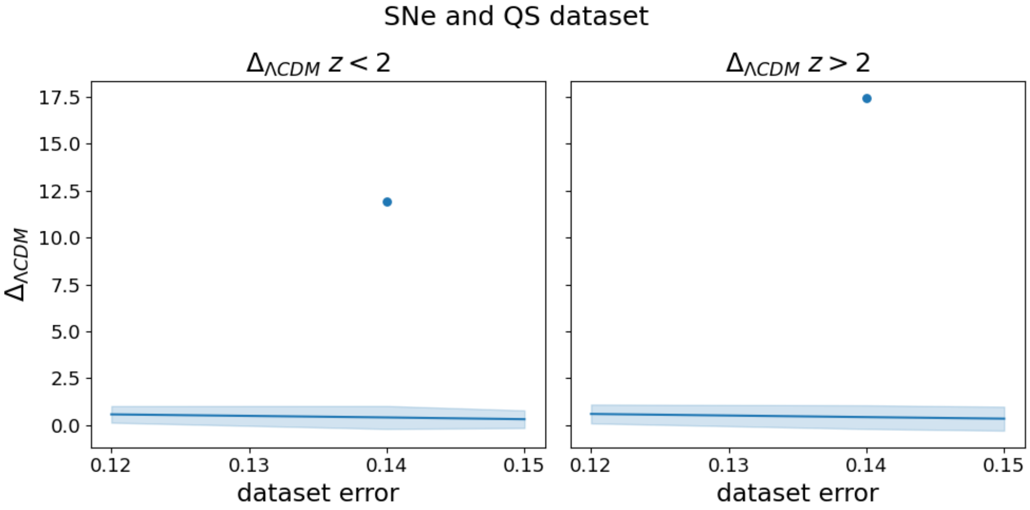

Figure 5.9: $\Delta_{\Lambda\text{CDM}}$ vs. the imposed error, for the combined supernovae + quasars sample at redshifts $z < 2$ (left panel) and $z > 2$ (right panel). In blue the reference mean and variance (represented as a shaded region) obtained with mock ΛCDM samples. The symbols are obtained upon processing the synthetic example generated via the CPL model.

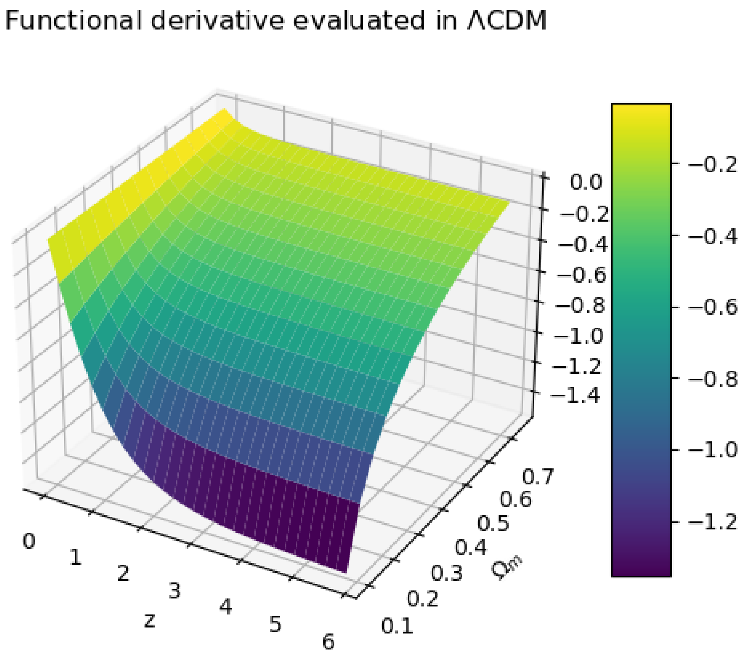

Figure 5.10: Plot of the functional derivative computed in (5.14) varying Ω_m and z.

As a final point we elaborate on the reason why different models appear indistinguishable at small z. Function I_{NN} is the argument of a functional that goes from the space of function I to the space of the predictions. The way those two spaces communicate (or rather how function I reverberates on every y_{pred}) is a non trivial function of the hyperparameters (as e.g. Ω_m and the integration steps) and the domain explored. To clarify this point we plot the functional derivative $\frac{\delta y_{\mathrm{pred}}}{\delta I}$ (evaluated at ΛCDM model) against Ω_m and z. By visual inspection of Figure 5.10 it is clear the relevant impact played by small z and large Ω_m. The functional derivative is hence very small for the portion of the dataset that is populated by the vast majority of SNe entries. This implies that different models (in terms of the associated $I(z)$) can yield very similar predictions. It is hence difficult to draw conclusions about the validity of different models, if one solely deals with data at small redshifts.

Chapter 6

Reverberation Mapping for cosmology

Reverberation mapping is a technique used to probe the structure and kinematics of the Broad Line Region (BLR) in AGN and to measure SMBH masses. The fundamental principle behind this method is the time delay or “lag” between variations in the central continuum source (attributed to the accretion disk surrounding the central black hole) and the subsequent response in the broad emission lines. The observed lag represents the light travel time from the central ionizing source to the BLR clouds that produce the reprocessed emission lines. This lag essentially gives a size scale for the BLR. By combining the BLR size obtained from the lag with the velocity width of the broad emission lines (typically measured from the line full width at half maximum or dispersion), one can estimate the virial mass of the central black hole. Furthermore, velocity-resolved measurements of the BLR response can offer insights into the physics and kinematics of the BLR. In essence, reverberation mapping uses time resolution to study the emissivity and position of gas in a spatially unresolved source (Blandford & McKee, 1982). In this chapter, we will explore reverberation mapping and its significance for cosmology and for our understanding of AGN physics. In the first section, we will outline the primary principles of the technique. In the second section, we will address the debate surrounding the tension on the H_0 constant, and discuss how the combination of Infrared reverberation mapping and interferometry can provide an independent estimate of it. In the third section, we will discuss the measuring of SMBH masses at high redshift. As we discussed in the Introduction, having precise, reliable mass estimates at high redshift is fundamental for our understanding of SMBH formation and accretion through cosmic time. We will discuss the campaign on a redshift $z = 2.8$ lensed quasar that can offer valuable perspectives on this issue.

6.1 The technique

The fundamental concept behind reverberation mapping is to discern the time lag between flux variations in an ionizing source (here, the AGN accretion disc) and the flux variations from the region irradiated by this ionizing flux (namely, the BLR). If we assume that the BLR emission is essentially a re-processing of the ionizing emission from the centre, there should be a strong correlation between the two. However, due to the additional distance the reprocessed BLR emission light must travel before reaching the observer, its variations will be detected after those originating from the disc. The lag, τ, is influenced by the system geometry, but generally, it is on the order of R/c, where R represents the average radius of the BLR, and c is the speed of light.

Three primary assumptions underpin reverberation mapping: 1) the ionizing, irradiating flux originates from a singular, central source; 2) the fundamental timescale in the process is the light travel time between the disc and the BLR, with other timescales (like the reprocessing time within the BLR clouds) being minimal in comparison; and 3) a linear relationship exists between the reprocessed and the ionizing flux. The function depicting flux in relation to varying time is termed the “light curve”. The ionizing and reprocessed light curves, $F_i(t)$ and $F_r(t)$, respectively, can be expressed as having a constant component plus a variable one: $F_i(t) = \bar{F}_i + \Delta F_i(t)$ and $F_r(t) = \bar{F}_r + \Delta F_r(t)$. From this, the connection between the variable components of the ionizing and reprocessed light can be described as $F_r(t) = \int_0^{\tau_{max}} \Psi(\tau)\Delta F_i(t-\tau)d\tau$, where $\Psi(\tau)$ is termed the “transfer function”. This function encompasses all data about the BLR geometry, which dictates the precise reprocessing of the incoming ionizing flux.

To effectively use this technique, one needs to monitor emissions from both the disc and the BLR over timescales at least three times the duration of the anticipated time lag. Given the BLR radius typically approximates several tens of light-days, monitoring should span a minimum of several months before a reliable time lag can be ascertained. Through this monitoring, light curves for both the disc and the BLR are obtained. Subsequently, these light curves are cross-correlated to deduce the time lag and thereby determine the BLR size. Reverberation mapping was initially employed to ascertain BLR sizes by evaluating lags between the UV/optical continuum from the disc and the emission of broad lines such as CIV and Hβ. This principle has since been expanded to examine the dusty torus in infrared, the accretion disk (in the UV/optical continuum), and the X-ray corona.

6.2 Measurement of the Hubble constant

As shown in the Introduction, the standard, flat ΛCDM is able to satisfactory explain a large number of cosmological data. At the same time, with the improvement of cosmological observations, tensions in the estimate of key cosmological models have emerged, with the H_0 tension being the most statistically significant. The Hubble constant, H_0, represents the rate at which the universe is currently expanding, and it is one of the most important parameters in cosmology. Historically, astronomers have used various methods to determine H_0, but recently two primary methods have produced conflicting results.

One method involves using Cepheid variable stars to calibrate Type Ia supernovae absolute magnitude in the nearby universe. This ’local’ approach, largely championed by the SH0ES team, has yielded a value of $H_0 = 73.2 \pm 1.3$ km s^{-1} Mpc^{-1} (Riess et al., 2009, 2011, 2016, 2019).

On the other hand, the Planck satellite, which measures the temperature and polarization spectra of the Cosmic Microwave Background (CMB) radiation, gives a value of $H_0 = 67.4 \pm 0.5$ km s^{-1} Mpc^{-1}(Hinshaw et al., 2013; Planck Collaboration et al., 2016, 2020). This method infers H_0 by studying the early universe and modeling its subsequent evolution. The tension between the two measurements is now at more than 4σ of significance. It is important to notice a difference between the two methods: SNIa measure the H_0 value locally and, in doing so, there are no cosmological assumptions (expect for the Cosmological Principle). The value derived from the CMB measurement, instead, is obtained assuming the standard flat ΛCDM cosmological model. If another cosmological model is assumed, the value we obtain from the CMB changes consequentially.

The discrepancy between these two values is significant and cannot be easily brushed off as mere observational error. It suggests either unknown systematic errors in one (or both) of the measurements or a hint of new physics beyond our current understanding of the Universe.

In this Section, we briefly sum up the topic of the tension, and then explain how Reverberation Mapping on local AGN, in synergy with interferometry measurements, could provide an independent estimate of H_0, without the need of calibration.

6.2.1 The tension

Figure 6.1, from Di Valentino et al. (2021), shows the current status (as in 2021) of the H_0 tension. The measurements are divided between "direct" and "indirect" ones. Indirect measurements are based on the analysis of the CMB data, and are derived by assuming a flat ΛCDM model. Direct measurements, instead, measure H_0 in the local Universe.

The CMB gives us information about the Universe at redshift $z \sim 1100$, when it was less than 500 Myr old. It is characterized by temperature anisotropies on a blackbody spectrum, which have arised from fluctuations in the early Universe. These temperature anisotropies are decomposed in their constituent angular scales using spherical harmonics. In this way we can derive the power spectrum, which is the plot of the temperature fluctuations as a function of their angular scale. This spectrum can be compared with the predictions of different cosmological models, and fit to derive the parameters inside such model. Using the assumption of the ΛCDM model, one can derive the Hubble parameter H at the redshift of the CMB, which means, the expansion rate of the Universe at redshift $z \sim 1100$. Using again the assumption of the ΛCDM model, one can extrapolate this value to derive H_0, the current expansion rate of the Universe.

The best-established method for this measurement comes from measuring the distance-redshift relation, a process which is often referred to as the building of the "distance ladder". The approach consists in using geometry (usually parallax measurements) to calibrate the absolute luminosities of pulsating Cepheid variables. Once Cepheid absolute luminosity is known, they become "standard candles", which is, by measuring their observed luminosity one can derive their distance without any cosmological assumption. Cepheids are observed up to 10 - 40 Mpc, which means we can use them to derive measurements inside our own and in nearby galaxies. However, their brightness is not sufficient to be seen at large cosmological distances, so we can not use them to probe objects that are in the "Hubble flow". Luckily, we can use Cepheids to calibrate the intrinsic brightness of SNIa, which are instead bright enough to be observed at cosmological distances. In other words, by first determining the distance to nearby galaxies using Cepheids and then observing SNIa in those same galaxies, we estimate the true luminosity of SNIa. Once calibrated, SNIa

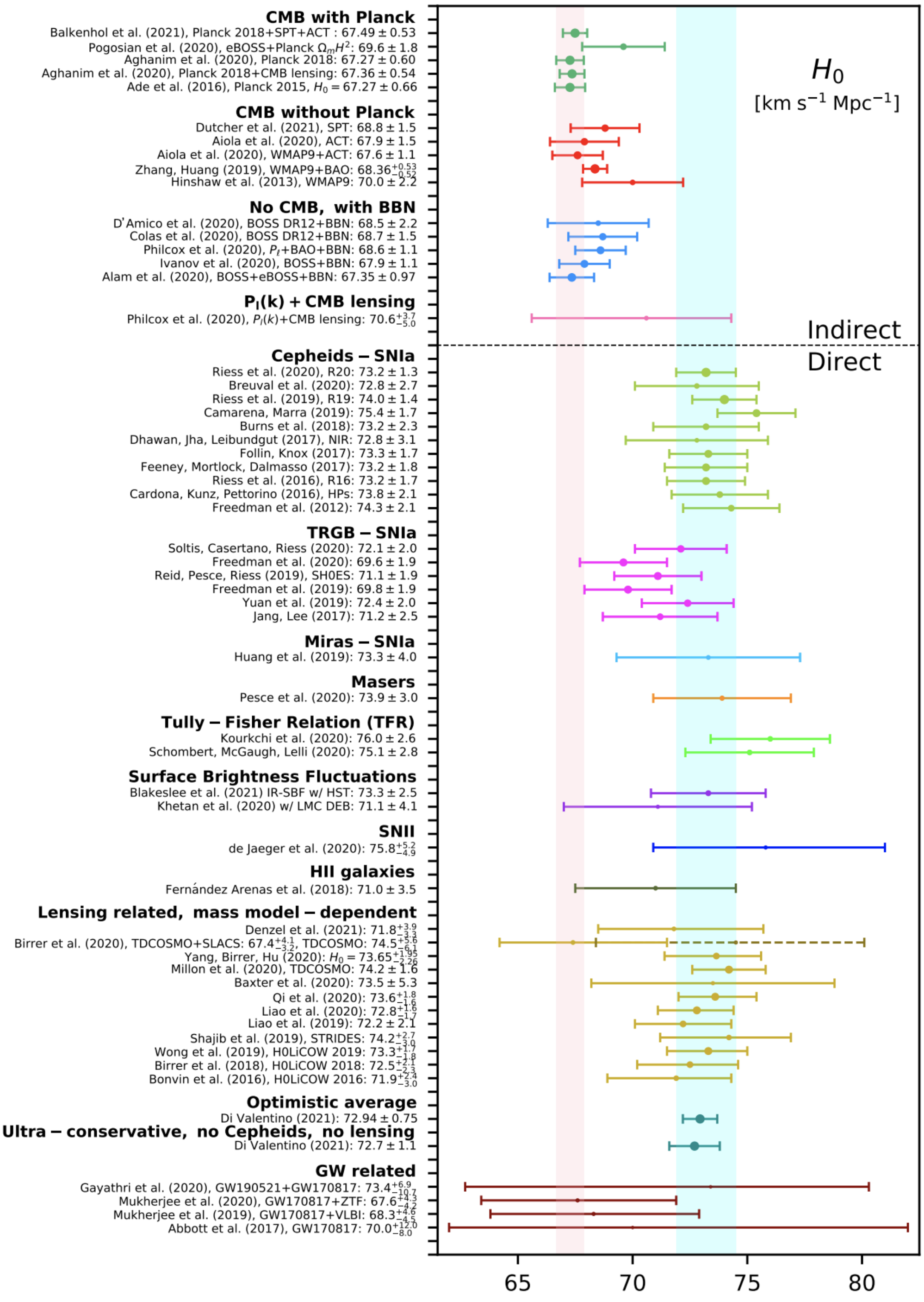

Figure 6.1: Figure from Di Valentino et al. (2021), which shows the current status of the H_0 tension as in 2021, distinguishing between direct (or "local") and indirect measurements. The cyan vertical band corresponds to the H_0 value from Riess et al. (2019) $H_0 = 73.2 \pm 1.3$ km s^{-1} Mpc^{-1}; the light pink vertical band corresponds to the H_0 value as reported by Planck 2018 team Planck Collaboration et al. (2020) within a flat ΛCDM scenario.

can be used to measure distances to much farther galaxies. Given a sample of objects for which we know the distance and the redshift (which means, the velocity at which they are receding from us), the fit of the "Hubble law" $v = H_0 * d$ allows us to estimate the Hubble constant.

The significance of the tension started the analysis and discussion of cosmological models alternative to the standard flat ΛCDM, which go from the simplest extensions of the ΛCDM paradigm (like the "Early Dark Energy" and "Late Dark Energy" models), to Inflationary Models, Modified gravity models and more. For a review on the topic, see Di Valentino et al. (2021).

On the other hand, it has also been suggested that the tension does not actually exists, but arises because of systematics in one (or more) measurements. Regarding the local measurements of H_0, errors in the parallax measurements might propagate up the ladder, causing the tensions. Others have suggested that the period-luminosity relationship that makes Cepheids standard candles can actually be affected by metallicity, age or other stellar properties in way that are not accurately taken into account, and/or that the corrections applied to Cepheid measurements for dust extinction are not reliable. The "Tip of the Red Giant Branch" (TRGB) has been proposed as an alternative to Cepheids for the SNIa calibration (e.g., Freedman et al., 2019; Anand et al., 2022). However, calibrating the absolute magnitude of the TRGB and understanding how metallicity and other properties influence it introduces new challenges. The "local flow" effect has also been proposed as a possible cause of the tension: peculiar velocities of the galaxies in which SNIa are observed might introduce systematic errors.

Overall, there are many factors and challenges to be considered in discussing the "distance ladder" measurement of H_0. One way to resolve the issue consists in providing another local H_0 measurement, completely independent of the "distance ladder" ones. If the two measurements were to coincide, it would be a strong evidence for the actual presence of the tension, and therefore the need for new cosmology beyond the standard ΛCDM model. This is where Reverberation Mapping comes into play.

6.2.2 SARM

To measure the local value of the H_0 constant, "standard candles" have usually been used. We remind that a standard candle is an object for which we can know its absolute magnitude without any cosmological assumption. By comparing it with the observe luminosity, we can derive its distance and, given the distance and the redshift, we can estimate the Hubble constant. Another way to do so is by using a "standard ruler", which is, an object whose size on the sky can be known without any cosmological assumption.

The combination of Infrared Reverberation Mapping and Spectroastrometry (the method has been called "SARM") allows us to use AGN as standard rulers by measuring the size of their BLR. In this way, we can estimate the H_0 constant. We have already discussed how Reverberation Mapping allows us to measure the linear size of the BLR of an AGN, ΔR, by tracking the time delay between variations in the disc continuum emission and the subsequent response in the BLR emission. Spectroastrometry, instead, uses interferometry to measure the angular size of an AGN BLR, $\Delta\theta$. By combining light from multiple telescopes, we achieve unparalleled angular resolution, permitting the direct measurement of minute angular scales (GRAVITY Collaboration et al., 2017). Once we have estimated both the linear and angular sizes of a given AGN BLR, we can treat it as a "standard ruler", in much the same way that Cepheids or SNIa serve as standard candles. Indeed, we know that an object angular distance is related to its linear and angular sizes through the relation $D_A = \Delta R/\Delta\theta$. With knowledge of the distance and the object redshift, we can derive an estimate of the H_0 constant.

The primary advantage of this approach is its pure geometric nature; it does not rely on any other measurement, calibration, or assumption, aside from positing that the RM and interferometry measurements quantify the same property. During the last year of my PhD I have been involved in the Reverberation Mapping side of the SARM project on the AGN Ark120. We will now briefly discuss Specroastrometry and then the on-going campaign on the AGN Ark120.

GRAVITY and 3C 273

GRAVITY is an instrument at the Very Large Telescope (VLT) Interferometer on Cerro Paranal, Chile. It combines light from up to four 8.2m VLT Unit Telescopes, achieving milliarcsecond-level angular resolution. GRAVITY can detect relative positional shifts as small as 10 microarcseconds. It has been designed primarily for precision astrometry and imaging in the near-infrared.

The spectroastrometry technique consists in combining high-resolution spectroscopic data to extract spatial (or astrometric) information on scales finer than the telescope's diffraction limit. This is achieved by examining the centroid position of spectral features as a function of wavelength, effectively mapping the spatial origin of different spectral components. Implementing this technique, GRAVITY can obtain both high-resolution spectra and precise astrometric measurements by combining light from multiple telescopes, allowing to probe the spatial distribution of emission regions around compact objects, such as the size of the BLR of an AGN. These observations can therefore measure the angular structure of the BLR in the direction perpendicular to the line-of-sight (GRAVITY Collaboration et al., 2017).

The SARM method has been firstly tested on the object 3C 273, an AGN with redshift $z \sim 0.158$. For the RM side of the project, a 10-yr RM campaign has been conducted with joint observations on the Bok 2.3m telescope at Steward Observatory, University of Arizona, and the Lijiang 2.4m telescope in Yunnan Observatory, Chinese Academy of Sciences (Zhang et al., 2019). The campaign generated light curves for the $H\beta$ line with an error of $\sim 2\%$. 3C 273 was observed with GRAVITY in July 2017, January, March, and May 2018 (Gravity Collaboration et al., 2018), with a spectral resolution of $\lambda/\Delta\lambda \sim 500$, which allows to constrain the BLR parameter using the Paschenα line at 1875 nm. The joinery of the RM and SA results gives an angular distance of $D_A = 551.5^{+97.3}_{-78.7}$ Mpc. Employing the redshift-distance relation, the obtained estimate for the Hubble constant is $H_0 = 71.5^{+11.9}_{-10.6}$km s^{-1} Mpc^{-1}. The uncertainties on this estimate are such that they are consistent both with the CMB predictions and the SNIa ones. However, the result is extremely encouraging as it shows that with a few tens of observations like this it would be possible to constrain the Hubble constant at a precision level that would be useful for the Hubble tension.

The Ark120 campaign

So far, all RM campaigns of GRAVITY targets have been conducted in the optical, focusing on broad Hβ, while GRAVITY observes in the near-infrared (NIR). Any difference in BLR emissivity between the optical and NIR is thus a major source of systematic uncertainty. To fully realize the potential of SARM, it is essential to monitor a GRAVITY target in the NIR, measuring infrared broad emission lines such as Paschenα, Paschenβ and Brackettγ. Ark120 is an AGN at redshift $z = 0.03271$ (Theureau et al., 2005), and it was supposed to be observed by GRAVITY between September 2022 and April 2023. After a feasibility test in the first part of 2022, a monitor of Ark120 at IRTF with the SpeX instrument was proposed and accepted for the period September 2022-January 2023. The Infrared Telescope Facility (IRTF) is a 3.0-meter telescope optimized for infrared observations, located at the Mauna Kea Observatory in Hawaii. SpeX is a medium-resolution infrared spectrograph designed for IRTF. It operates in the near-infrared range, covering wavelengths from 0.8 to 5.5 microns. The instrument boasts multiple spectroscopic modes. The one used for Ark120 observations is the Short Wavelength (SXD) mode, which offers medium-resolution (R750-2000) spectroscopy on a wavelength range of 0.8 to 2.4 microns. The goals of the campaign were to (i) measure the infrared time lag, delivering a black hole mass measurement and determining the size of the NIR BLR, (ii) map the geometry and kinematics of the NIR BLR, by applying dynamical model methods (Pancoast et al., 2014) to the velocity resolved RM light curves, (iii) study the relation between optical and near infrared emissivity, by comparing the results of the IR campaign with past and ongoing optical RM efforts on the same targets, and with the GRAVITY view, and (iv) compare RM and SA, measure geometric distances, and eventually the Hubble constant.
Together with the IRTF monitoring, the object was monitored by the Las Cumbres Observatory (LCO) telescope network every other day, providing photometric information.
Observations started in September 2022. When observing an in the infrared from ground-based telescopes, the obtained spectrum is affected by the Earth atmosphere, which introduces telluric absorption lines primarily due to water vapour, carbon dioxide, and other molecular species. To correct for these telluric absorptions, we employ a process using a reference star, often termed a “telluric standard star”. A star, typically a hot, featureless A or B-type star, is chosen because its intrinsic spectrum is well-known and predominantly free of complex molecular absorption features in the infrared. This allows any observed features in the star’s spectrum to be attributed primarily to atmospheric effects. The AGN and the telluric standard star are observed sequentially, preferably close in time and at similar air masses. This ensures that both observations experience similar atmospheric conditions. The spectra of both the AGN and the telluric standard star are then extracted, and the observed spectrum of the AGN is divided by the spectrum of the telluric standard star. This removes the telluric absorption features introduced by the Earth atmosphere. However, this also divides out the intrinsic spectrum of the star. To compensate for dividing

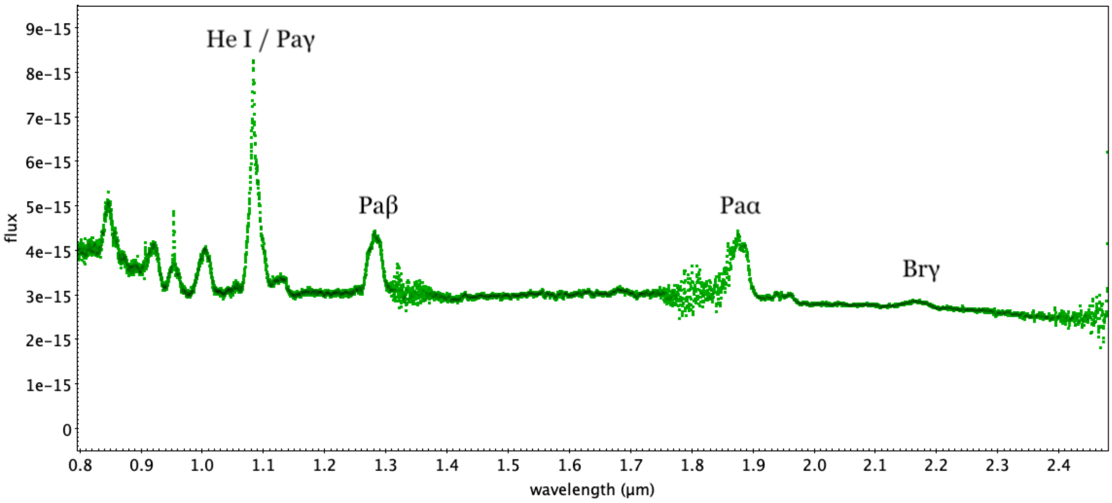

Figure 6.2: NIR spectrum of Ark120 obtained at IRTF in 2022 on October the 16th. The relevant lines are highlighted: HeI/Paγ, Paβ, Paα, and Brγ, respectively.

out the star intrinsic spectrum, the result from the previous step is multiplied by a model or template spectrum of the telluric standard star. This ensures that the final spectrum represents the AGN intrinsic emission.

For this campaign, the star HD 34317 was chosen, and it was observed at every observations right before or after the AGN observation. In Figure 6.2 we show, as an example, the final spectrum for the 16th October 2022 night. The lines whose signal to noise ratio is enough for them to be relevant for Reverberation Mapping are HeI/Paγ, Paβ, Paα, and Brγ. Unfortunately, IRTF stopped operating in the last days of October 2022 because the dome shutter broke down, and was not operating again until spring 2023. Therefore, the campaign could not continue. Furthermore, during the time of the 12 observations that took place between September and October 2022, Ark120 did not show any visible variability, so it was not possible to recover a time-lag and, therefore, not even a rough estimate of the BLR size with the infrared lines.

The GRAVITY side of the project was unfortunate too, as bad weather nights prevented Ark120 from being observed.

Although the campaign has been unlucky, its scientific relevance is undeniable. Efforts are being made so that what could not be done in 2022, will be done between 2023 and 2024. The GRAVITY collaboration has planned to observe Ark120 between September 2023 and April 2024; so, unless another unfortunate series of bad weather nights makes it unviable, the interferometric observations of the BLR would be happing soon. At the moment of this thesis (October 2023), new Infrared monitoring campaigns of Ark120 have been proposed, both at IRTF and at the Gemini South facility (with the instrument Flamingos2), and we hope that they will be approved. The LCO telescope networks continue its monitoring of Ark120, and we will also have the spectral coverage between 0.95 – 2.4 microns from the 3.5 meter ARC telescope at the APO facility between October 2023 and March 2024. Regarding the ARC monitoring, due to the difference in atmospheric transmission, the Paα line will not be recovered as confidently as in the IRTF observations. However, Brγ coverage will still exist. Furthermore, the ARC spectra includes the [S III] line that will be

used for traditional flux scaling as is done in the optical, allowing for exploration of the efficacy of the narrow-line flux scaling technique next to using a comparison star for NIR RM work.
To sum up, observations in the next six months should give us the opportunity to test the SARM technique on the AGN Ark120, to give a measurement of the H_0 constant and to establish SARM as a cosmological probe. The relevance of this result for cosmology is huge, as it can provide a purely geometrical estimate of H_0, independent from any kind of calibration, providing a fundamental step in the discussion around the H_0 tension.

6.3 High-redshift SMBH masses

Given the BLR radius and the line width for a broad emission line, if we assume the motion of the BLR around the SMBH to be virialized, we can use the results of reverberation mapping obtain an estimate of the SMBH mass as:

$$M_{BH} = f\frac{R\Delta V^2}{G} \tag{6.1}$$

where G is the gravitational constant, and f is an order-unity scaling factor that accounts for the geometry and kinematics of the BLR emitting gas. The need for this scaling factor is one of the primary sources of uncertainty for reverberation mapping mass measurements. The population average has been obtained with comparison of the $M_{BH} - \sigma_*$ relation of quiescent galaxies with AGN (Gültekin et al., 2009; Kormendy & Ho, 2013; McConnell & Ma, 2013). The assumption behind the comparison is that all galaxies (active and non) come from the same parent sample. When the line velocity width is measured as the second moment of the line, $< f >$ values range around 4-5 (Park et al., 2012; Grier et al., 2013; Batiste et al., 2017). If we assume inclination to be the main contribution to the value of the scale factor f, a value of $< f >\sim 4 - 5$ implies that the average AGN in the reverberation mapping sample is viewed with an inclination angle of ~30°, which is consistent with the ideas behind the AGN unification mode (Urry & Padovani, 1995). However, the single AGN might be viewed at angles that deviate from the population average. As a consequence, reverberation mapping SMBH masses for individual objects that make use of $< f >$ have an additional factor of 2-3 uncertainty.

Reverberation mapping campaigns are very time consuming, Hence, this technique can not be used to measure the masses of hundreds of thousands of AGN. To do so, one can however use another important product of reverberation mapping, which is the relation between the BLR radius and the AGN luminosity (the $R_{BLR} - L_{AGN}$ relation, Koratkar & Gaskell, 1991; Kaspi et al., 2005). This relationship, whose existence is consistent with the predictions made with photoionization arguments, can be calibrated and then used to estimate M_{BH} for large numbers of AGN with only one spectrum per objects (e.g., Shen et al., 2011). Most of the BLR sizes have been measured using the $H\beta$ line, so this emitting region of the BLR is the one better constrained at the moment. For high redshift quasars, MgII and CIV have been used (Vestergaard & Peterson, 2006; Onken & Kollmeier, 2008; Woo et al., 2018). However, the relationship between the $H\beta$, CIV and MgII BLRs is not

completely understood, and the limited luminosity and redshift range of UV reverberation mapping measurements means that the relation for these lines relies on extrapolations from local AGNs. Furthermore, studies have shown that many AGN, including some of the ones with highest accretion rate, fall below the $R_{BLR} - L_{AGN}$ relation (e.g., Fonseca Alvarez et al., 2020). The matter is being investigated, with studies suggesting that it is related to differences in the ionising continuum properties, while others suggest it is caused by physical changes that come with the high accretion rate (Dalla Bontà et al., 2020).
In addition to this, differences in the line profiles between CIV and the Hβ lines have called into question the validity of CIV as a line for single-epoch mass measurements (Baskin & Laor, 2005; Shen & Kelly, 2012). At the same, time, studies suggest that this issue can be mitigated with a proper data quality selection (e.g., Vestergaard & Peterson, 2006).

Although difficult, understanding the precise masses of Supermassive Black Holes (SMBH) at high redshift is paramount to cosmology and our overarching comprehension of the Universe's evolution. The characterization of these objects is fundamental to understand the early stages of structure formation and of the evolution os galaxies and their central SMBH. It is still not clear how these objects came to be and how they evolved together with their galaxies - measuring their masses, we can distinguish between different formation and growth models, and have a clearer understanding of how they influenced star formation rates and induced galactic outflows.
In addition to this, mergers of SMBH are predicted to be potent sources of gravitational waves. Knowledge of their mass distribution at high redshift can pave the way for future gravitational wave detectors by forecasting potential signatures.

In order to gain a deeper understanding of high-redshift SMBH mass measurements, additional reverberation mapping measurements are required. Over the past few decades, several campaigns have been undertaken to enhance the CIV $R_{BLR} - L_{AGN}$ relation by increasing the sample size (Kaspi et al., 2007; Grier et al., 2019; Hoormann et al., 2019). However, high-redshift reverberation mapping campaigns face several challenges:

- Reverberation mapping necessitates high signal-to-noise ratios. At higher redshifts, we encounter fainter objects, making this requirement more challenging to meet.
- For objects with satisfactory signal-to-noise ratios, they typically exhibit high luminosity. However, more luminous AGNs are known to be less variable than their less luminous counterparts. As a result, obtaining a reliable time lag in a feasible time frame becomes more difficult when using a high-luminosity object.
- Reverberation mapping is inherently time-consuming. When examining high-redshift objects, the $(1+z)$ time dilation effect becomes significant. For instance, an object with $z \sim 2$ demands thrice the monitoring duration to ascertain the same lag as it would if it were a local entity.
- These campaigns are further extended because more luminous objects possess a larger BLR, as dictated by the $R_{BLR} - L_{AGN}$ relation. This implies a longer average monitoring duration.

- variability is a stochastic feature in AGN. This means that we are never certain that the object that we are monitoring is going to vary enough during the campaign. The extended time needed for high redshift campaign exacerbates this issue.

In the subsequent section, we will explore a campaign where these challenges are addressed through gravitational lensing.

6.3.1 The case for strong lensing: SDSS J2222+2745

SDSS J2222+2745 is a quasar at redshift $z = 2.8$, lensed by a foreground galaxy cluster. It was discovered by Dahle et al. (2013), as part of the Sloan Giant Arcs Survey (Hennawi et al., 2008). Figure 6.3 shows the field of view with the multiply imaged quasar and the galaxy cluster in the foreground. The image magnification was characterized by Sharon et al. (2017). The presence of the strong lensing by the galaxy cluster makes this object an incredible target for a reverberation mapping campaign. First of all, thanks to the magnification, the quasar is detectable at $g \sim 21$, despite being intrinsically dimmer. This makes is possible to study an object that otherwise would have been undetectable with present technologies at redshift $z = 2.8$. Furthermore, the quasar has three bright images (A, B, and C in the Figure). This allows us to put together the monitoring of the three distinct images, therefore obtaining more reliable and precise estimates. In addition to this, photometric monitoring on the three images has shown that image C leads the other by nearly two years, with $\Delta\tau_{AB}$=-42.44 days and $\Delta\tau_{AC}$ = 696.65 days (Dahle et al., 2015). This means that with any observations, we can basically extend the lightcurve of the object by two years, shortening the required campaign time. Furthermore, because image C was shown to have an extreme flux variation between 2014 and 2016, it was guaranteed that the other images would have followed. Lensing allows us to “predict the future” and be sure that the images we are going to observe are going to vary. This is crucial, especially for an object at this high redshift, as we have seen that the higher the redshift, the longer the reverberation mapping campaign has to be.

This object was observed from 2016 June to 2020 September using the Multi-Object Spectrograph at the Gemini Observatory North (GMOS), for a total of 47 epochs. For each epoch, three standard stars were also observed, to provide flux calibration. During the same period, it was observed with high-cadence photometry with the Alhambra Faint Object Spectrograph and Camera (ALFOSC) at the 2.56m Nordic Optical Telescope (NOT), with observations using the SDSS$'_g$, SDSS$'_i$, and He I 588$_6$ filters, with mean cadences of 18.3, 18.7, and 17.7 days, respectively.

The results for this part of the campaign were published in Williams et al. (2021). Here, we summarize the results of that paper, then we move to explain while a further extension of the campaign was needed, and present the “work in progress” results.

Results of the initial campaign The GMOS spectra were calibrated using the three standard stars, and then fitted with a model that included a powerlaw continuum, narrow and broad gaussians for the emission lines, and a fourth-order Gauss Hermite poynomial for the CIV and the CIII] lines. The integrated and velocity-resolved CIV emission line

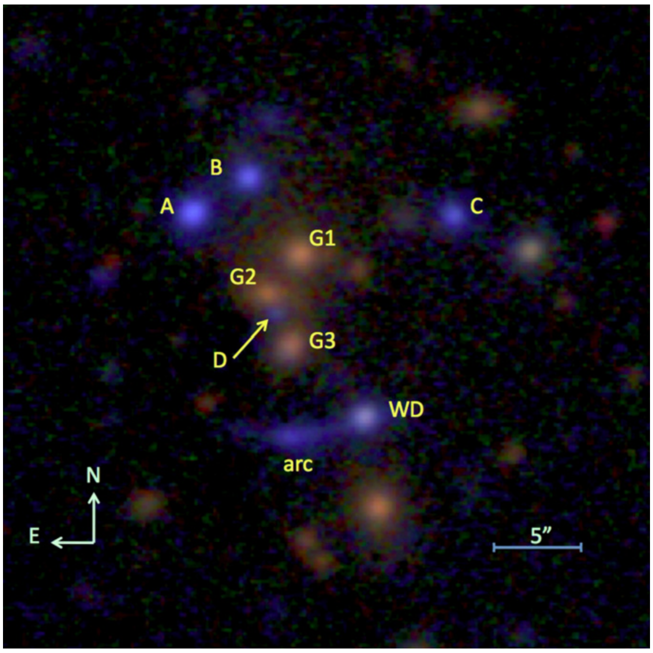

Figure 6.3: Figure from Dahle et al. (2013); NOT/MOSCA gri color composite image of SDSS J2222+2745. The four brightest quasar images are labeled A–D, the three bright foreground cluster galaxies are labeled G1–G3, and a foreground white dwarf is labeled WD.

lags were measured with respect to the g-band continuum, and then cross-correlated with it to obtain the time lag. The obtained integrated CIV lag is $\tau_{cen} = 36.5^{+2.9}_{-3.9}$ in the rest frame, which corresponds to 139^{+11}_{-15} days in the observed frame. Evidence was found for a velocity-resolved lag structure, in which the core of the line responds the slowest ($86.2^{+4.5}_{-5.0}$ days) and the wings the fastest (25^{+11}_{-15} and $7.5^{+4.2}_{-3.5}$). This behaviour is consistent with BLR gas moving in circular Keplerian motion.

The results for this object were compared with others with CIV reverberation mapping measurements for the $R_{BLR} - L_{AGN}$ relation. SDSS J2222+2745 lies 0.26 dex above the mean relation. However, the results is consistent with the relation when campaign variability, intrinsic scatter, and all sources of uncertainty are considered. The uncertainties on the BLR size that are achieved for this object are incredibly small, a results which is unprecedented for objects with similar luminosity and redshift. This is thanks to the lens strong magnification and due to the combination of three different light curves for the same objects for constraining the BLR size.

Calculating the emission line widths and using a scale factor f derived from a sample of objects with both Hβ and CIV line measurements (Dalla Bontà et al., 2020), the mass of the SMBH was computed. When the FWHM of the line is used, the combined result from the three images is $\log(M_{BH}/M_{\odot}) = 9.35 \pm 0.51$, while when the line dispersion σ is used, the result is $\log(M_{BH}/M_{\odot}) = 8.63 \pm 0.27$.

Extended campaign The campaign up to 2020 was able to measure the velocity-resolved lag of the CIV line. The velocity-resolved time lags are fundamental to better understand the structure and kinematics of the BLR, as they can disentangle the geometry (spherical, disk-like, etc.) and kinematics (rotational, radial, or complex motions) of the BLR. Standard reverberation mapping gives a size estimate, but velocity-resolved data can reveal whether different portions of the BLR are moving towards or away from us, or even rotating around the central black hole. Furthermore, some parts of the BLR might respond to changes in the continuum emission faster than others, indicating a stratified BLR structure. This stratification provides insights into the physical conditions, ionization processes, and distribution of the gas around the AGN. For the case of SDSS J2222+2745, there is evidence in the 2016-2020 data for the core responding on longer time scale than the wings. This is what is expected from a pure circular keplerian motion. [1] However, the lags in the wings are not robustly determined due to the lower signal to noise ratio. With simulations it was shown that two additional years of data would have provided enough information to give conclusive velocity-resolved lag measurements, which brought to the writing of a proposal for additional monitoring of the source. Figure 6.4 shows the cross-correlation function for the 5650 Å -5780 Å wavelength bin as with the 2016-2020 data (upper left panel), and as simulation predict it to be after two more years of data (lower left panel). With the additional data, improvements like the one showed in the Figure were expected for all the velocity bins.

The proposal was accepted, which brought to SDSS J2222+2745 to be observed for 9 additional epochs between 2021 and 2022.

Preliminary results for the extended campaign The goals of this second part of the campaign are (i) to provide smaller uncertainties on the BLR size and the SMBH mass, given the extended lightcurve, (ii) to obtain reliable constrains on the velocity-resolved lags previously discussed, and (iii) to extend the analysis to the CIII] line.
The CIII] line has rarely been used for reverberation mapping campaigns (e.g., Trevese et al., 2014), given its much lower flux compare to CIV. The amazing signal to noise ratio achievable thanks to lensing allows us have a reliable estimate of the CIII] flux for our lensed quasar. It is of great interest to compare the results of the reverberation mapping for different emission lines for the same source, when possible. Indeed, this allows us to test the idea for which emission that requires higher ionization is found closer to the central SMBH compared to lines corresponding to lower ionizations. Furthermore, we can use it to understand if the geometry and kinematics of different parts of the BLR act differently inside of the same source.

The spectra for the campaign extension have been modelled and fit as done by Williams et al. (2021). Starting from image A, the mean spectrum for the campaign was fit, and then used as the starting guess for the fitting of the mean spectrum for each year, an then again

---

[1] Indeed, when the central source (close to the black hole) varies in brightness, the innermost, faster-moving clouds will respond first due to their proximity to the central source. However, their emission (being highly Doppler-shifted) contributes to the wings of the line profile. The outer, slower-moving clouds are farther away, so it takes longer for the light from the varying central source to reach them. When they respond, their emission is closer to the line centre.

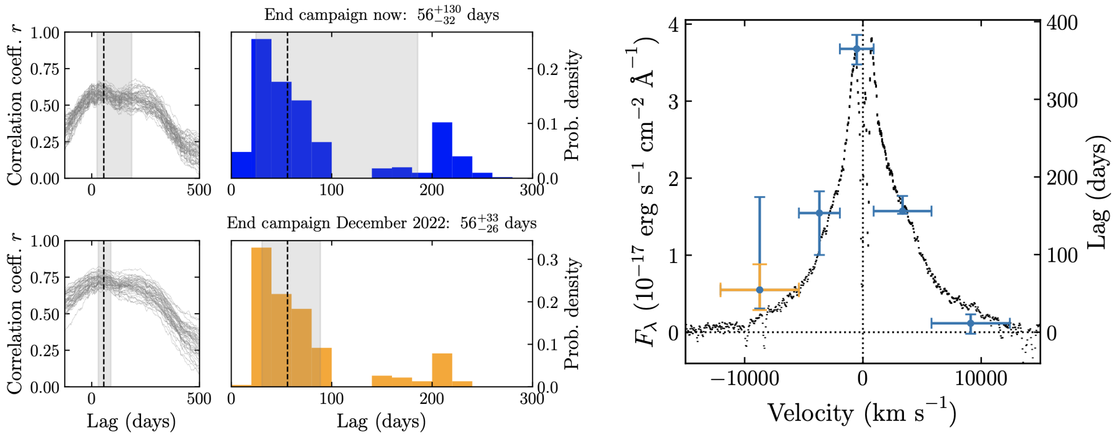

Figure 6.4: Figure from the 2021 GMOS proposal. Left: Cross-correlation function (CCF) for the 5650 Å -5780 Å wavelength bin vs. the g-band continuum light curve for the data obtained so far (top) and including simulated data through December 2022 (bottom). The histograms show the median peak lag (vertical dashed line) and 68% confidence interval (shaded area). The additional two seasons are crucial to suppress the secondary peak which is most likely due to aliasing Right: Time lags (blue) calculated for the five emission line light curves of Figure 2 vs. the g-band continuum. The x-axis error bars indicate the width of the window. There is clear velocity structure in the lag profile, with longer lags near the core of the emission line profile and shorter lags in the wings. The orange error bars show the improvement to the 5650 Å -5780 Å wavelength measurement when we add data through 2022. Improvements are of course expected for all the bins.

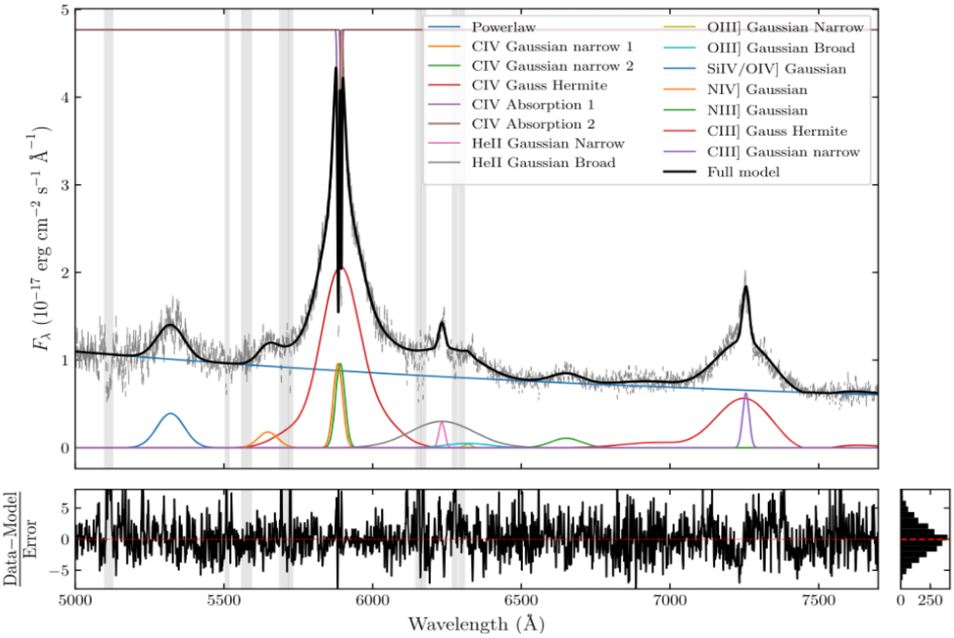

Figure 6.5: Figure from Williams et al. (2021); example model fit to the quasar image A spectrum from 2017 June 21. The top panel shows the data in grey, the total model fit in black, and the individual components with different colours. Gray shaded bands indicate wavelength ranges that were masked out during the fitting procedure due to either bad pixels or absorption features. The normalized residuals are shown in the bottom panel.

this was used as the starting guess for the fit of each individual epoch. In the fit for the individual epochs, the narrow components of the lines were kept fixed, as their emission can be considered constant given the time range of the observations. For more detail about the fitting procedure, refer to Section 4 of Williams et al. (2021). An example fit for image A (from 2017 June 21) is shown in Figure 6.5.

Compared to that method, we implemented one major change about the CIII] line. As can be seen in Figure 6.5, the previous analysis considered a 4th order Gauss-Hermite polynomial for the CIII] line. However, it can be seen that the polynomial ended up modelling parts of the continuum, spreading out in a wavelength range of ~100 Å. This is likely to be due due the presence of an additional feature, at wavelengths shorter than the CIII] line, that is instead modelled by the Gauss-Hermite polynomial. As the previous analysis was only interested in the CIV shape, this was not a big issue, as the only important thing was that the fit correctly recovered the spectrum shape for all those features that were not the CIV line. However, as we are now interested not only in having a more reliable estimates for the CIV line, but also in extending our analysis to the CIII] line, we need to more accurately model the line. After some tests, we found that the best strategy was to add a broad gaussian line to model the bump on the left side of the CIIII] line, and fit the CIII] line as a narrow plus a broad gaussian. In this way, a better covering of the line shape is achieved, as can be seen in Figure 6.6.

The analysis on how to correctly model the lines, and on what to attribute the additional

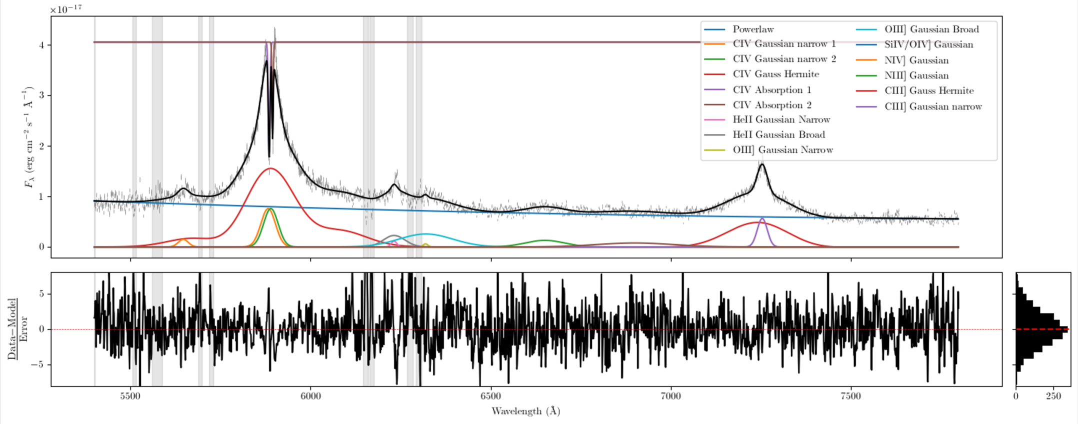

Figure 6.6: New modelling for the CIII] line that includes the presence of an additional gaussian component on the left side of the CIII] line.

Gaussian component, is still ongoing . As a preliminary results, we show the lightcurves in Figure 6.7. We see that the CIII] lightcurve traces the CIV one, as expected, although uncertainties on each flux estimate are visibly higher. From the cross-correlation of the CIII] line light curve with the g band light curve, we obtain a time lag of $\tau_{CIII]} = 52^{+7}_{-5}$. This is roughly consistent with the prediction of $R_{CII]}/R_{CIV}$ ~2.
Given that the choosing of the right line shape is still a work in progress, we do not provide here the results of the velocity-resolved analysis for the two lines, although we point out that the uncertainties on the CIV line estimates do tend to be smaller than with the 2016-2020 data, as predicted.

To sum up, the current work-in-progress analysis on the extended campaign on SDSS J2222+2745 aims at recovering reliable lag estimates for the CIV and the CIII] BLRs, and compare the two, together with refining the velocity-resolved lag analysis, that allows us to investigate the kinematics, geometry and stratification of the BLR of this quasar. It is important to point out that a better understanding of the BLR geometry, kinematics and stratification is fundamental in the scheme of the quest for the understanding of SMBH at high redshift. Indeed, with a better understanding of the BLR we might get a more detailed view on the scale factor f, which is still a big source of uncertainty for mass estimates. Furthermore, a more detailed and reliable view of the CIV BLR can help us position this object on the $R_{BLR} - L_{AGN}$ relation with more confidence, and better the calibration of the relation for single epoch mass measurements.
In the coming future, surveys like LSST are going to observe tens of objects like this one (Taak & Treu, 2023). Although this kind of analysis is very time-consuming, having even 10 more point on the $R_{BLR} - L_{AGN}$ would be fundamental to have better calibrations and therefore a better understating of SMBH at all redshift and their role in the Universe structure formation history.

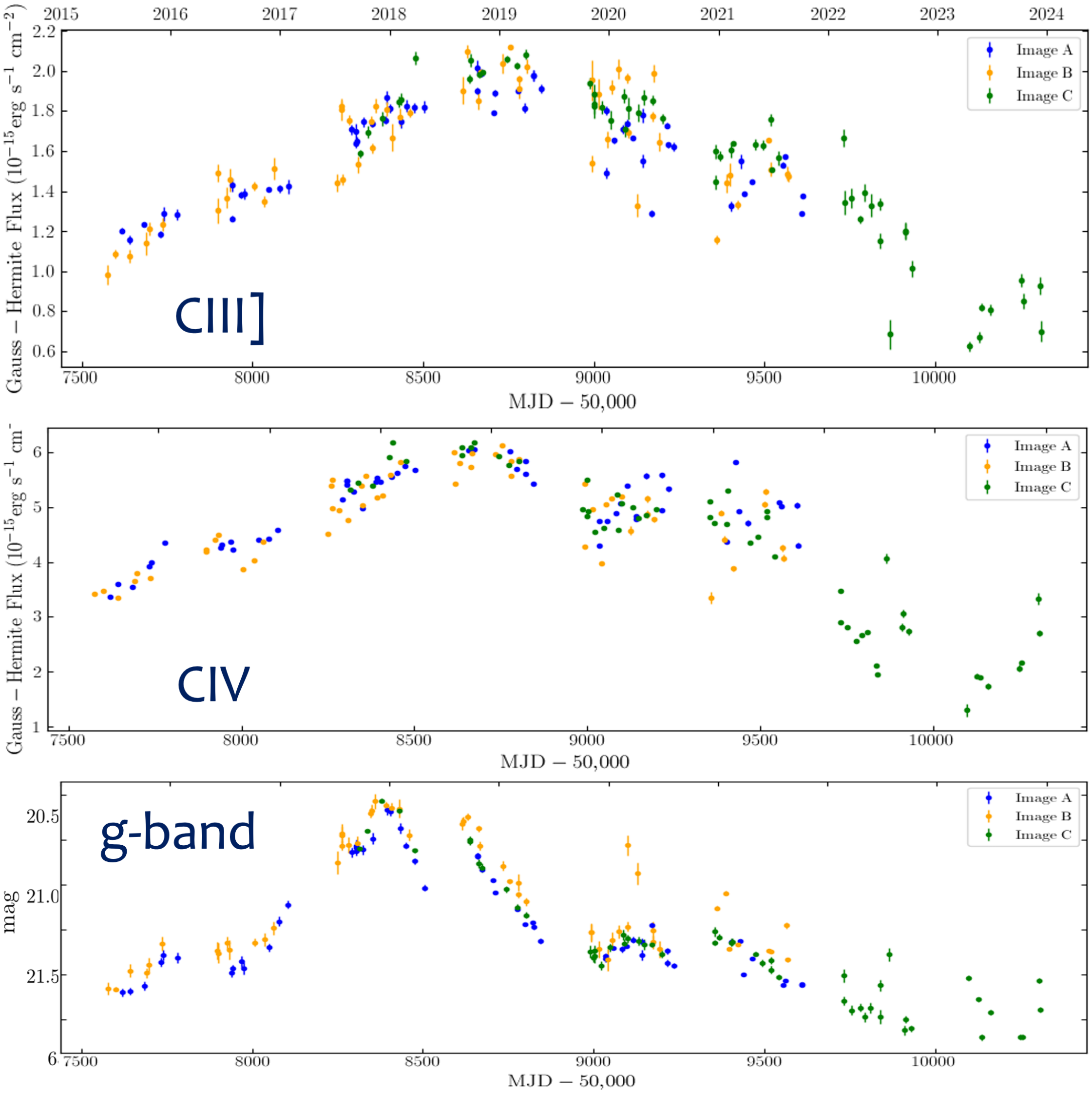

Figure 6.7: Preliminary Figure - The C IV , CIII] emission lines light curves and g-band continuum light curve. The blue, orange, and green points correspond to the quasar images A, B, and C. All data points have been shifted to match the trailing image B observed frame and scaled to match the brighter image A flux using the measured gravitational lensing time delays and relative magnifications.

Chapter 7

The obscured fraction of AGN through cosmic epochs

The characterization of AGN demographics and their evolution is crucial to understanding the history of the accretion onto SMBH and their relation with their host galaxies. We know that the masses of SMBHs residing in the centers of most galaxies correlate with the host properties, such as stellar luminosity, stellar mass, and bulge velocity dispersion (e.g., Marconi & Hunt, 2003; Ferrarese et al., 2006; Kormendy & Ho, 2013; de Nicola et al., 2019). These correlations indicate a co-evolution scenario of SMBHs and galaxies across cosmic epochs that has been observationally investigated and theoretically modeled (e.g., Croton et al., 2006; Somerville et al., 2008; Fabian, 2012; Habouzit et al., 2019; Ricarte et al., 2019), but is still far from being understood in its entirety. For example, in the early Universe the very formation and accretion processes leading to the first SMBHs are still debated. A simple accretion history on stellar masss black holes formed by the first stars is challenged by the discoveries of SMBHs of 1-10 billion solar masses at redshifts higher than 6 (Mortlock et al., 2011; Wu et al., 2015; Bañados et al., 2016; Farina et al., 2022; see also Fan et al., 2022, for a recent review). To match these masses, the accretion process needs to be Eddington-limited or even super-Eddington for long times, or we need to have very massive black hole ”seeds” to start with.
Although both the accretion process and the masses of the seeds are still debated, the majority of galaxies are thought to have undergone an active nuclear phase, in which they can be detected as AGN (Kormendy & Ho, 2013). This makes investigations of AGN at different cosmic epochs crucial so that we can understand the growth and evolution of both SMBHs and galaxies.

However, the presence of gas and dust, both in the innermost nuclear regions and across the whole host galaxy, poses a significant challenge to AGN detection and characterization, given the damping of the emission in the optical-UV band, where AGN intrinsic power peaks. AGN population synthesis models agree that most SMBH growth is hidden to our view by high gas column densities (see, e.g., Gilli et al., 2007; Ueda et al., 2014; Ananna et al., 2019). This scenario has been confirmed by several observational works (e.g., Lanzuisi et al., 2018; Vito et al., 2018), which further show that, at high redshifts (z >3–4), the fraction of luminous AGN obscured by column densities $N_H > 10^{23}$ cm^{-2} is particularly high, ~80 %, as opposed to 20–30 % measured in the local Universe (see, e.g., Torres-Albà et al., 2021).

X-ray surveys provide one of the most effective ways to detect obscured AGN over a wide range of redshifts and luminosities (see, e.g., Brandt & Alexander, 2015; Xue, 2017; Hickox & Alexander, 2018, for an extensive review), and are therefore key to finding and characterizing accreting supermassive black holes in the early Universe. While the AGN emission in the X-rays is <10 % of the overall AGN luminosity (e.g., Lusso et al., 2012; Duras et al., 2020), it undergoes very little contamination from non-AGN processes (e.g., X-ray binaries, diffuse gas emission), and is significantly less biased against obscuration than optical emission. These reasons make X-ray surveys a great and efficient way to detect AGN and to characterize them and their obscuration. Several works have used X-ray data to investigate the evolution of AGN obscuration with cosmic times (La Franca et al., 2005; Tozzi et al., 2006; Treister & Urry, 2006; Ueda et al., 2014; Buchner et al., 2015; Aird et al., 2015; Liu et al., 2017; Vito et al., 2018; Lanzuisi et al., 2018; Iwasawa et al., 2020; Peca et al., 2023). These works find that the fraction of obscured objects increases with redshift, but the physical origin of this trend is not completely understood (see, e.g., Iwasawa et al., 2012). There are indeed arguments that suggest that the properties of the obscuring torus do not evolve significantly: for example, the spectral energy distributions (SEDs) of AGN are the same at very different redshifts (Richards et al., 2006; Bianchini et al., 2019). This means that the covering factor of the dusty torus is unlikely to change with time. The properties of the interstellar medium (ISM) of the host galaxies, instead, vary significantly with time. The content of gas is higher at early cosmic times (see, e.g., Scoville et al., 2017; Tacconi et al., 2018; Aravena et al., 2020), and galaxies are also smaller in size (Allen et al., 2017; Fujimoto et al., 2017). This means that, as the ISM density increases at high redshift, its column density can reach very high values and be the principal contribution to the obscuration of AGN, as shown in several recent works (e.g., Circosta et al., 2019; D'Amato et al., 2020; Gilli et al., 2022). This has also been shown by hydrodynamic (Trebitsch et al., 2019) and cosmological (Ni et al., 2020) simulations.

In this chapter we investigate the X-ray properties and derive the obscured fraction of the AGN sample in the J1030 *Chandra* deep survey. In 2017 the *Chandra* telescope observed a 355 arcmin^2 region around the z = 6.31 quasar SDSS J1030+0525 for ~500 ks. The field around it has dense multi-wavelength coverage, being observed with MUSYC-DEEP, HST/WFC3, HST/ACS, VLT/MUSE, WIRCam, IRAC (see, e.g., Peca et al., 2021). The *Chandra* survey has a 0.5-2 keV flux limit $f_{0.5-2keV} = 6 \times 10^{-17}$erg s^{-1}cm^{-2} in the central square arcmin and it is, to date, the fifth deepest extragalactic X-ray field (Nanni et al., 2020). The survey resulted in the detection of 256 sources, of which 3 are identified as stars based on their spectra, and 4 more based on their brightness in the K band and low X-ray-to-optical rate (Nanni et al., 2020; Marchesi et al., 2021). Among the remaining 249 extragalactic sources, Marchesi et al. (2021) were able to compute a photo-z for 243 of them, which make the sample considered here.

Multiple spectroscopic campaigns allowed the determination of the spectroscopic redshifts for 135 objects out of these 243 (i.e., 56 % of the extragalactic sample; Marchesi et al., 2021, 2023). Here we present the complete spectral analysis of the X-ray spectra of the 243 *Chandra* J1030 extragalactic objects. Our goal is to determine the physical properties of these sources and to study the evolution of the obscured AGN fraction with luminosity and redshift. We assume a flat ΛCDM cosmology with $H_0 = 69.6$ km s^{-1}Mpc^{-1}, $\Omega_m = 0.29$ and $\Omega_\Lambda = 0.71$ (Bennett et al., 2014).

7.1 Sample description and X-ray spectral extraction

The *Chandra* J1030 extragalactic sample is made of 243 objects, for which we have a redshift estimate that can be either photometric or spectroscopic. In Figure 7.1 we show the redshift distribution of the objects in the catalog; a spectroscopic redshift estimate is available for 135 of them. For 20 out of the 108 photometric estimates, the redshift probability distribution is flat (see Marchesi et al., 2021).[1] In Figure 7.1 we show their minimum redshift estimate. The 135 objects with spectroscopic redshift also have spectral classification and are divided into four categories (see Marchesi et al., 2021): 20 narrow-line AGN (NL-AGN), 43 broad-line AGN (BL-AGN), 32 early type galaxies (ETG), and 40 emission line galaxies (ELG). The numbers above are updated with respect to the values reported in Marchesi et al. (2021), following new spectroscopic observations (Marchesi et al., 2023).

Regarding the objects with only photometric redshift, recent works have proposed a way to derive an additional redshift estimate from the X-ray spectra (e.g., Simmonds et al., 2018; Sicilian et al., 2022). The method was tested for a subsample of the catalogue in Peca et al. (2021). However, this method, requires highly obscured objects with a large number of counts ($N > 150$, Sicilian et al. (2022)) to give redshift estimates that are more accurate and reliable than the photometric values. Given the average properties of the sources in our sample, the X-ray spectrum is likely to provide a more precise redshift estimate only when the Fe Kα line is detected. This will be discussed in Section 7.2.

The spectra are extracted using the software *Chandra* Interactive Analysis of Observations (CIAO) v.4.13. For the choice of the extraction radius, we performed a preliminary ad hoc analysis. We argue that, as the PSF broadens with the increasing of the off-axis angle of the object, the best choice for the extraction radius might be different for objects at different off-axis angles. Furthermore, we expect to include more signal in a larger radius when the signal to noise ratio (S/N) is higher; given this, we might need different radii between low- and high-count objects.

To investigate this issue, we performed an analysis on a randomly chosen subsample of 35 objects that span the off-axis–count plane in the same way as the whole sample. We extracted and fitted the spectra obtained with different extraction radii, corresponding to the 75% 80%, 85%, 90%, and 95% of the encircled energy. We then compared the S/N obtained with each encircled energy choice. The S/N varies significantly between the different choices and, more importantly, there is no clear trend of the maximum of the S/N with the off-axis angle and/or the object counts. Therefore, we deemed an extraction radius R that corresponds to 90% of the encircled energy to be a good choice for all the objects in our sample, consistent with what is already present in the literature (e.g., Marchesi et al., 2016).

---

[1] The photometric SED and redshift probability distributions can be found on the website: `http://j1030-field.oas.inaf.it/xray_redshift_J1030.html`

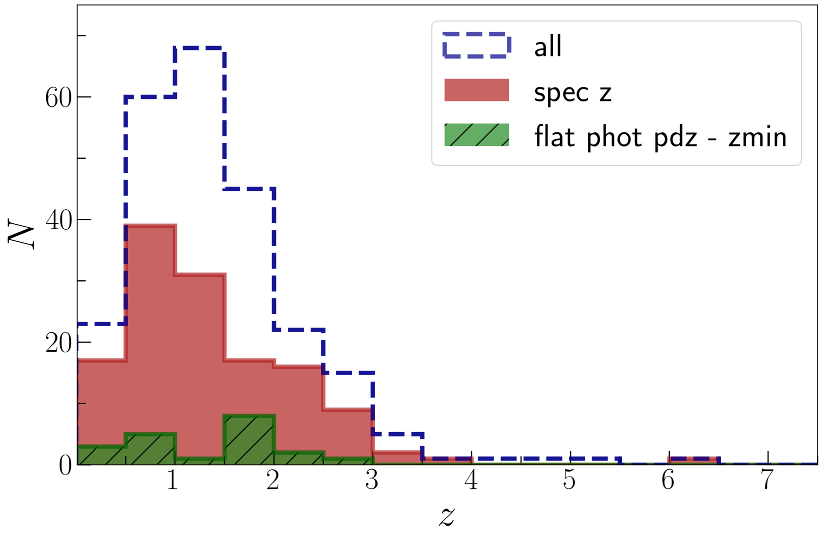

Figure 7.1: Redshift distribution of the J1030 *Chandra* catalog. The histograms show the distribution for the whole sample (blue dashed) and the distribution for the objects with spectroscopic redshift (red filled). A spectroscopic redshift is available for 135 out of 243 objects. Also shown is the redshift lower limit for the 20 objects for which the photometric redshift probability distribution is flat (green striped).

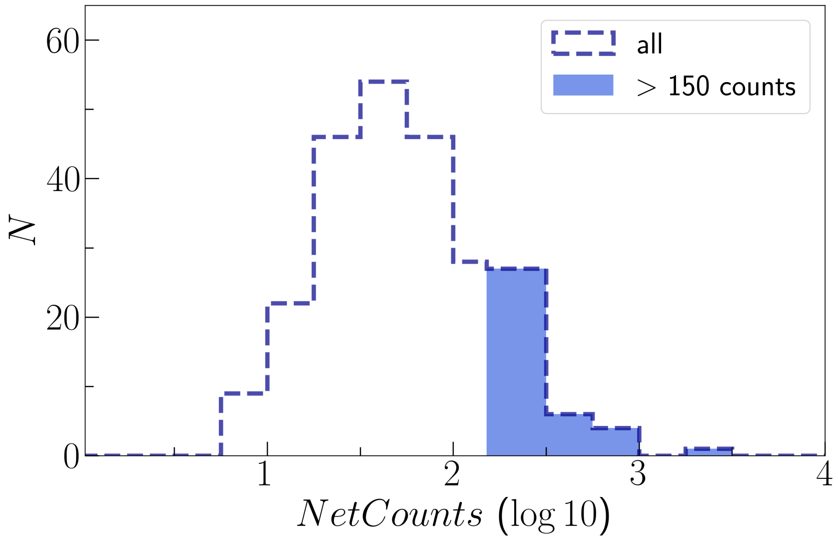

Figure 7.2: Count histogram for the J1030 field *Chandra* catalog. There are 39 out of 243 objects that have more than 150 net counts in the 0.5-7 keV range, shown as the light blue filled part of the histogram.

For each object, we also extracted a background spectrum in an annulus of radii R+2.5″ and R+20″, manually removing from the annulus any possible contaminating source. Given each source background, we chose not to model it locally as (i) we want to minimize dependencies on sharp background variations and (ii) the number of counts in background spectra extracted in small regions around each source can be very small and therefore the parameters uncertainties very high. We therefore want to characterize the whole background and then use the fitted background model to simultaneously fit each source with its local background, adding a background normalization parameter to the source fit to rescale the normalization to that of the local background.
We selected three regions to do that, all centered on the center of the field: one circular region of 3 arcmin radius, one annulus with 3 and 6 arcmin radii and an annulus with 6 and 9 arcmin radii. In each region we excluded circular regions of 5 arcsec radius around all the X-ray detected sources. We extracted the spectrum of each region in the energy range 0.8-7 keV, which is the one we use to fit the sources. We modeled each background spectrum with a power-law and four Gaussian lines, following the model used by Fiore et al. (2012) for the *Chandra* Deep Field South survey. Following the same model, we also tried to (i) add a second power-law component, and (ii) add a thermal component, but both turned out to be non-significant. We therefore excluded these components from the final background shape, which ends up being composed of a power law and four Gaussians.

In Figure 7.3 we can see the spectra and the resulting best-fit. Given the best fit parameters of the modeled background, we used them as ”frozen” parameters in the source+background fit analysis, only adding a multiplicative constant as a free parameter to re-scale the background spectrum to that of each object.

The selected background regions provided a sufficient sample for background estimation, allowing the spectral fitting analysis to proceed. For each object, we used the CIAO command *specextract* to extract the source and the background spectrum and to build the response matrix (RMF) and the ancillary response files (ARFs). This was done for each of the ten observations and the results were then combined with the CIAO tool *combine_spectra*. To avoid empty channels, the resulting spectra were binned to a minimum of one count per bin. In the end, for each object, we produced the combined source spectrum, the combined background spectrum, and the combined RMF and ARF files.

7.2 Spectral analysis

Once the spectra and the ancillary files were derived, we fitted them using *sherpa* (Freeman et al., 2001), fitting the background together with the source.

The source spectral shape is modeled with an absorbed power law. The Galactic absorption ($N_{\mathrm{H,Gal}} = 2.5\times10^{20}\,\mathrm{cm}^{-2}$) and the redshift are fixed parameters, while the column density at the source redshift, N_{H}, is always assumed to be a free parameter. In principle, the power-law photon index Γ should also be left free to vary. However, given the well-known degeneracy between Γ and $\mathrm{N_H}$, in low-statistic spectra, a fit with both parameters free to vary can lead to unreliable results. For this reason, we decided to fix the photon index Γ and leave the column density N_{H} as the only free parameter in sources

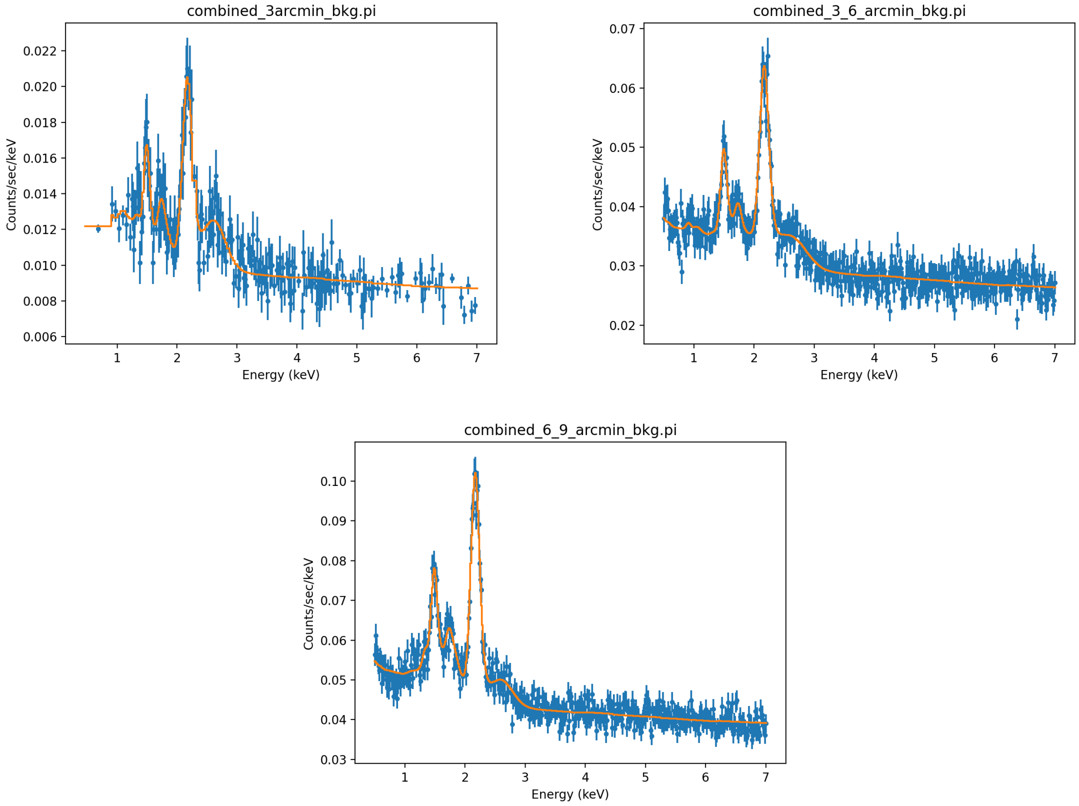

Figure 7.3: Spectrum (blue) and best fit (yellow) of the background spectra for different regions of the field: (a) 3’ circle; (b) annulus of radii 3’ and 6’; (c) annulus of radii 6’ and 9’

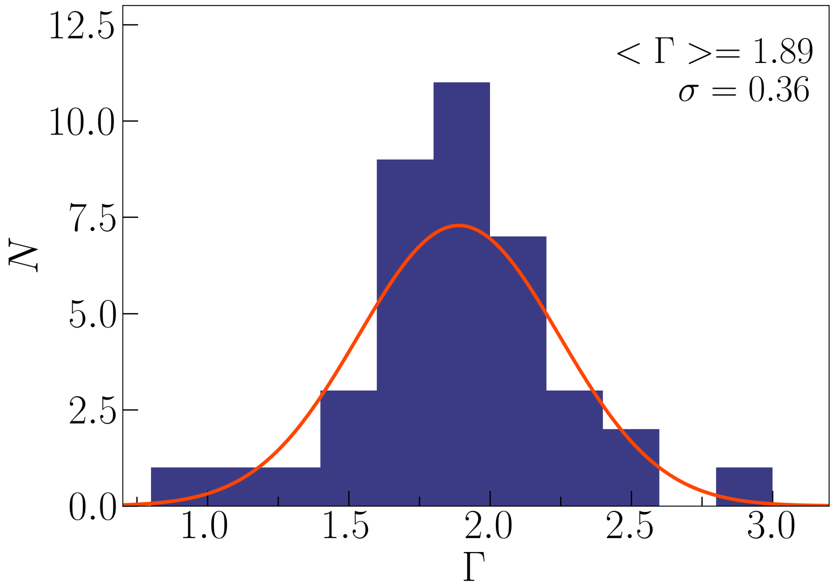

Figure 7.4: Photon index distribution for the 39 objects with more than 150 net counts. A Gaussian fit gives $\langle\Gamma\rangle = 1.89, \sigma = 0.36$.

with 0.5–7 keV net counts below a given threshold. After some tests, we decided to put the threshold at 150 net counts in the 0.5–7 keV band. In Figure 7.2 we show the net counts distribution (in the 0.5-7 keV band) for the objects in the catalog. Given the Poissonian nature of the data, we used the C statistic to perform the fit (Cash, 1979).

We first performed the fit procedure for the 39 objects with more than 150 counts. In Figure 7.4 we show the resulting photon index distribution; when fitted as a Gaussian, we found $\langle\Gamma\rangle = 1.89$ and a standard deviation $\sigma_\Gamma = 0.36$. So we assume $\Gamma = 1.9$ as a fixed parameter in fitting the objects with less than 150 counts. This value is also consistent with average values of the photon index Γ found in the literature (Mainieri et al., 2007; Lanzuisi et al., 2013; Marchesi et al., 2016; Liu et al., 2017). The free parameters of the fit are then three (the power-law slope, the normalization, and the column density) when the objects have more than 150 counts, and two (the power-law normalization and the column density) otherwise. We derived 90% uncertainties on the fitted parameters with *sherpa get_conf*(). In Fig. 7.5 we show four representative X-ray spectra in our sample. For 131 objects out of 243 the fit procedure returned only upper limits for the column density N_H; for the others, we obtained N_H estimates with upper and lower bounds. We note that an absorbed power-law model may not be an accurate representation of the X-ray spectra of the most heavily obscured, Compton-thick AGN ($N_\mathrm{H} > 10^{24}\mathrm{cm}^{-2}$), where reflection components may dominate over the transmitted ones (Comastri et al., 2010; Marchesi et al., 2018). Nonetheless, the primary objective of this study is to determine the fraction of obscured AGN using absorption thresholds of 10^{22} or $10^{23}\mathrm{cm}^{-2}$. In this regard, we contend that an absorbed power-law model is well suited for discerning whether the obscuration of an object exceeds the aforementioned thresholds.

The Fe $K\alpha$ line at 6.4 keV is a common feature of X-ray AGN spectra; the more obscured an object is, the more prominent this feature becomes, given the suppression of the primary continuum. Therefore, we expect to find it in a fraction of objects. To check for this presence, we performed the spectral fitting again, adding a new component to the source model, to search for the presence of a line at 6.4 keV. We applied one of two different strategies depending on whether the source has a spectroscopic redshift determination or a photometric one. For objects with spectroscopic redshifts, we performed a new fit with the same model as before but with the addition of a single Gaussian line with 0.05 keV width. We considered the presence of the line to be significant when compared to the statistic of the best-fitting simple absorbed power-law model; we obtain $\Delta C > 2.7$ as we are adding one free parameter to the fit, the line normalization. This corresponds to a 90% confidence level for one parameter of interest (see, e.g., Avni, 1976; Tozzi et al., 2006; Brightman et al., 2014). This happens for 9 objects out of 135: XID 2, XID 4, XID 8, XID 31, XID 44, XID 70, XID 73, XID 114, XID 115. For these objects, we also derived the rest-frame equivalent width of the Fe Kα line. The results are shown in Table 7.1.

XID	z	CNTS	N_H	Γ	EW (keV)
2	0.628	827	$1.6^{+0.6}_{-0.6}$	$1.9^{+0.2}_{-0.2}$	$0.14^{+0.09}_{-0.06}$
4	2.013	259	< 4	$1.9^{+0.4}_{-0.3}$	$0.19^{+0.12}_{-0.10}$
8	2.78	102	$31.53^{+15.91}_{-14.82}$	1.9	$0.21^{+0.12}_{-0.08}$
31	2.377	154	< 2.7	$2.0^{+0.3}_{-0.3}$	$0.20^{+0.17}_{-0.11}$
44	1.486	195	$3.2^{+3.2}_{-2.9}$	$2.4^{+0.5}_{-0.5}$	$0.24^{+0.42}_{-0.12}$
70	0.764	228	$4.1^{+2.4}_{-2.2}$	$1.7^{+0.5}_{-0.5}$	$0.21^{+0.21}_{-0.12}$
73	2.171	226	< 4.3	$2.0^{+0.4}_{-0.3}$	$0.24^{+0.12}_{-0.10}$
114	0.533	159	< 0.9	$2.2^{+0.6}_{-0.4}$	$0.74^{+0.50}_{-0.42}$
115	0.76	76	$7.7^{+4.0}_{-3.0}$	1.9	$0.41^{+0.24}_{-0.17}$

Table 7.1: Best-fit parameters for the nine objects with spectroscopic redshifts where a significant Fe Kα line is detected at 6.4 keV.
The counts are the net full counts; the photon index Γ is free to vary if there are more than 150 net counts of the source, while it is fixed if there are fewer than 150. The column density N_H is shown in units of $10^{22}\mathrm{cm}^{-2}$. The equivalent width of the emission line is shown in keV. Errors are at the 90% confidence level.

For objects for which we only have a photometric redshift estimate, the uncertainties on the redshift value are much bigger. Therefore, in searching for a significant Fe Kα line, we let the redshift of the model be a free parameter. We performed the fit with a single power-law model with the addition of a single Gaussian line with a fixed 6.4 keV energy and a fixed 0.05 keV width, imposing the line redshift to be the same as the absorbed power-law. In this case, there are two additional parameters to the fit, which are the redshift and the line normalization. Therefore, we consider the presence of the emission line significant if the difference in the statistic is $\Delta C > 5.4$. We found this to be true for 7 objects: XID 46, XID 137, XID 167, XID 193, XID 200, XID 205, and XID 345, whose properties are shown in Table 7.2. From this fit, we derive a redshift estimate, which in all cases is consistent with the photometric one, but provides a much smaller uncertainty. The average uncertainty on the redshift estimate for these objects goes from 0.94 in the photometric case to 0.07. We

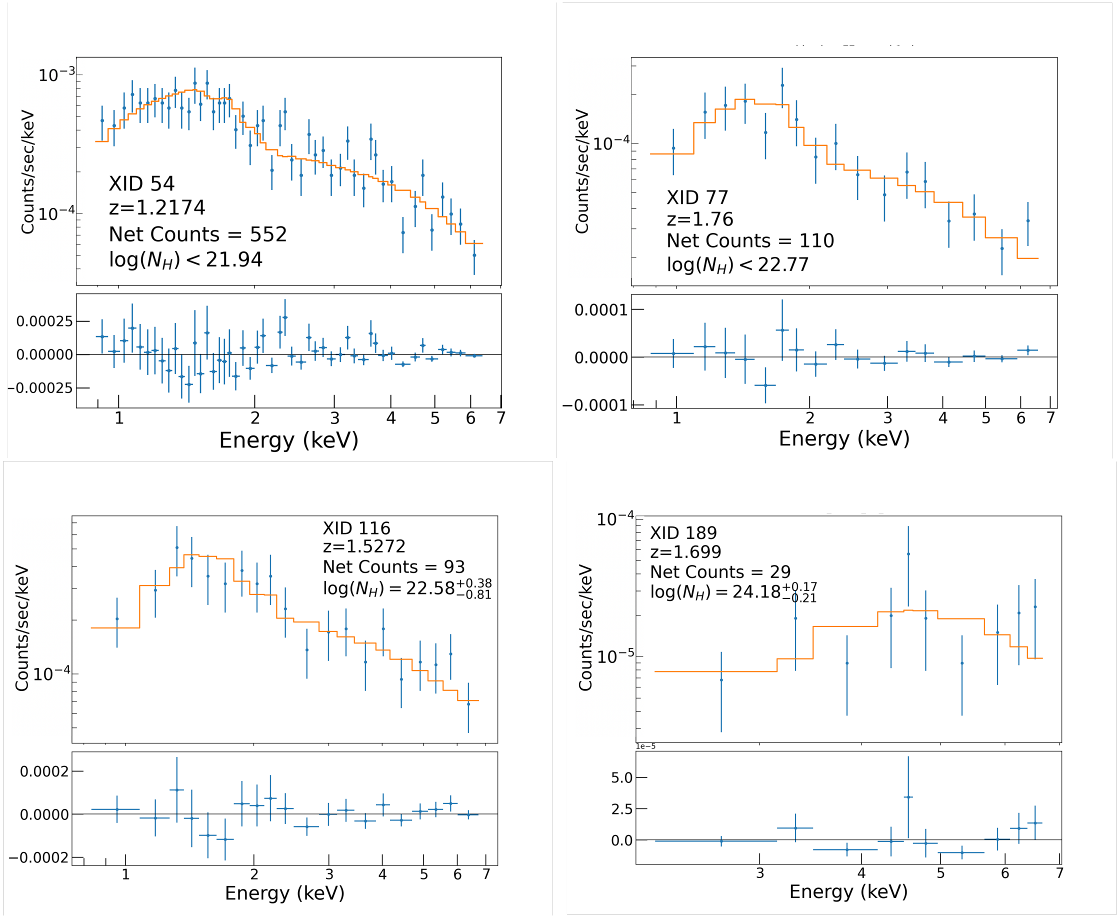

Figure 7.5: X-ray spectra in the 0.5-7 keV energy range (blue points) and best-fit models (orange solid lines) for four representative objects in the sample. In the lower panel, residuals are shown. The obscuration levels range from unobscured to heavily obscured. The lower panels show the residuals (data-model) of the fit.

XID	z_{phot}	CNTS	$N_{\rm H}$	Γ	EW	z_{line}
46	$1.64^{+1.08}_{-0.44}$	148	$3.0^{+3.1}_{-2.7}$	1.9	$0.38^{+0.22}_{-0.17}$	$1.45^{+0.05}_{-0.05}$
137	$1.98^{+3.2}_{-0.48}$	55	$10.3^{+11.3}_{-8.3}$	1.9	$0.6^{+0.3}_{-0.2}$	$2.23^{+0.08}_{-0.07}$
167	$2.94^{+0.72}_{-0.72}$	22	< 237	1.9	$0.92^{+2.63}_{-0.50}$	$3.32^{+0.13}_{-0.12}$
193	$0.61^{+0.13}_{-0.15}$	35	$5.5^{+4.0}_{-2.7}$	1.9	$0.54^{+0.44}_{-0.33}$	$0.68^{+0.04}_{-0.06}$
200	$1.49^{+1.57}_{-0.21}$	112	< 2.1	1.9	$0.3^{+0.1}_{-0.1}$	$1.43^{+0.07}_{-0.06}$
205	$2.34^{+2.44}_{-1.78}$	38	104^{+254}_{-61}	1.9	$12.8^{+12.2}_{-7.4}$	$2.83^{+0.15}_{-0.10}$
345	$0.44^{+0.18}_{-0.01}$	180	< 9	$0.9^{+2.1}_{-0.9}$	$7.2^{+14.0}_{-5.7}$	$0.62^{+0.03}_{-0.02}$

Table 7.2: Same as Table 1, with the addition of the best-fitting line redshift (z_{line}) and associated 90% uncertainties.

note that for the object XID 205, which has a photometric redshift estimate of $z = 2.34^{+2.44}_{-1.78}$, we get an X-ray redshift estimate of $z = 2.82^{+0.15}_{-0.11}$, which is consistent with the redshift of the large-scale structure discovered in the field, $z = 2.78$ (Marchesi et al., 2023).

Often a double power-law component is needed to fit AGN X-ray spectra, to model scattered emission which is typically found in obscured sources (Ueda et al., 2007). To test for the presence of this component, we assumed a phenomenological model and we performed again the fit adding a power-law component with the same photon index as the main one, with no absorption and with a multiplicative constant whose maximum value was fixed at 0.3. Therefore, we only have one additional parameter, the multiplicative constant. We looked for objects for which $\Delta C > 2.7$ but we found none. This differs from the results in previous studies, where at least a few percent of objects are usually found to have a significant double power-law component (e.g., Marchesi et al., 2016).

This might be caused by the decrease in the effective area of the Chandra telescope at energies below ~1.5 keV, mostly caused by the deposition of materials on the *Advanced CCD Imager Spectrometer* (ACIS) detector.

7.2.1 Column density probability distributions

From the spectral fit, we obtain for each object a $N_{\rm H}$ estimate; for 131 out of 243 objects, the estimate is an upper limit for the $N_{\rm H}$ value, while for the others we have a best-fit $N_{\rm H}$ value with upper and lower bounds. We can better understand the $N_{\rm H}$ estimates by deriving the $N_{\rm H}$ probability distributions for the objects in our sample. We used the *sherpa* command *int_proj* to compute the fit statistic C as the $N_{\rm H}$ parameter is varied from 10^{19} to $10^{26}{\rm cm}^{-2}$, using a logarithmic step of $\Delta\log(N_H) = 0.07$. Given the statistic values, we derived the probability distribution $p(\log(N_{\rm H}) \propto exp(-C/2))$ and normalized its integral to one. In Figure 7.6 we show, as an example, the $N_{\rm H}$ probability distributions of the objects shown in Fig. 7.5. For XID 54 and XID 77, the fitting procedure returns an upper limit for the $N_{\rm H}$ estimate. However, we can see that the probability distributions are very different: for XID 54, each $N_{\rm H}$ value below $\sim 10^{22}{\rm cm}^{-2}$ is more or less equally likely; instead for XID 77 and XID 116, there is a clear peak of the probability distribution around $\sim 3 \times 10^{22}{\rm cm}^{-2}$, although the fit was not able to retrieve a lower bound to the $N_{\rm H}$ estimate. This is true for

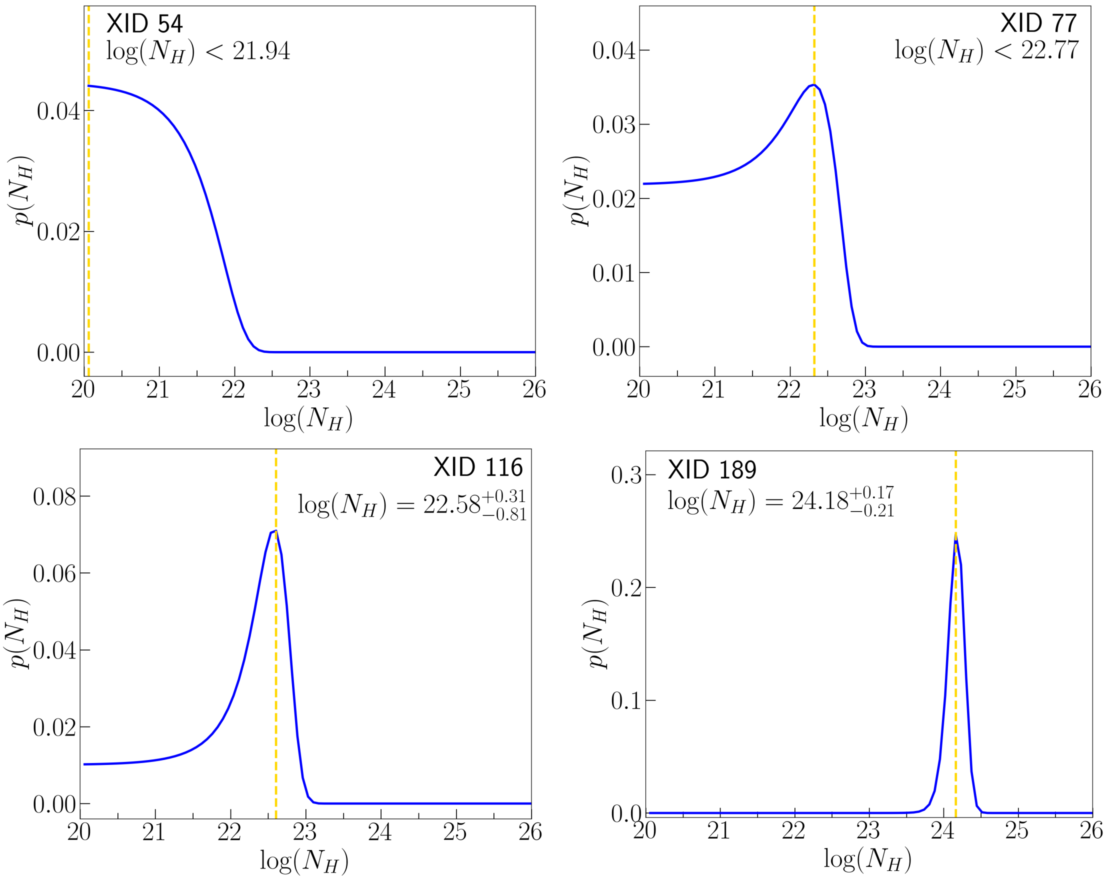

Figure 7.6: $N_{\rm H}$ probability distributions for the four objects shown in Fig 7.5. The yellow dashed lines show the values of NH at which the minimum of the fit statistic is found. For the two objects in the upper panels, we could only derive upper limits to the $N_{\rm H}$ measurements, whereas for the two objects in the lower panels, a significant (>90% c.l.) column density was measured.

more than half of the objects in the sample. The N_{H} probability distributions are in many cases asymmetric, with low N_{H} values having a higher probability even when the peak of the distribution is at high N_{H} values. Instead, XID 189[2] presents a case in which the best-fitting N_{H} is well constrained, with a 90% lower limit higher than zero. The fit in this case is indeed able to retrieve both an upper and a lower bound for N_{H}. As can be seen with these examples, the probability distribution is a more accurate way to describe the column density of a source, compared with the nominal N_{H} value that we obtain from the fit. We therefore chose to use the $p(\log(N_{\mathrm{H}}))$ to derive the obscured fractions, as we discuss in Section 7.3. The N_{H} probability distributions for all the objects in the sample are available on the project webpage[3].

7.2.2 Results

At the end of our spectral analysis, we derived the column density N_{H} for the 243 AGN in the J1030 *Chandra* field. The catalog with the basic physical properties derived from our analysis is available online;[4] in Table 7.3 we show a portion of it. For each object, we provide the column density, the photon index, the (de-absorbed) rest-frame 2-10 keV luminosity, and relative 90% uncertainties. In Figure 7.7 we show the global N_{H}–redshift distribution for the sample; the objects are shown with different symbols and colors depending on their spectral identification (Marchesi et al., 2021). A trend of N_{H} with redshift can be seen: objects at higher redshift have on average higher N_{H} values. This is partly due to a selection effect. When moving toward higher redshifts, the photoelectric absorption cutoff moves outside the limit of the observing band, and it is therefore more difficult to constrain lower N_{H} values (Civano et al., 2005; Lanzuisi et al., 2013). A thorough analysis of the obscured fraction trend with redshift that takes this factor into account is provided in the next section.

It should be noted that the column densities we obtain for objects for which we have a classification from the optical spectrum are consistent with the optical classifications themselves: BL-AGN (in blue) have low column densities, and for 90% of them the spectral fit can only obtain an upper limit for N_{H}, while NL-AGN (in red) have higher average column densities and the fraction for which we get upper limit for N_{H} is 40%. This fraction is 51% for ELGs and 52% for ETGs. The sources for which we obtain the higher N_{H} values are more likely to be those without an optical spectral classification (in gray), which is consistent with them being obscured and therefore not easily observed in the UV-optical. In Figure 7.8 we show the intrinsic (i.e de-absorbed) rest-frame 2-10 keV luminosity versus redshift, with the same classification code.

7.3 Obscured fraction

Our goal here is to investigate the dependence of the column density N_{H} on redshift and luminosity. We have to consider that at different redshifts we sample different average

---

[2] Which, we note, is also the central object of the protocluster described in Gilli et al. (2019)

[3] `http://j1030-field.oas.inaf.it/xray_redshift_J1030.html`

[4] `http://j1030-field.oas.inaf.it/chandra_1030`

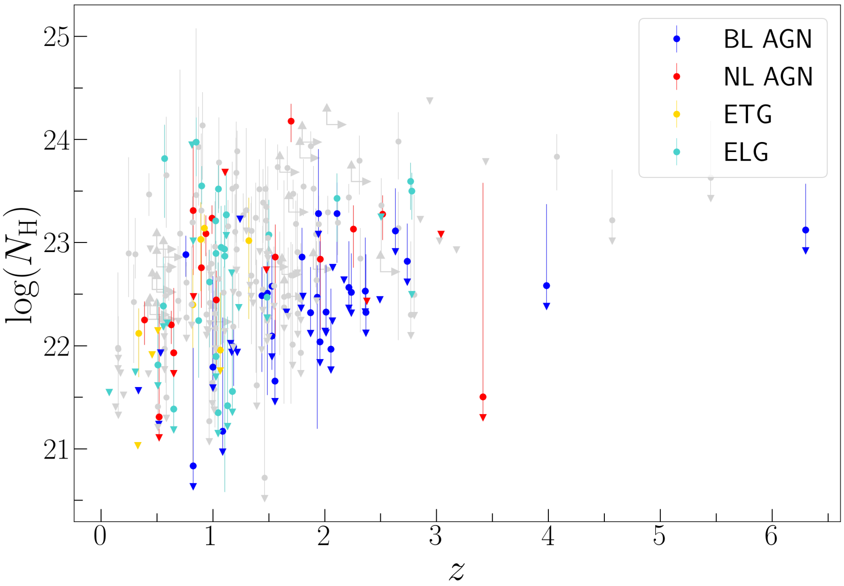

Figure 7.7: N_{H}–redshift distribution of all the objects in the catalog. Up- and right-pointing arrows show the redshift and N_{H} lower limits for objects with a flat photometric redshift probability curve. Upper limits are shown as down-pointing triangles. The spectral types are color-coded: red, NL-AGN, blue, BL-AGN, yellow, Early Type Galaxies, aquamarine, Emission Line Galaxies, gray, no spectral identification

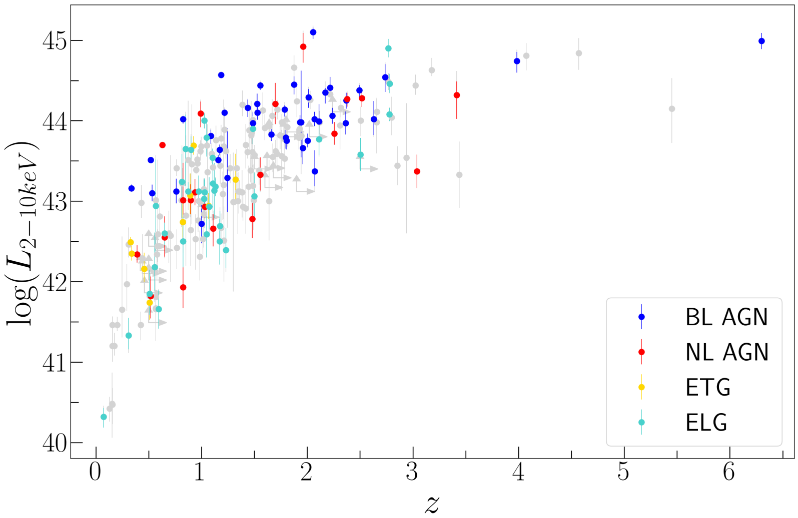

Figure 7.8: Distribution of the intrinsic (de-absorbed) rest frame 2-10 keV luminosity for the 243 objects in the catalog. Up- and right-pointing triangles show the redshift and luminosity lower limits for objects with a flat photometric redshift probability curve. The color-coding for the spectral types is the same as in Figure 7.7.

XID	z	CNTS	$\log(N_H)$	Γ	$\log(L_{2-10keV})$	cstat/dof
			cm^{-2}		erg/s	
1	$3.18^{+1.3}_{-1.3}$	252^{+17}_{-16}	< 22.93	$1.52^{+0.33}_{-0.23}$	$44.63^{+0.15}_{-0.08}$	144.1/172
2	0.6279	827^{+30}_{-29}	$22.20^{+0.14}_{-0.19}$	$1.93^{+0.22}_{-0.22}$	$43.70^{+0.04}_{-0.04}$	288.8/296
3	1.0974	164^{+14}_{-13}	$22.52^{+0.25}_{-0.49}$	$2.12^{+0.53}_{-0.48}$	$43.59^{+0.13}_{-0.11}$	99.6/134
4	2.0133	259^{+17}_{-16}	< 22.79	$1.85^{+0.36}_{-0.32}$	$42.29^{+0.13}_{-0.10}$	127.5/171
5	0.9679	37^{+7}_{-6}	< 22.40	1.9	$42.76^{+0.18}_{-0.15}$	265.7/289
6	0.5181	993^{+33}_{-32}	< 21.24	$1.903^{+0.12}_{-0.10}$	$43.51^{+0.03}_{-0.03}$	59.79/72

Table 7.3: *Chandra* J1030 spectral catalog. For each object, we provide the redshift (which is derived from spectroscopy when provided without uncertainties, from photometry otherwise), the (0.5-7 keV) counts (see Nanni et al., 2018) the logarithm of the column density, the photon index (which is fixed to 1.9 when there are fewer than 150 counts), the intrinsic rest-frame 2-10 keV luminosity and the value of the C-statistic over degrees of freedom. Six objects also have a redshift estimate derived from the presence of the Fe Kα line (see Table 7.2). This table is available in its entirety in machine-readable form on the website.

luminosities. In the literature there is evidence of the obscuring fraction being a function of both redshift and luminosity (see, e.g., Ueda et al., 2014; Aird et al., 2015; Ananna et al., 2019). Therefore, we need to perform our analysis at a fixed luminosity to derive the evolution of $N_{\rm H}$ with redshift, and at a fixed redshift to derive the $N_{\rm H}$ dependence on luminosity. We considered the intrinsic luminosity–redshift plane, which can be seen in Figure 7.9, and we selected only the objects that have a hard band detection, which is 203 out of 243, to get a uniform selection function and to apply reliable correction to go from observed to intrinsic obscured fractions (see Section 7.3.1).

For the luminosity-dependence analysis, we selected the subsample shown in blue, where the average redshift is ~ 1.2 in each bin. This subsample can be divided into three luminosity bins, with $42.8 < \log(L_{2-10keV}) < 43.3$, $43.3 < \log(L_{2-10keV}) < 43.8$, and $43.8 < \log(L_{2-10keV}) < 44.5$, respectively. In each bin, we have 38, 32, and 18 objects, respectively. Of these objects, the ones with a spectroscopic redshift estimate are 19 out of 38 in the first bin, 11 out of 32 in the second bin, and 11 out of 18 in the third bin.
For the redshift dependence analysis, we selected a subsample of objects, shown in green, with an average luminosity of 10^{44} erg/s. This subsample is then divided into three subsamples with redshift $0.8 < z < 1.6$, $1.6 < z < 2.2$, and $2.2 < z < 2.8$. In each redshift bin, the average luminosity is ~ 10^{44} erg/s, and we have 18, 24, and 20 objects per bin, respectively. Out of these objects, the ones with a spectroscopic redshift estimate are 11 out of 18 in the first bin, 11 out of 24 in the second bin, and 14 out of 20 in the third bin. These bins were selected to maximize the source statistics, while keeping the best completeness in each bin. We note that for the few objects with a flat redshift probability distribution (4 out of a total of 60 objects in the five different bins), we used their best redshift estimate to determine whether they belong to a certain bin.

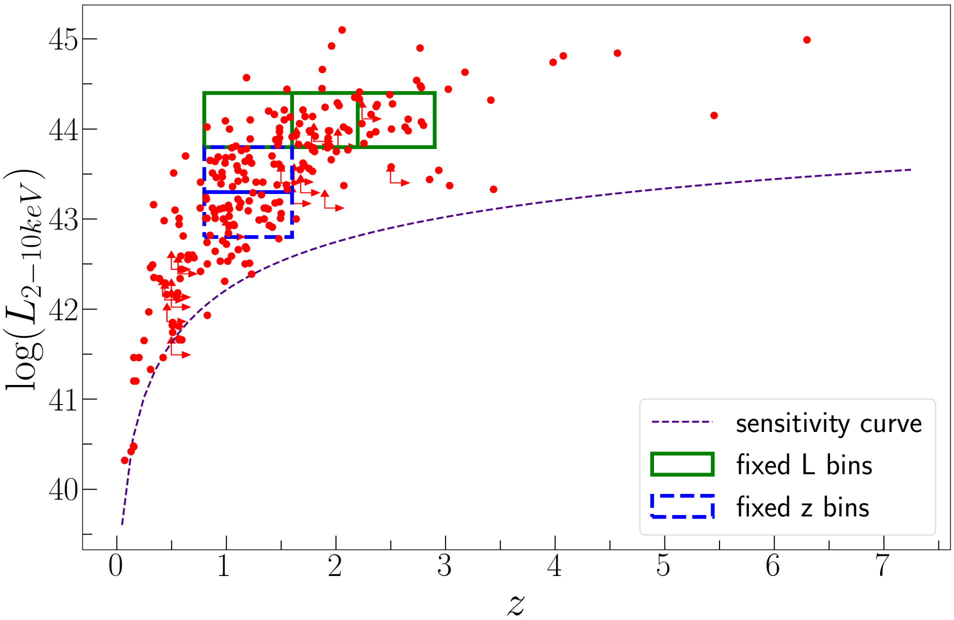

Figure 7.9: Intrinsic rest-frame 2-10 keV luminosity as a function of redshift for the 203 Chandra J1030 sources detected in the 2-7 keV band. Up- and right- pointing triangles show the redshift and luminosity lower limits for objects with a flat photometric redshift probability curve. The subsample used for the analysis of the $N_{\rm H}$–redshift evolution (Section 7.3.2) is in green; the subsample used for the analysis of the $N_{\rm H}$–luminosity evolution (Section 7.3.1) is in blue. The dashed purple line represents the survey sensitivity curve, at 50% of the field coverage (Nanni et al., 2020)

.

7.3.1 Obscured fraction dependence on 2–10 keV luminosity

We want to derive the obscured fraction f_{22} (the fraction of objects with a column density $N_H > 10^{22}\mathrm{cm}^{-2}$) and f_{23} (the fraction of objects with a column density $N_H > 10^{23}\mathrm{cm}^{-2}$). For each object, we could simply use the best-fit value of N_{H} as the N_{H} estimate. However, this does not take into account how likely it is for the true N_{H} value for a given object to be that of the nominal result of the fit. Furthermore, for objects with similar N_{H} values (or upper limits), the probability distributions can vary significantly from one object to another, as shown before.

Considering all of this, we derived the obscured fractions using the probability distribution functions described in Section 7.2. For each object, we considered the fraction of $p(\log(N_{\mathrm{H}}))$ at N_{H} values higher than $10^{22}\mathrm{cm}^{-2}$ ($10^{23}\mathrm{cm}^{-2}$). We summed all the fractions for the objects in a given luminosity bin and got an estimate of the number of obscured sources that correctly takes into account the probability distribution functions. By dividing this number by the total number of objects in the bin, we obtain the observed obscured fraction. We performed this for the two different obscuration thresholds ($10^{22}\mathrm{cm}^{-2}$ and $10^{23}\mathrm{cm}^{-2}$) and for each luminosity bin.

In Figure 7.10 we show the comparison between the results obtained via this procedure, which uses the $p(\log(N_{\mathrm{H}}))$, and using the nominal values of N_{H}. It can be seen that using the $p(\log(N_{\mathrm{H}}))$ we get obscured fractions f that are systematically lower than the others (even by $\Delta f \sim 0.18$). This is expected, as most N_{H} probability distributions are skewed toward lower N_{H} values. Therefore, there are objects for which the nominal N_{H} value can be higher than the obscuration threshold, but that does not overall contribute much to the obscured population in terms of its probability distribution. The asymmetry of the probability distributions mainly depends on the lack of information at soft X-ray energies. Because of this, it is often possible in the fitting procedure to get a high obscuration level excluded, but it is not possible to distinguish between a non-obscured and a mildly obscured object.

Table 7.4: Number of objects, average redshift, and fraction of AGN with $\log(N_{\mathrm{H}}) > 22$ (f_{22}) and $\log(N_{\mathrm{H}}) > 23$ (f_{23}) in three luminosity bin and relative uncertainties

Bin	N	$\bar{z}$	f_{22}	f_{23}
$42.8 < \log(L) < 43.3$	38	1.15	0.80 ± 0.11	0.65 ± 0.11
$43.3 < \log(L) < 43.8$	32	1.16	0.80 ± 0.10	0.78 ± 0.10
$43.8 < \log(L) < 44.5$	18	1.35	0.78 ± 0.08	0.39 ± 0.08

For the uncertainties on these obscured fractions, we know that confidence intervals on sample proportions are usually derived using the binomial distribution. We can, for example, use the Wilson score interval (Wilson, 1927) to derive confidence intervals, which will depend, in each bin, on the number of objects in the bin and on the obscured fraction derived using the probability distributions. When doing so for the three luminosity bins,

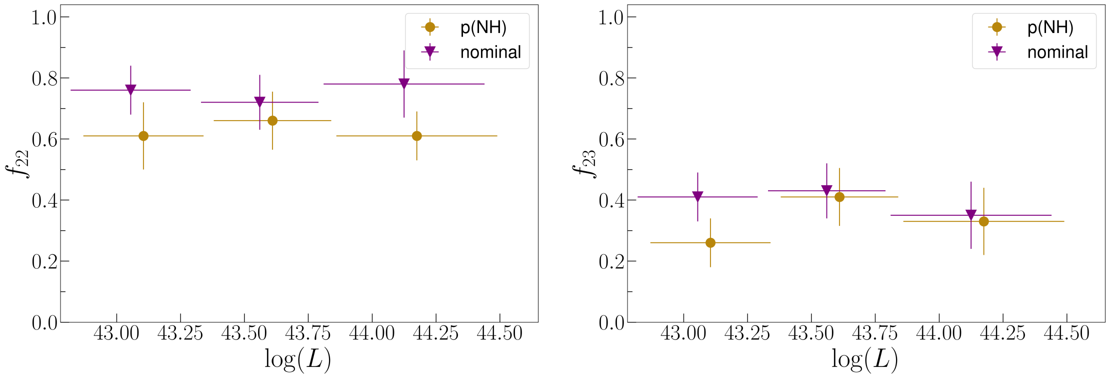

Figure 7.10: Fraction of obscured $z \sim 1.2$ AGN as a function of intrinsic 2-10 keV luminosity. The purple triangles show the observed obscured fractions derived using the nominal $N_{\rm H}$ value derived from the fit. The gold circles are the values obtained using the probability distribution of $N_{\rm H}$. The second set of points is shifted by 0.05 on the log(L) axis for visual clarity. The 1-σ uncertainties, derived with the bootstrapping procedure, are shown. This figure highlights the relevance of using the probability distributions in deriving the obscured fraction of AGN. These results are still not corrected for the survey sky coverage (for those, see Figure 7.12). *Left*: Obscured fraction derived using $N_{\rm H} > 10^{22}{\rm cm}^{-2}$ as the threshold (f_{22}). *Right*: Obscured fraction derived using $N_{\rm H} > 10^{23}{\rm cm}^{-2}$ as the threshold (f_{23}).

we get lower uncertainties around ~0.5 and upper uncertainties around ~0.10. However, this method takes into account the uncertainties related to the finiteness of the sample only and does not consider that the $N_{\rm H}$ estimates are not exact. To deal with this, we derived the uncertainties with a bootstrap procedure: for each bin, we randomly extract, from the bin, with re-entry, a number of objects equal to the bin size. Then, for each object, we extract a value for $N_{\rm H}$ from its probability distribution. We then compute the obscured fraction as the number of objects with $N_{\rm H} > 10^{22}{\rm cm}^{-2}$ over the total. We repeat this 10000 times and we obtain a f_{22} (or f_{23}) distribution, from which we extract the peak and the 16% and 84% quantiles as the values for f_{22} (or f_{23}) and the corresponding uncertainties. In this way, both the finiteness of the bin and the uncertainties on each $N_{\rm H}$ estimate are taken into account.

We now must consider that our survey is flux-limited. This means that we are likely to miss preferentially obscured (i.e. fainter) objects rather than unobscured ones. Therefore, the obscured fractions that we derive are only lower limits to the intrinsic obscured fraction, and the true value is higher. We need to correct the obtained values for the number of objects that we are not observing (called the Malmquist bias). To do so, we proceeded in the following way for each luminosity bin and for each obscuration threshold (10^{22} and 10^{23} cm^{-2}): we considered the *intrinsic* number of obscured and unobscured sources in a given redshift and luminosity range (N_O^{int} and N_U^{int}, respectively) expected in the population synthesis model of the cosmic X-ray background (XRB) of Gilli et al. (2007). To derive them, we used the online tool [5] to compute the surface density, or integral number counts,

---

[5] http://www.bo.astro.it/~gilli/counts.html

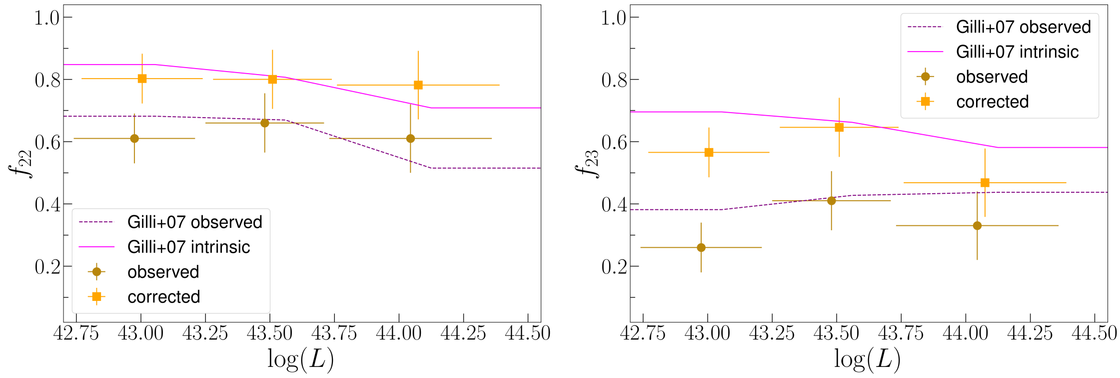

Figure 7.11: Fraction of $z \sim 1.2$ obscured AGN with as a function of intrinsic 2-10 keV luminosity. The gold circles are the observed fraction, while the orange squares are the values obtained once corrected for the survey sky coverage. The second set of points is shifted by 0.05 on the log(L) axis for visual clarity. The solid magenta line represents the predictions from the Gilli et al. (2007) model for the intrinsic obscured fraction; the dashed purple line shows the prediction for the observed fraction accounting for the survey sky coverage. *Left*: Obscured fraction derived using $N_{\rm H} > 10^{22}$cm^{-2} as the threshold (f_{22}). *Right*: Obscured fraction derived using $N_{\rm H} > 10^{23}$cm^{-2} as the threshold (f_{23}).

$N(> S)$, above any given 2-10 keV flux limit S of both obscured and unobscured AGN. The expected intrinsic number of obscured and unobscured AGN in J1030 N_O^{int}, N_U^{int} were obtained by multiplying the source surface density at $f_{2-10keV} = 10^{-20}$ cgs (i.e. at ≈ zero flux) by the geometric area of J1030. From the integral number counts, we then obtained the differential source counts dN/dS and folded them with the sky coverage A(S) of the J1030 survey (Nanni et al., 2020) as $\int dn/dS\,A(S)dS$. Because the sky coverage is given in the 2-7 keV flux range, we convert it to the 2-10 keV range by assuming a power-law spectrum with a photon index of 1.4, which is the average *observed* index for the AGN population.

In this way, we obtain N_O^{obs} and N_U^{obs}, which are the expected *observed* number of obscured and unobscured objects. We then derived the intrinsic and the observed ratios of the number of obscured objects to unobscured objects, $R_{int} = N_O^{int}/N_U^{int}$ and $R_{obs} = N_O^{obs}/N_U^{obs}$. As we lose more obscured objects than unobscured ones when in the presence of a flux limit, R_{int} will always be higher than R_{obs}. We can now derive $p = R_{obs}/R_{int}$ as the corrective parameter that we need to implement to go from our *observed* obscured fraction to the *intrinsic one*. This number is always smaller than one.

If we now define the observed obscured fraction(s) as $f_{22} = N[10^{22} - 10^{26}]/N[10^{20} - 10^{26}]$ and $f_{23} = N[10^{23} - 10^{26}]/N[10^{20} - 10^{26}]$, we can derive the corrected fractions as:

$$f_{22}^{corrected} = \frac{f_{22}}{f_{22}(1-p)+p} \tag{7.1}$$

and we can derive $f_{23}^{corrected}$ in the same way.

We did this for each luminosity bin, starting from the fractions derived with the $p(\log(N_{\rm H}))$, and the resulting corrected fractions are shown in Table 7.4. In Figure 7.11 we show the observed fractions, in dark gold, and the corrected fractions, in orange. We also

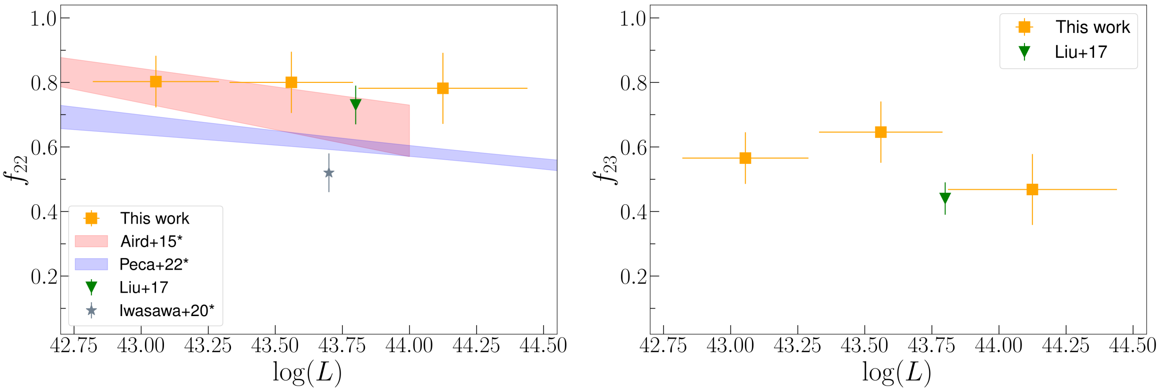

Figure 7.12: Fraction of $z \sim 1.2$ obscured AGN, corrected for completeness, as a function of intrinsic 2-10 keV luminosity. The orange squares show the results from this thesis. *Left*: Obscured fraction derived using $N_{\rm H} > 10^{22}\text{cm}^{-2}$ as the threshold (f_{22}). The results from this thesis are compared with those of Aird et al. (2015) (red shaded), Liu et al. (2017) (green triangle), Iwasawa et al. (2020) (gray star) and Peca et al. (2023) (blue shade). The Aird et al. (2015), Iwasawa et al. (2020) and Peca et al. (2023) obscured fraction consider column densities up to 10^{24}cm^{-2}. Aird et al. (2015) data is centered at z ~1; the Iwasawa et al. (2020) data is centered at z ~1.35. Given the different definitions of f_{22} and the redshift differences, some scatter among the results is expected. *Right*: Obscured fraction derived using $N_{\rm H} > 10^{23}\text{cm}^{-2}$ as the threshold (f_{23}). The results of this thesis are compared with those of Liu et al. (2017) (in green). The f_{23} obscured fraction at log(L) ~ 44.1 is in good agreement with that of Liu et al. (2017) at log(L) ~ 43.8.

show the magenta solid line, which is the predicted *intrinsic* obscured fraction of the Gilli et al. (2007) model, and the dashed purple line which is the *observed* obscured fraction, given the J1030 X-ray sky coverage. Overall, the high uncertainties make it hard to see a clear trend of the obscured fraction with the luminosity. We can compare our results with others in the literature, with the caveat of only considering those samples of objects with a redshift similar to our ($z \sim 1.2$). For the f_{22} fraction, we can compare it with the work of Aird et al. (2015) (considering the subsample of objects in that work that are found at redshift $z \sim 1$), Liu et al. (2017) (considering the objects found at redshift $z \sim 1.2$), Iwasawa et al. (2020), and Peca et al. (2023). For the f_{23} fraction we only have the Liu et al. (2017) data to compare with. These comparisons can be seen in Figure 7.12.

The f_{22} that we obtain at $\log(L) \sim 44$ are consistent with those of Aird et al. (2015) and Liu et al. (2017), while they are higher than those of Iwasawa et al. (2020) and Peca et al. (2023). Overall, the J1030 f_{22} does not show a significant decline with increasing luminosity as commonly found in the literature, but, given the large error bars, it cannot be ruled out either. For f_{23}, the estimate at $\log(L) \sim 44$ obtained in this thesis is consistent with the results of Liu et al. (2017), while we lack data at different luminosities for an additional comparison. We also note that our obscured fractions at $z \sim 1.2$ are on average higher than those measured by Aird et al. (2015), Iwasawa et al. (2020) and Peca et al. (2023), as expected: in these works the obscured fraction is derived as the number of objects with $10^{22}\mathrm{cm}^{-2} < N_H < 10^{24}\mathrm{cm}^{-2}$ over the number of objects with $10^{20}\mathrm{cm}^{-2} < N_H < 10^{24}\mathrm{cm}^{-2}$, while we considered the probability distributions of N_H from $10^{22}\mathrm{cm}^{-2}$ to $10^{26}\mathrm{cm}^{-2}$, that is, we included a correction for an additional population of C-thick AGN.

7.3.2 Obscured fraction dependence on with redshift

To investigate the redshift evolution of the obscured fraction, we performed the same analysis as described in Section 7.3.1, but for the three redshift bins with the same average luminosity of $\log(L) \sim 44$ (see Fig 7.9). We used the bootstrap procedure to derive, for each bin, the f_{22} and f_{23} and the corresponding uncertainties.

Table 7.5: Number of objects, average 2-10 keV luminosity, and fraction of AGN with $\log(N_\mathrm{H}) > 22$ (f_{22}) and $\log(N_\mathrm{H}) > 23$ (f_{23}), corrected for the completeness of the survey, in three redshift bin and relative uncertainties

Bin	N	$\overline{\log(L)}$	f_{22}	f_{23}
$0.8 < z < 1.6$	18	44.03	0.78 ± 0.11	0.47 ± 0.11
$1.6 < z < 2.2$	23	43.96	0.76 ± 0.11	0.61 ± 0.09
$2.2 < z < 3.2$	18	44.15	0.74 ± 0.11	0.51 ± 0.11

We then corrected the observed obscured fractions and recovered the intrinsic ones in each redshift bin using the same correction method described in Section 4.1. In Table 7.5

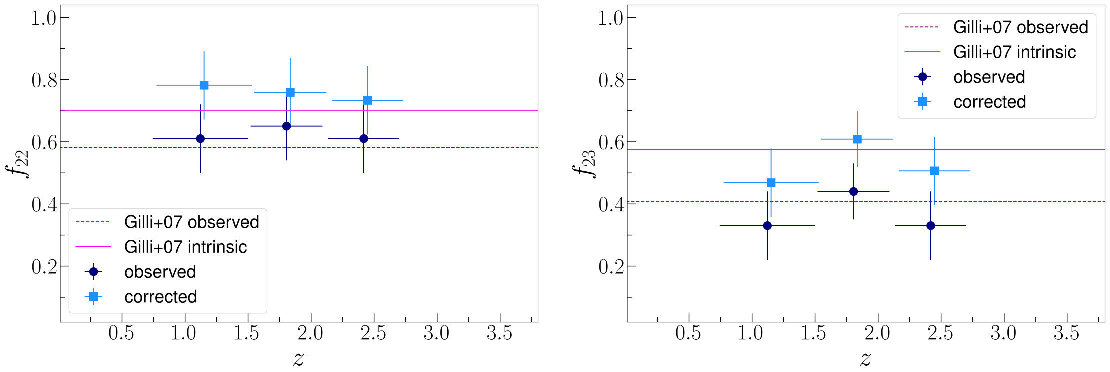

Figure 7.13: Fractions of $\log(L_{2-10}) \sim 44$ obscured AGN as a function of redshift. The navy circles are the observed fractions, while the light blue squares are those obtained once corrected for the presence of the survey sky coverage. The second set of points is shifted by 0.05 on the z-axis for visual clarity. The solid magenta line represents the predictions from the Gilli et al. (2007) model for the intrinsic obscured fraction; the dashed purple line is the prediction for the observed fraction once the sky coverage is taken into account. *Left:* Obscured fraction derived using $N_{\rm H} > 10^{22}{\rm cm}^{-2}$ as the threshold (f_{22}). *Right*: Obscured fraction derived using $N_{\rm H} > 10^{23}{\rm cm}^{-2}$ as the threshold (f_{23}).

we show the results of the corrected obscured fractions and their uncertainties in the three redshift bins. The results are shown in Figure 7.13. As in Section 7.3.1, we note that, as expected, the corrected fractions are higher than the observed ones, because the presence of a flux limit preferentially removes obscured sources from the sample.

We can now compare our results with those of other works. It is important to note that we should compare our obscured fractions with others obtained from samples with similar average luminosity. In Figure 7.14 we show our results (in light blue) together with those of Burlon et al. (2011), Aird et al. (2015), Liu et al. (2017), Vito et al. (2018), Iwasawa et al. (2020), and Peca et al. (2023), which are a representative sample of the trends in the literature. The obscured fractions in Liu et al. (2017) were obtained in redshift ranges similar to those used in this study; the results are very consistent for the first two redshift bins, while they are more distant for the higher redshift points. The number of objects per bin in Liu et al. (2017) is roughly the same as in the J1030 sample; our uncertainties of the obscured fraction estimates are significantly higher, given the lower quality of the data and given that we took both the binomial error and the $N_{\rm H}$ uncertainties into account.

It should be noted that that Aird et al. (2015), Iwasawa et al. (2020), and Peca et al. (2023) obscured fractions consider column densities up to $10^{24}{\rm cm}^{-2}$. Given the different definitions of f_{22} and the redshift differences, some discrepancy between the results is expected. When also considering the low-redshift results of Burlon et al. (2011), and the high-redshift Vito et al. (2018) estimate for f_{23}, there is evidence of a clear redshift trend, with bins at higher redshift showing a higher obscured fraction, both for f_{22} and f_{23}. Overall, our results are consistent with those in the literature.

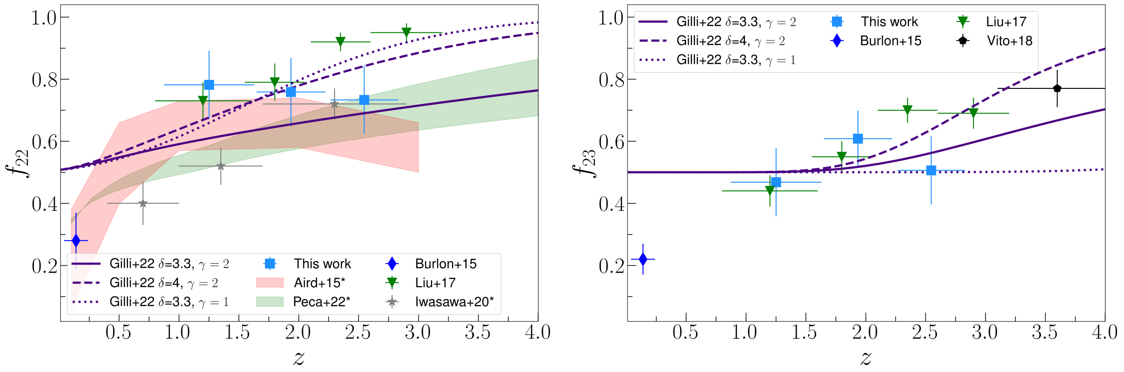

Figure 7.14: Fraction of log(L) ~ 44 obscured AGN with $N_H > 10^{22}\text{cm}^{-2}$ (f_{22}) as a function of redshift. The light blue squares show the results of this thesis. The prediction of the Gilli et al. (2022) model for the evolution of the obscured fraction with the redshift is shown as the indigo lines, with different styles representing the different parameters of the model (see Section 7.4.2). *Left*: Obscured fraction derived using $N_H > 10^{22}\text{cm}^{-2}$ as the threshold (f_{22}). The results from this thesis are compared with those of Burlon et al. (2011) (blue diamond), Liu et al. (2017) (green triangle), Aird et al. (2015) (red shaded), Iwasawa et al. (2020) (gray star), and Peca et al. (2023) (green shaded). The Aird et al. (2015), Liu et al. (2017), Iwasawa et al. (2020), and Peca et al. (2023) obscured fraction consider column densities up to 10^{24}cm^{-2}. *Right:* Obscured fraction derived using $N_H > 10^{23}\text{cm}^{-2}$ as the threshold (f_{23}). The results from work are compared with those of Burlon et al. (2011) (blue diamond), Liu et al. (2017) (green triangle), and Vito et al. (2018) (black pentagon). The Vito et al. (2018) obscured fraction consider column densities up to 10^{25}cm^{-2}.

7.4 Discussion

In this section, we discuss the results of our analysis and their interpretation, as well as possible limitations and biases.

7.4.1 Limitations and biases

The first limitation affecting our work is related to the sample statistics. Although the J1030 *Chandra* survey is one of the deepest X-ray surveys, about 30% of the objects have less than 30 net counts, which affects our ability to derive accurate parameters from the spectral fit. Furthermore, the progressive degradation of the *Chandra* detectors affects the soft X-ray response in a non-negligible way. As the spectral shape at low energies is more informative and allows us to distinguish between different levels of obscuration, our ability to derive reliable N_{H} estimates is reduced. The low sensitivity at low energies also significantly skews the N_{H} probability distributions toward low N_{H} values, even when a significant probability peak around a certain value is found (see e.g., XID 77 and XID 116 in Fig.4). These uncertainties clearly affect the accuracy with which we can estimate the obscured AGN fractions, compared with surveys with longer exposures and with surveys where observations have been carried out in earlier years of the *Chandra* satellite life.

Another source of uncertainty comes from the fact that 44% of the objects in our sample only have a photometric redshift estimate. In the fit procedure, we considered the redshift as a fixed parameter. However, the errors on the photometric redshifts can be significant. This again affects the accuracy of the N_{H} estimates. Furthermore, in the obscured AGN fraction analysis, some objects that fall in a given luminosity–redshift bin might actually belong to other, adjacent bins. Observational campaigns aimed at improving the spectroscopic redshift completeness of the J1030 *Chandra* sample are being planned.

We measured the obscured AGN fractions in different bins of X-ray luminosity and redshift. The main source of errors on these fractions is related to (i) the limited sample statistics in each luminosity–redshift bin, and (ii) the uncertainties in the column density estimate of each source. Given that the uncertainties we derive from the Wilson score values are in the ~0.05-0.10 range, compared to total uncertainties derived from the bootstrapping procedure of ~ 0.11, we can say that the first contribution is generally more significant than the second. When compared with the results obtained from other surveys, our uncertainties are significantly higher. This depends on the fact that, in general, previous studies do not take *both* sources of uncertainties into account, on the lower data quality of our X-ray spectra when compared with other samples (e.g., Liu et al. 2017), and on the higher statistics of other surveys.

The uncertainties in our obscured fraction estimates are such that we do not have clear evidence of a redshift or a luminosity trend with the J1030 data alone (see Figures 7.12 and 7.13). At the same time, as shown in Fig. 7.14, our results are generally consistent with those in the literature for AGN with similar luminosities and at similar redshifts. Furthermore, when combined with samples of X-ray selected AGN covering a broader range of redshifts, our results follow the general literature trends, where the obscured AGN fraction increases toward higher redshifts.

Another bias that might be affecting our results is the *classification bias*. When an object has a small number of counts, heavily obscured objects can be misclassified as mildly obscured ones (for more details, see Brightman & Ueda, 2012; Lanzuisi et al., 2018). The low-luminosity objects are the ones with a smaller number of counts, and therefore the most affected by this bias. In terms of the obscured fraction trend with the luminosity, this means that we are probably underestimating the obscured fraction in the first luminosity bin, which might be preventing us from seeing a clear trend.

7.4.2 Evolution of the obscured AGN fraction

While the decrease in the obscured AGN fraction with luminosity is generally interpreted in the framework of the so-called receding-torus models (Lusso et al., 2013; Ricci et al., 2017), the physics behind the increasing trend of the obscured fraction with redshift is not completely understood.

We compared our results with the model recently proposed in Gilli et al. (2022). In that work, the authors argue that the evolution of the obscured AGN fraction is produced by the increasing density of the ISM in the host galaxies, and give an analytic model for that. In Figure 7.14 we show the predictions of the baseline model of Gilli et al. (2022) as the solid indigo lines. The other lines reflect different assumptions in the model parameters that we discuss below. Considering the baseline model, we see that there is a good agreement for f_{23}, while for f_{22} our values are higher than the prediction, although consistent at a 1.5σ level. Our measurements are generally in better agreement with the model curves than the measurements of Liu et al. (2017), who found larger obscured AGN fractions at $z > 2$. We recall, however, that the model curves from Gilli et al. (2022) are an example of how the increased ISM density may provide a good representation of the observed trend, but they were not derived through best-fit procedures to any specific dataset. Here we explore the parameter space of that model further, trying to determine, for instance, how the ISM properties should change with redshift to reproduce the steeper trend observed by Liu et al. (2017).

By considering a number of tracers for the total mass and volume of the ISM in galaxy samples at different redshifts, mainly from ALMA, and simple assumptions on the gas density profiles, Gilli et al. (2022) measured the cosmic evolution of the ISM column density toward the nuclei of massive galaxies. This was parameterized as $N_{\rm H,ISM} \propto (1+z)^{\delta}$. They also assumed that the ISM is composed of individual gas clouds with surface densities and radii distributed as a Schechter function and that the characteristic cloud surface density $\Sigma_{c,*}$ may evolve with redshift as $(1+z)^{\gamma}$. The redshift evolution of the ISM-obscured AGN fraction above a given $N_{\rm H,ISM}$ threshold depends on both δ and γ (see Eqs. 40 and 41 in Gilli et al. 2022). Broadly speaking, a rapid increase in the total column density with redshift would imply a correspondingly rapid increase in the obscured AGN fraction. This increase is nonetheless softened if ISM clouds are significantly denser at earlier cosmic epochs as fewer clouds would then be needed to reproduce the same total gas density, reducing in turn the chances that galaxy nuclei are hidden by one of these. The baseline model in Gilli et al. (2022) assumed $\delta = 3.3$, as driven by the results from ALMA observations, and $\gamma = 2$, which, when combined with the obscuration from a small-scale component (i.e. the torus) was found to produce f_{22} and f_{23} trends in good agreement with the observations. Clearly, the uncertainties on δ and γ are large, as we still have limited knowledge of the overall ISM properties of distant galaxies. In Fig. 7.14 we show the expected trends for f_{22} and f_{23} when first increasing δ and then decreasing γ, leaving all the other model parameters unchanged. A faster increase in the total ISM density with redshift ($\delta = 4$) is needed to explain the steep trend observed by Liu et al. (2017) for f_{22} and f_{23} and the large f_{23} value measured by Vito et al. (2018) at $z \sim 3.6$. On the other hand, interestingly, a milder evolution of the characteristic gas surface density of ISM clouds ($\gamma = 1$) would only explain the steeper trend in f_{22}, but not in f_{23}, because, below $z \sim 4 - 5$ the distribution of cloud surface densities would be rich in clouds with $\Sigma_{c,*} > 10^{22}\text{cm}^{-2}$, but still short of high-density clouds with $\Sigma_{c,*} > 10^{23}\text{cm}^{-2}$. It is only at $z \sim 6$ and above that $\Sigma_{c,*}$ would increase enough to return significant fractions of very dense clouds.

To summarize, current measurements of the obscured AGN fractions as a function of cosmic time, including ours, are in agreement with an evolving ISM model in which the total gas column density of massive galaxies evolves as fast as $N_{\text{H,ISM}} \propto (1+z)^{3.3-4}$, and in which the individual gas clouds become progressively denser toward early epochs $[\Sigma_{c,*} \propto (1+z)^2]$. Such a scenario will likely be tested soon with improved accuracy by new ALMA observations.

7.4.3 Compton-thick AGN

Our work only considers the X-ray spectral fitting as an obscuration diagnostic. This means that, it is likely that we are not able to correctly characterize heavily obscured objects, especially Compton-thick (CT) AGN, which also tend to have a small number of counts. In addition to this, absorption models like the one we used (*phabs*) do not work well in a very high column density regime. We find eight objects with a nominal N_{H} higher than 10^{24}cm^{-2} out of 243, which means that we have a CT fraction of 3.3%. If we consider the $p(\log(N_{\text{H}}))$ and sum all the fractions with $N_{\text{H}} > 10^{24}\text{cm}^{-2}$, we get an observed fraction $f_{24} = 3\%$, close to that we obtained from nominal values. This value is smaller than the $\sim 8\%$ CT fraction that is found by Liu et al. (2017) for the *Chandra* Deep Field South. However, in that work, the authors use additional criteria other than the X-ray spectral fitting to determine if a source is Compton thick. In Lanzuisi et al. (2018), instead, where the only diagnostic is again the X-ray spectral analysis, the CT fraction found in the COSMOS *Chandra* survey was 2.2%, similar to our result.

Based on these previous results, it is therefore likely that if additional multi-band diagnostics were implemented, we would get a larger number of CT objects. Therefore, the CT fraction that we get is to be considered as a lower limit for the intrinsic value.

7.5 Conclusions

In this chapter, we showed the analysis of the X-ray spectra of the 243 extragalactic sources of the J1030 *Chandra* catalog and used the results to derive the obscured fraction of AGN at different redshift and luminosities. Here we outline the main results of our work and future perspectives.

- We fitted the Chandra X-ray spectra with absorbed power laws, and checked for the presence of the Fe Kα line and a soft excess. We could use spectroscopic redshift information for 44% of the sample, while we relied on photometric redshift estimates for the rest. For 7 objects with a photometric redshift only, we were able to refine the redshift estimate via X-ray spectroscopy. The best-fit spectral parameters derived for the whole sample are available at the J1030 website [6].

- We measured the obscured fractions f_{22} and f_{23} (i.e. the fraction of AGN with $N_{\rm H} > 10^{22}{\rm cm}^{-2}$ and $10^{23}{\rm cm}^{-2}$, respectively) using the full column density probability distributions derived from the spectral fits $p(\log(N_{\rm H}))$. We measured f_{22} and f_{23} in three redshift bins ($0.8 < z < 1.6, 1.6 < z < 2.2$ and $2.2 < z < 2.8$) for AGN with $\log(L) \sim 44$, and in three luminosity bins ($42.8 < \log(L_{2-10keV}) < 43.3, 43.3 < \log(L_{2-10keV}) < 43.8$, and $43.8 < \log(L_{2-10keV}) < 44.5$), for AGN at $z \sim 1.2$. We corrected these *observed* fractions for the sky coverage of the survey and derived accurate measurement errors through a bootstrapping procedure that accounts for both the finite size of the sample and the uncertainties on the $N_{\rm H}$ estimates.

- We measured average values of $f_{22} \sim 0.7 - 0.8$ and $f_{23} \sim 0.5 - 0.6$. While these average values are in broad agreement with those in other works (Aird et al., 2015; Liu et al., 2017), we did not see clear trends with luminosity or redshift, as opposed to what is often found in the literature. This might, at least partially, depend on residual, uncorrected biases, and/or on the limited dynamical range in luminosity and redshift spanned by our data. Nonetheless, when combined with measurements performed in the local Universe, our data point to an increase in the obscured AGN fractions with redshift, in agreement with other findings.

- We finally considered predictions from recent analytic models that ascribe the redshift evolution of the obscured AGN fraction to the increased density of the ISM in high-z hosts, which adds significant obscuration to that of the parsec-scale torus (Gilli et al., 2022). When combined with literature measurements, our results favor a scenario in which the total ISM column density grows with redshift as $N_{\rm H,ISM} \propto (1+z)^{3.3-4}$, and in which the characteristic surface density of individual gas clouds in the ISM evolves as $\Sigma_{c,*} \propto (1+z)^2$.

To gain a deeper understanding of nuclear obscuration at different cosmic epochs, and as a function of the various AGN physical properties, large object samples are needed, that would go significantly beyond those available from current X-ray probes. What is believed to be the bulk of the AGN population (low-luminosity, possibly obscured objects) is now partly missed at medium-high redshift values, and completely lost beyond redshift $z \sim 6$. Next-generation X-ray imaging surveys, such as those proposed with the Survey and Time-domain Astrophysical Research eXplorer (*STAR-X*[7]), a Medium Explorer mission selected by NASA for Phase A study, the Advanced X-ray Imaging Satellite (*AXIS*, Mushotzky et al., 2019; Marchesi et al., 2020), a probe-class mission proposed to NASA, and the

---

[6] `http://j1030-field.oas.inaf.it/chandra_1030`

[7] `http://star-x.xraydeep.org/`

L-class mission *Athena* under scrutiny at ESA (Nandra et al., 2013), would offer new opportunities to detect and characterize highly obscured sources. These observatories are expected to discover a few thousand heavily obscured ($N_{\mathrm{H}} > 10^{23}$ cm^{-2}) AGN at $z > 3$, shedding light on the overall growth of SMBHs before cosmic noon.

Chapter 8

Conclusions and future projects

The aim of this thesis has been to investigate different ways in which AGN can be implemented in cosmology. Given the immense power of AGN and their presence across all cosmological epochs, they offer a valuable tool to delve deeper into the Universe expansion history, which is inextricably tied to our comprehension of the Universe on a large-scale. Moreover, the activity of AGN is closely associated with the processes of SMBH accretion and interaction with the host galaxies. Therefore, by examining AGN, we can enhance our understanding of the history of structure formation and galaxy evolution. Simultaneously, many fundamental questions about AGN emission physics remain unanswered, so that addressing these can refine their role and implementation in cosmological contexts.

In Chapter 3, 4, and 5, we studied the $L_X - L_{UV}$ relation in quasars. This non-linear relation allows us to use quasars as standard candles, measuring cosmological distances at redshift values where SNIa or other probes are absent.
In Chapter 3 we presented an analysis of the relation that was, for the first time, carried out on a sample of ∼1800 objects with a complete UV spectroscopic analysis and an X-ray spectroscopic analysis at redshfits higher than 1.9. We investigated which quantities worked best as the UV and X-ray emission proxies, aiming at (i) further reducing the observed dispersion and (ii) getting a better understanding of the physics behind the relation itself. We found out that the MgII line flux, when present, works as a better proxy in terms of the obtained dispersion than the monochromatic 2500 Å one. At the same time, we found that the MgII line flux traces the photometrically derived 2500 Å flux, while the spectroscopic 2500Å flux does not, giving us an even more non-linear relation with the X-ray flux. This allows use to use the spectroscopic 2500 Å to derive more precise distance measurements for all the redshift range of our sample. Furthermore, we determined that the "toy-model" for disc-corona interaction suggested by Lusso & Risaliti (2017) is incompatible with our sample spectroscopic data.
In Chapter 4 we investigated another important aspect of the $L_X - L_{UV}$ relation, its dispersion. We examined those factors known to contribute to the observed dispersion but that we can not remove with the sample selection. Focusing on a subset of roughly 300 objects with multiple serendipitous observations via the XMM-*Newton* telescope, we discerned that variability contributes as much as 0.08 dex to the observed dispersion. Ideally, having numerous observations for the majority of our sample would enable us to average these

observations, thus reducing this contribution. However, this is not possible at current times. We also employed a mock sample of quasars to ascertain the effect of quasar inclination on the observed dispersion, determining its contribution to be approximately 0.06 dex. Regrettably, we currently lack a method to measure inclination across all redshifts, rendering us unable to make corrections for our quasar samples. Yet, these findings are paramount, suggesting that the intrinsic dispersion of the relationship is minimal, potentially approaching zero. This implies a robust physical connection between the UV disc and the X-ray corona over a broad luminosity spectrum. This is crucial both for validating the results of the cosmological implementation of quasars as standard candles, which point out a significant tension with the predictions of th flat ΛCDM model, and for our comprehension of the physics underlying the relation.

Future plans about this research topic include revisiting the said “toy model”, using the insights from the findings presented here. In doing this, we will also consider the latest developments in corona physics, including the polarization results from Pal et al. (2023), which challenge the traditional spherical lamp-post geometry, and the recent simulation outcomes of corona emission (e.g., Gronkiewicz et al. (2023)). Any physical model of the relation should also consider the results on the intrinsic dispersion, and predict a tight relation on a wide range of quasars luminosities.

In Chapter 5 we estimated the shape of the Hubble Diagram of quasars and Supernovae Ia with the use of a Neural Network, aiming at deriving a non-parametric fit of the shape of the diagram. We found that the presence of a strong tension with the prediction of the flat ΛCDM model are confirmed, and the data suggest to go and look at models in the family of the Interacting Dark Sector.

In Chapter 6, we focused on reverberation mapping and how this technique can be applied for cosmology in two distinct ways. Firstly, we discussed the “SARM” project for the implementation of reverberation mapping together with spectroastrometry to obtain an estimate of the H_0 constant independent from any calibration. Such a result can potentially shed light on the problem of the H_0 tension. We discussed the campaign targeting the local AGN Ark120. This campaign represent the first time where we are attempting at deriving the interferometry and the reverberation mapping measurements for the same lines. Although the campaign was unlucky in the past year, both on the reverberation mapping and on the spectroastrometry side, we are hopeful that in the coming year we will obtain the measurements needed, and that we will obtain an estimate of the H_0 constant from Ark120.

Reverberation mapping also allows us to measure SMBH masses, which is essential to understand SMBH formation and accretion and their interaction with the host galaxy over time. Presently, obtaining accurate and reliable mass estimates at high redshift poses substantial challenges. We introduced the ongoing project centred on the reverberation mapping campaign for SDSS J2222+2745, a lensed quasar at redshift z =2.8. The lensing allows us to overcome the most severe issues of RM mass measurements at high redshift (signal to noise ratio requirements, variability prediction, extended monitoring time needed). We analysed the data from the last campaign extension, which consisted of one year of additional monitoring, focusing both on the CIV and CIII] lines. The goal was to use the extended light curve to obtain better constraints on the velocity-resolved time lags for the CIV line, and to extend the analysis to the CIII] line as well. Preliminary results show that the CIII] lag is consistent with its emission being further away from the SMBH, as predicted by the corresponding ionizing potential, and that the velocity-resolved time lags for the CIV line are consistent with a mostly circular motion. Together with completing the analysis, future plans include employing the CARAMEL code Pancoast et al. (2014) for direct BLR modelling. This approach allows mass estimates without assuming a scale factor (see next paragraph) and gives important information about the BLR geometry, inclination, and kinematics. Object like SDSS J2222+2745 are very rare (this campaign is the first one of its kind), but future surveys, like LSST, promise to tens more, enhancing our understanding of SMBH and the physics of the BLR by (i) providing precise mass measurements and (ii) allowing us to calibrate the $R_{BLR} - L$ relation for "single epoch" mass measurements at otherwise unreachable redshift values. The results from this campaign will be useful as a test for the kind of data that will be available in the coming future, and to prepare future campaigns in the best, most efficient ways.

In Chapter 7, we investigated the obscured fraction of AGN in the J1030 *Chandra* deep field, the 5th deepest X-ray field to date. Studying the obscured fraction of AGN is fundamental for understanding the evolution of AGN and galaxies properties over time. Moreover, it is crucial for constraining the total number of AGN present at different epochs, which is needed to reconstruct the accretion history of SMBH. We performed the X-ray data reduction and spectroscopic analysis of the 243 AGN detected in the field. The primary objectives were to create an X-ray catalogue for the AGN in the field and to determine the obscured fraction of AGN in the field as a function of redshift and luminosity. The findings show an increasing obscured fraction at redshift compared to local measurements. These results, juxtaposed with those of other surveys, align with analytic models that attribute the greater obscured fraction at high redshifts to the increasingly dense ISM of their hosts. Future plans about this topic include harnessing the extensive multi-wavelength analysis of the J1030 field for the Compton Thick characterization, especially in the infrared and radio. IR observations capture the emission from the AGN dusty torus, which reprocesses the absorbed emission originating from the disc, while radio emission traces the cold molecular gas in the host galaxy, associated with both star formation and AGN fuelling. By utilising both archival data and selecting objects for future proposals, we aim to obtain a more accurate understanding of the "Compton-Thick candidates" in the J1030 *Chandra* field. This extension to multi-wavelength probes will also benefit the other AGN in the field for which the obscuration level is poorly constrained.

List of Publications

Concerning this thesis:
Quasars as Standard Candles IV. Analysis of the X-ray and UV indicators of the disc-corona relation - A&A, 676, A143 - **M. Signorini**, G. Risaliti, E. Lusso, E. Nardini, G. Bargiacchi, A. Sacchi, B. Trefoloni
`https://ui.adsabs.harvard.edu/abs/2023A%26A...676A.143S/abstract`

Quasars as Standard Candles V. Evaluation of a ≤ 0.06 *dex intrinsic dispersion in the* $L_X - L_{UV}$ *relation* - accepted for publication (A&A) - **M. Signorini**, G. Risaliti, E. Lusso, E. Nardini, G. Bargiacchi, A. Sacchi, B. Trefoloni
`https://ui.adsabs.harvard.edu/abs/2023arXiv231208448S/abstract`

X-ray properties and obscured fraction of AGN in the J1030 Chandra field - A&A, 676, A49 - **M. Signorini**, S. Marchesi, R. Gilli, M. Brusa, A. Comastri, Q. D'Amato, K. Iwasawa, G. Lanzuisi, G. Mazzolari, M. Mignoli, A. Peca, I. Prandoni, P. Tozzi, C. Vignali, F. Vito, C. Norman
`https://ui.adsabs.harvard.edu/abs/2023A%26A...676A..49S/abstract`

Non-parametric analysis of the Hubble Diagram with Neural Networks - A&A, 678, A13 - L. Giambagli, D. Fanelli, G. Risaliti, **M. Signorini**
`https://ui.adsabs.harvard.edu/abs/2023A%26A...678A..13G/abstract`

Quasars as standard candles: III. Validation of a new sample for cosmological studies – A&A, 642, 150 - Lusso E.; Risaliti G.; Nardini E.; Bargiacchi G.; Benetti M.; Bisogni S.; Capozziello S.; Civano F.; Eggleston L.; Elvis M.; Fabbiano G.; Gilli R.; Marconi A.; Paolillo M.; Piedipalumbo E.; Salvestrini F; **Signorini M.**; Vignali C.
`https://ui.adsabs.harvard.edu/abs/2020A%26A...642A.150L/abstract`

Other publications:
Quasars as high-redshift standard candles - Astronomische Nachrichten proceeding, doi: 10.1002/asna.20230054 – G. Risaliti, E. Lusso, E. Nardini, G. Bargiacchi, S. Bisogni, A. Sacchi, **M. Signorini**, B. Trefoloni
`https://ui.adsabs.harvard.edu/abs/2023AN....34430054R/abstract`

Quasars as standard candles VI: spectroscopic validation of the cosmological sample – submitted to A&A - B. Trefoloni, E. Lusso, E. Nardini, G. Risaliti, A. Marconi, G.

Bargiacchi, A. Sacchi, **M. Signorini**

AGN STORM 2. VI. Mapping Temperature Fluctuations in the Accretion Disk of Mrk 817 - accepted (ApJ), arXiv:2310.01497- J. M. M. Neustadt, C. S. Kochanek, J. Montano, J. Gelbord, A. J. Barth, G. De Rosa, G. A. Kriss, E. M. Cackett, K. Horne, E. A. Kara, H. Landt, H. Netzer, N. Arav, M. C. Bentz, E. Dalla Bonta, M. Dehghanian, P. Du, R. Edelson, G. J. Ferland, C. Fian, T. Fischer, M. R. Goad, D. H. Gonzalez Buitrago, V. Gorjian, C. J. Grier, P. B. Hall, Y. Homayouni, C. Hu, D. Ilic, M. D. Joner, J. Kaastra, S. Kaspi, K. T. Korista, A. B. Kovacevic, C. Lewin, Y. Li, I. M. McHardy, M. Mehdipour, J. A. Miller, C. Panagiotou, E. Partington, R. Plesha, R. W. Pogge, L. C. Popovic, D. Proga, T. Storchi-Bergmann, D. Sanmartim, M. R. Siebert, **M. Signorini**, M. Vestergaard, F. Zaidouni, Y. Zu
`https://ui.adsabs.harvard.edu/abs/2023arXiv231001497N/abstract`

The most luminous blue quasars at 3.0< z <3.3 III. Enhanced iron and faint [O iii] emission in X-ray weak quasars - A&A, 677, A111 - B. Trefoloni, E. Lusso, E. Nardini, G. Risaliti, G. Bargiacchi, S. Bisogni, F. M. Civano, M. Elvis, G. Fabbiano, R. Gilli, A. Marconi, G. T. Richards, A. Sacchi, F. Salvestrini, **M. Signorini**, C. Vignali
`https://ui.adsabs.harvard.edu/abs/2023A%26A...677A.111T/abstract`

LBT-MODS spectroscopy of high redshift candidates in the Chandra J1030 field. A newly discovered z ∼ 2.8 large scale structure - A&A, 673, A97 - S. Marchesi, M. Mignoli, R. Gilli, G. Mazzolari, **M. Signorini**, M. Brienza, S. Bisogni, M. Bolzonella, O. Cucciati, Q. D'Amato, A. Peca, I. Prandoni, P. Tozzi, C. Vignali, F. Vito, A. Comastri
`https://ui.adsabs.harvard.edu/abs/2023A%26A...673A..97M/abstract`

Quasars as high-redshift standard candles - A&A, 663, L7 - Sacchi A., Risaliti G., **Signorini M.**, Lusso E., Nardini E., Bargiacchi G., Bisogni S., Civano F., Elvis M., Fabbiano G., Gilli R., Trefoloni B., Vignali C.
`https://ui.adsabs.harvard.edu/abs/2022A%26A...663L...7S/abstract`

Quasar Cosmology: dark energy evolution and spatial curvature - MNRAS, v. 515, 2, pp. 1795-1806 - Bargiacchi G., Benetti M., Capozziello S., Lusso E., Risaliti G., **Signorini M.**
`https://ui.adsabs.harvard.edu/abs/2022MNRAS.515.1795B/abstract`

Cosmography by orthogonalized logarithmic polynomials – A&A, 649, A65 - Bargiacchi G., Risaliti G., Benetti M., Capozziello S., Lusso E., Saccardi A., **Signorini M.**
`https://ui.adsabs.harvard.edu/abs/2021A%26A...649A..65B/abstract`

The most luminous blue quasars at 3.0 < z < 3.3 II. CIV/X-ray emission and accretion disc physics – A&A, 653, A158 - E. Lusso, E. Nardini, S. Bisogni, G. Risaliti, R. Gilli, G. T. Richards, F. Salvestrini, C. Vignali, G. Bargiacchi, F. Civano, M. Elvis, G. Fabbiano, A. Marconi, A. Sacchi, **M. Signorini**
`https://ui.adsabs.harvard.edu/abs/2021A%26A...653A.158L/abstract`

Acknowledgements

Mi trovo per la terza e ultima volta a ringraziare la persone che hanno fatto sì che una mia tesi venisse scritta. Il fatto che sia la terza volta, mi fa dire di non essere troppo prolissa e ripetitiva. D'altra parte, è anche l'ultima, nonché la chiusura di un percorso di otto e più anni all'Università degli Studi di Firenze. In più, il numero di persone da ringraziare, con mia sorpresa, aumenta negli anni. Quindi, come si intuisce da questa introduzione, temo che invece sarò prolissa e ripetitiva.

Ringrazio Guido, per la fiducia in me che costantemente ha mostrato negli anni, anche in momenti in cui io stessa ne ho avuta poca, per le ore passate su Skype anche a orari improbabili per palrare di scienza e non, e per la passione che riesce sempre a trasmettere.
Ringrazio Tommaso per avermi fatto sentire accolta a UCLA e avermi introdotto al magico mondo del Reverberation Mapping, per i consigli, per aver creduto che potessi far convergere quelle chain, per i caffè Lavazza e i pranzi di gruppo. E' stata un'esperienza fantastica, scientificamente e umanamente parlando.
Grazie a Beta per aver provato a farmi fare dei plot decenti, e Ema per aver provato a far sì che le frasi nei miei paper avessero senso compiuto. Temo che entrambi gli sforzi siano stati parzialmente vani, ma ho imparato tanto lavorando con voi.
Grazie a Roberto e Stefano per avermi fatto prendere parte al mondo di J1030, per tutto l'aiuto, e per aver accolto nel dizionario la parola "rettangolarizzazione". Roberto, un grazie in più per aver sopportato le mie mail piene di emoji.

Grazie a tutta la mia famiglia, per sostenermi costantemente con un affetto così grande che passa gli oceani, e per essere da sempre i miei *fan* numero uno a ogni svolta della mia vita. Vi voglio tantissimo bene e vi sento sempre vicini.
Menzione speciale a Andrea, per essere il meraviglioso cugino che è. So che non te lo dico abbastanza ma ti voglio un gran bene, e sono così fiera di te. Scusa se ti ho abbandonato a Natale, e se ho mancato la tua laurea. Non lo faccio più.
Un grazie particolare anche ai miei genitori, perché tutto il mio percorso universitario, compreso questo dottorato, non sarebbe stato possibile senza il loro supporto e il loro grande affetto. Starvi lontana per un po' mi ha fatto realizzare che vi assomiglio più di quanto pensassi, e non potevo sperare di assomigliare a persone migiliori.

Grazie a tutti gli abitanti di Arcetri con cui ho condiviso pranzi, partite a GeoGuessr o worlde, e pause caffè, già mi rattrista aver dovuto riconsegnare le chiavi, mi mancherete molto. Essendo questi i ringraziamenti per la tesi, *I own* una menzione speciale a Elena

Bertola per avermi portata a mangiare ravioli e bere vino la sera della consegna, al posto di stare a rileggerla settecento volte, e per essersi accollata le mie ansie sui postdoc. Grazie a Giulia Tozzi, per aver condiviso le fatiche dei corsi online e di tutte le cose poco chiare di questo dottorato. Grazie a Andrea Sacchi, miglior creatore di sticker telegram che il mondo abbia visto, e gran compagno di lamentele. Grazie a Bartolometo Trefoloni, per gli sfoghi condivisi, le bevute alle conferenze, e per avermi fatto scroccare la chiavetta del caffè cento volte. Grazie per aver sempre dato ascolto a tutti gli 'scusa Bart ma secondo te...?', e per aver anche tu cercato di insegnarmi a fare le 'figure belline'. Chissà se imparerò mai. Scherzi a parte, mi hai fatto sentire molto meno sola in questi anni, grazie.

Grazie a tutti i miei splendidi amici incontrati al Polo Scientifico, per aver reso tale landa desolata un posto in cui mi sento sempre a casa e in cui ho sempre voglia di tornare, e per i tanti bei momenti passati insieme dentro e fuori dal Polo. Se guardo indietro, mi sembra impossibile pensare che prima non faceste parte della mia vita. *Whatever the future brings*, so che questi otto anni saranno sempre valsi la pena per il fatto di avervi potuto incontrare e per averli potuti condividere con voi.

Thanks to all the people I've met and re-met at conferences and schools; you've made the roaming around so much more pleasant. Thanks to Elsa, Richard and Davide for being the best.
Thanks to all my LA friends, inside and outside of UCLA. In even my wildest dreams, I couldn't have imagined meeting so many amazing people in such a short time. Thanks for welcoming me in your lives and thanks to anyone with whom I shared karaoke night, climbing sessions, hikes, game nights, beers, camping trips, pizzas, hang-gliding lessons, and so on. You've made my time in LA awesome. Thanks to all the UCLA Grad students for welcoming me there, and for the grunches, the groffes, for the strike, and for the nights out. Honourable mention goes to Benji, for sharing the bureaucracy pain. Thanks to the 'Swifties', I was so happy to have such amazing neighbours.
I love you all, I'll miss you, and I wish you all the best. As one good friend of mine always likes to remember, 'I wish there was a way to know you're in the good old days before you've actually left them.' I bet there are more to come.
And big thanks to all the people that drove me around, be sure I know that's true love language in LA.

Grazie alle mie Mujeres, perché nonostante gli anni che passano e il delirio di dover fare i doodle, riusciamo comunque imperterrite a vederci per i nostri aperelli e cenini; mi scalda sempre il cuore.
Grazie a Celli, per i reel dei gatti che fanno cose buffe, gli audio di 3 minuti e le chiacchierate di quattro ore. Mi sembra impossibile che siamo così vecchi da aver accettato di fare delle *chiamate*, e che come dei veri e propri adulti le *programmiamo* a seconda dei nostri *impegni*. L'età e la distanza sono proprio brutte bestie. Grazie per esserci *regardless of that*.
Grazie a Olga e Sofia, che neanche a questo giro assisteranno a una mia discussione di tesi, ma a cui devo così tanto e a cui voglio tanto bene. Questi ultimi tre anni sono stati pieni di partenze e lontananze, ma nonostante questo anche sempre pieni di abbracci e di

chiacchierate. Che possa essere sempre così.

Concludo con la persona che più di tutte si merita i miei ringraziamenti, e senza la quale sarebbe stato cento volte più difficile arrivare in fondo a questo dottoranto. Grazie a Lorenzo, per essere l'incredibile compagno di viaggio che è, nella vita universitaria e in quella vera, per essermi sempre accanto, per tutte le risate, per essersi accollato viaggi aerei improbabili per venirmi a trovare, per riuscire sempre a tirarmi su quando sono giù, e per tutto ciò che non riuscirò mai a mettere su carta. Come cantava Mark Knopfler, *'I love you like the stars above'*.

Bibliography

Aird, J., Coil, A. L., Georgakakis, A., Nandra, K., Barro, G., & Pérez-González, P. G. (2015). The evolution of the X-ray luminosity functions of unabsorbed and absorbed AGNs out to z∼ 5. *MNRAS*, *451*, 1892–1927. doi:10.1093/mnras/stv1062. arXiv:1503.01120.

Akritas, M. G., & Bershady, M. A. (1996). Linear Regression for Astronomical Data with Measurement Errors and Intrinsic Scatter. *ApJ*, *470*, 706. doi:10.1086/177901. arXiv:astro-ph/9605002.

Allen, R. J., Kacprzak, G. G., Glazebrook, K., Labbé, I., Tran, K.-V. H., Spitler, L. R., Cowley, M., Nanayakkara, T., Papovich, C., Quadri, R., Straatman, C. M. S., Tilvi, V., & van Dokkum, P. (2017). The Size Evolution of Star-forming Galaxies since z ∼ 7 Using ZFOURGE. *ApJl*, *834*, L11. doi:10.3847/2041-8213/834/2/L11. arXiv:1612.05262.

Anand, G. S., Tully, R. B., Rizzi, L., Riess, A. G., & Yuan, W. (2022). Comparing Tip of the Red Giant Branch Distance Scales: An Independent Reduction of the Carnegie-Chicago Hubble Program and the Value of the Hubble Constant. *ApJ*, *932*, 15. doi:10.3847/1538-4357/ac68df. arXiv:2108.00007.

Ananna, T. T., Treister, E., Urry, C. M., Ricci, C., Kirkpatrick, A., LaMassa, S., Buchner, J., Civano, F., Tremmel, M., & Marchesi, S. (2019). The Accretion History of AGNs. I. Supermassive Black Hole Population Synthesis Model. *ApJ*, *871*, 240. doi:10.3847/1538-4357/aafb77. arXiv:1810.02298.

Aravena, M., Boogaard, L., Gónzalez-López, J., Decarli, R., Walter, F., Carilli, C. L., Smail, I., Weiss, A., Assef, R. J., Bauer, F. E., Bouwens, R. J., Cortes, P. C., Cox, P., da Cunha, E., Daddi, E., Díaz-Santos, T., Inami, H., Ivison, R., Novak, M., Popping, G., Riechers, D., van der Werf, P., & Wagg, J. (2020). The ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: The Nature of the Faintest Dusty Star-forming Galaxies. *ApJ*, *901*, 79. doi:10.3847/1538-4357/ab99a2. arXiv:2006.04284.

Arnaud, K. A. (1996). XSPEC: The First Ten Years. In G. H. Jacoby, & J. Barnes (Eds.), *Astronomical Data Analysis Software and Systems V* (p. 17). volume 101 of *Astronomical Society of the Pacific Conference Series*.

Aviles, A., Bravetti, A., Capozziello, S., & Luongo, O. (2014). Precision cosmology with padé rational approximations: Theoretical predictions versus observational limits. *Physical Review D*, *90*, 043531.

Avni, Y. (1976). Energy spectra of X-ray clusters of galaxies. *ApJ*, *210*, 642–646. doi:`10.1086/154870`.

Bañados, E., Venemans, B. P., Decarli, R., Farina, E. P., Mazzucchelli, C., Walter, F., Fan, X., Stern, D., Schlafly, E., Chambers, K. C., Rix, H. W., Jiang, L., McGreer, I., Simcoe, R., Wang, F., Yang, J., Morganson, E., De Rosa, G., Greiner, J., Baloković, M., Burgett, W. S., Cooper, T., Draper, P. W., Flewelling, H., Hodapp, K. W., Jun, H. D., Kaiser, N., Kudritzki, R. P., Magnier, E. A., Metcalfe, N., Miller, D., Schindler, J. T., Tonry, J. L., Wainscoat, R. J., Waters, C., & Yang, Q. (2016). The Pan-STARRS1 Distant z ¿ 5.6 Quasar Survey: More than 100 Quasars within the First Gyr of the Universe. *ApJs*, *227*, 11. doi:`10.3847/0067-0049/227/1/11`. `arXiv:1608.03279`.

Baldwin, J. A. (1977). Luminosity Indicators in the Spectra of Quasi-Stellar Objects. *ApJ*, *214*, 679–684. doi:`10.1086/155294`.

Bargiacchi, G., Benetti, M., Capozziello, S., Lusso, E., Risaliti, G., & Signorini, M. (2022). Quasar cosmology: dark energy evolution and spatial curvature. *MNRAS*, *515*, 1795–1806. doi:`10.1093/mnras/stac1941`. `arXiv:2111.02420`.

Bargiacchi, G., Risaliti, G., Benetti, M., Capozziello, S., Lusso, E., Saccardi, A., & Signorini, M. (2021). Cosmography by orthogonalized logarithmic polynomials. *A&A*, *649*, A65. URL: `https://doi.org/10.1051/0004-6361/202140386`. doi:`10.1051/0004-6361/202140386`.

Baskin, A., & Laor, A. (2005). What controls the CIV line profile in active galactic nuclei? *MNRAS*, *356*, 1029–1044. doi:`10.1111/j.1365-2966.2004.08525.x`. `arXiv:astro-ph/0409196`.

Batiste, M., Bentz, M. C., Raimundo, S. I., Vestergaard, M., & Onken, C. A. (2017). Recalibration of the M $_{BH}$-$\sigma_\star$ Relation for AGN. *ApJ*, *838*, L10. doi:`10.3847/2041-8213/aa6571`. `arXiv:1612.02815`.

Battye, R. A., Charnock, T., & Moss, A. (2015). Tension between the power spectrum of density perturbations measured on large and small scales. *Phys. Rev. D*, *91*, 103508. URL: `https://link.aps.org/doi/10.1103/PhysRevD.91.103508`. doi:`10.1103/PhysRevD.91.103508`.

Bennett, C. L., Larson, D., Weiland, J. L., & Hinshaw, G. (2014). The 1% Concordance Hubble Constant. *ApJ*, *794*, 135. doi:`10.1088/0004-637X/794/2/135`. `arXiv:1406.1718`.

Bianchi, S., Maiolino, R., & Risaliti, G. (2012). AGN Obscuration and the Unified Model. *Advances in Astronomy*, *2012*, 782030. doi:`10.1155/2012/782030`. `arXiv:1201.2119`.

Bianchini, F., Fabbian, G., Lapi, A., Gonzalez-Nuevo, J., Gilli, R., & Baccigalupi, C. (2019). Broadband Spectral Energy Distributions of SDSS-selected Quasars and of Their Host Galaxies: Intense Activity at the Onset of AGN Feedback. *ApJ*, *871*, 136. doi:`10.3847/1538-4357/aaf86b`. `arXiv:1806.06516`.

Bisogni, S., Marconi, A., & Risaliti, G. (2017). Orientation effects on spectral emission features of quasars. *MNRAS*, *464*, 385–397. doi:`10.1093/mnras/stw2324`. `arXiv:1609.06705`.

Blandford, R. D., & McKee, C. F. (1982). Reverberation mapping of the emission line regions of Seyfert galaxies and quasars. *ApJ*, *255*, 419–439. doi:`10.1086/159843`.

Boyle, B. J., Shanks, T., Croom, S. M., Smith, R. J., Miller, L., Loaring, N., & Heymans, C. (2000). The 2dF QSO Redshift Survey - I. The optical luminosity function of quasi-stellar objects. *MNRAS*, *317*, 1014–1022. doi:`10.1046/j.1365-8711.2000.03730.x`. `arXiv:astro-ph/0005368`.

Brandt, W. N., & Alexander, D. M. (2015). Cosmic X-ray surveys of distant active galaxies. The demographics, physics, and ecology of growing supermassive black holes. *A&Ar*, *23*, 1. doi:`10.1007/s00159-014-0081-z`. `arXiv:1501.01982`.

Brightman, M., Nandra, K., Salvato, M., Hsu, L.-T., Aird, J., & Rangel, C. (2014). Compton thick active galactic nuclei in Chandra surveys. *MNRAS*, *443*, 1999–2017. doi:`10.1093/mnras/stu1175`. `arXiv:1406.4502`.

Brightman, M., & Ueda, Y. (2012). The evolution of the Compton thick fraction and the nature of obscuration for active galactic nuclei in the Chandra Deep Field South. *MNRAS*, *423*, 702–717. doi:`10.1111/j.1365-2966.2012.20908.x`. `arXiv:1203.1045`.

Buchner, J., & Bauer, F. E. (2017). Galaxy gas as obscurer - II. Separating the galaxy-scale and nuclear obscurers of active galactic nuclei. *MNRAS*, *465*, 4348–4362. doi:`10.1093/mnras/stw2955`. `arXiv:1610.09380`.

Buchner, J., Georgakakis, A., Nandra, K., Brightman, M., Menzel, M.-L., Liu, Z., Hsu, L.-T., Salvato, M., Rangel, C., Aird, J., Merloni, A., & Ross, N. (2015). Obscuration-dependent Evolution of Active Galactic Nuclei. *ApJ*, *802*, 89. doi:`10.1088/0004-637X/802/2/89`. `arXiv:1501.02805`.

Burlon, D., Ajello, M., Greiner, J., Comastri, A., Merloni, A., & Gehrels, N. (2011). Three-year Swift-BAT Survey of Active Galactic Nuclei: Reconciling Theory and Observations? *ApJ*, *728*, 58. doi:`10.1088/0004-637X/728/1/58`. `arXiv:1012.0302`.

Calderone, G., Nicastro, L., Ghisellini, G., Dotti, M., Sbarrato, T., Shankar, F., & Colpi, M. (2017). QSFIT: automatic analysis of optical AGN spectra. *MNRAS*, *472*, 4051–4080. doi:`10.1093/mnras/stx2239`. `arXiv:1612.01580`.

Capozziello, S., D'Agostino, R., & Luongo, O. (2020). High-redshift cosmography: auxiliary variables versus Padé polynomials. *Monthly Notices of the Royal Astronomical Society*, *494*, 2576–2590. URL: `https:`

//doi.org/10.1093/mnras/staa871. doi:10.1093/mnras/staa871. arXiv:https://academic.oup.com/mnras/article-pdf/494/2/2576/33113700/staa871.pdf.

Cárdenas, V. H. (2015). Exploring hints for dark energy density evolution in light of recent data. *Physics Letters B*, *750*, 128–134. doi:10.1016/j.physletb.2015.08.064. arXiv:1405.5116.

Cash, W. (1979). Parameter estimation in astronomy through application of the likelihood ratio. *ApJ*, *228*, 939–947. doi:10.1086/156922.

Cattoën, C., & Visser, M. (2007). The Hubble series: convergence properties and redshift variables. *Classical and Quantum Gravity*, *24*, 5985–5997. doi:10.1088/0264-9381/24/23/018. arXiv:0710.1887.

Chevallier, M., & Polarski, D. (2001). Accelerating universes with scaling dark matter. *International Journal of Modern Physics D*, *10*, 213–223. URL: http://dx.doi.org/10.1142/S0218271801000822. doi:10.1142/s0218271801000822.

Chiaraluce, E., Vagnetti, F., Tombesi, F., & Paolillo, M. (2018). The X-ray/UV ratio in active galactic nuclei: dispersion and variability. *Aap*, *619*, A95. doi:10.1051/0004-6361/201833631. arXiv:1808.06964.

Circosta, C., Vignali, C., Gilli, R., Feltre, A., Vito, F., Calura, F., Mainieri, V., Massardi, M., & Norman, C. (2019). X-ray emission of z ¿ 2.5 active galactic nuclei can be obscured by their host galaxies. *A&A*, *623*, A172. doi:10.1051/0004-6361/201834426. arXiv:1901.07108.

Civano, F., Comastri, A., & Brusa, M. (2005). X-ray spectral analysis of optically faint sources in the Chandra deep fields. *MNRAS*, *358*, 693–704. doi:10.1111/j.1365-2966.2005.08821.x. arXiv:astro-ph/0501397.

Colgáin, E. Ó., Sheikh-Jabbari, M. M., Solomon, R., Dainotti, M. G., & Stojkovic, D. (2022). Putting Flat ΛCDM In The (Redshift) Bin. *arXiv e-prints*, (p. arXiv:2206.11447). doi:10.48550/arXiv.2206.11447. arXiv:2206.11447.

Comastri, A., Iwasawa, K., Gilli, R., Vignali, C., Ranalli, P., Matt, G., & Fiore, F. (2010). Suzaku Observations of Hard X-ray-selected Seyfert 2 Galaxies. *ApJ*, *717*, 787–794. doi:10.1088/0004-637X/717/2/787. arXiv:1005.3253.

Croom, S. M., Richards, G. T., Shanks, T., Boyle, B. J., Strauss, M. A., Myers, A. D., Nichol, R. C., Pimbblet, K. A., Ross, N. P., Schneider, D. P., Sharp, R. G., & Wake, D. A. (2009). The 2dF-SDSS LRG and QSO survey: the QSO luminosity function at 0.4 ¡ z ¡ 2.6. *MNRAS*, *399*, 1755–1772. doi:10.1111/j.1365-2966.2009.15398.x. arXiv:0907.2727.

Croton, D. J., Springel, V., White, S. D. M., De Lucia, G., Frenk, C. S., Gao, L., Jenkins, A., Kauffmann, G., Navarro, J. F., & Yoshida, N. (2006). The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies. *MNRAS*, *365*, 11–28. doi:10.1111/j.1365-2966.2005.09675.x. arXiv:astro-ph/0508046.

Dahle, H., Gladders, M. D., Sharon, K., Bayliss, M. B., & Rigby, J. R. (2015). Time Delay Measurements for the Cluster-lensed Sextuple Quasar SDSS J2222+2745. *ApJ*, *813*, 67. doi:10.1088/0004-637X/813/1/67. arXiv:1505.06187.

Dahle, H., Gladders, M. D., Sharon, K., Bayliss, M. B., Wuyts, E., Abramson, L. E., Koester, B. P., Groeneboom, N., Brinckmann, T. E., Kristensen, M. T., Lindholmer, M. O., Nielsen, A., Krogager, J. K., & Fynbo, J. P. U. (2013). SDSS J2222+2745: A Gravitationally Lensed Sextuple Quasar with a Maximum Image Separation of 15.”1 Discovered in the Sloan Giant Arcs Survey. *ApJ*, *773*, 146. doi:10.1088/0004-637X/773/2/146. arXiv:1211.1091.

Dalla Bontà, E., Peterson, B. M., Bentz, M. C., Brandt, W. N., Ciroi, S., De Rosa, G., Fonseca Alvarez, G., Grier, C. J., Hall, P. B., Hernández Santisteban, J. V., Ho, L. C., Homayouni, Y., Horne, K., Kochanek, C. S., Li, J. I. H., Morelli, L., Pizzella, A., Pogge, R. W., Schneider, D. P., Shen, Y., Trump, J. R., & Vestergaard, M. (2020). The Sloan Digital Sky Survey Reverberation Mapping Project: Estimating Masses of Black Holes in Quasars with Single-epoch Spectroscopy. *ApJ*, *903*, 112. doi:10.3847/1538-4357/abbc1c. arXiv:2007.02963.

D’Amato, Q., Gilli, R., Vignali, C., Massardi, M., Pozzi, F., Zamorani, G., Circosta, C., Vito, F., Fritz, J., Cresci, G., Casasola, V., Calura, F., Feltre, A., Manieri, V., Rigopoulou, D., Tozzi, P., & Norman, C. (2020). Dust and gas content of high-redshift galaxies hosting obscured AGN in the Chandra Deep Field-South. *A&A*, *636*, A37. doi:10.1051/0004-6361/201936175. arXiv:2003.08631.

de Nicola, S., Marconi, A., & Longo, G. (2019). The fundamental relation between supermassive black holes and their host galaxies. *MNRAS*, *490*, 600–612. doi:10.1093/mnras/stz2472. arXiv:1909.01749.

de Vries, W. H., Becker, R. H., White, R. L., & Loomis, C. (2005). Structure Function Analysis of Long-Term Quasar Variability. *Aj*, *129*, 615–629. doi:10.1086/427393. arXiv:astro-ph/0411348.

Demianski, M., Piedipalumbo, E., Sawant, D., & Amati, L. (2017). Cosmology with gamma-ray bursts. I. The Hubble diagram through the calibrated $E_{p,I}$-E_{iso} correlation. *A&A*, *598*, A112. doi:10.1051/0004-6361/201628909. arXiv:1610.00854.

Di Valentino, E., Melchiorri, A., & Silk, J. (2021). Investigating Cosmic Discordance. *Astrophys. J. Lett.*, *908*, L9. doi:10.3847/2041-8213/abe1c4. arXiv:2003.04935.

Di Valentino, E., Mena, O., Pan, S., Visinelli, L., Yang, W., Melchiorri, A., Mota, D. F., Riess, A. G., & Silk, J. (2021). In the realm of the Hubble tension-a review of solutions. *Classical and Quantum Gravity*, *38*, 153001. doi:10.1088/1361-6382/ac086d. arXiv:2103.01183.

Dialektopoulos, K., Said, J. L., Mifsud, J., Sultana, J., & Zarb Adami, K. (2022). Neural network reconstruction of late-time cosmology and null tests. *jcap*, *2022*, 023. doi:10.1088/1475-7516/2022/02/023. arXiv:2111.11462.

Duras, F., Bongiorno, A., Ricci, F., Piconcelli, E., Shankar, F., Lusso, E., Bianchi, S., Fiore, F., Maiolino, R., Marconi, A., Onori, F., Sani, E., Schneider, R., Vignali, C., & La Franca, F. (2020). Universal bolometric corrections for active galactic nuclei over seven luminosity decades. *A&A*, *636*, A73. doi:10.1051/0004-6361/201936817. arXiv:2001.09984.

Elvis, M., Hao, H., Civano, F., Brusa, M., Salvato, M., Bongiorno, A., Capak, P., Zamorani, G., Comastri, A., Jahnke, K., Lusso, E., Mainieri, V., Trump, J. R., Ho, L. C., Aussel, H., Cappelluti, N., Cisternas, M., Frayer, D., Gilli, R., Hasinger, G., Huchra, J. P., Impey, C. D., Koekemoer, A. M., Lanzuisi, G., Le Floc'h, E., Lilly, S. J., Liu, Y., McCarthy, P., McCracken, H. J., Merloni, A., Roeser, H. J., Sanders, D. B., Sargent, M., Scoville, N., Schinnerer, E., Schiminovich, D., Silverman, J., Taniguchi, Y., Vignali, C., Urry, C. M., Zamojski, M. A., & Zatloukal, M. (2012). Spectral Energy Distributions of Type 1 Active Galactic Nuclei in the COSMOS Survey. I. The XMM-COSMOS Sample. *ApJ*, *759*, 6. doi:10.1088/0004-637X/759/1/6. arXiv:1209.1478.

Fabian, A. C. (2012). Observational Evidence of Active Galactic Nuclei Feedback. *Araa*, *50*, 455–489. doi:10.1146/annurev-astro-081811-125521. arXiv:1204.4114.

Fan, X., Banados, E., & Simcoe, R. A. (2022). Quasars and the Intergalactic Medium at Cosmic Dawn. *arXiv e-prints*, (p. arXiv:2212.06907). doi:10.48550/arXiv.2212.06907. arXiv:2212.06907.

Farina, E. P., Schindler, J.-T., Walter, F., Bañados, E., Davies, F. B., Decarli, R., Eilers, A.-C., Fan, X., Hennawi, J. F., Mazzucchelli, C., Meyer, R. A., Trakhtenbrot, B., Volonteri, M., Wang, F., Worseck, G., Yang, J., Gutcke, T. A., Venemans, B. P., Bosman, S. E. I., Costa, T., De Rosa, G., Drake, A. B., & Onoue, M. (2022). The X-shooter/ALMA Sample of Quasars in the Epoch of Reionization. II. Black Hole Masses, Eddington Ratios, and the Formation of the First Quasars. *ApJ*, *941*, 106. doi:10.3847/1538-4357/ac9626. arXiv:2207.05113.

Fernandez, R., Bryan, G. L., Haiman, Z., & Li, M. (2014). H_2 suppression with shocking inflows: testing a pathway for supermassive black hole formation. *MNRAS*, *439*, 3798–3807. doi:10.1093/mnras/stu230. arXiv:1401.5803.

Ferrarese, L., Côté, P., Dalla Bontà, E., Peng, E. W., Merritt, D., Jordán, A., Blakeslee, J. P., Haşegan, M., Mei, S., Piatek, S., Tonry, J. L., & West, M. J. (2006). A Fundamental Relation between Compact Stellar Nuclei, Supermassive Black Holes, and Their Host Galaxies. *ApJl*, *644*, L21–L24. doi:10.1086/505388. arXiv:astro-ph/0603840.

Fiore, F., Puccetti, S., Grazian, A., Menci, N., Shankar, F., Santini, P., Piconcelli, E., Koekemoer, A. M., Fontana, A., Boutsia, K., Castellano, M., Lamastra, A., Malacaria, C., Feruglio, C., Mathur, S., Miller, N., & Pannella, M. (2012). Faint high-redshift

AGN in the Chandra deep field south: the evolution of the AGN luminosity function and black hole demography. *A&A*, *537*, A16. doi:10.1051/0004-6361/201117581. arXiv:1109.2888.

Fonseca Alvarez, G., Trump, J. R., Homayouni, Y., Grier, C. J., Shen, Y., Horne, K., Li, J. I. H., Brandt, W. N., Ho, L. C., Peterson, B. M., & Schneider, D. P. (2020). The Sloan Digital Sky Survey Reverberation Mapping Project: The Hβ Radius-Luminosity Relation. *Apj*, *899*, 73. doi:10.3847/1538-4357/aba001. arXiv:1910.10719.

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. (2013). emcee: The MCMC Hammer. *pasp*, *125*, 306. doi:10.1086/670067. arXiv:1202.3665.

Fossati, G., Maraschi, L., Celotti, A., Comastri, A., & Ghisellini, G. (1998). A unifying view of the spectral energy distributions of blazars. *MNRAS*, *299*, 433–448. doi:10.1046/j.1365-8711.1998.01828.x. arXiv:astro-ph/9804103.

Freedman, W. L., Madore, B. F., Hatt, D., Hoyt, T. J., Jang, I. S., Beaton, R. L., Burns, C. R., Lee, M. G., Monson, A. J., Neeley, J. R., Phillips, M. M., Rich, J. A., & Seibert, M. (2019). The Carnegie-Chicago Hubble Program. VIII. An Independent Determination of the Hubble Constant Based on the Tip of the Red Giant Branch. *Apj*, *882*, 34. doi:10.3847/1538-4357/ab2f73. arXiv:1907.05922.

Freeman, P., Doe, S., & Siemiginowska, A. (2001). Sherpa: a mission-independent data analysis application. In J.-L. Starck, & F. D. Murtagh (Eds.), *Astronomical Data Analysis* (pp. 76–87). volume 4477 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. doi:10.1117/12.447161. arXiv:astro-ph/0108426.

Fujimoto, S., Ouchi, M., Shibuya, T., & Nagai, H. (2017). Demonstrating a New Census of Infrared Galaxies with ALMA (DANCING-ALMA). I. FIR Size and Luminosity Relation at z = 0-6 Revealed with 1034 ALMA Sources. *ApJ*, *850*, 83. doi:10.3847/1538-4357/aa93e6. arXiv:1703.02138.

Gallagher, R., Maiolino, R., Belfiore, F., Drory, N., Riffel, R., & Riffel, R. A. (2019). Widespread star formation inside galactic outflows. *MNRAS*, *485*, 3409–3429. doi:10.1093/mnras/stz564. arXiv:1806.03311.

Gallagher, S. C., Brandt, W. N., Chartas, G., & Garmire, G. P. (2002). X-Ray Spectroscopy of Quasi-Stellar Objects with Broad Ultraviolet Absorption Lines. *Apj*, *567*, 37–41. doi:10.1086/338485. arXiv:astro-ph/0110579.

Gianolli, V. E., Kim, D. E., Bianchi, S., Agís-González, B., Madejski, G., Marin, F., Marinucci, A., Matt, G., Middei, R., Petrucci, P. O., Soffitta, P., Tagliacozzo, D., Tombesi, F., Ursini, F., Barnouin, T., De Rosa, A., Di Gesu, L., Ingram, A., Loktev, V., Panagiotou, C., Podgorny, J., Poutanen, J., Puccetti, S., Ratheesh, A., Veledina, A., Zhang, W., Agudo, I., Antonelli, L. A., Bachetti, M., Baldini, L., Baumgartner, W. H., Bellazzini, R., Bongiorno, S. D., Bonino, R., Brez, A., Bucciantini, N., Capitanio, F., Castellano, S., Cavazzuti, E., Chen, C. T., Ciprini, S., Costa, E., Del Monte, E., Di Lalla, N., Di Marco, A., Donnarumma, I., Doroshenko, V., Dovčiak, M., Ehlert, S. R., Enoto, T., Evangelista, Y., Fabiani, S., Ferrazzoli, R., García, J. A., Gunji, S., Heyl, J., Iwakiri, W., Jorstad, S. G., Kaaret, P., Karas, V., Kislat, F., Kitaguchi, T., Kolodziejczak, J. J., Krawczynski, H., La Monaca, F., Latronico, L., Liodakis, I., Maldera, S., Manfreda, A., Marscher, A. P., Marshall, H. L., Massaro, F., Mitsuishi, I., Mizuno, T., Muleri, F., Negro, M., Ng, C. Y., O'Dell, S. L., Omodei, N., Oppedisano, C., Papitto, A., Pavlov, G. G., Peirson, A. L., Perri, M., Pesce-Rollins, M., Pilia, M., Possenti, A., Ramsey, B. D., Rankin, J., Roberts, O. J., Romani, R. W., Sgrò, C., Slane, P., Spandre, G., Swartz, D. A., Tamagawa, T., Tavecchio, F., Taverna, R., Tawara, Y. et al. (2023). Uncovering the geometry of the hot X-ray corona in the Seyfert galaxy NGC 4151 with IXPE. *MNRAS*, *523*, 4468–4476. doi:10.1093/mnras/stad1697. arXiv:2303.12541.

Gilli, R., Comastri, A., & Hasinger, G. (2007). The synthesis of the cosmic X-ray background in the Chandra and XMM-Newton era. *A&A*, *463*, 79–96. doi:10.1051/0004-6361:20066334. arXiv:astro-ph/0610939.

Gilli, R., Mignoli, M., Peca, A., Nanni, R., Prandoni, I., Liuzzo, E., D'Amato, Q., Brusa, M., Calura, F., Caminha, G. B., Chiaberge, M., Comastri, A., Cucciati, O., Cusano, F., Grandi, P., Decarli, R., Lanzuisi, G., Mannucci, F., Pinna, E., Tozzi, P., Vanzella, E., Vignali, C., Vito, F., Balmaverde, B., Citro, A., Cappelluti, N., Zamorani, G., & Norman, C. (2019). Discovery of a galaxy overdensity around a powerful, heavily obscured FRII radio galaxy at z = 1.7: star formation promoted by large-scale AGN feedback? *A&A*, *632*, A26. doi:10.1051/0004-6361/201936121. arXiv:1909.00814.

Gilli, R., Norman, C., Calura, F., Vito, F., Decarli, R., Marchesi, S., Iwasawa, K., Comastri, A., Lanzuisi, G., Pozzi, F., D'Amato, Q., Vignali, C., Brusa, M., Mignoli, M., & Cox, P. (2022). Supermassive black holes at high redshift are expected to be obscured by their massive host galaxies' interstellar medium. *A&A*, *666*, A17. doi:10.1051/0004-6361/202243708. arXiv:2206.03508.

GRAVITY Collaboration, Abuter, R., Accardo, M., Amorim, A., Anugu, N., Ávila, G., Azouaoui, N., Benisty, M., Berger, J. P., Blind, N., Bonnet, H., Bourget, P., Brandner, W., Brast, R., Buron, A., Burtscher, L., Cassaing, F., Chapron, F., Choquet, É., Clénet, Y., Collin, C., Coudé Du Foresto, V., de Wit, W., de Zeeuw, P. T., Deen, C., Delplancke-Ströbele, F., Dembet, R., Derie, F., Dexter, J., Duvert, G., Ebert, M., Eckart, A., Eisenhauer, F., Esselborn, M., Fédou, P., Finger, G., Garcia, P., Garcia Dabo, C. E., Garcia Lopez, R., Gendron, E., Genzel, R., Gillessen, S., Gonte, F., Gordo, P., Grould, M., Grözinger, U., Guieu, S., Haguenauer, P., Hans, O., Haubois, X., Haug, M., Haussmann, F., Henning, T., Hippler, S., Horrobin, M., Huber, A., Hubert, Z., Hubin, N., Hummel, C. A., Jakob, G., Janssen, A., Jochum, L., Jocou, L., Kaufer, A., Kellner, S., Kendrew, S., Kern, L., Kervella, P., Kiekebusch, M., Klein, R., Kok, Y., Kolb, J., Kulas, M., Lacour, S., Lapeyrère, V., Lazareff, B., Le Bouquin, J. B., Lèna, P., Lenzen, R., Lévêque, S., Lippa, M., Magnard, Y., Mehrgan, L., Mellein, M., Mérand, A., Moreno-Ventas, J., Moulin, T., Müller, E., Müller, F., Neumann, U., Oberti, S., Ott, T., Pallanca, L., Panduro, J., Pasquini, L., Paumard, T., Percheron, I., Perraut, K., Perrin, G. et al. (2017). First light for GRAVITY: Phase referencing optical interferometry for the Very Large Telescope Interferometer. *A&A*, *602*, A94. doi:10.1051/0004-6361/201730838. arXiv:1705.02345.

Gravity Collaboration, Sturm, E., Dexter, J., Pfuhl, O., Stock, M. R., Davies, R. I., Lutz, D., Clénet, Y., Eckart, A., Eisenhauer, F., Genzel, R., Gratadour, D., Hönig, S. F., Kishimoto, M., Lacour, S., Millour, F., Netzer, H., Perrin, G., Peterson, B. M., Petrucci, P. O., Rouan, D., Waisberg, I., Woillez, J., Amorim, A., Brandner, W., Förster Schreiber, N. M., Garcia, P. J. V., Gillessen, S., Ott, T., Paumard, T., Perraut, K., Scheithauer, S., Straubmeier, C., Tacconi, L. J., & Widmann, F. (2018). Spatially resolved rotation of the broad-line region of a quasar at sub-parsec scale. *Nature*, *563*, 657–660. doi:10.1038/s41586-018-0731-9. arXiv:1811.11195.

Grier, C. J., Martini, P., Watson, L. C., Peterson, B. M., Bentz, M. C., Dasyra, K. M., Dietrich, M., Ferrarese, L., Pogge, R. W., & Zu, Y. (2013). Stellar Velocity Dispersion Measurements in High-luminosity Quasar Hosts and Implications for the AGN Black Hole Mass Scale. *ApJ*, *773*, 90. doi:10.1088/0004-637X/773/2/90. arXiv:1305.2447.

Grier, C. J., Shen, Y., Horne, K., Brandt, W. N., Trump, J. R., Hall, P. B., Kinemuchi, K., Starkey, D., Schneider, D. P., Ho, L. C., Homayouni, Y., I-Hsiu Li, J., McGreer, I. D., Peterson, B. M., Bizyaev, D., Chen, Y., Dawson, K. S., Eftekharzadeh, S., Guo, Y., Jia, S., Jiang, L., Kneib, J.-P., Li, F., Li, Z., Nie, J., Oravetz, A., Oravetz, D., Pan, K., Petitjean, P., Ponder, K. A., Rogerson, J., Vivek, M., Zhang, T., & Zou, H. (2019). The Sloan Digital Sky Survey Reverberation Mapping Project: Initial C IV Lag Results from Four Years of Data. *APJ*, *887*, 38. doi:10.3847/1538-4357/ab4ea5. arXiv:1904.03199.

Gronkiewicz, D., Różańska, A., Petrucci, P.-O., & Belmont, R. (2023). Thermal instability as a constraint for warm X-ray coronas in active galactic nuclei. *A&A*, *675*, A198. doi:10.1051/0004-6361/202244410. arXiv:2303.17384.

Gültekin, K., Richstone, D. O., Gebhardt, K., Lauer, T. R., Tremaine, S., Aller, M. C., Bender, R., Dressler, A., Faber, S. M., Filippenko, A. V., Green, R., Ho, L. C., Kormendy, J., Magorrian, J., Pinkney, J., & Siopis, C. (2009). The M-σ and M-L Relations in Galactic Bulges, and Determinations of Their Intrinsic Scatter. *ApJ*, *698*, 198–221. doi:10.1088/0004-637X/698/1/198. arXiv:0903.4897.

Haardt, F., & Maraschi, L. (1991). A Two-Phase Model for the X-Ray Emission from Seyfert Galaxies. *ApJl*, *380*, L51. doi:10.1086/186171.

Haardt, F., & Maraschi, L. (1993). X-Ray Spectra from Two-Phase Accretion Disks. *ApJ*, *413*, 507. doi:10.1086/173020.

Habouzit, M., Genel, S., Somerville, R. S., Kocevski, D., Hirschmann, M., Dekel, A., Choi, E., Nelson, D., Pillepich, A., Torrey, P., Hernquist, L., Vogelsberger, M., Weinberger, R., & Springel, V. (2019). Linking galaxy structural properties and star formation activity to black hole activity with IllustrisTNG. *MNRAS*, *484*, 4413–4443. doi:10.1093/mnras/stz102. arXiv:1809.05588.

Haiman, Z. (2004). Constraints from Gravitational Recoil on the Growth of Supermassive Black Holes at High Redshift. *Apj*, *613*, 36–40. doi:`10.1086/422910`. `arXiv:astro-ph/0404196`.

Hamuy, M., Phillips, M. M., Suntzeff, N. B., Schommer, R. A., Maza, J., & Aviles, R. (1996). The Absolute Luminosities of the Calan/Tololo Type IA Supernovae. *Aj*, *112*, 2391. doi:`10.1086/118190`. `arXiv:astro-ph/9609059`.

Harrison, C. M. (2017). Impact of supermassive black hole growth on star formation. *Nature Astronomy*, *1*, 0165. doi:`10.1038/s41550-017-0165`. `arXiv:1703.06889`.

Hennawi, J. F., Gladders, M. D., Oguri, M., Dalal, N., Koester, B., Natarajan, P., Strauss, M. A., Inada, N., Kayo, I., Lin, H., Lampeitl, H., Annis, J., Bahcall, N. A., & Schneider, D. P. (2008). A New Survey for Giant Arcs. *ApJ*, *135*, 664–681. doi:`10.1088/0004-6256/135/2/664`. `arXiv:astro-ph/0610061`.

Heymans, C., Tröster, T., Asgari, M., Blake, C., Hildebrandt, H., Joachimi, B., Kuijken, K., Lin, C., Sánchez, A. G., van den Busch, J. L., Wright, A. H., Amon, A., Bilicki, M., de Jong, J., Crocce, M., Dvornik, A., Erben, T., Fortuna, M. C., Getman, F., Giblin, B., Glazebrook, K., Hoekstra, H., Joudaki, S., Kannawadi, A., Köhlinger, F., Lidman, C., Miller, L., Napolitano, N. R., Parkinson, D., Schneider, P., Shan, H., Valentijn, E. A., Verdoes Kleijn, G., & Wolf, C. (2021). Kids-1000 cosmology: Multi-probe weak gravitational lensing and spectroscopic galaxy clustering constraints. *A&A*, *646*, A140. URL: `https://doi.org/10.1051/0004-6361/202039063`. doi:`10.1051/0004-6361/202039063`.

Hickox, R. C., & Alexander, D. M. (2018). Obscured Active Galactic Nuclei. *Annual Review of Astronomy and Astrophysics*, *56*, 625–671. doi:`10.1146/annurev-astro-081817-051803`. `arXiv:1806.04680`.

Hinshaw, G., Larson, D., Komatsu, E., Spergel, D. N., Bennett, C. L., Dunkley, J., Nolta, M. R., Halpern, M., Hill, R. S., Odegard, N., Page, L., Smith, K. M., Weiland, J. L., Gold, B., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Tucker, G. S., Wollack, E., & Wright, E. L. (2013). Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. *Apjs*, *208*, 19. doi:`10.1088/0067-0049/208/2/19`. `arXiv:1212.5226`.

Holsclaw, T., Alam, U., Sansó, B., Lee, H., Heitmann, K., Habib, S., & Higdon, D. (2010). Nonparametric reconstruction of the dark energy equation of state. *Phys. Rev. D*, *82*, 103502. URL: `https://link.aps.org/doi/10.1103/PhysRevD.82.103502`. doi:`10.1103/PhysRevD.82.103502`.

Hook, I. M., McMahon, R. G., Boyle, B. J., & Irwin, M. J. (1994). The variability of optically selected quasars. *MNRAS*, *268*, 305–320. doi:`10.1093/mnras/268.2.305`.

Hoormann, J. K., Martini, P., Davis, T. M., King, A., Lidman, C., Mudd, D., Sharp, R., Sommer, N. E., Tucker, B. E., Yu, Z., Allam, S., Asorey, J., Avila, S., Banerji, M., Brooks, D., Buckley-Geer, E., Burke, D. L., Calcino, J., Carnero Rosell, A., Carollo, D., Carrasco Kind, M., Carretero, J., Castander, F. J., Childress, M., De Vicente, J., Desai, S., Diehl, H. T., Doel, P., Flaugher, B., Fosalba, P., Frieman, J., García-Bellido, J., Gerdes, D. W., Gruen, D., Gutierrez, G., Hartley, W. G., Hinton, S. R., Hollowood, D. L., Honscheid, K., Hoyle, B., James, D. J., Krause, E., Kuehn, K., Kuropatkin, N., Lewis, G. F., Lima, M., Macaulay, E., Maia, M. A. G., Menanteau, F., Miller, C. J., Miquel, R., Möller, A., Plazas, A. A., Romer, A. K., Roodman, A., Sanchez, E., Scarpine, V., Schubnell, M., Serrano, S., Sevilla-Noarbe, I., Smith, M., Smith, R. C., Soares-Santos, M., Sobreira, F., Suchyta, E., Swann, E., Swanson, M. E. C., Tarle, G., Uddin, S. A., & DES Collaboration (2019). C IV black hole mass measurements with the Australian Dark Energy Survey (OzDES). *MNRAS*, *487*, 3650–3663. doi:10.1093/mnras/stz1539. arXiv:1902.04206.

Iwasawa, K., Comastri, A., Vignali, C., Gilli, R., Lanzuisi, G., Brandt, W. N., Tozzi, P., Brusa, M., Carrera, F. J., Ranalli, P., Mainieri, V., Georgantopoulos, I., Puccetti, S., & Paolillo, M. (2020). The xmm deep survey in the cdfs - xi. x-ray spectral properties of 185 bright sources. *A&A*, *639*, A51. URL: https://doi.org/10.1051/0004-6361/201937246. doi:10.1051/0004-6361/201937246.

Iwasawa, K., Gilli, R., Vignali, C., Comastri, A., Brandt, W. N., Ranalli, P., Vito, F., Cappelluti, N., Carrera, F. J., Falocco, S., Georgantopoulos, I., Mainieri, V., & Paolillo, M. (2012). The XMM deep survey in the CDF-S. II. A 9-20 keV selection of heavily obscured active galaxies at z ¿ 1.7. *A&A*, *546*, A84. doi:10.1051/0004-6361/201220036. arXiv:1209.0916.

Kaspi, S., Brandt, W. N., Maoz, D., Netzer, H., Schneider, D. P., & Shemmer, O. (2007). Reverberation Mapping of High-Luminosity Quasars: First Results. *ApJ*, *659*, 997–1007. doi:10.1086/512094. arXiv:astro-ph/0612722.

Kaspi, S., Maoz, D., Netzer, H., Peterson, B. M., Vestergaard, M., & Jannuzi, B. T. (2005). The Relationship between Luminosity and Broad-Line Region Size in Active Galactic Nuclei. *ApJ*, *629*, 61–71. doi:10.1086/431275. arXiv:astro-ph/0504484.

Kelly, B. C., Bechtold, J., & Siemiginowska, A. (2009). Are the Variations in Quasar Optical Flux Driven by Thermal Fluctuations? *ApJ*, *698*, 895–910. doi:10.1088/0004-637X/698/1/895. arXiv:0903.5315.

Koratkar, A. P., & Gaskell, C. M. (1991). Radius-Luminosity and Mass-Luminosity Relationships for Active Galactic Nuclei. *ApJ*, *370*, L61. doi:10.1086/185977.

Kormendy, J., & Ho, L. C. (2013). Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies. *Annual Review of Astronomy and Astrophysics*, *51*, 511–653. doi:10.1146/annurev-astro-082708-101811. arXiv:1304.7762.

Krolik, J. H., & Kallman, T. R. (1988). The Effects of Thermal Accretion Disk Spectra on the Emission Lines from Active Galactic Nuclei. *ApJ*, *324*, 714. doi:10.1086/165931.

La Franca, F., Fiore, F., Comastri, A., Perola, G. C., Sacchi, N., Brusa, M., Cocchia, F., Feruglio, C., Matt, G., Vignali, C., Carangelo, N., Ciliegi, P., Lamastra, A., Maiolino, R., Mignoli, M., Molendi, S., & Puccetti, S. (2005). The HELLAS2XMM Survey. VII. The Hard X-Ray Luminosity Function of AGNs up to z = 4: More Absorbed AGNs at Low Luminosities and High Redshifts. *ApJ*, *635*, 864–879. doi:`10.1086/497586`. `arXiv:astro-ph/0509081`.

Lanzuisi, G., Civano, F., Elvis, M., Salvato, M., Hasinger, G., Vignali, C., Zamorani, G., Aldcroft, T., Brusa, M., Comastri, A., Fiore, F., Fruscione, A., Gilli, R., Ho, L. C., Mainieri, V., Merloni, A., & Siemiginowska, A. (2013). The Chandra-COSMOS survey - IV. X-ray spectra of the bright sample. *MNRAS*, *431*, 978–996. doi:`10.1093/mnras/stt222`. `arXiv:1302.1062`.

Lanzuisi, G., Civano, F., Marchesi, S., Comastri, A., Brusa, M., Gilli, R., Vignali, C., Zamorani, G., Brightman, M., Griffiths, R. E., & Koekemoer, A. M. (2018). The Chandra COSMOS Legacy Survey: Compton thick AGN at high redshift. *MNRAS*, *480*, 2578–2592. doi:`10.1093/mnras/sty2025`. `arXiv:1803.08547`.

Lanzuisi, G., Ponti, G., Salvato, M., Hasinger, G., Cappelluti, N., Bongiorno, A., Brusa, M., Lusso, E., Nandra, P. K., Merloni, A., Silverman, J., Trump, J., Vignali, C., Comastri, A., Gilli, R., Schramm, M., Steinhardt, C., Sanders, D., Kartaltepe, J., Rosario, D., & Trakhtenbrot, B. (2014). Active Galactic Nucleus X-Ray Variability in the XMM-COSMOS Survey. *ApJ*, *781*, 105. doi:`10.1088/0004-637X/781/2/105`. `arXiv:1312.2264`.

Lin, W., & Ishak, M. (2017). Cosmological discordances: A new measure, marginalization effects, and application to geometry versus growth current data sets. *Phys. Rev. D*, *96*, 023532. URL: `https://link.aps.org/doi/10.1103/PhysRevD.96.023532`. doi:`10.1103/PhysRevD.96.023532`.

Liu, T., Tozzi, P., Wang, J.-X., Brandt, W. N., Vignali, C., Xue, Y., Schneider, D. P., Comastri, A., Yang, G., Bauer, F. E., Paolillo, M., Luo, B., Gilli, R., Wang, Q. D., Giavalisco, M., Ji, Z., Alexander, D. M., Mainieri, V., Shemmer, O., Koekemoer, A., & Risaliti, G. (2017). X-Ray Spectral Analyses of AGNs from the 7Ms Chandra Deep Field-South Survey: The Distribution, Variability, and Evolutions of AGN Obscuration. *Apjs*, *232*, 8. doi:`10.3847/1538-4365/aa7847`. `arXiv:1703.00657`.

Lusso, E., Comastri, A., Simmons, B. D., Mignoli, M., Zamorani, G., Vignali, C., Brusa, M., Shankar, F., Lutz, D., Trump, J. R., Maiolino, R., Gilli, R., Bolzonella, M., Puccetti, S., Salvato, M., Impey, C. D., Civano, F., Elvis, M., Mainieri, V., Silverman, J. D., Koekemoer, A. M., Bongiorno, A., Merloni, A., Berta, S., Le Floc'h, E., Magnelli, B., Pozzi, F., & Riguccini, L. (2012). Bolometric luminosities and Eddington ratios of X-ray selected active galactic nuclei in the XMM-COSMOS survey. *MNRAS*, *425*, 623–640. doi:`10.1111/j.1365-2966.2012.21513.x`. `arXiv:1206.2642`.

Lusso, E., Comastri, A., Vignali, C., Zamorani, G., Brusa, M., Gilli, R., Iwasawa, K., Salvato, M., Civano, F., Elvis, M., Merloni, A., Bongiorno, A., Trump, J. R., Koekemoer,

A. M., Schinnerer, E., Le Floc'h, E., Cappelluti, N., Jahnke, K., Sargent, M., Silverman, J., Mainieri, V., Fiore, F., Bolzonella, M., Le Fèvre, O., Garilli, B., Iovino, A., Kneib, J. P., Lamareille, F., Lilly, S., Mignoli, M., Scodeggio, M., & Vergani, D. (2010). The X-ray to optical-UV luminosity ratio of X-ray selected type 1 AGN in XMM-COSMOS. *A&A*, *512*, A34. doi:`10.1051/0004-6361/200913298`. `arXiv:0912.4166`.

Lusso, E., Hennawi, J. F., Comastri, A., Zamorani, G., Richards, G. T., Vignali, C., Treister, E., Schawinski, K., Salvato, M., & Gilli, R. (2013). The Obscured Fraction of Active Galactic Nuclei in the XMM-COSMOS Survey: A Spectral Energy Distribution Perspective. *ApJ*, *777*, 86. doi:`10.1088/0004-637X/777/2/86`. `arXiv:1309.0814`.

Lusso, E., Nardini, E., Bisogni, S., Risaliti, G., Gilli, R., Richards, G. T., Salvestrini, F., Vignali, C., Bargiacchi, G., Civano, F., Elvis, M., Fabbiano, G., Marconi, A., Sacchi, A., & Signorini, M. (2021). The most luminous blue quasars at 3.0 ¡ z ¡ 3.3. II. C IV/X-ray emission and accretion disc physics. *A&A*, *653*, A158. doi:`10.1051/0004-6361/202141356`. `arXiv:2107.02806`.

Lusso, E., Piedipalumbo, E., Risaliti, G., Paolillo, M., Bisogni, S., Nardini, E., & Amati, L. (2019). Tension with the flat ΛCDM model from a high-redshift Hubble diagram of supernovae, quasars, and gamma-ray bursts. *A&A*, *628*, L4. doi:`10.1051/0004-6361/201936223`. `arXiv:1907.07692`.

Lusso, E., & Risaliti, G. (2016). The Tight Relation between X-Ray and Ultraviolet Luminosity of Quasars. *ApJ*, *819*, 154. doi:`10.3847/0004-637X/819/2/154`. `arXiv:1602.01090`.

Lusso, E., & Risaliti, G. (2017). Quasars as standard candles. I. The physical relation between disc and coronal emission. *A&A*, *602*, A79. doi:`10.1051/0004-6361/201630079`. `arXiv:1703.05299`.

Lusso, E., Risaliti, G., Nardini, E., Bargiacchi, G., Benetti, M., Bisogni, S., Capozziello, S., Civano, F., Eggleston, L., Elvis, M., Fabbiano, G., Gilli, R., Marconi, A., Paolillo, M., Piedipalumbo, E., Salvestrini, F., Signorini, M., & Vignali, C. (2020). Quasars as standard candles. III. Validation of a new sample for cosmological studies. *A&A*, *642*, A150. doi:`10.1051/0004-6361/202038899`. `arXiv:2008.08586`.

Macaulay, E., Wehus, I. K., & Eriksen, H. K. (2013). Lower growth rate from recent redshift space distortion measurements than expected from planck. *Phys. Rev. Lett.*, *111*, 161301. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.111.161301`. doi:`10.1103/PhysRevLett.111.161301`.

MacCrann, N., Zuntz, J., Bridle, S., Jain, B., & Becker, M. R. (2015). Cosmic discordance: are Planck CMB and CFHTLenS weak lensing measurements out of tune? *MNRAS*, *451*, 2877–2888. doi:`10.1093/mnras/stv1154`. `arXiv:1408.4742`.

Mainieri, V., Hasinger, G., Cappelluti, N., Brusa, M., Brunner, H., Civano, F., Comastri, A., Elvis, M., Finoguenov, A., Fiore, F., Gilli, R., Lehmann, I., Silverman, J., Tasca, L., Vignali, C., Zamorani, G., Schinnerer, E., Impey, C., Trump, J., Lilly, S., Maier, C., Griffiths, R. E., Miyaji, T., Capak, P., Koekemoer, A., Scoville, N., Shopbell, P., & Taniguchi, Y. (2007). The XMM-Newton Wide-Field Survey in the COSMOS Field. IV. X-Ray Spectral Properties of Active Galactic Nuclei. *ApJs*, *172*, 368–382. doi:10.1086/516573. arXiv:astro-ph/0612361.

Maiolino, R., Russell, H. R., Fabian, A. C., Carniani, S., Gallagher, R., Cazzoli, S., Arribas, S., Belfiore, F., Bellocchi, E., Colina, L., Cresci, G., Ishibashi, W., Marconi, A., Mannucci, F., Oliva, E., & Sturm, E. (2017). Star formation inside a galactic outflow. *Nature*, *544*, 202–206. doi:10.1038/nature21677. arXiv:1703.08587.

Marchesi, S., Ajello, M., Marcotulli, L., Comastri, A., Lanzuisi, G., & Vignali, C. (2018). Compton-thick AGNs in the NuSTAR Era. *ApJ*, *854*, 49. doi:10.3847/1538-4357/aaa410. arXiv:1801.03166.

Marchesi, S., Gilli, R., Lanzuisi, G., Dauser, T., Ettori, S., Vito, F., Cappelluti, N., Comastri, A., Mushotzky, R., Ptak, A., & Norman, C. (2020). Mock catalogs for the extragalactic X-ray sky: Simulating AGN surveys with ATHENA and with the AXIS probe. *A&A*, *642*, A184. doi:10.1051/0004-6361/202038622. arXiv:2008.09133.

Marchesi, S., Lanzuisi, G., Civano, F., Iwasawa, K., Suh, H., Comastri, A., Zamorani, G., Allevato, V., Griffiths, R., Miyaji, T., Ranalli, P., Salvato, M., Schawinski, K., Silverman, J., Treister, E., Urry, C. M., & Vignali, C. (2016). The Chandra COSMOS-Legacy Survey: Source X-Ray Spectral Properties. *ApJ*, *830*, 100. doi:10.3847/0004-637X/830/2/100. arXiv:1608.05149.

Marchesi, S., Mignoli, M., Gilli, R., Mazzolari, G., Signorini, M., Brienza, M., Bisogni, S., Bolzonella, M., Cucciati, O., D'Amato, Q., Peca, A., Prandoni, I., Tozzi, P., Vignali, C., Vito, F., & Comastri, A. (2023). LBT-MODS spectroscopy of high-redshift candidates in the Chandra J1030 field. A newly discovered z~2.8 large scale structure. *arXiv e-prints*, (p. arXiv:2303.13575). doi:10.48550/arXiv.2303.13575. arXiv:2303.13575.

Marchesi, S., Mignoli, M., Gilli, R., Peca, A., Bolzonella, M., Nanni, R., Annunziatella, M., Balmaverde, B., Brusa, M., Calura, F., Cassarà, L. P., Chiaberge, M., Comastri, A., Cusano, F., D´Amato, Q., Iwasawa, K., Lanzuisi, G., Marchesini, D., Morishita, T., Prandoni, I., Rossi, A., Tozzi, P., Vignali, C., Vito, F., Zamorani, G., & Norman, C. (2021). Redshift identification of x-ray-selected active galactic nuclei in the j1030 field: searching for large-scale structures and high-redshift sources. *A&A*, *656*, A117. URL: https://doi.org/10.1051/0004-6361/202141416. doi:10.1051/0004-6361/202141416.

Marconi, A., & Hunt, L. K. (2003). The Relation between Black Hole Mass, Bulge Mass, and Near-Infrared Luminosity. *ApJl*, *589*, L21–L24. doi:10.1086/375804. arXiv:astro-ph/0304274.

Markowitz, A., & Edelson, R. (2004). An Expanded Rossi X-Ray Timing Explorer Survey of X-Ray Variability in Seyfert 1 Galaxies. *ApJ*, *617*, 939–965. doi:10.1086/425559. arXiv:astro-ph/0409045.

McConnell, N. J., & Ma, C.-P. (2013). Revisiting the Scaling Relations of Black Hole Masses and Host Galaxy Properties. *ApJ*, *764*, 184. doi:10.1088/0004-637X/764/2/184. arXiv:1211.2816.

Menzel, M. L., Merloni, A., Georgakakis, A., Salvato, M., Aubourg, E., Brandt, W. N., Brusa, M., Buchner, J., Dwelly, T., Nandra, K., Pâris, I., Petitjean, P., & Schwope, A. (2016). A spectroscopic survey of X-ray-selected AGNs in the northern XMM-XXL field. *MNRAS*, *457*, 110–132. doi:10.1093/mnras/stv2749. arXiv:1511.07870.

Meyer, F., Liu, B. F., & Meyer-Hofmeister, E. (2000). Evaporation: The change from accretion via a thin disk to a coronal flow. *A&A*, *361*, 175–188. arXiv:astro-ph/0007091.

Mortlock, D. J., Warren, S. J., Venemans, B. P., Patel, M., Hewett, P. C., McMahon, R. G., Simpson, C., Theuns, T., Gonzáles-Solares, E. A., Adamson, A., Dye, S., Hambly, N. C., Hirst, P., Irwin, M. J., Kuiper, E., Lawrence, A., & Röttgering, H. J. A. (2011). A luminous quasar at a redshift of z = 7.085. *Nature Astronomy*, *474*, 616–619. doi:10.1038/nature10159. arXiv:1106.6088.

Mushotzky, R., Aird, J., Barger, A. J., Cappelluti, N., Chartas, G., Corrales, L., Eufrasio, R., Fabian, A. C., Falcone, A. D., Gallo, E., Gilli, R., Grant, C. E., Hardcastle, M., Hodges-Kluck, E., Kara, E., Koss, M., Li, H., Lisse, C. M., Loewenstein, M., Markevitch, M., Meyer, E. T., Miller, E. D., Mulchaey, J., Petre, R., Ptak, A. J., Reynolds, C. S., Russell, H. R., Safi-Harb, S., Smith, R. K., Snios, B., Tombesi, F., Valencic, L., Walker, S. A., Williams, B. J., Winter, L. M., Yamaguchi, H., Zhang, W. W., Arenberg, J., Brandt, N., Burrows, D. N., Georganopoulos, M., Miller, J. M., Norman, C. A., & Rosati, P. (2019). The Advanced X-ray Imaging Satellite. In *Bulletin of the American Astronomical Society* (p. 107). volume 51. doi:10.48550/arXiv.1903.04083. arXiv:1903.04083.

Nandra, K., Barret, D., Barcons, X., Fabian, A., den Herder, J.-W., Piro, L., Watson, M., Adami, C., Aird, J., Afonso, J. M., Alexander, D., Argiroffi, C., Amati, L., Arnaud, M., Atteia, J.-L., Audard, M., Badenes, C., Ballet, J., Ballo, L., Bamba, A., Bhardwaj, A., Stefano Battistelli, E., Becker, W., De Becker, M., Behar, E., Bianchi, S., Biffi, V., Bîrzan, L., Bocchino, F., Bogdanov, S., Boirin, L., Boller, T., Borgani, S., Borm, K., Bouché, N., Bourdin, H., Bower, R., Braito, V., Branchini, E., Branduardi-Raymont, G., Bregman, J., Brenneman, L., Brightman, M., Brüggen, M., Buchner, J., Bulbul, E., Brusa, M., Bursa, M., Caccianiga, A., Cackett, E., Campana, S., Cappelluti, N., Cappi, M., Carrera, F., Ceballos, M., Christensen, F., Chu, Y.-H., Churazov, E., Clerc, N., Corbel, S., Corral, A., Comastri, A., Costantini, E., Croston, J., Dadina, M., D'Ai, A., Decourchelle, A., Della Ceca, R., Dennerl, K., Dolag, K., Done, C., Dovciak, M., Drake, J., Eckert, D., Edge, A., Ettori, S., Ezoe, Y., Feigelson, E., Fender, R., Feruglio, C., Finoguenov, A., Fiore, F., Galeazzi, M., Gallagher, S., Gandhi, P., Gaspari, M., Gastaldello, F., Georgakakis, A., Georgantopoulos, I., Gilfanov, M., Gitti, M., Gladstone, R., Goosmann, R., Gosset, E., Grosso, N., Guedel, M., Guerrero, M., Haberl, F., Hardcastle, M. et al. (2013). The Hot and Energetic Universe: A White Paper presenting the science theme motivating the Athena+ mission. *arXiv e-prints*, (p. arXiv:1306.2307). doi:10.48550/arXiv.1306.2307. arXiv:1306.2307.

Nanni, R., Gilli, R., Vignali, C., Mignoli, M., Comastri, A., Vanzella, E., Zamorani, G., Calura, F., Lanzuisi, G., Brusa, M., Tozzi, P., Iwasawa, K., Cappi, M., Vito, F., Balmaverde, B., Costa, T., Risaliti, G., Paolillo, M., Prandoni, I., Liuzzo, E., Rosati, P., Chiaberge, M., Caminha, G. B., Sani, E., Cappelluti, N., & Norman, C. (2018). The 500 ks chandra observation of the z = 6.31 qso sdss j1030 + 0524. *A&A*, *614*, A121. URL: https://doi.org/10.1051/0004-6361/201832694. doi:10.1051/0004-6361/201832694.

Nanni, R., Gilli, R., Vignali, C., Mignoli, M., Peca, A., Marchesi, S., Annunziatella, M., Brusa, M., Calura, F., Cappelluti, N., Chiaberge, M., Comastri, A., Iwasawa, K., Lanzuisi, G., Liuzzo, E., Marchesini, D., Prandoni, I., Tozzi, P., Vito, F., Zamorani, G., & Norman, C. (2020). The deep chandra survey in the sdss j1030+0524 field. *A&A*, *637*, A52. URL: https://doi.org/10.1051/0004-6361/202037914. doi:10.1051/0004-6361/202037914.

Nardini, E., Lusso, E., Risaliti, G., Bisogni, S., Civano, F., Elvis, M., Fabbiano, G., Gilli, R., Marconi, A., Salvestrini, F., & Vignali, C. (2019). The most luminous blue quasars at 3.0 ¡ z ¡ 3.3. I. A tale of two X-ray populations. *Aap*, *632*, A109. doi:10.1051/0004-6361/201936911. arXiv:1910.04604.

Netzer, H. (1980). Excitation of MG II and Fe II lines in quasars and Seyfert galaxies. *ApJ*, *236*, 406–418. doi:10.1086/157757.

Ni, Y., Di Matteo, T., Gilli, R., Croft, R. A. C., Feng, Y., & Norman, C. (2020). QSO obscuration at high redshift ($z \gtrsim 7$): predictions from the BLUETIDES simulation. *MNRAS*, *495*, 2135–2151. doi:10.1093/mnras/staa1313. arXiv:1912.03780.

Nunes, R. C., & Vagnozzi, S. (2021). Arbitrating the S8 discrepancy with growth rate measurements from redshift-space distortions. *Monthly Notices of the Royal Astronomical Society*, *505*, 5427–5437. URL: https://doi.org/10.1093/mnras/stab1613. doi:10.1093/mnras/stab1613. arXiv:https://academic.oup.com/mnras/article-pdf/505/4/5427/38846505/stab1613.pdf.

Omukai, K. (2001). Primordial Star Formation under Far-Ultraviolet Radiation. *Apj*, *546*, 635–651. doi:10.1086/318296. arXiv:astro-ph/0011446.

Onken, C. A., & Kollmeier, J. A. (2008). An Improved Method for Using Mg II to Estimate Black Hole Masses in Active Galactic Nuclei. *ApJ*, *689*, L13. doi:10.1086/595746. arXiv:0810.1950.

Pal, I., Stalin, C. S., Chatterjee, R., & Agrawal, V. K. (2023). X-ray polarization observations of IC 4329A with IXPE: Constraining the geometry of the X-ray corona. *arXiv e-prints*, (p. arXiv:2305.09365). doi:10.48550/arXiv.2305.09365. arXiv:2305.09365.

Pancoast, A., Brewer, B. J., & Treu, T. (2014). Modelling reverberation mapping data - I. Improved geometric and dynamical models and comparison with cross-correlation results. *MNRAS*, *445*, 3055–3072. doi:10.1093/mnras/stu1809. arXiv:1407.2941.

Paolillo, M., Papadakis, I., Brandt, W. N., Luo, B., Xue, Y. Q., Tozzi, P., Shemmer, O., Allevato, V., Bauer, F. E., Comastri, A., Gilli, R., Koekemoer, A. M., Liu, T., Vignali, C., Vito, F., Yang, G., Wang, J. X., & Zheng, X. C. (2017). Tracing the accretion history of supermassive black holes through X-ray variability: results from the ChandraDeep Field-South. *MNRAS*, *471*, 4398–4411. doi:10.1093/mnras/stx1761. arXiv:1707.05332.

Park, D., Kelly, B. C., Woo, J.-H., & Treu, T. (2012). Recalibration of the Virial Factor and M $_{BH}$-σ_* Relation for Local Active Galaxies. *ApJ*, *203*, 6. doi:10.1088/0067-0049/203/1/6. arXiv:1209.3773.

Peca, A., Cappelluti, N., Urry, C. M., LaMassa, S., Marchesi, S., Ananna, T. T., Baloković, M., Sanders, D., Auge, C., Treister, E., Powell, M., Turner, T. J., Kirkpatrick, A., & Tian, C. (2023). On the Cosmic Evolution of AGN Obscuration and the X-Ray Luminosity Function: XMM-Newton and Chandra Spectral Analysis of the 31.3 deg^2 Stripe 82X. *ApJ*, *943*, 162. doi:10.3847/1538-4357/acac28. arXiv:2210.08030.

Peca, A., Vignali, C., Gilli, R., Mignoli, M., Nanni, R., Marchesi, S., Bolzonella, M., Brusa, M., Cappelluti, N., Comastri, A., Lanzuisi, G., & Vito, F. (2021). X-Ray Redshifts for Obscured AGN: A Case Study in the J1030 Deep Field. *Apj*, *906*, 90. doi:10.3847/1538-4357/abc9c7. arXiv:2011.05983.

Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E., Hook, I. M., Kim, A. G., Kim, M. Y., Lee, J. C., Nunes, N. J., Pain, R., Pennypacker, C. R., Quimby, R., Lidman, C., Ellis, R. S., Irwin, M., McMahon, R. G., Ruiz-Lapuente, P., Walton, N., Schaefer, B., Boyle, B. J., Filippenko, A. V., Matheson, T., Fruchter, A. S., Panagia, N., Newberg, H. J. M., Couch, W. J., & Project, T. S. C. (1999). Measurements of Ω and Λ from 42 High-Redshift Supernovae. *Apj*, *517*, 565–586. doi:10.1086/307221. arXiv:astro-ph/9812133.

Petrosian, V., Singal, J., & Mutchnick, S. (2022). Can the Distance−Redshift Relation be Determined from Correlations between Luminosities? *ApJl*, *935*, L19. doi:10.3847/2041-8213/ac85ac. arXiv:2205.07981.

Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Bartlett, J. G., Bartolo, N., Battaner, E., Battye, R., Benabed, K., Benoît, A., Benoit-Lévy, A., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bonaldi, A., Bonavera, L., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J. F., Catalano, A., Challinor, A., Chamballu, A., Chary, R. R., Chiang, H. C., Chluba, J., Christensen, P. R., Church, S., Clements, D. L., Colombi, S., Colombo, L. P. L., Combet, C., Coulais, A., Crill, B. P., Curto, A., Cuttaia, F., Danese, L., Davies, R. D., Davis, R. J., de Bernardis, P., de Rosa, A., de Zotti, G., Delabrouille, J., Désert, F. X., Di Valentino, E., Dickinson, C., Diego, J. M., Dolag, K., Dole, H., Donzelli, S., Doré, O., Douspis, M., Ducout, A., Dunkley, J., Dupac, X., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Farhang, M., Fergusson, J., Finelli, F., Forni, O., Frailis, M., Fraisse, A. A., Franceschi, E., Frejsel, A., Galeotta, S., Galli, S., Ganga, K., Gauthier, C., Gerbino, M., Ghosh, T., Giard, M., Giraud-Héraud, Y., Giusarma, E., Gjerløw, E., González-Nuevo, J., Górski, K. M., Gratton, S., Gregorio, A., Gruppuso, A., Gudmundsson, J. E., Hamann, J., Hansen, F. K., Hanson, D., Harrison, D. L., Helou, G. et al. (2016). Planck 2015 results. XIII. Cosmological parameters. *A&A*, *594*, A13. doi:10.1051/0004-6361/201525830. arXiv:1502.01589.

Planck Collaboration, Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Battye, R., Benabed, K., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J. F., Carron, J., Challinor, A., Chiang, H. C., Chluba, J., Colombo, L. P. L., Combet, C., Contreras, D., Crill, B. P., Cuttaia, F., de Bernardis, P., de Zotti, G., Delabrouille, J., Delouis, J. M., Di Valentino, E., Diego, J. M., Doré, O., Douspis, M., Ducout, A., Dupac, X., Dusini, S., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Fantaye, Y., Farhang, M., Fergusson, J., Fernandez-Cobos, R., Finelli, F., Forastieri, F., Frailis, M., Fraisse, A. A., Franceschi, E., Frolov, A., Galeotta, S., Galli, S., Ganga, K., Génova-Santos, R. T., Gerbino, M., Ghosh, T., González-Nuevo, J., Górski, K. M., Gratton, S., Gruppuso, A., Gudmundsson, J. E., Hamann, J., Handley, W., Hansen, F. K., Herranz, D., Hildebrandt, S. R., Hivon, E., Huang, Z., Jaffe, A. H., Jones, W. C., Karakci, A., Keihänen, E., Keskitalo, R., Kiiveri, K., Kim, J., Kisner, T. S., Knox, L., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lagache, G., Lamarre, J. M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Le Jeune, M., Lemos, P., Lesgourgues, J., Levrier, F. et al. (2020). Planck 2018 results. VI. Cosmological parameters. *A&A*, *641*, A6. doi:10.1051/0004-6361/201833910. arXiv:1807.06209.

Ponti, G., Papadakis, I., Bianchi, S., Guainazzi, M., Matt, G., Uttley, P., & Bonilla, N. F. (2012). CAIXA: a catalogue of AGN in the XMM-Newton archive. III. Excess variance analysis. *A&A*, *542*, A83. doi:10.1051/0004-6361/201118326. arXiv:1112.2744.

Rakshit, S., Stalin, C. S., & Kotilainen, J. (2020). Spectral Properties of Quasars from Sloan Digital Sky Survey Data Release 14: The Catalog. *ApJs*, *249*, 17. doi:10.3847/1538-4365/ab99c5. arXiv:1910.10395.

Ricarte, A., Tremmel, M., Natarajan, P., & Quinn, T. (2019). Tracing black hole and galaxy co-evolution in the ROMULUS simulations. *MNRAS*, *489*, 802–819. doi:10.1093/mnras/stz2161. arXiv:1904.10116.

Ricci, C., Koss, M., Trakhtenbrot, B., Bauer, F., Treister, E., Ueda, Y., Schawinski, K., Oh, K., & Lamperti, I. (2017). The relation between nuclear obscuration, galaxy interactions and accretion properties of AGN. In J.-U. Ness, & S. Migliari (Eds.), *The X-ray Universe 2017* (p. 190).

Richards, G. T., Fan, X., Newberg, H. J., Strauss, M. A., Vanden Berk, D. E., Schneider, D. P., Yanny, B., Boucher, A., Burles, S., Frieman, J. A., Gunn, J. E., Hall, P. B., Ivezić, Ž., Kent, S., Loveday, J., Lupton, R. H., Rockosi, C. M., Schlegel, D. J., Stoughton, C., SubbaRao, M., & York, D. G. (2002). Spectroscopic Target Selection in the Sloan Digital Sky Survey: The Quasar Sample. *Aj*, *123*, 2945–2975. doi:`10.1086/340187`. `arXiv:astro-ph/0202251`.

Richards, G. T., Lacy, M., Storrie-Lombardi, L. J., Hall, P. B., Gallagher, S. C., Hines, D. C., Fan, X., Papovich, C., Vanden Berk, D. E., Trammell, G. B., Schneider, D. P., Vestergaard, M., York, D. G., Jester, S., Anderson, S. F., Budavári, T., & Szalay, A. S. (2006). Spectral Energy Distributions and Multiwavelength Selection of Type 1 Quasars. *ApJs*, *166*, 470–497. doi:`10.1086/506525`. `arXiv:astro-ph/0601558`.

Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. (2019). Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond ΛCDM. *ApJ*, *876*, 85. doi:`10.3847/1538-4357/ab1422`. `arXiv:1903.07603`.

Riess, A. G., Macri, L., Casertano, S., Lampeitl, H., Ferguson, H. C., Filippenko, A. V., Jha, S. W., Li, W., & Chornock, R. (2011). A 3% Solution: Determination of the Hubble Constant with the Hubble Space Telescope and Wide Field Camera 3. *ApJ*, *730*, 119. doi:`10.1088/0004-637X/730/2/119`. `arXiv:1103.2976`.

Riess, A. G., Macri, L., Casertano, S., Sosey, M., Lampeitl, H., Ferguson, H. C., Filippenko, A. V., Jha, S. W., Li, W., Chornock, R., & Sarkar, D. (2009). A Redetermination of the Hubble Constant with the Hubble Space Telescope from a Differential Distance Ladder. *ApJ*, *699*, 539–563. doi:`10.1088/0004-637X/699/1/539`. `arXiv:0905.0695`.

Riess, A. G., Macri, L. M., Hoffmann, S. L., Scolnic, D., Casertano, S., Filippenko, A. V., Tucker, B. E., Reid, M. J., Jones, D. O., Silverman, J. M., Chornock, R., Challis, P., Yuan, W., Brown, P. J., & Foley, R. J. (2016). A 2.4% Determination of the Local Value of the Hubble Constant. *ApJ*, *826*, 56. doi:`10.3847/0004-637X/826/1/56`. `arXiv:1604.01424`.

Risaliti, G., & Lusso, E. (2015). A Hubble Diagram for Quasars. *ApJ*, *815*, 33. doi:`10.1088/0004-637X/815/1/33`. `arXiv:1505.07118`.

Risaliti, G., & Lusso, E. (2019). Cosmological Constraints from the Hubble Diagram of Quasars at High Redshifts. *Nature Astronomy*, *3*, 272–277. doi:`10.1038/s41550-018-0657-z`. `arXiv:1811.02590`.

Ross, N. P., McGreer, I. D., White, M., Richards, G. T., Myers, A. D., Palanque-Delabrouille, N., Strauss, M. A., Anderson, S. F., Shen, Y., Brandt, W. N., Yèche, C., Swanson, M. E. C., Aubourg, É., Bailey, S., Bizyaev, D., Bovy, J., Brewington, H., Brinkmann, J., DeGraf, C., Di Matteo, T., Ebelke, G., Fan, X., Ge, J., Malanushenko, E., Malanushenko, V., Mandelbaum, R., Maraston, C., Muna, D., Oravetz, D., Pan, K., Pâris, I., Petitjean, P., Schawinski, K., Schlegel, D. J., Schneider, D. P., Silverman, J. D., Simmons, A., Snedden, S., Streblyanska, A., Suzuki, N., Weinberg, D. H., & York, D. (2013). The SDSS-III Baryon Oscillation Spectroscopic Survey: The Quasar Luminosity Function from Data Release Nine. *ApJ*, *773*, 14. doi:`10.1088/0004-637X/773/1/14`. `arXiv:1210.6389`.

Sacchi, A., Risaliti, G., Signorini, M., Lusso, E., Nardini, E., Bargiacchi, G., Bisogni, S., Civano, F., Elvis, M., Fabbiano, G., Gilli, R., Trefoloni, B., & Vignali, C. (2022). Quasars as high-redshift standard candles. *A&A*, *663*, L7. doi:10.1051/0004-6361/202243411. arXiv:2206.13528.

Salvestrini, F., Risaliti, G., Bisogni, S., Lusso, E., & Vignali, C. (2019). Quasars as standard candles II. The non-linear relation between UV and X-ray emission at high redshifts. *A&A*, *631*, A120. doi:10.1051/0004-6361/201935491. arXiv:1909.12309.

Sanders, D. B., Phinney, E. S., Neugebauer, G., Soifer, B. T., & Matthews, K. (1989). Continuum Energy Distributions of Quasars: Shapes and Origins. *ApJ*, *347*, 29. doi:10.1086/168094.

Scolnic, D., Brout, D., Carr, A., Riess, A. G., Davis, T. M., Dwomoh, A., Jones, D. O., Ali, N., Charvu, P., Chen, R., Peterson, E. R., Popovic, B., Rose, B. M., Wood, C. M., Brown, P. J., Chambers, K., Coulter, D. A., Dettman, K. G., Dimitriadis, G., Filippenko, A. V., Foley, R. J., Jha, S. W., Kilpatrick, C. D., Kirshner, R. P., Pan, Y.-C., Rest, A., Rojas-Bravo, C., Siebert, M. R., Stahl, B. E., & Zheng, W. (2022). The Pantheon+ Analysis: The Full Data Set and Light-curve Release. *ApJ*, *938*, 113. doi:10.3847/1538-4357/ac8b7a. arXiv:2112.03863.

Scolnic, D. M., Jones, D. O., Rest, A., Pan, Y. C., Chornock, R., Foley, R. J., Huber, M. E., Kessler, R., Narayan, G., Riess, A. G., Rodney, S., Berger, E., Brout, D. J., Challis, P. J., Drout, M., Finkbeiner, D., Lunnan, R., Kirshner, R. P., Sanders, N. E., Schlafly, E., Smartt, S., Stubbs, C. W., Tonry, J., Wood-Vasey, W. M., Foley, M., Hand, J., Johnson, E., Burgett, W. S., Chambers, K. C., Draper, P. W., Hodapp, K. W., Kaiser, N., Kudritzki, R. P., Magnier, E. A., Metcalfe, N., Bresolin, F., Gall, E., Kotak, R., McCrum, M., & Smith, K. W. (2018). The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample. *ApJ*, *859*, 101. doi:10.3847/1538-4357/aab9bb. arXiv:1710.00845.

Scoville, N., Lee, N., Vanden Bout, P., Diaz-Santos, T., Sanders, D., Darvish, B., Bongiorno, A., Casey, C. M., Murchikova, L., Koda, J., Capak, P., Vlahakis, C., Ilbert, O., Sheth, K., Morokuma-Matsui, K., Ivison, R. J., Aussel, H., Laigle, C., McCracken, H. J., Armus, L., Pope, A., Toft, S., & Masters, D. (2017). Evolution of Interstellar Medium, Star Formation, and Accretion at High Redshift. *ApJ*, *837*, 150. doi:10.3847/1538-4357/aa61a0. arXiv:1702.04729.

Seikel, M., Clarkson, C., & Smith, M. (2012). Reconstruction of dark energy and expansion dynamics using gaussian processes. *arXiv: Cosmology and Nongalactic Astrophysics*, .

Seyfert, C. K. (1943). Nuclear Emission in Spiral Nebulae. *Apj*, *97*, 28. doi:10.1086/144488.

Shaban, F., Siemiginowska, A., Suleiman, R. M., El-Nawawy, M. S., & Ali, A. (2022). X-ray properties of high-redshift Radio Loud and Radio Quiet Quasars observed by Chandra. *Journal of High Energy Astrophysics*, *36*, 152–161. doi:10.1016/j.jheap.2022.10.002. arXiv:2301.02866.

Shafieloo, A., Kim, A. G., & Linder, E. V. (2012). Gaussian process cosmography. *Phys. Rev. D*, *85*, 123530. URL: https://link.aps.org/doi/10.1103/PhysRevD.85.123530. doi:10.1103/PhysRevD.85.123530.

Shakura, N. I., & Sunyaev, R. A. (1973). Black holes in binary systems. Observational appearance. *A&A*, *24*, 337–355.

Sharon, K., Bayliss, M. B., Dahle, H., Florian, M. K., Gladders, M. D., Johnson, T. L., Paterno-Mahler, R., Rigby, J. R., Whitaker, K. E., & Wuyts, E. (2017). Lens Model and Time Delay Predictions for the Sextuply Lensed Quasar SDSS J2222+2745. *ApJ*, *835*, 5. doi:10.3847/1538-4357/835/1/5. arXiv:1609.08848.

Shen, Y., & Kelly, B. C. (2012). The Demographics of Broad-line Quasars in the Mass-Luminosity Plane. I. Testing FWHM-based Virial Black Hole Masses. *ApJ*, *746*, 169. doi:10.1088/0004-637X/746/2/169. arXiv:1107.4372.

Shen, Y., Richards, G. T., Strauss, M. A., Hall, P. B., Schneider, D. P., Snedden, S., Bizyaev, D., Brewington, H., Malanushenko, V., Malanushenko, E., Oravetz, D., Pan, K., & Simmons, A. (2011). A Catalog of Quasar Properties from Sloan Digital Sky Survey Data Release 7. *ApJ*, *194*, 45. doi:10.1088/0067-0049/194/2/45. arXiv:1006.5178.

Sicilian, D., Civano, F., Cappelluti, N., Buchner, J., & Peca, A. (2022). X-Ray Redshifts of Obscured Chandra Source Catalog Active Galactic Nuclei. *ApJ*, *936*, 39. doi:10.3847/1538-4357/ac82f4. arXiv:2203.13825.

Silk, J. (2013). Unleashing Positive Feedback: Linking the Rates of Star Formation, Supermassive Black Hole Accretion, and Outflows in Distant Galaxies. *Apj*, *772*, 112. doi:10.1088/0004-637X/772/2/112. arXiv:1305.5840.

Silk, J., & Rees, M. J. (1998). Quasars and galaxy formation. *Aap*, *331*, L1–L4. doi:10.48550/arXiv.astro-ph/9801013. arXiv:astro-ph/9801013.

Simmonds, C., Buchner, J., Salvato, M., Hsu, L. T., & Bauer, F. E. (2018). XZ: Deriving redshifts from X-ray spectra of obscured AGN. *A&A*, *618*, A66. doi:10.1051/0004-6361/201833412. arXiv:1807.01782.

Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., & Hernquist, L. (2008). A semi-analytic model for the co-evolution of galaxies, black holes and active galactic nuclei. *MNRAS*, *391*, 481–506. doi:10.1111/j.1365-2966.2008.13805.x. arXiv:0808.1227.

Steffen, A. T., Strateva, I., Brandt, W. N., Alexander, D. M., Koekemoer, A. M., Lehmer, B. D., Schneider, D. P., & Vignali, C. (2006). The X-Ray-to-Optical Properties of Optically Selected Active Galaxies over Wide Luminosity and Redshift Ranges. *ApJ*, *131*, 2826–2842. doi:10.1086/503627. arXiv:astro-ph/0602407.

Svensson, R. (1982). The pair annihilation process in relativistic plasmas. *ApJ*, *258*, 321–334. doi:10.1086/160081.

Svensson, R., & Zdziarski, A. A. (1994). Black hole accretion disks with coronae. *ApJ*, *436*, 599–606. doi:10.1086/174934.

Taak, Y. C., & Treu, T. (2023). Strong lensed QSOs with variability detectable by LSST: How many are there? *MNRAS*, *524*, 5446–5453. doi:10.1093/mnras/stad2201. arXiv:2304.02784.

Tacconi, L. J., Genzel, R., Saintonge, A., Combes, F., García-Burillo, S., Neri, R., Bolatto, A., Contini, T., Förster Schreiber, N. M., Lilly, S., Lutz, D., Wuyts, S., Accurso, G., Boissier, J., Boone, F., Bouché, N., Bournaud, F., Burkert, A., Carollo, M., Cooper, M., Cox, P., Feruglio, C., Freundlich, J., Herrera-Camus, R., Juneau, S., Lippa, M., Naab, T., Renzini, A., Salome, P., Sternberg, A., Tadaki, K., Übler, H., Walter, F., Weiner, B., & Weiss, A. (2018). PHIBSS: Unified Scaling Relations of Gas Depletion Time and Molecular Gas Fractions. *ApJ*, *853*, 179. doi:10.3847/1538-4357/aaa4b4. arXiv:1702.01140.

Tanaka, T. L. (2014). Driving the growth of the earliest supermassive black holes with major mergers of host galaxies. *Classical and Quantum Gravity*, *31*, 244005. doi:10.1088/0264-9381/31/24/244005. arXiv:1406.3023.

Tananbaum, H., Avni, Y., Branduardi, G., Elvis, M., Fabbiano, G., Feigelson, E., Giacconi, R., Henry, J. P., Pye, J. P., Soltan, A., & Zamorani, G. (1979). X-ray studies of quasars with the Einstein Observatory. *ApJl*, *234*, L9–L13. doi:10.1086/183100.

Theureau, G., Coudreau, N., Hallet, N., Hanski, M., Alsac, L., Bottinelli, L., Gouguenheim, L., Martin, J. M., & Paturel, G. (2005). Kinematics of the local universe . XII. 21-cm line measurements of 586 galaxies with the new Nançay receiver. *A&A*, *430*, 373–383. doi:10.1051/0004-6361:20047152.

Torres-Albà, N., Marchesi, S., Zhao, X., Ajello, M., Silver, R., Ananna, T. T., Baloković, M., Boorman, P. B., Comastri, A., Gilli, R., Lanzuisi, G., Murphy, K., Urry, C. M., & Vignali, C. (2021). Compton-thick AGN in the NuSTAR Era VI: The Observed Compton-thick Fraction in the Local Universe. *ApJ*, *922*, 252. doi:10.3847/1538-4357/ac1c73. arXiv:2109.00599.

Tozzi, P., Gilli, R., Mainieri, V., Norman, C., Risaliti, G., Rosati, P., Bergeron, J., Borgani, S., Giacconi, R., Hasinger, G., Nonino, M., Streblyanska, A., Szokoly, G., Wang, J. X., & Zheng, W. (2006). X-ray spectral properties of active galactic nuclei in the Chandra Deep Field South. *A&A*, *451*, 457–474. doi:10.1051/0004-6361:20042592. arXiv:astro-ph/0602127.

Trebitsch, M., Volonteri, M., & Dubois, Y. (2019). Black hole obscuration and duty-cycles mediated by AGN feedback in high-redshift galaxies. *MNRAS*, *487*, 819–831. doi:10.1093/mnras/stz1280. arXiv:1901.01261.

Treister, E., & Urry, C. M. (2006). The Evolution of Obscuration in Active Galactic Nuclei. *ApJl*, *652*, L79–L82. doi:10.1086/510237. arXiv:astro-ph/0610525.

Trevese, D., Perna, M., Vagnetti, F., Saturni, F. G., & Dadina, M. (2014). C IV and C III] Reverberation Mapping of the Luminous Quasar PG 1247+267. *ApJ*, *795*, 164. doi:10.1088/0004-637X/795/2/164. arXiv:1409.5448.

Turner, E. L. (1991). Quasars and Galaxy Formation. I. The Z ¿ 4 Objects. *Aj*, *101*, 5. doi:10.1086/115663.

Ueda, Y., Akiyama, M., Hasinger, G., Miyaji, T., & Watson, M. G. (2014). Toward the Standard Population Synthesis Model of the X-Ray Background: Evolution of X-Ray Luminosity and Absorption Functions of Active Galactic Nuclei Including Compton-thick Populations. *ApJ*, *786*, 104. doi:10.1088/0004-637X/786/2/104. arXiv:1402.1836.

Ueda, Y., Eguchi, S., Terashima, Y., Mushotzky, R., Tueller, J., Markwardt, C., Gehrels, N., Hashimoto, Y., & Potter, S. (2007). Suzaku Observations of Active Galactic Nuclei Detected in the Swift BAT Survey: Discovery of a "New Type" of Buried Supermassive Black Holes. *ApJl*, *664*, L79–L82. doi:10.1086/520576. arXiv:0706.1168.

Urry, C. M., & Padovani, P. (1995). Unified Schemes for Radio-Loud Active Galactic Nuclei. *Publications of the Astronomical Society of the Pacific*, *107*, 803. doi:10.1086/133630. arXiv:astro-ph/9506063.

Vagnetti, F., Turriziani, S., & Trevese, D. (2011). Ensemble X-ray variability of active galactic nuclei from serendipitous source catalogues. *A&A*, *536*, A84. doi:10.1051/0004-6361/201118072. arXiv:1110.4768.

Vanden Berk, D., Wilhite, B., Kron, R., Ivezic, Z., Pereyra, N., & SDSS Collaboration (2004). Photometric and Spectroscopic Variability of Quasars. In *American Astronomical Society Meeting Abstracts* (p. 120.02). volume 205 of *American Astronomical Society Meeting Abstracts*.

Vestergaard, M., & Peterson, B. M. (2006). Determining Central Black Hole Masses in Distant Active Galaxies and Quasars. II. Improved Optical and UV Scaling Relationships. *ApJ*, *641*, 689–709. doi:10.1086/500572. arXiv:astro-ph/0601303.

Vignali, C., Brandt, W. N., & Schneider, D. P. (2003). X-Ray Emission from Radio-Quiet Quasars in the Sloan Digital Sky Survey Early Data Release: The α_{ox} Dependence upon Ultraviolet Luminosity. *ApJ*, *125*, 433–443. doi:10.1086/345973. arXiv:astro-ph/0211125.

Vito, F., Brandt, W. N., Bauer, F. E., Calura, F., Gilli, R., Luo, B., Shemmer, O., Vignali, C., Zamorani, G., Brusa, M., Civano, F., Comastri, A., & Nanni, R. (2019). The X-ray properties of z ¿ 6 quasars: no evident evolution of accretion physics in the first Gyr of the Universe. *A&A*, *630*, A118. doi:10.1051/0004-6361/201936217. arXiv:1908.09849.

Vito, F., Brandt, W. N., Yang, G., Gilli, R., Luo, B., Vignali, C., Xue, Y. Q., Comastri, A., Koekemoer, A. M., Lehmer, B. D., Liu, T., Paolillo, M., Ranalli, P., Schneider, D. P., Shemmer, O., Volonteri, M., & Wang, J. (2018). High-redshift AGN in the Chandra Deep Fields: the obscured fraction and space density of the sub-L_* population. *MNRAS*, *473*, 2378–2406. doi:10.1093/mnras/stx2486. arXiv:1709.07892.

Webb, N. A., Coriat, M., Traulsen, I., Ballet, J., Motch, C., Carrera, F. J., Koliopanos, F., Authier, J., de la Calle, I., Ceballos, M. T., Colomo, E., Chuard, D., Freyberg, M., Garcia, T., Kolehmainen, M., Lamer, G., Lin, D., Maggi, P., Michel, L., Page, C. G., Page, M. J., Perea-Calderon, J. V., Pineau, F. X., Rodriguez, P., Rosen, S. R., Santos Lleo, M., Saxton, R. D., Schwope, A., Tomás, L., Watson, M. G., & Zakardjian, A. (2020). The XMM-Newton serendipitous survey. IX. The fourth XMM-Newton serendipitous source catalogue. *A&A*, *641*, A136. doi:10.1051/0004-6361/201937353. arXiv:2007.02899.

Williams, P. R., Treu, T., Dahle, H., Valenti, S., Abramson, L., Barth, A. J., Dyrland, K., Gladders, M., Horne, K., & Sharon, K. (2021). The Black Hole Mass of the z = 2.805 Multiply Imaged Quasar SDSS J2222+2745 from Velocity-resolved Time Lags of the C IV Emission Line. *ApJ*, *911*, 64. doi:10.3847/1538-4357/abe943. arXiv:2011.02007.

Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, *22*, 209–212. URL: https://www.tandfonline.com/doi/abs/10.1080/01621459.1927.10502953. doi:10.1080/01621459.1927.10502953. arXiv:https://www.tandfonline.com/doi/pdf/10.1080/01621459.1927.10502953.

Wise, J. H., Regan, J. A., O'Shea, B. W., Norman, M. L., Downes, T. P., & Xu, H. (2019). Formation of massive black holes in rapidly growing pre-galactic gas clouds. *Nat*, *566*, 85–88. doi:10.1038/s41586-019-0873-4. arXiv:1901.07563.

Wong, K. C., Suyu, S. H., Chen, G., Rusu, C. E., Millon, M., Sluse, D., Bonvin, V., Fassnacht, C. D., Taubenberger, S., Auger, M. W., Birrer, S., Chan, J. H. H., Courbin, F., Hilbert, S., Tihhonova, O., Treu, T., Agnello, A., Ding, X., Jee, I., Komatsu, E., Shajib, A. J., Sonnenfeld, A., Blandford, R. D., Koopmans, L. V. E., Marshall, P. J., & Meylan, G. (2019). H0LiCOW – XIII. A 2.4 per cent measurement of H0 from lensed quasars: 5.3sigma tension between early- and late-Universe probes. *Monthly Notices of the Royal Astronomical Society*, *498*, 1420–1439. URL: https://doi.org/10.1093/mnras/stz3094. doi:10.1093/mnras/stz3094. arXiv:https://academic.oup.com/mnras/article-pdf/498/1/1420/33755111/stz3094.pdf.

Woo, J.-H., Le, H. A. N., Karouzos, M., Park, D., Park, D., Malkan, M. A., Treu, T., & Bennert, V. N. (2018). Calibration and Limitations of the Mg II Line-based Black Hole Masses. *ApJ*, *859*, 138. doi:10.3847/1538-4357/aabf3e. arXiv:1804.02798.

Worrall, D. M., Giommi, P., Tananbaum, H., & Zamorani, G. (1987). X-Ray Studies of Quasars with the Einstein Observatory. IV. X-Ray Dependence on Radio Emission. *Apj*, *313*, 596. doi:10.1086/164999.

Wu, Q., & Shen, Y. (2022). A Catalog of Quasar Properties from Sloan Digital Sky Survey Data Release 16. *arXiv e-prints*, (p. arXiv:2209.03987). arXiv:2209.03987.

Wu, X.-B., Wang, F., Fan, X., Yi, W., Zuo, W., Bian, F., Jiang, L., McGreer, I. D., Wang, R., Yang, J., Yang, Q., Thompson, D., & Beletsky, Y. (2015). An ultraluminous quasar with a twelve-billion-solar-mass black hole at redshift 6.30. *Nature Astronomy*, *518*, 512–515. doi:10.1038/nature14241. arXiv:1502.07418.

Xue, Y. Q. (2017). The Chandra deep fields: Lifting the veil on distant active galactic nuclei and X-ray emitting galaxies. *New Astronomy Reviews*, *79*, 59–84. doi:10.1016/j.newar.2017.09.002. arXiv:1709.04601.

Yoshida, N., Abel, T., Hernquist, L., & Sugiyama, N. (2003). Simulations of Early Structure Formation: Primordial Gas Clouds. *Apj*, *592*, 645–663. doi:10.1086/375810. arXiv:astro-ph/0301645.

Young, M., Elvis, M., & Risaliti, G. (2009). The Fifth Data Release Sloan Digital Sky Survey/XMM-Newton Quasar Survey. *Apj*, *183*, 17–32. doi:10.1088/0067-0049/183/1/17. arXiv:0905.0496.

Young, M., Elvis, M., & Risaliti, G. (2010). The X-ray Energy Dependence of the Relation Between Optical and X-ray Emission in Quasars. *ApJ*, *708*, 1388–1397. doi:10.1088/0004-637X/708/2/1388. arXiv:0911.0474.

Zhang, Z.-X., Du, P., Smith, P. S., Zhao, Y., Hu, C., Xiao, M., Li, Y.-R., Huang, Y.-K., Wang, K., Bai, J.-M., Ho, L. C., & Wang, J.-M. (2019). Kinematics of the Broad-line Region of 3C 273 from a 10 yr Reverberation Mapping Campaign. *ApJ*, *876*, 49. doi:10.3847/1538-4357/ab1099. arXiv:1811.03812.

Zubovas, K., Nayakshin, S., King, A., & Wilkinson, M. (2013). AGN outflows trigger starbursts in gas-rich galaxies. *MNRAS*, *433*, 3079–3090. doi:10.1093/mnras/stt952. arXiv:1306.0684.