



<b>Publication Year</b>	2021
<b>Acceptance in OA @INAF</b>	2022-06-07T14:06:41Z
<b>Title</b>	The Chandra view of the relation between X-ray and UV emission in quasars
<b>Authors</b>	Bisogni, Susanna; Lusso, E.; Civano, F.; NARDINI, EMANUELE; Risaliti, G.; et al.
<b>DOI</b>	10.1051/0004-6361/202140852
<b>Handle</b>	<a href="http://hdl.handle.net/20.500.12386/32221">http://hdl.handle.net/20.500.12386/32221</a>
<b>Journal</b>	ASTRONOMY & ASTROPHYSICS
<b>Number</b>	655

# The *Chandra* view of the relation between X-ray and UV emission in quasars<sup>★</sup>

S. Bisogni<sup>1,2</sup> , E. Lusso<sup>3,4</sup>, F. Civano<sup>2</sup>, E. Nardini<sup>3,4</sup>, G. Risaliti<sup>3,4</sup>, M. Elvis<sup>2</sup>, and G. Fabbiano<sup>2</sup>

<sup>1</sup> INAF – Istituto di Astrofisica Spaziale e Fisica Cosmica Milano, Via Corti 12, 20133 Milano, Italy  
e-mail: [susanna.bisogni@inaf.it](mailto:susanna.bisogni@inaf.it)

<sup>2</sup> Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA

<sup>3</sup> Dipartimento di Fisica e Astronomia, Università di Firenze, Via G. Sansone 1, 50019 Sesto Fiorentino, Firenze, Italy

<sup>4</sup> INAF – Osservatorio Astrofisico di Arcetri, 50125 Florence, Italy

Received 22 March 2021 / Accepted 1 September 2021

## ABSTRACT

We present a study of the relation between X-rays and ultraviolet emission in quasars for a sample of broad-line, radio-quiet objects obtained from the cross-match of the Sloan Digital Sky Survey DR14 with the latest *Chandra* Source Catalog 2.0 (2332 quasars) and the *Chandra* COSMOS Legacy survey (273 quasars). The non-linear relation between the ultraviolet (at 2500 Å,  $L_{UV}$ ) and the X-ray (at 2 keV,  $L_X$ ) emission in quasars has been proved to be characterised by a smaller intrinsic dispersion than the observed one, as long as a homogeneous selection, aimed at preventing the inclusion of contaminants in the sample, is fulfilled. By leveraging on the low background of *Chandra*, we performed a complete spectral analysis of all the data available for the SDSS-CSC2.0 quasar sample (i.e. 3430 X-ray observations), with the main goal of reducing the uncertainties on the source properties (e.g. flux, spectral slope). We analysed whether any evolution of the  $L_X-L_{UV}$  relation exists by dividing the sample in narrow redshift intervals across the redshift range spanned by our sample,  $z \approx 0.5-4$ . We find that the slope of the relation does not evolve with redshift and it is consistent with the literature value of 0.6 over the explored redshift range, implying that the mechanism underlying the coupling of the accretion disc and hot corona is the same at the different cosmic epochs. We also find that the dispersion decreases when examining the highest redshifts, where only pointed observations are available. These results further confirm that quasars are ‘standardisable candles’, that is we can reliably measure cosmological distances at high redshifts where very few cosmological probes are available.

**Key words.** galaxies: active – galaxies: nuclei – quasars: general – quasars: supermassive black holes

## 1. Introduction

Although the non-linear relation of ultraviolet ( $L_{UV}$ ) and X-ray ( $L_X$ ) luminosities in quasars is well known (Tananbaum et al. 1979; Zamorani et al. 1981; Avni & Tananbaum 1986; Wilkes et al. 1994; Steffen et al. 2006; Just et al. 2007; Young et al. 2010; Lusso et al. 2010), we still lack the understanding of the physical mechanism connecting the accretion disc and the hot corona. Photons from the accretion disc (emitting in the UV) are scattered to higher energies (X-rays) by the relativistic electrons present in a hot plasma (the so-called corona) through inverse Compton (e.g. Haardt & Maraschi 1991, 1993). Nonetheless, for the hot corona not to cool down very fast, there must be a mechanism transferring part of the gravitational energy lost during the accretion from the disc to the corona (e.g. Merloni & Fabian 2001), which likely involves magnetic reconnection (Di Matteo 1998; Liu et al. 2002), but whose details are hardly known.

An increase by a factor of 10 in the UV emission corresponds to an increase by only a factor of 4 at X-ray energies, meaning that more luminous quasars in the UV are relatively less luminous in the X-rays. This effect is parametrised as a non-linear relation, which is usually expressed in terms of the logarithms:

$$\log(L_X) = \gamma \log(L_{UV}) + \beta, \quad (1)$$

<sup>★</sup> Full Table 2 is only available at the CDS via anonymous ftp to [cdsarc.u-strasbg.fr](http://cdsarc.u-strasbg.fr) (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/cat/J/A+A/655/A109>

where the proxies used for the UV and X-ray emissions are the (rest-frame) 2500 Å and 2 keV monochromatic luminosities, respectively, and the slope has been found to be  $\gamma \approx 0.6 \pm 0.1$  in all the works mentioned above. This relation holds for several decades in both UV and X-ray luminosity (e.g. Steffen et al. 2006; Lusso et al. 2010; Salvestrini et al. 2019) and it does not appear to evolve with redshift (e.g. Vignali et al. 2003; Vagnetti et al. 2010), suggesting that the physical mechanism behind the disc-corona synergy must be universal, namely the same from the faintest to the brightest objects of the active galactic nuclei (AGN) population and at different ages of the Universe. Still, we have very few clues on what the physical process at work is.

Our group has investigated the  $L_X-L_{UV}$  relation, with the twofold aim of understanding its physical meaning (Lusso & Risaliti 2016, 2017) and testing its application for cosmological purposes (Risaliti & Lusso 2015, 2019; Bisogni et al. 2017a; Lusso et al. 2019, 2020; Bargiacchi et al. 2021). Thanks to the non-linearity of this relation, we can have an independent measurement of the AGN cosmological distances once we convert luminosities into fluxes as:

$$\log(f_X) = \gamma \log(f_{UV}) + (2\gamma - 2) \log(D_L) + \beta' \quad (2)$$

where the luminosity distance  $D_L$  is not canceled out if  $\gamma \neq 1$  and  $\beta'$  depends on slope and intercept as  $\beta' = \beta + (\gamma - 1)\log 4\pi$ . On these grounds, it is then possible to build a Hubble diagram for quasars (Risaliti & Lusso 2015, 2019) and use these objects as

cosmological probes, bridging the gap from the farthest observed supernovae Ia (at  $z \approx 2$ , Pantheon sample; Scolnic et al. 2018) to considerably higher redshifts ( $z \approx 7.5$ , Bañados et al. 2018).

In both the physical and the cosmological applications, the analysis of the dispersion in the  $L_X-L_{UV}$  relation is crucial. In the first case, it can constrain the coupling of the AGN disc and corona. In the second, it determines the precision achievable in the measure of cosmological distances, that is in the extraction of cosmological information.

A substantial progress came with the discovery that most of the observed dispersion,  $\delta_{\text{obs}}$ , is not intrinsic (Lusso & Risaliti 2016, 2017). When the contributions to the dispersion ascribable to observational issues can be properly removed – or at least reduced –  $\delta_{\text{obs}}$  significantly decreases (see also Lusso et al. 2020). Several contributors to the dispersion have been found, such as (1.) dust reddening and X-ray absorption, which can prevent the measure of the intrinsic fluxes at 2500 Å and 2 keV, respectively; (2.) calibration issues, especially in the X-ray band, related to the serendipitous detection of many of the sources in the X-ray catalogues, with a range of distances from the aim-point of the observation; (3.) the inclusion in the sample of sources detected following a positive fluctuation with respect to their average emission (Eddington bias); (4.) the inclination of the accretion disc with respect the line of sight, affecting the measure of the intrinsic UV flux; (5.) emission variability in both bands; (6.) non-simultaneity of the UV and X-ray emission measurements. Some of these factors are in principle correctable (1–5), while we can only be aware of others (e.g. 6).

In a recent work, Risaliti & Lusso (2019) selected  $\sim 1600$  quasars from a parent sample of  $\sim 8000$  objects, obtained through the cross-match of the SDSS DR7+DR12 and 3XMM DR7. By excluding from the sample the quasars affected by the observational problems listed above, and by using, mainly, X-ray photometric data, they found a dispersion  $\delta = 0.24$  dex, to be compared to the initial value of  $\delta \sim 0.40$  dex of the parent sample. The latter value is consistent with the one found since the earliest works, which had deterred, until very recently, from using the  $L_X-L_{UV}$  relation for any cosmological application. Such a reduction in the dispersion allowed the detection, to a  $\sim 3-4\sigma$  significance level, of a discrepancy between the cosmological parameters measured at  $z < 1.4$  and at  $z > 1.4$ , which is compatible with the tension in the value of the Hubble constant  $H_0$  measured with the Distance Ladder methods (Riess et al. 2019) and that measured from the temperature fluctuations of the Cosmic Microwave Background assuming the concordance  $\Lambda$ CDM model (Planck Collaboration VI 2020).

In the wake of these achievements, here we present the analysis of a brand new sample of 7036 SDSS DR14–*Chandra* Source Catalog 2.0 quasars. The main goals of this work are (1.) exploiting the low background level of *Chandra* to further decrease the observed dispersion in the  $L_X-L_{UV}$  relation by measuring the 2 keV fluxes spectroscopically; (2.) examining the *non*-evolution of the slope of the relation with redshift, crucial to both its employment in cosmological applications and the understanding of quasar evolution across cosmic time; and (3.) accounting for the contribution to the dispersion of the X-ray variability through the use of multiple observations of the sources in the sample.

This paper is organised as follows: in Sect. 2 we present the sample and the preliminary selection criteria. In Sect. 3 we describe how the fluxes in the UV and X-ray bands are retrieved. For the X-ray band, the procedure used to infer the flux limit for each observation and the criteria for assuring that no Eddington bias is present in the sample are presented. In Sect. 4 we describe the analysis on the  $f_X-f_{UV}$  relation and its possible

evolution with redshift, and the analysis of the  $L_X-L_{UV}$  relation for the total sample. We present our results and conclusions in Sects. 5 and 6, respectively. Throughout the paper, when computing luminosities, we assume a concordance flat  $\Lambda$ CDM model, with  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.3$  and  $\Omega_\Lambda = 1 - \Omega_M$ .

## 2. Data and sample selection

Our parent sample includes sources selected from the following catalogues:

1. The cross-match of the fourteenth data release of the Sloan Digital Sky Survey quasar catalogue<sup>1</sup> (Pâris et al. 2018) and the latest release of the *Chandra* Source Catalog (CSC 2.0<sup>2</sup>; Evans et al. 2010). The SDSS DR14 quasar catalogue includes the previous releases (DR7 and DR12) and expands them with newly observed sources, for a total of more than 500 000 spectroscopically confirmed quasars. The CSC 2.0 contains information on  $\sim 315\,000$  X-ray sources detected in 10 382 *Chandra* ACIS and HRC-I imaging observations publicly released prior of 2015. We obtained 7036 unique matches within 3'' from the SDSS DR14 coordinates, corresponding to more than 10 000 single *Chandra* observations in the X-ray band, when multiple observations of the same source are taken into account.
2. Unobscured (type-1) AGN from the *Chandra* COSMOS Legacy Survey (Civano et al. 2016), a 4.6-Ms *Chandra* programme providing a coverage of  $\sim 160$  ks for the central  $1.5 \text{ deg}^2$  and of  $\sim 80$  ks for the outer region of the  $2.2 \text{ deg}^2$  COSMOS field<sup>3</sup>. We selected type-1 AGN following the spectroscopic and photometric identification presented in Marchesi et al. (2016). As a first requirement, we asked for secure associations with the near-infrared counterparts: this selected 3910 of the initial 4016 sources. When a spectroscopic observation is available, the sources are identified as Broad Line AGN (BLAGN) if they present at least one broad line with  $FWHM > 2000 \text{ km s}^{-1}$  (635 type-1 AGN out of the total 2262/3910 with a spectrum), whilst 3798/3910 sources have a spectral energy distribution (SED) fitting classification (880/3798 SEDs are reproduced by a type-1 AGN template in the optical/UV). When a spectrum is available, the spectroscopic classification is preferred over the one obtained from the SED analysis. Our selection yielded a sample of 1001 type-1 AGN (635 spectroscopically identified and 366 classified only through SED fitting),  $\sim 25\%$  of the whole *Chandra* COSMOS Legacy catalogue, with an unambiguous match between X-ray detection and NIR counterpart and either a spectral or a photometric classification. Only 943 out of the 1001 X-ray selected type-1 objects have a detection in the Ultra-VISTA survey (McCracken et al. 2012) and the required photometric information to perform the analysis of their SED and to infer their 2500 Å flux. Quasars in the COSMOS sample that overlapped with the SDSS-CSC2.0 one were excluded from the final COSMOS sample, yielding 882 objects. Throughout this work we made use of the photometric information listed in the Near Infrared selected catalogue COSMOS2015 (Laigle et al. 2016).

<sup>1</sup> [https://www.sdss.org/dr14/algorithms/qso\\_catalog/](https://www.sdss.org/dr14/algorithms/qso_catalog/)

<sup>2</sup> <http://cxc.harvard.edu/csc/>

<sup>3</sup> This programme is an extension of the previous *Chandra* COSMOS survey (C-COSMOS, Elvis et al. 2009; Puccetti et al. 2009; Civano et al. 2012), which already covered with 1.8 Ms of *Chandra* observations  $\sim 1/4$  of the COSMOS field at the same depth plus  $0.5 \text{ deg}^2$  at  $\sim 80$  ks depth.

**Table 1.** Summary of the observations.

<i>SDSS – CSC 2.0</i>				
UV pre-selection	DR7	DR12	DR14	TOT sources
SDSS vs CSC 2.0 (3'')	2348	3276	1412	7036
Non-BAL (flags)	2217	3111	1412	6740
Non-BAL (Gibson09)	2211	3111	1412	6734
Non-BAL (Allen11)	2206	3111	1412	6731
Non-RL (flags)	1922	3111	1412	6445
Non-RL (Mingo16 MIXR)	1919	3096	1412	6427
Non-RL (Mingo16 crit.)	1919	3011	1376	6306
Dust-free	1729	1993	1027	4749
Host-free	1600	1965	1013	4578
X-ray pre-selection				TOT observations (sources)
SDSS vs CSC 2.0 obs	2294 (1559)	3318 (1911)	1265 (976)	6877 (4446)
Soft band	1988 (1189)	2636 (1141)	1000 (567)	5624 (2897)
$\Theta < 10$ arcmin	1510 (975)	1998 (1014)	704 (477)	4212 (2466)
Available spectra	1362 (922)	1642 (995)	565 (475)	3569 (2392)
Flux limit	1322 (891)	1560 (967)	548 (474)	3430 (2332)
X-ray selection				TOT observations (sources)
$\Gamma$	998 (683)	1066 (697)	378 (337)	2442 (1717)
Eddington bias	706 (477)	535 (350)	144 (131)	1385 (958)
<i>Chandra COSMOS Legacy</i>				
UV pre-selection				TOT sources
COSMOS Type-1 AGN				1001
Ultra-vista data				882
Non-BAL (flags)				875
Non-BAL (Gibson09)				875
Non-BAL (Allen11)				875
Non-RL (flags)				872
Non-RL ( $R > 10$ )				823
Dust-free				289
Host-free				288
X-ray pre-selection				TOT sources
Soft band				273
X-ray selection				TOT sources
$\Gamma$				151
Eddington bias				140

**Notes.** Summary of the observations statistics among the different sub-samples and categories. Numbers in parentheses are sources corresponding to observations.

### 2.1. Sample selection in the optical/UV band

For the sample selection, we followed the approach described in Lusso et al. (2020). Briefly, we excluded from the sample broad absorption line (BAL) and radio-loud (RL) quasars. The presence of obscuration (Murray et al. 1995; Elvis 2000) or an additional contribution to the X-ray emission due to winds and jets (e.g. Wilkes & Elvis 1987) require a complex modelling of the AGN spectra to disentangle these processes from the intrinsic emission of the X-ray corona, which in turn increases the

uncertainty on the X-ray flux measurement. We finally selected sources that are not dust-absorbed or host-galaxy contaminated in the UV/optical band.

#### 2.1.1. Broad absorption line quasars

*SDSS-CSC2.0.* To exclude BALs, we used the DR7 and DR12 dedicated flags provided by the quasar catalogue published by Shen et al. (2011) and the visual BAL flag provided by Pâris et al. (2017), respectively. To exclude as many BALs as possible among sources observed for the first time in DR14, we used the Balnicity Index (BI) as defined in Weymann et al. (1991) and provided by Pâris et al. (2018). Since BAL quasars are traditionally defined as having  $BI(CIV) > 0$ , we required  $BI(CIV) = 0$ . The exclusion of sources through the BI is not as accurate as a visual inspection, but it assures to not exclude non-BAL objects from the sample. With this selection we obtained 6740 objects. We then cross-matched the sample with the BAL sample by Gibson et al. (2009) and with the one by Allen et al. (2011), finding 6 and 5 more BALs, respectively. As discussed in Pâris et al. (2018), these two works adopt slightly different fitting strategies in the evaluation of quasar emission, implying differences in the BAL distributions of the samples. Some objects can therefore be classified as BAL by Gibson et al. (2009) and/or Allen et al. (2011), but not as such in Pâris et al. (2018). To be conservative, we excluded all the objects classified as BAL in at least one of these three works. The final ‘BAL-free’ sample is composed by 6729 quasars, 96% of the parent sample (Table 1).

*COSMOS.* For the objects present in either the DR7 (10 sources) or DR12 (19) SDSS quasar catalogues, we used the dedicated flag and exclude 7 BALs. The cross-match with the Gibson et al. (2009) and Allen et al. (2011) did not give any result. The BAL free sample is composed by 875 sources.

#### 2.1.2. Radio-Loud quasars

*SDSS-CSC2.0.* We first excluded 284 DR7 quasars with a value of the radio loudness parameter,  $R$ , higher than 10 (Kellermann et al. 1989), where  $R = f_{\nu,6\text{cm}}/f_{\nu,2500}$  is listed in the Shen et al. (2011) catalogue. We then cross-matched the sample with the MIXR catalogue of 2753 objects, obtained through the cross-match of the largest catalogues available in the Mid-Infrared (WISE), X-ray (3XMM-DR5) and Radio (FIRST/NVSS) and published by Mingo et al. (2016). This yielded 44 sources within 3'' of positional error, 18 of which are classified as RL and hence excluded from our sample. The analysis of Mingo et al. (2016) on the multi-wavelength MIXR sample shows also that, regardless of their properties in other bands, sources with a  $L_{1.4\text{GHz}} > 5 \times 10^{41}$  can be classified as RL. We therefore cross-matched our sample with the FIRST/NVSS catalogue and retrieved the rest-frame  $L_{1.4\text{GHz}}$  by assuming a radio spectral slope of 0.5 (Hao et al. 2014). Within a maximum positional error of 30'', we found 346 matches, of which 121 are classified as RL following the criterion on  $L_{1.4\text{GHz}}$ . The ‘RL-free’ sample consists of 6306 quasars.

*COSMOS.* For the sources in the SDSS DR7 quasars catalogue, we used the dedicated flag, excluding 3 RL sources. Once the UV SED analysis was performed (Appendix A) and the 4400 Å luminosity was available, we computed the radio-loudness parameter  $R = L_{5\text{GHz}}/L_{4400} > 10$ , using the Smolčić et al. (2017) catalogue at 3 GHz to infer the luminosity at 5 GHz, and excluded 49 more sources, leaving 823 sources in the ‘RL-free’ COSMOS sample.

### 2.1.3. Dust reddening and host galaxy contamination

**SDSS-CSC2.0.** Dust along the line of sight and contamination by the host galaxy are the main reasons for an incorrect measure of the intrinsic 2500 Å flux, proxy for the accretion disc emission. In order to keep in the sample only sources with a measure of the 2500 Å close to the intrinsic one, we adopted the same criteria as in [Lusso & Risaliti \(2016\)](#), by computing the spectral slopes of the optical spectra so to evaluate to what degree the continuum is affected by dust absorption or by galaxy contamination (see Appendix A and Fig. A.1 for details). We included in the sample only the sources with all the five SDSS bands magnitudes available (6282/6306) and selected 4749 blue, host-galaxy- and dust-free quasars. Finally, by choosing only sources with  $z > 0.48$ , we excluded 171 more objects, obtaining a selection of 4578 quasars. Sources at lower redshifts are more affected by host-galaxy contamination for two main reasons: first, low-redshift AGN are on average intrinsically fainter than their higher-redshift companions and do not outshine their hosts to the same extent; second, the BOSS spectrograph wavelength coverage (3650–10 400 Å) sets a lower limit for the 2500 Å to be inferred without need for extrapolation. The combination of the two effects could lead to an overestimation of the 2500 Å flux measure for sources at  $z < 0.48$ .

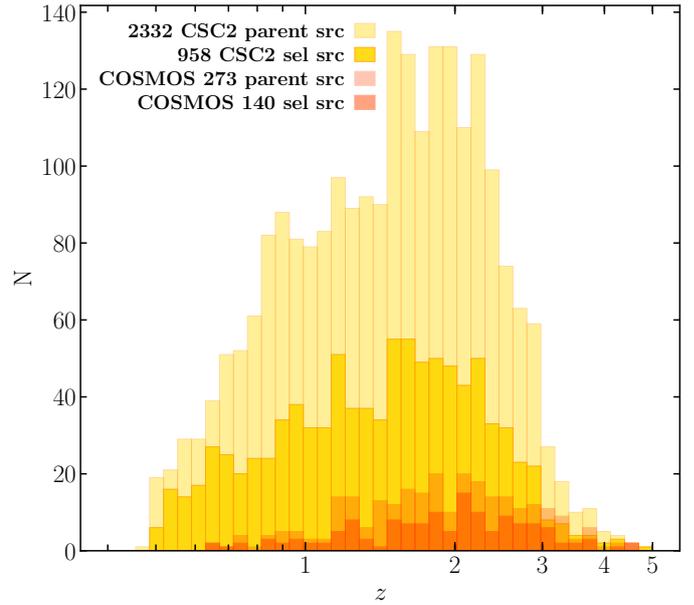
**COSMOS.** By applying the same criteria on the spectral slopes computed with the UV SED analysis, we excluded 534 reddened sources. By considering only  $z > 0.48$  sources, the sample reduced to 288 objects. The *Chandra* COSMOS Legacy sources are X-ray selected, and therefore biased towards a more obscured population of quasars, as it is evident comparing the spectral slopes in the UV/optical/NIR bands for the COSMOS sample with those of the SDSS (optically selected) one (Fig. A.1). The fraction of discarded sources through this last selection is large, about 60% of the initial sample (882 quasars).

### 2.2. Sample selection in the X-ray band

**SDSS-CSC2.0.** Along with tabulated source properties, the CSC 2.0 provides data products – ready to be used for scientific analysis – that can be directly downloaded using the *CSCview* application<sup>4</sup>. The precision of the 2 keV flux measurement is one of the major concerns in this study. For this reason, we performed a complete spectral analysis using the X-ray spectra available for all the observations of X-ray sources and retrievable from the CSC 2.0 archive as data products.

As mentioned above, CSC 2.0 is built on the 10 382 *Chandra* ACIS and HRC-I imaging observations publicly released prior of 2015. The CSC 2.0 performs source detection on the sum of overlapping observations (stacks), to increase the sensitivity, but includes source properties computed both at the level of the stack and also at the level of the single observations contributing to the stack. As we are interested in estimating the contribution of the X-ray variability to the observed dispersion of the relation, including in the sample those sources with multiple observations is a basic requirement. For the 4446 sources in our sample, we found 6877 observations (Table 1) and the corresponding spectra. With a preliminary selection, we restricted the observations to those with:

1. A measurement of the flux in the soft (0.5–2 keV) or in both the soft and the hard band (2–7 keV) and



**Fig. 1.** Distribution in redshift of the sample of 2332 SDSS-CSC2.0 and 273 *Chandra*-COSMOS Legacy non-BAL, radio-quiet, dust-free selected sources. The same is shown for the final sample, selected based on the analysis in the X-ray band.

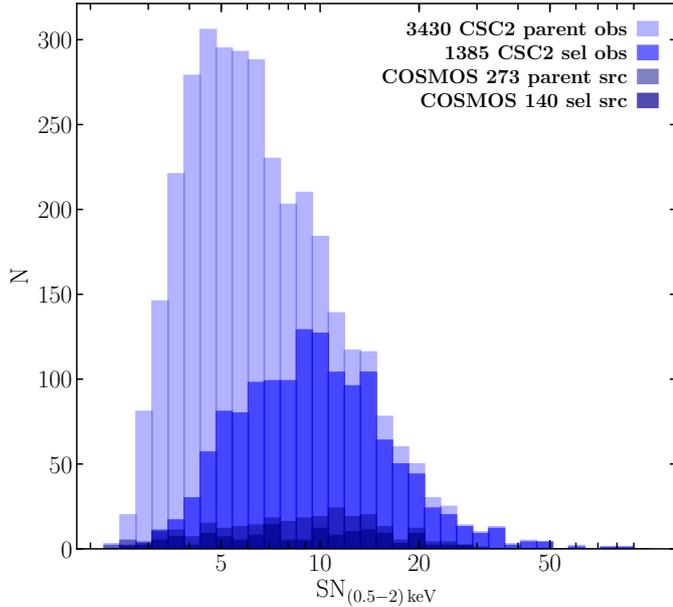
2. An off-axis angle (the distance of the source from the aim-point)  $< 10'$  in that observation, which provides the best of spatial resolution.

Due to the quality of the X-ray measurements for sources that are mainly serendipitous in X-ray catalogues, the largest among the contributions to the dispersion are ascribable to the Eddington bias and to the accuracy and precision of the X-ray fluxes ([Risaliti & Lusso 2015, 2019](#); [Lusso & Risaliti 2016, 2017](#)). By using *Chandra* spectra to directly measure the X-ray flux and limiting our selection to on-axis sources, we are partially addressing these issues. Given the importance of the determination of the flux limit for each observation – enabling us to discard the sources included in the sample because of a positive fluctuation compared to their average intensity – we kept only those observations for which such a measure is possible. This selected 3430 observations (Table 1), corresponding to 2332 sources, of which 439 have multiple (i.e. two or more) observations<sup>5</sup>. The distributions of soft-band signal-to-noise ratio and redshift are shown in Figs. 1 and 2, respectively.

**COSMOS.** In this case, with a median exposure time of  $\sim 156$  ( $\sim 159$ ) ks in the soft (hard) band, compared to a median effective exposure time of  $\sim 20$  ks for CSC 2.0 observations, we chose to use the X-ray photometric measurements already available in the *Chandra* COSMOS Legacy catalogue ([Civano et al. 2016](#)). Requiring the availability of at least the soft band selected 273 sources, 230 of which also have the hard band available. The flux limit in this survey is quite uniform and is reported in [Civano et al. \(2016\)](#).

<sup>5</sup> We note that we examined all the X-ray observations available for the sources we pre-selected in the UV band. As a result, we analysed 3430 spectra extracted at the source region from *Chandra* imaging observations, surviving the selection criteria described above in terms of off-axis angle and of availability of flux measurement. The forthcoming X-ray analysis was also carried out at the observation-level, unless otherwise specified.

<sup>4</sup> <http://cda.harvard.edu/cscview/>



**Fig. 2.** Distribution of the signal-to-noise in the soft band for the 3430 CSC 2.0 observations, corresponding to 2332 non-BAL, radio-quiet, dust free sources and for the 273 *Chandra* COSMOS Legacy sources. The same is shown for the final sample, selected based on the analysis in the X-ray band.

### 3. Data analysis

In this section, we report how the monochromatic fluxes have been computed from the multi-wavelength data collected, and how the sample has been corrected for the Eddington bias.

#### 3.1. Computation of UV fluxes

The flux density at 2500 Å for all the sources in the sample was obtained through an analysis of the SED, which made use of the multi-wavelength photometric information from the UV to near-infrared bands, available for both the SDSS-CSC2.0 and the *Chandra* COSMOS Legacy samples. The determination of SED slopes allowed the exclusion of the dust-affected sources in the sample, as mentioned in Sect. 2.1. For the SDSS-DR7 sources, we chose to use the 2500 Å flux densities provided by Shen et al. (2011), inferred from a complete spectral fitting procedure that took into account continuum and emission lines for quasar spectra. A detailed description of the analysis of the UV/optical/NIR SED is given in Appendix A. Table 2 lists the final 2500-Å fluxes used for the statistical analysis.

In the following, we focus on the analysis in the X-ray band.

#### 3.2. SDSS-CSC2.0 X-ray spectral analysis

##### 3.2.1. Spectral fitting

The *CSCview* application provides access to the CSC 2.0 catalogue allowing the user to download uniformly extracted spectra<sup>6</sup> and ancillary files (response matrices and background spectra) for each source in each observation. A complete spectral analysis was carried out with the fitting package XSPEC v12.10.1b (Arnaud 1996) for each source in the selected obser-

vations. All of the spectra analysed have at least five counts in the soft band, while only 20 have no counts in the hard band. 3420/3430 have >10 counts in the full band and ~34% have >100 counts. Figure 3 shows the distributions of the full, soft and hard spectral counts for the sample. No minimum number of counts or of net vs background counts (SN) has been set for performing the fit.

We assumed a power-law with Galactic absorption  $N_{\text{H}}$ , whose values were drawn from the CSC 2.0 catalog, as a fitting model to the spectrum. The best fit to the data allowed the determination of the integrated flux in the soft and hard energy bands ( $F_{0.5-2\text{keV}}$  and  $F_{2-7\text{keV}}$ ), and of the rest-frame flux at 2 keV, tracer of the corona emission, and its error – once the *cflux* component<sup>7</sup> was included in the model – leaving flux and photon index  $\Gamma$  as free parameters of the fit. The flux (luminosity) values at 2 keV were determined with uncertainties of  $0.10^{+0.10}_{-0.05}$  dex (median, 16th and 84th percentile of the error distribution). Figures 4 and 5 show the distribution of fluxes at 2 keV and 2500 Å as a function of redshift and that of the photon indexes  $\Gamma$  for the single observations, respectively, as obtained from the data analysis.

Due to the presence of sources with an unexpectedly high ( $\Gamma > 3$ ) or low ( $\Gamma < 1.4$ ) photon index, we checked our spectral analysis looking at the hardness ratios (HR, see Appendix B). The comparison between the HRs computed using the net counts from the CSC 2.0 catalog, and those derived from our spectral analysis shows an excellent agreement, especially below  $\Gamma = 2.8$ , the upper threshold we adopted for the photon index. This is a model-independent confirmation of the reliability of our spectral analysis (see Appendix B for details).

##### 3.2.2. Computation of the flux limit

Active galactic nuclei with an X-ray flux measurement close to the flux limit of the observation will be observed only in case of a positive fluctuation. This well known effect, named Eddington bias, introduces a systematic and also redshift dependent bias at the faint end of the X-ray flux distribution, which has the effect to flatten the  $L_X$ – $L_{UV}$  relation (see also Sect. 5.3 of Lusso et al. 2020). We therefore computed the rest-frame 2 keV flux limit for each observation to exclude the sources caught just as a positive fluctuation. This flux was computed by interpolating – or extrapolating, depending on the redshift of the source – the monochromatic flux limits  $f_S$  and  $f_H$  in the broad soft and hard bands. This same method was used for inferring the rest-frame 2 keV flux when only photometric data are available, and was employed in the case of the *Chandra* COSMOS Legacy sample (see Sect. 3.3 for a detailed explanation of the method).

The flux limit was then derived as follows: for each observation, we made use of the percentage of net counts – measured by XSPEC as the number of source over total (source+background) counts – in the soft 0.5–2 keV and hard 2–7 keV bands to compute the significance of the source ( $S_{\text{soft}}$  and  $S_{\text{hard}}$ ), and a factor taking into account the background level in the integrated bands,  $P_{\text{bkg}} = \frac{1-S}{S}$ . By multiplying  $F_{0.5-2\text{keV}}$  and  $F_{2-7\text{keV}}$  for  $P_{\text{bkg}}$ , we obtained an approximation of the background flux in the soft and hard band respectively. In doing so, even if we were incorrectly assuming the same spectral shape for source and background, the background flux in the soft/hard band turned out to be slightly higher than the one obtained with the less steep spectral shape

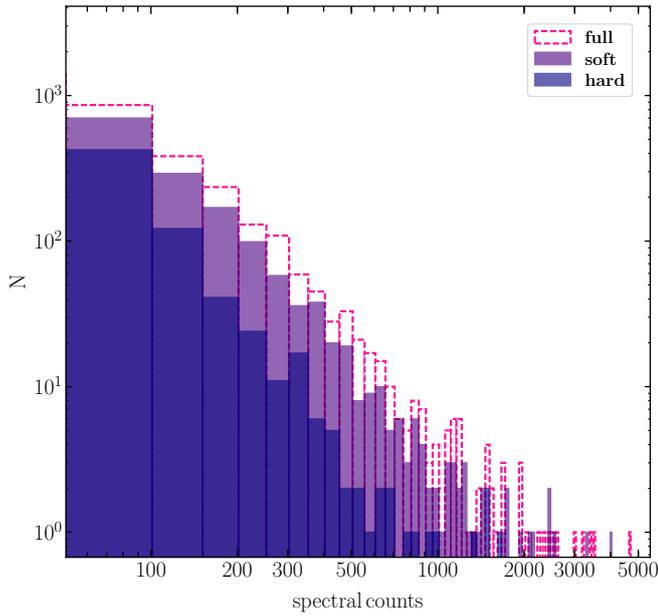
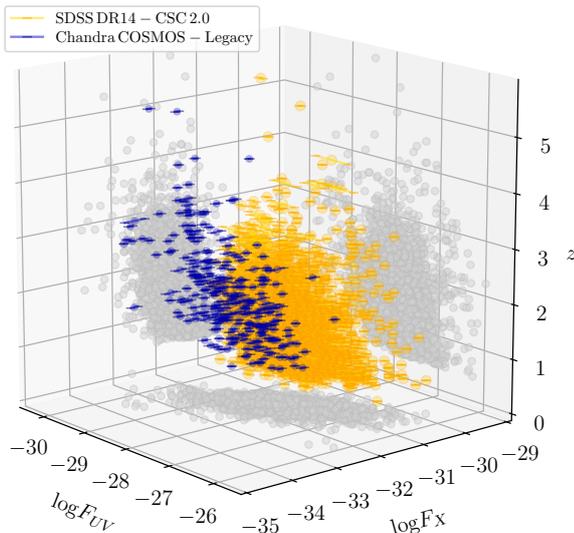
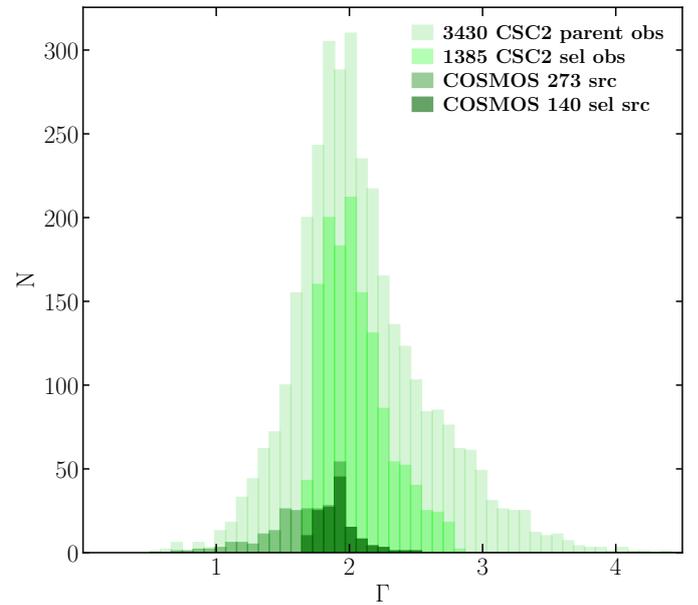
<sup>6</sup> Software and calibration database versions used for each observation are specified in the header of the .fits files.

<sup>7</sup> The *cflux* (calculate flux) model component in XSPEC allowed us to treat the intrinsic power-law flux in a given energy range of interest as a fit parameter (in lieu of the power-law normalisation).

**Table 2.** Summary of optical and X-ray properties.

SDSS name	Redshift	OBSID <sup>(a)</sup>	$\Theta$ <sup>(b)</sup>	Exp <sup>(c)</sup>	DRflag <sup>(d)</sup>	$f_{2500}$ <sup>(e)</sup>	$f_{2\text{keV}}$ <sup>(f)</sup>	$\Gamma$ <sup>(g)</sup>	$f_{\text{lim } 2\text{keV}}$ <sup>(h)</sup>	$C_{\text{soft}}$ <sup>(i)</sup>	$C_{\text{hard}}$ <sup>(j)</sup>	$\text{SN}_{\text{soft}}$ <sup>(k)</sup>	$\text{SN}_{\text{hard}}$ <sup>(l)</sup>
000102.74+023503.2	0.76489	4837	4.80644	5788.17	DR14	$-28.21 \pm 0.02$	$-31.12^{+0.06}_{-0.07}$	$1.7^{+0.3}_{-0.2}$	-31.79	42	22	6.4	4.6
000130.63+233443.5	2.93	14898	5.82812	39465.1	DR12	$-28.01 \pm 0.05$	$-32.04^{+0.11}_{-0.12}$	$1.5^{+0.3}_{-0.2}$	-32.80	63	34	7.4	4.9
000150.39+023830.2	0.949	4837	8.79053	5788.17	DR12	$-27.833 \pm 0.010$	$-31.20^{+0.08}_{-0.09}$	$3.4^{+0.5}_{-0.5}$	-31.59	28	8	5.2	2.4
000151.58+232857.6	0.884201	14898	2.47162	39461.9	DR14	$-27.747 \pm 0.014$	$-31.74^{+0.06}_{-0.06}$	$2.0^{+0.2}_{-0.2}$	-32.48	60	25	7.7	4.9
000151.95+232421.0	2.03318	14898	5.64085	39461.9	DR12	$-28.46 \pm 0.05$	$-31.99^{+0.13}_{-0.13}$	$2.2^{+0.4}_{-0.3}$	-32.52	30	12	5.2	2.5

**Notes.** Summary of optical and X-ray properties inferred from the analysis for the first five observations in the sample. This table is available in its entirety at the CDS. <sup>(a)</sup>*Chandra* observation identifier. <sup>(b)</sup>Off-axis angle [arcmin]. <sup>(c)</sup>Exposure time [s]. <sup>(d)</sup>SDSS Data Release. <sup>(e)</sup>Flux at 2500 Å as inferred from the UV SED analysis [ $\text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1}$ ]. <sup>(f)</sup>Flux at 2 keV as inferred from the spectral analysis assuming a power-law model and Galactic absorption  $N_{\text{H}}$  [ $\text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1}$ ]. <sup>(g)</sup>Photon index. <sup>(h)</sup>Flux limit at 2 keV per observation [ $\text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1}$ ]. <sup>(i)</sup>Raw counts in the soft band (0.5–2 keV). <sup>(j)</sup>Raw counts in the hard band (2–7 keV). <sup>(k)</sup>Signal-to-noise in the soft band (0.5–2 keV). <sup>(l)</sup>Signal-to-noise in the hard band (2–7 keV).


**Fig. 3.** Distribution of spectral counts in the full (0.5–7 keV), soft (0.5–2 keV) and hard band (2–7 keV).

**Fig. 4.** Distribution of the 2-keV and 2500-Å fluxes for the 2332 CSC 2.0 (yellow) and 273 *Chandra* COSMOS Legacy (blue) non-BAL, radio-quiet, dust-free sources. For CSC 2.0 sources with multiple observations, the mean of the  $\log F_X$  values is plotted.

**Fig. 5.** Distribution of the photon index  $\Gamma$  for the 3430 selected SDSS-CSC2.0 observations, corresponding to 2322 non-BAL, radio-quiet, dust-free sources and for the 273 *Chandra* COSMOS Legacy sources. The same is shown after for the final sample, selected based on the analysis in the X-ray band.

characterising the background. Since, as explained below, we were indirectly using this value to exclude from the sample the sources close to the flux limit – proportional to the background flux – a higher value resulted in a more conservative selection. The monochromatic background fluxes  $f_{\text{bkgs}}$  and  $f_{\text{bkgsH}}$  in the soft and hard bands were then used to infer the flux limit in the bands, as shown by the following expression for the soft-band case:

$$f_{\text{lim S}} = \frac{\eta_S \text{SN}_{\text{MIN}}^2 + \text{SN}_{\text{MIN}} \sqrt{\eta_S^2 \text{SN}_{\text{MIN}}^2 + 8 \eta_S f_{\text{bkgs}}}}{2}, \quad (3)$$

where  $\eta_S$ , the conversion factor between counts and fluxes ( $\eta [\text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1} \text{counts}^{-1}] = f/C$ ), is given by

$$\eta_S = \frac{f_S + \sqrt{f_S^2 + 8 C_{\text{raw S}} \Delta f_S^2}}{4 C_{\text{raw S}}}, \quad (4)$$

where  $C_{\text{raw S}}$  includes net and background counts, and both Eqs. (3) and (4) are obtained, after simple algebra, from the

equation

$$\text{SN}_{\text{MIN}} = \frac{C_{\text{MIN}}}{\sqrt{C_{\text{MIN}} + 2 C_{\text{bkg}}}}, \quad (5)$$

where the choice of a minimum signal-to-noise in the formula,  $\text{SN}_{\text{MIN}}$  (in our case  $\text{SN}_{\text{MIN}} = 3$ ), implies the other quantities. The  $f_{\text{lim}2\text{keV}}$  was then inferred from interpolation or extrapolation of  $f_{\text{lim}S}$  and  $f_{\text{lim}H}$ , depending on the redshift of the source.

The X-ray properties inferred from this analysis are listed in Table 2.

### 3.3. COSMOS X-ray photometric analysis

We decided to consider the background-subtracted, aperture-corrected fluxes values already listed in the *Chandra* COSMOS Legacy catalogue. The photometric data were estimated from the count rates  $R$  in each band – obtained with the *Chandra* EmlDetect (CMLDetect) maximum likelihood algorithm – using the relation  $F = R \times (CF \times 10^{-11})$ , with  $CF$  energy conversion factor.  $CF$  was computed with the CIAO tool *srcflux*, assuming a power-law spectrum with  $\Gamma = 1.4$  and a Galactic column density of  $N_{\text{H}} = 2.6 \times 10^{20} \text{ cm}^{-2}$  (for details, see, Civano et al. 2016).

To infer the flux at rest-frame 2 keV, we interpolated – or extrapolated, depending on the redshift of the source – the monochromatic flux in the broad soft and hard bands that we estimated from the integrated values available from the *Chandra* COSMOS Legacy catalogue. In order to choose the energies in the soft and hard bands at which computing the monochromatic fluxes, we applied a new method developed by our group (Risaliti & Lusso 2019, see Sect. 1 of their Supplementary Material). Instead of using the more conventional geometric mean or the mean energy of a band, we defined the pivot point  $E_p$  as the specific energy at which the monochromatic flux  $f(E_p)$  and the photon index  $\Gamma$  of the spectrum have a null covariance, that is the errors on the two quantities are independent. To compute this energy, a spectrum of high signal-to-noise was simulated, assuming typical calibration and an average background for *Chandra* observations, along with a typical photon index of 1.7 for the spectrum. We performed fits to the simulated data for different values of the energy  $E_p$ , assuming a spectrum  $f(E) = f(E_p)(E/E_p)^\Gamma$  and  $f(E_p)$  and  $\Gamma$  as free parameters. Error contours of  $f(E_p)$  versus  $\Gamma$  were plotted within XSPEC. When the two quantities are co-dependent, the ellipsoid of the error contours appears tilted with respect to the  $x$ - and  $y$ -axis, as expected when the covariance is different from zero. When, instead, the axes of the ellipsoid are parallel to the  $x$  and  $y$  axes, the errors are independent. We found the values at which this occurs to be  $E_S = 1.05 \text{ keV}$  and  $E_H = 3.1 \text{ keV}$  for the soft and the hard band respectively. The pivot energies have many advantages: (1) by definition, at these energies the covariance between  $f(E_p)$  and  $\Gamma$  is zero; (2) the flux  $f(E_p)$  is independent from the value of  $\Gamma$  assumed for the energy spectrum in the band, that is the accuracy of the flux estimate is independent from the value assumed by the photometric catalogue used (for CSC 2.0,  $\Gamma = 2$ ); (3) following the relation between the total flux in the band ( $F$ ) and at a monochromatic energy ( $f(E)$ )<sup>8</sup>, we can write:

$$f(E) = \frac{(2 - \Gamma)E^{1-\Gamma}}{E_2^{2-\Gamma} - E_1^{2-\Gamma}} F. \quad (6)$$

<sup>8</sup>  $F = \int_{E_1}^{E_2} f(E) dE$ .

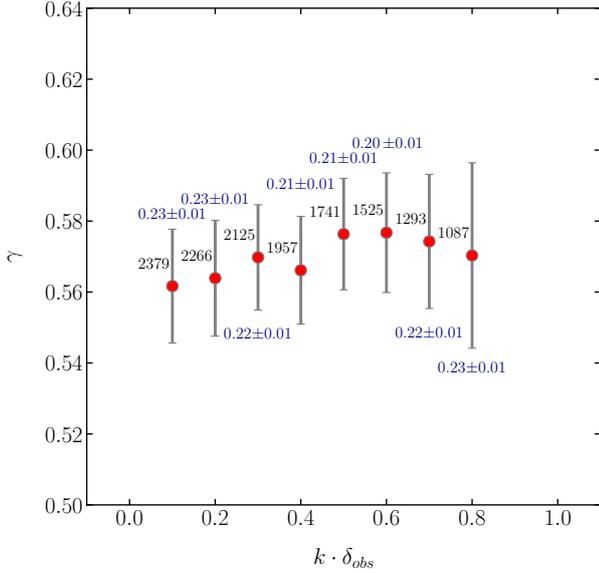
When  $E = E_p$ , the relative error on the monochromatic flux is the same as the one in the band, that is  $\Delta(f(E_p))/f(E_p) = \Delta(F)/F$ , and the absolute error  $\Delta(f(E_p))$  is the smallest possible. We then computed the value of the rest 2-keV flux by interpolation (or extrapolation) of the monochromatic soft and hard fluxes estimated at the pivot frequencies from the integrated fluxes in the bands. When only the soft band was available, we inferred the 2-keV flux by assuming a typical photon index  $\Gamma = 1.9$ . A comparison of the COSMOS and SDSS-CSC2.0 rest-frame 2 keV and 2500 Å fluxes as a function of redshift is shown in Fig. 4. The *Chandra* COSMOS Legacy sample (blue data points) covers a fainter locus of the distribution between the fluxes ( $\langle \log f_{2\text{keV}} \rangle = -32.40$ ,  $\langle \log f_{2500\text{Å}} \rangle = -28.92$  versus  $\langle \log f_{2\text{keV}} \rangle = -31.59$ ,  $\langle \log f_{2500\text{Å}} \rangle = -27.82$  of the SDSS-CSC2.0 sample), and, on average, a slightly higher redshift ( $\langle z \rangle \sim 1.8$  vs.  $\langle z \rangle \sim 1.5$ ). The same procedure was employed for the determination of the COSMOS flux limit at 2 keV rest-frame. This time, flux limits in the soft and hard bands were already available from the sensitivity maps of the survey.

### 3.4. Eddington bias threshold estimation

A quasar whose average X-ray emission is close to the flux limit of the observation will be detected only if caught on a positive fluctuation (Eddington bias). The inclusion of such sources corresponds to the presence of non-representative data points in the  $L_X$ – $L_{UV}$  relation. In order to clean the sample from this effect, we excluded those sources with an *expected* X-ray flux lower than a threshold defined by the flux limit of their observation. This was done as follows: (1.) from the  $L_{2500}$  obtained through the UV SED analysis and the assumption of the concordance  $\Lambda$ CDM model, we estimated the expected  $L_X$ , assuming slope and intercept of the  $L_X$ – $L_{UV}$  relation from Lusso & Risaliti (2016); (2.) we derived the expected X-ray flux and define a threshold  $f_{\text{thr}}$ . We kept in the sample only those sources for which

$$\log(f_{\text{lim}}) < \log(f_{\text{thr}}) - k \delta_{\text{obs}}, \quad (7)$$

where  $k$  is a number and  $\delta_{\text{obs}}$  is the observed dispersion in the sample. The number of rejected sources depends on the choice of  $k$ , which we evaluated with the following procedure: we divided the sample in redshift bins (see details in Sect. 4) and analysed the relation between the fluxes in the UV and X-ray bands. If the size of the redshift bin is small enough to make the difference in the luminosity distances negligible, that is if the contribution to the dispersion in the relation due to the difference in the luminosity distance among the objects is smaller than the observed  $\delta$ , we can use the flux as a proxy for the luminosity and examine the  $f_X$ – $f_{UV}$  relation, making our analysis cosmology-independent. The inclusion of sources in the sample having a positive fluctuation of their X-ray emission above the flux limit, but with an observed X-ray flux value very close to the flux limit of the given observation, causes a ‘flattening’ of the slope in the relation between the fluxes. We therefore analysed the behaviour of the mean slope and dispersion for the  $f_X$ – $f_{UV}$  relation – in all the redshift bins with a number of sources  $>5$  – at increasing values of  $k$ , that is at more conservative cuts for the Eddington bias. This was done as follows: 1. we divided the sample in redshift bins with size  $\Delta \log z = 0.06$  (similar results are found for  $\log z = 0.05$  and  $\log z = 0.07$ ); 2. we computed the fit to the data for the  $f_X$ – $f_{UV}$  relation, obtaining a slope and intercept, in each redshift bin; 3. we then computed the arithmetic mean for all the slopes in the bins, obtaining an average value. We then repeated the steps above for different choices of  $k \delta_{\text{obs}}$ , with the



**Fig. 6.** Mean slope for the  $f_X - f_{UV}$  relation for the case of  $\Gamma = 1.7-2.8$ , evaluated as the arithmetic mean of all the slopes in the redshift bins with more than five objects, as a function of  $k \delta_{obs}$ , that is of the amount of sources excluded from the sample because of the selection for the Eddington bias (see text for details). The mean slope is flatter for low values of  $k$ , reaches a maximum of  $\gamma \sim 0.58$  and then drops back while  $k$  increases. Mean dispersion as a function of  $k \delta_{obs}$  is listed in blue, while in black the number of sources surviving the selection for each  $k \delta_{obs}$  choice.

intention of determining the  $k \delta_{obs}$  value for which the trade-off between number of sources kept in the sample and the flattening of the relation between  $\gamma$  and  $k \delta_{obs}$  was met. Figure 6 show the eight average slopes result of this analysis for a range of  $k \delta_{obs}$  values from 0.1 to 0.8. This analysis was also performed for different choices of the photon index. Here we show the results for the final choice of the range of accepted photon index values ( $\Gamma = 1.7-2.8$ , see Appendix C for details).

As shown in Fig. 6 for the case of  $\Gamma = 1.7-2.8$ , the mean slope of the relation is flatter for low values of  $k$ , reaches a maximum at  $\gamma \sim 0.57-0.58$  and then drops back while  $k$  increases. The statistical significance of the difference among the slopes based on the  $k \delta_{obs}$  value is small, given their uncertainties, meaning that the presence of the Eddington bias in *Chandra* observations is softened by a very low, basically zero, level of background. This represents a marked difference with respect to the SDSS-XMM sample, for which the dependence of the mean slope of the  $f_X - f_{UV}$  relation in the redshift bins on the chosen cut in  $k \delta_{obs}$  was much stronger (Risaliti & Lusso 2019). We chose  $k \delta_{obs} = 0.6$  as the best compromise between a steady value for the slope of the relation and the number of sources included in the sample in all the cases we examined.

#### 4. Statistical analysis: the $L_X - L_{UV}$ relation

We performed a complete analysis to understand which combination of selection filters in photon index, that is lack of obscuration in the X-rays, and flux threshold, namely Eddington bias, represents the best compromise between the size of the sample and the minimisation of the observed dispersion (see discussion in Appendix C). Our final choice ( $\Gamma = 1.7-2.8$ ,  $k \delta_{obs} = 0.6$ ) reduced the sample to 1385 CSC 2.0 observations (i.e. less than a half of the original 3430 observations) and 140 *Chandra*

COSMOS Legacy sources, corresponding to a total of 1098 single sources (see Table 1). When a source was observed more than once, we chose to compute the X-ray flux as a simple arithmetic mean of the multiple measures (see discussion below).

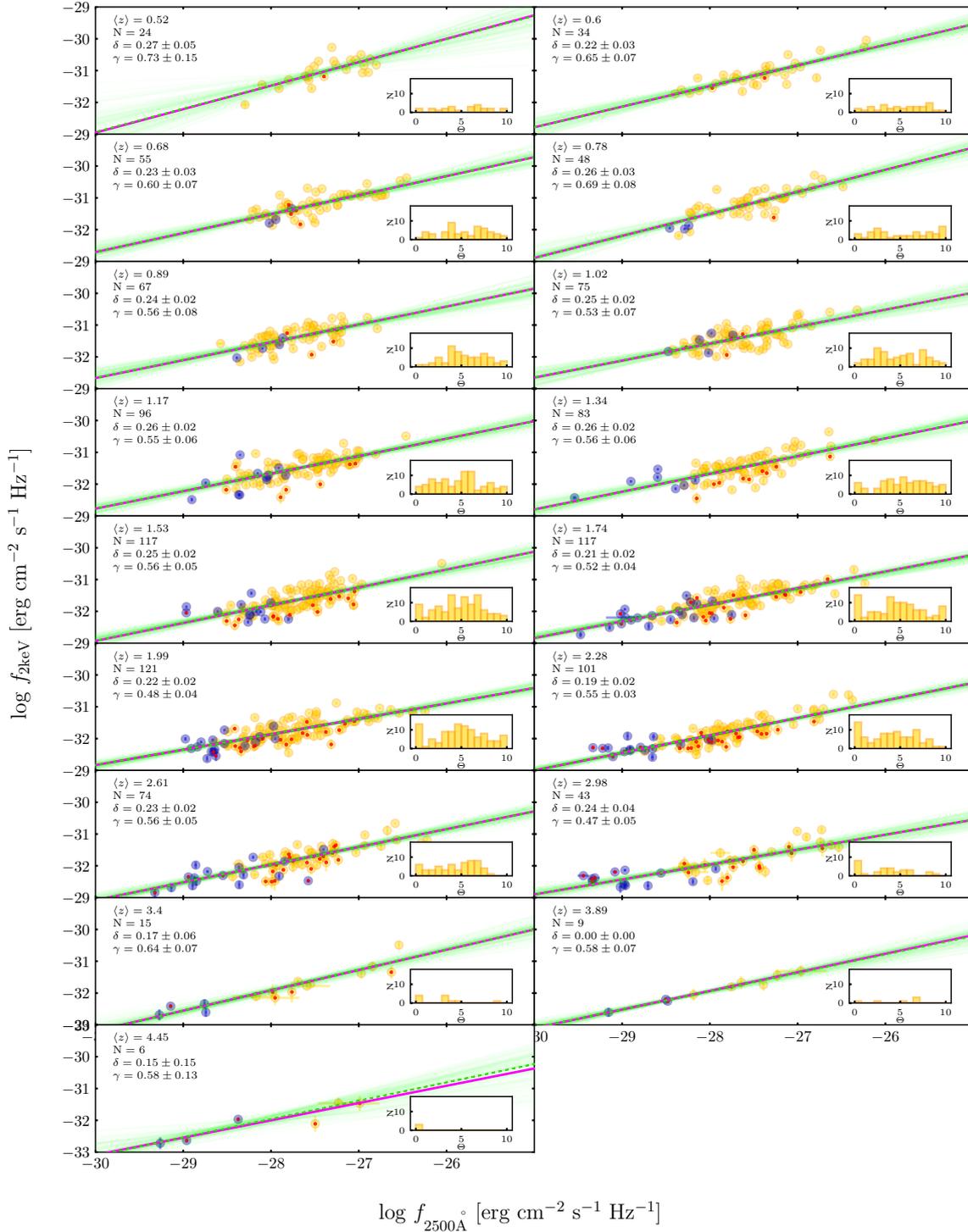
The statistics and the quality of the data involved allowed an examination of the behaviour of the relation with redshift within narrow, and yet well populated, redshift bins. This test served two fundamental purposes. On the one hand, checking the non-evolution of the slope in the relation between the fluxes, the one we are relying on for measuring cosmological distances, is crucial for the cosmological employment of the relation. On the other hand, it gives us the opportunity of examining the physics connecting accretion disc and hot corona up to a redshift at which AGN were younger, and see if the process linking the two components appears to be the same over time.

When the size of the logarithmic redshift bin is small enough, we can use fluxes in place of luminosities, performing a test on the (non-)evolution with redshift that is completely independent from any assumption on cosmology. Risaliti & Lusso (2019) analysed in detail the choice of the bin size and verified that, as long as  $\Delta \log(z) \leq 0.1$ , the slope in the relation does not depend on it. Thanks to the statistics available, we chose  $\Delta \log(z) = 0.06$  and we limited our analysis to the redshift bins with more than five objects. We performed the same analysis for bins of size  $\Delta \log(z) = 0.05$  and  $\Delta \log(z) = 0.07$ , finding no significant difference (see Fig. 9 for a comparison of the results for different sizes of the redshift bins). For our selection in  $\Gamma$  and the choice of the threshold for the Eddington bias, the division yielded 17 redshift bins. To perform the fitting to the data, we adopted the Python package emcee (Foreman-Mackey et al. 2013), a pure-Python implementation of Goodman & Weare’s Affine Invariant Markov chain Monte Carlo (MCMC) Ensemble sampler. To check that the results were independent from the employed method, we also performed the analysis using the Linmix package (Kelly 2007), an algorithm that makes use of a Bayesian approach to linear regression and takes into account the errors in both the  $x$  and  $y$  variable. We performed the fit a first time, then applied a  $3\sigma$  clipping to the data, repeating this sequence for a total of three times. This yielded no significant difference with respect to the analysis without  $\sigma$  clipping. The results are shown in Figs. 7 and 9, and summarised in Table 3. In Fig. 7, red points indicate when the observations are characterised by an  $SN < 5$  in the soft band. Most of them are (X-ray) fainter objects at intermediate redshifts. This confirms that data points that passed our selection criteria, even if with a low SN, follow the relation and are representative of the population of blue quasars.

For the relation between the luminosities on the selected sample, we adopted the same fitting methods and applied a  $3\sigma$  clipping as described above. The result is shown in Fig. 8, where the final selection of CSC 2.0 and COSMOS sources is in gold and blue, respectively, while the original sample of 3430 observations with a  $\delta_{obs} = 0.32$  dex, is in grey. The exclusion of 14/1098 sources at more than  $3\sigma$  from the best-fit led to a final dispersion  $\delta = 0.24 \pm 0.01$  (compared to  $\delta = 0.25 \pm 0.01$  if no sigma-clipping is applied), the same found by Risaliti & Lusso (2019) with the SDSS DR7/DR12 – XMM-Newton sample. The slope is slightly flatter compared to the one of the XMM sample ( $\gamma = 0.59 \pm 0.01$  vs  $\gamma = 0.633 \pm 0.002$ , Risaliti & Lusso 2019).

#### 5. Discussion

Figures 7–9 and Table 3 show the main results of our analysis. First, the slope of the  $f_X - f_{UV}$  relation does not evolve with



**Fig. 7.** Regression analysis of the final selected sample ( $\Gamma = 1.7$ – $2.8$ ,  $k \delta_{\text{obs}} = 0.6$ ). The division of the sample in logarithmic redshift bins of size  $\Delta \log(z) = 0.06$  yields 17 bins with more than five objects. The results for two fitting methods adopted, emcee and linmix, are plotted in (dashed) lime green and magenta, respectively. The inset shows the distribution of off-axis angles  $\Theta$  for the sources in the redshift bin. Sources with at least one observation with  $\text{SN} < 5$  in the soft band are marked with a red point. For each redshift bin, we list the median redshift ( $\langle z \rangle$ ), the number of data points  $N$  and dispersion and slope from the emcee algorithm.

redshift, with an average value around  $\langle \gamma \rangle = 0.58 \pm 0.06$  up to  $z \approx 4.5$ . Second, the mean intrinsic dispersion of the relation has an average around  $\delta \approx 0.20$ , with a decreasing trend with redshift.

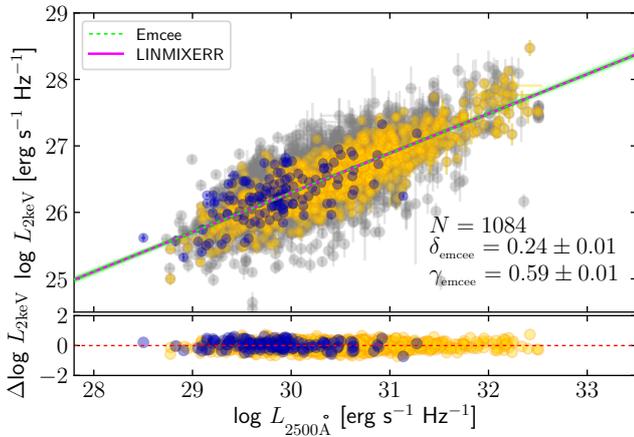
Figure 9 summarises the results on the slope and intrinsic dispersion of the  $f_X$ – $f_{UV}$  relation as a function of the median

redshift for the 17 bins. The regression analysis has been performed on the final selected sample ( $\Gamma = 1.7$ – $2.8$ ,  $k \delta_{\text{obs}} = 0.6$ ). As for the slope, it stays around the reference value ( $\gamma = 0.6$ , orange line) within  $1\sigma$  for most of the bins (13/17) and within  $2\sigma$  for all but one point ( $\langle z \rangle = 1.99$ ), which is however consistent within  $3\sigma$  with the expected value. This result not only implies

**Table 3.** Results of the fitting to the final selected sample.

$\langle z \rangle$	$N$	$\gamma_{\text{emcee}}$	$\beta_{\text{emcee}}$	$\delta_{\text{emcee}}$	$\gamma_{\text{linmix}}$	$\beta_{\text{linmix}}$	$\delta_{\text{obs}}$
0.52	24	$0.73 \pm 0.15$	$-31.3 \pm 0.07$	$0.27 \pm 0.05$	$0.73^{+0.15}_{-0.15}$	$-31.3^{+0.08}_{-0.08}$	$0.26 \pm 0.04$
0.60	34	$0.65 \pm 0.07$	$-31.3 \pm 0.04$	$0.22 \pm 0.03$	$0.65^{+0.08}_{-0.08}$	$-31.3^{+0.04}_{-0.04}$	$0.22 \pm 0.03$
0.68	56	$0.60 \pm 0.07$	$-31.3 \pm 0.03$	$0.23 \pm 0.03$	$0.60^{+0.07}_{-0.07}$	$-31.3^{+0.04}_{-0.04}$	$0.23 \pm 0.02$
0.78	48	$0.69 \pm 0.08$	$-31.3 \pm 0.04$	$0.26 \pm 0.03$	$0.69^{+0.08}_{-0.08}$	$-31.3^{+0.04}_{-0.04}$	$0.26 \pm 0.03$
0.89	67	$0.56 \pm 0.08$	$-31.4 \pm 0.03$	$0.24 \pm 0.02$	$0.56^{+0.08}_{-0.08}$	$-31.4^{+0.03}_{-0.03}$	$0.25 \pm 0.02$
1.02	77	$0.53 \pm 0.07$	$-31.4 \pm 0.03$	$0.25 \pm 0.02$	$0.53^{+0.07}_{-0.07}$	$-31.4^{+0.03}_{-0.03}$	$0.25 \pm 0.02$
1.16	98	$0.55 \pm 0.06$	$-31.5 \pm 0.03$	$0.26 \pm 0.02$	$0.55^{+0.06}_{-0.06}$	$-31.5^{+0.03}_{-0.03}$	$0.27 \pm 0.02$
1.33	84	$0.56 \pm 0.06$	$-31.5 \pm 0.03$	$0.26 \pm 0.02$	$0.56^{+0.06}_{-0.06}$	$-31.5^{+0.03}_{-0.03}$	$0.27 \pm 0.02$
1.53	117	$0.56 \pm 0.05$	$-31.6 \pm 0.02$	$0.25 \pm 0.02$	$0.56^{+0.05}_{-0.05}$	$-31.6^{+0.03}_{-0.03}$	$0.26 \pm 0.02$
1.74	117	$0.52 \pm 0.04$	$-31.6 \pm 0.02$	$0.21 \pm 0.02$	$0.52^{+0.05}_{-0.05}$	$-31.6^{+0.02}_{-0.02}$	$0.23 \pm 0.02$
1.99	123	$0.48 \pm 0.04$	$-31.7 \pm 0.02$	$0.22 \pm 0.02$	$0.48^{+0.04}_{-0.04}$	$-31.7^{+0.02}_{-0.02}$	$0.24 \pm 0.02$
2.28	104	$0.55 \pm 0.03$	$-31.7 \pm 0.02$	$0.19 \pm 0.02$	$0.55^{+0.03}_{-0.03}$	$-31.7^{+0.02}_{-0.02}$	$0.22 \pm 0.02$
2.6	74	$0.56 \pm 0.05$	$-31.8 \pm 0.03$	$0.23 \pm 0.02$	$0.56^{+0.05}_{-0.05}$	$-31.8^{+0.03}_{-0.03}$	$0.25 \pm 0.02$
2.98	43	$0.47 \pm 0.05$	$-31.8 \pm 0.05$	$0.24 \pm 0.04$	$0.47^{+0.05}_{-0.05}$	$-31.8^{+0.05}_{-0.05}$	$0.27 \pm 0.03$
3.4	15	$0.64 \pm 0.07$	$-31.7 \pm 0.06$	$0.17 \pm 0.06$	$0.64^{+0.08}_{-0.08}$	$-31.7^{+0.07}_{-0.08}$	$0.20 \pm 0.04$
3.89	9	$0.58 \pm 0.07$	$-31.8 \pm 0.05$	$0.00 \pm 0.01$	$0.58^{+0.10}_{-0.09}$	$-31.8^{+0.08}_{-0.07}$	$0.06 \pm 0.01$
4.45	6	$0.58 \pm 0.13$	$-31.8 \pm 0.15$	$0.15 \pm 0.15$	$0.56^{+0.24}_{-0.24}$	$-31.8^{+0.22}_{-0.23}$	$0.21 \pm 0.06$

**Notes.** Results of the fitting to the final selected sample ( $\Gamma = 1.7\text{--}2.8$ ,  $k \delta_{\text{obs}} = 0.6$ ) in the 17 redshift bins of size  $\Delta \log(z) = 0.06$ . The fitting was performed with the two Bayesian MCMC methods emcee and linmix (see the text for details).

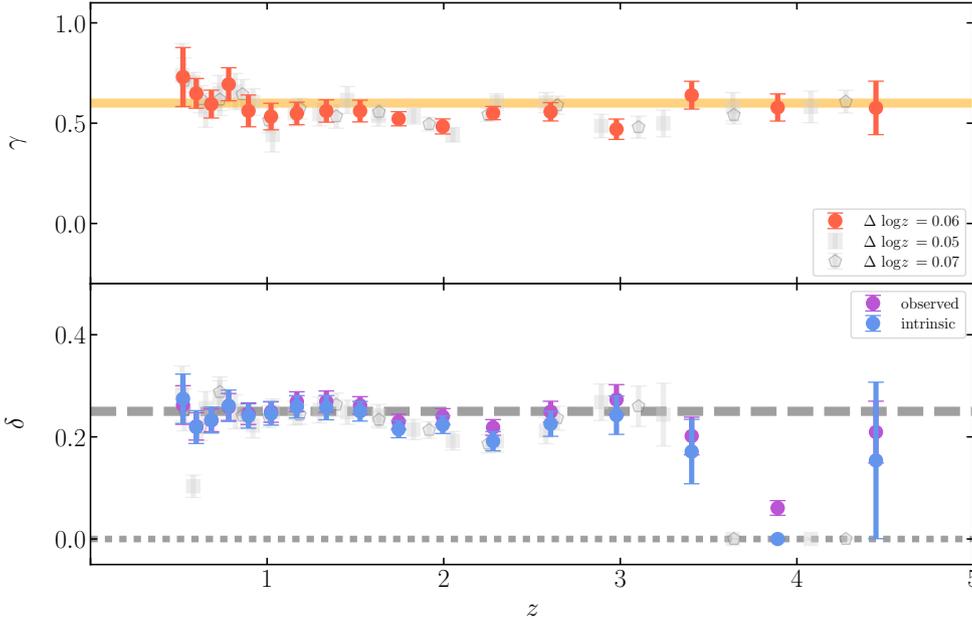


**Fig. 8.** Regression analysis of the  $L_X\text{--}L_{UV}$  relation for the final, entire selected SDSS-CSC2.0 (yellow) and *Chandra* COSMOS Legacy (blue) sample ( $\Gamma = 1.7\text{--}2.8$ ,  $k \delta_{\text{obs}} = 0.6$ ), while in grey is shown the pre-selected parent sample (2332 SDSS-CSC2.0 (3430 observations) and 273 *Chandra* COSMOS Legacy sources), before the selection in the X-ray band is applied (the observed dispersion in this case is  $\delta_{\text{obs}} = 0.32$  dex). The use of spectroscopic data leads to a final dispersion of 0.24 dex, in agreement with the one found in previous works on samples of similar size using photometric data.

that the same physical process empowering the hot corona for the X-ray emission has to be present since when quasars were younger objects (at least at  $z \sim 4$ ), but also that we can safely rely on the  $f_X\text{--}f_{UV}$  relation for measuring cosmological distances and build the Hubble diagram for quasars, the tool that allows us to probe for the first time the space of cosmological parameters in the redshift range  $z \sim 1.4\text{--}5$ .

One of the most important results that we obtained is shown in the bottom panel of Fig. 9, where the violet data points show

the observed dispersion for the  $\Delta \log z = 0.06$  redshift bins, while the light blue ones show the corresponding intrinsic dispersion, output of the emcee regression analysis (the same legend of the upper panel applies here for the grey data points). While moving towards higher redshifts, the dispersion drops constantly to smaller values compared to the average value of lower redshift bins. Given the limited number of sources in the high redshift bins, and the presence of a couple of points with large error bars in  $L_{UV}$  that ease the tightness in the relation, we conservatively take as a reference the observed dispersion. There is a clear indication of a decrease in the dispersion at high redshift: the observed dispersion (in violet) in the last three bins ranges around 0.06–0.20 dex (Table 3). Similar results are obtained for  $\Delta \log z = 0.05$  and 0.07, with the remarkable exception of the second  $\Delta \log z = 0.05$  bin for which the relation has significantly lower dispersion ( $\delta_{\text{obs}} = 0.10$  dex) with respect to the other bin sizes. This behaviour is the result of the inclusion of distant sources that have been detected with dedicated observation, as opposed to most closer sources, whose presence in the catalogue is based on serendipitous observations. This is shown in Fig. 7, where the inset histograms in the right corner of each  $z$ -bin subplot show the distribution of the off-axis angle  $\Theta$  for the sources in the bin. For higher  $z$ , the contribution of sources with  $\Theta < 5'$  becomes progressively more important. As proved by a sub-sample of 18 sources at  $z \sim 3$  observed with *XMM-Newton* (Risaliti & Lusso 2019; Nardini et al. 2019; Lusso et al. 2020), performing X-ray dedicated observations is crucial to reducing the dispersion associated with calibration issues in the X-ray band. For this sub-sample of 18 sources only, our group performed a spectral analysis both in both the UV and the X-ray band, obtaining a final dispersion of  $\delta = 0.12$ , to be compared to the  $\delta = 0.24$  of the entire, photometric sample (Risaliti & Lusso 2019). The very low background level of the spectra in the CSC 2.0, where pointed *Chandra* observations were performed,



**Fig. 9.** Analysis of the  $f_X-f_{UV}$  relation in redshift bins. *Top panel:* slope  $\gamma$  as a function of redshift. The slope stays around the reference value of 0.6 for the entire redshift range explored. *Bottom panel:* dispersion as a function of redshift. The dispersion stays around the median value  $\sim 0.25$  dex, showing a decreasing trend with redshift. This is easily explained by the fact that, at the highest redshifts, only pointed, dedicated observations, unaffected by the observational issues characterising serendipitous observations, are available. Both intrinsic dispersion (output of the emcee regression analysis) and observed one are shown.

allowed us to improve the fits of the relation well beyond  $z \simeq 3$ . In fact, the uncertainties in the X-ray measures are consistent with those in the UV (obtained through a fitting of the DR7 spectra and through an interpolation of the photometric SED for the remaining quasars). Quasars from the *XMM-Newton* sample that were drawn from the catalogue of serendipitous sources do not go beyond redshift  $z \sim 3.3$ . Measures beyond this redshift are mostly from pointed observations (see Lusso et al. 2020 for further details) available in the literature, mostly *Chandra* observations covering a redshift range  $z = 4.01-7.08$ . These have been studied by our group with the aim of examining the relation at the highest redshifts (Salvestrini et al. 2019). For the latter sample, we found a slope  $\gamma = 0.53 \pm 0.11$ , flatter than those in our higher redshift bins, but fully consistent (within  $1\sigma$ ) with our results. In the present study, leveraging on *Chandra*'s capabilities, we are therefore able to significantly extend the use of catalogue sources with respect to the *XMM-Newton* sample, up to  $z \sim 4.5$ .

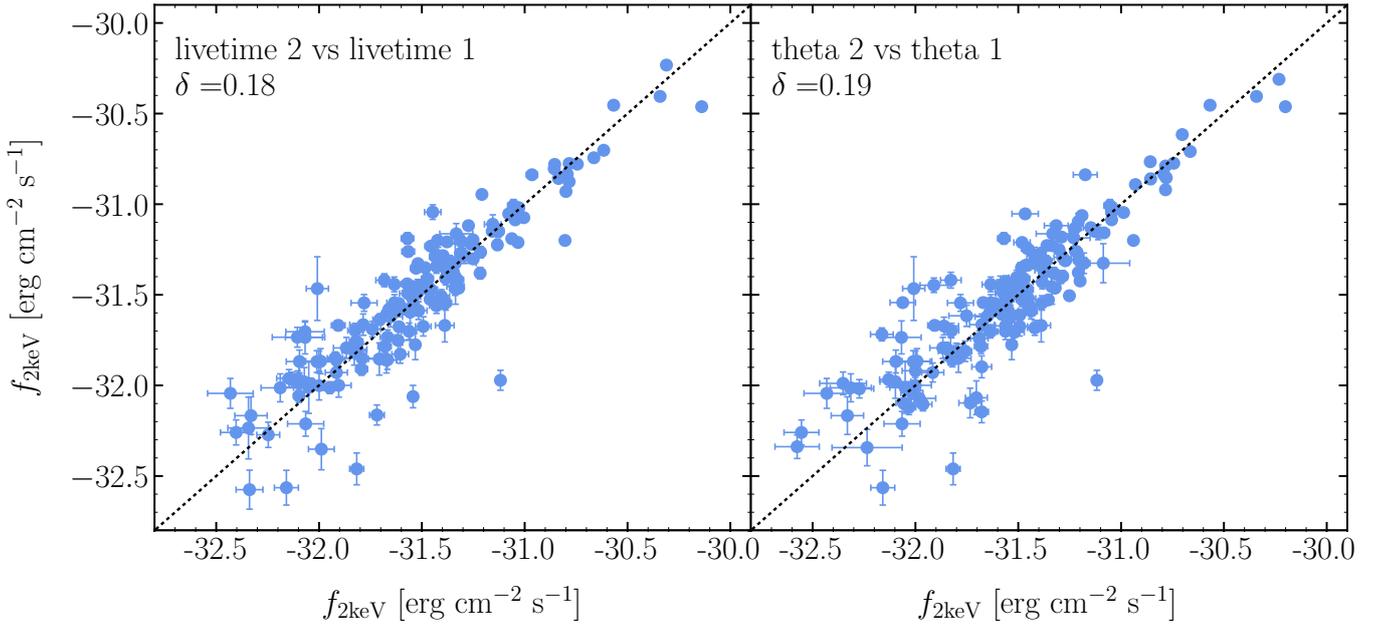
This result is crucial for the cosmological application of quasars through the  $f_X-f_{UV}$  relation, that is for establishing whether or not we can consider quasars as standard candles. In principle, with such a small dispersion, we can measure cosmological distances with an even better precision than with supernovae ( $\delta_{SNe} \sim 0.06$ ), and at redshifts that supernovae could never probe in similar numbers ( $z > 2$ ). Even when the dispersion is higher, if we consider for example the observed dispersion in the last three bins ( $\langle\delta\rangle \sim 0.15$ ), the amount of cosmological information provided by one supernova Ia can be achieved with  $(0.15/0.06)^2 \sim 7$  quasars, with the important difference that no supernova Ia has ever been spectroscopically confirmed at redshift higher than 2.26 (Scolnic et al. 2018), whereas hundreds of thousands of quasars are instead being discovered and observed by extragalactic all-sky survey in the last years, and even the most distant objects have already been targeted with dedicated X-ray observations up to redshift  $z \sim 7.5$  (Bañados et al. 2018). While the decrease in the dispersion with respect to previous works is fully appreciable in the analysis of the  $f_X-f_{UV}$  relation, especially in the high-redshift bins, it appears to be diluted when we analyse the relation between luminosities (Fig. 8). As a result, the dispersion is comparable to the one in the final SDSS-XMM sample ( $\delta \sim 0.24$ ). The similarity in the dispersion found for

the *Chandra* and the *XMM-Newton* samples, even if for the first one a complete spectroscopical analysis was performed while for the second one the flux at 2 keV was inferred from photometry, is not surprising if we consider that the *Chandra* sample is, on average, fainter than the XMM one<sup>9</sup>. A thorough comparison of these samples can be found in Lusso et al. (2020), where the overlap of  $\sim 200$  sources at  $z < 3$  between the two is taken into account.

### 5.1. Contribution of X-ray variability to the dispersion

We have reduced the dispersion in the  $f_X-f_{UV}$  relation by using spectroscopic data for the X-ray band, especially in the high redshift bins, where all the observations are dedicated. However, we still have to account for many other contributors, among which there are some we can address (e.g. variability in the X-rays and in the UV band) and some instead we can only acknowledge (e.g. radiations emitted by accretion disc and hot corona are intrinsically delayed, preventing an actual simultaneous monitoring of the two structures). As already mentioned, the CSC 2.0 makes available all the observations carried out for the sources in the catalogue. After having applied all of our filters in both the UV and X-ray band, we therefore have a subset of sources for which multiple observations ( $N_{\text{obs}} \geq 2$ ) are available (164/958, 17% of the sample). By analysing the  $L_X-L_{UV}$  relation for the sources with multiple observations, we can give an estimate of the contribution of X-ray variability to the dispersion in our sample. The best value for the 2 keV flux can be chosen using two different criteria: we can opt for the observation with the longest exposure time or the one in which the source has the smallest off-axis angle  $\Theta$ . In both cases we are looking for an optimisation of the source against the background flux, and for the most accurate and precise measure of the X-ray flux. Both, however, will

<sup>9</sup> For the final selection, the distribution of  $\log f_{UV}$  has a mean (median) value of  $-27.72 \pm 0.57$  ( $-27.72_{-0.35}^{+0.36}$ ) and  $-27.48 \pm 0.45$  ( $-27.50_{-0.28}^{+0.29}$ ) for the *Chandra* and *XMM-Newton* samples, respectively. Similarly, the  $\log f_X$  distribution has a mean (median) value of  $-31.60 \pm 0.46$  ( $-31.59_{0.35}^{+0.30}$ ) and  $-31.36 \pm 0.44$  ( $-31.41_{-0.26}^{+0.29}$ ) for the *Chandra* and *XMM-Newton* samples, respectively. The values listed as confidence intervals for the median are the 25th and 75th percentiles.



**Fig. 10.** Comparison of the rest-frame 2-keV flux as inferred from the longest versus the second longest observation (*left panel*), and from the observation characterised by the smallest versus the second smallest off-axis angle  $\Theta$ . The dispersions with respect to the 1:1 relation are a combination of the contribution due to the variability in the X-ray band and to observational issues, especially to flux calibration, related to different off-axis angle of the source in different observations.

be necessarily dependent on the flux level at the moment of the observation, namely on the variability of the source. We compared the 2-keV flux for the observations with the longest and second longest exposure times (Fig. 10, left panel), and those for the smallest and second smallest off-axis angle (right panel). The dispersion in the relation with respect to the bisector is, in both cases, the combination of the contribution due to the variability in the X-ray band with the contribution due to the off-axis angle of the source, that is a flux-calibration issue in the X-ray band. To minimise these contributions, we performed an arithmetic mean of the 2-keV fluxes for all the observations of each source. The result is shown in Fig. 11, where we compare the relation using the observation with the longest exposure (top panel), the smallest off-axis angle  $\Theta$  (mid panel) and the arithmetic mean of all the observations available. The dispersion for the sub-sample of sources with multiple observations is characterised by an overall higher dispersion ( $\delta \sim 0.27$  dex) if compared with the total sample ( $\delta \sim 0.24$  dex). In the case of the mean, however, we observe a significant decrease in the dispersion ( $\delta \sim 0.23$  dex), from which we can give an estimate of the contribution of the variability of

$$\sqrt{0.27^2 - 0.23^2} \sim 0.14 \text{ dex.} \quad (8)$$

This result is of the same order of magnitude of what was found from the comparison in Fig. 10, and it is in agreement with Lusso & Risaliti (2016).

### 5.2. Additional contributions to the dispersion

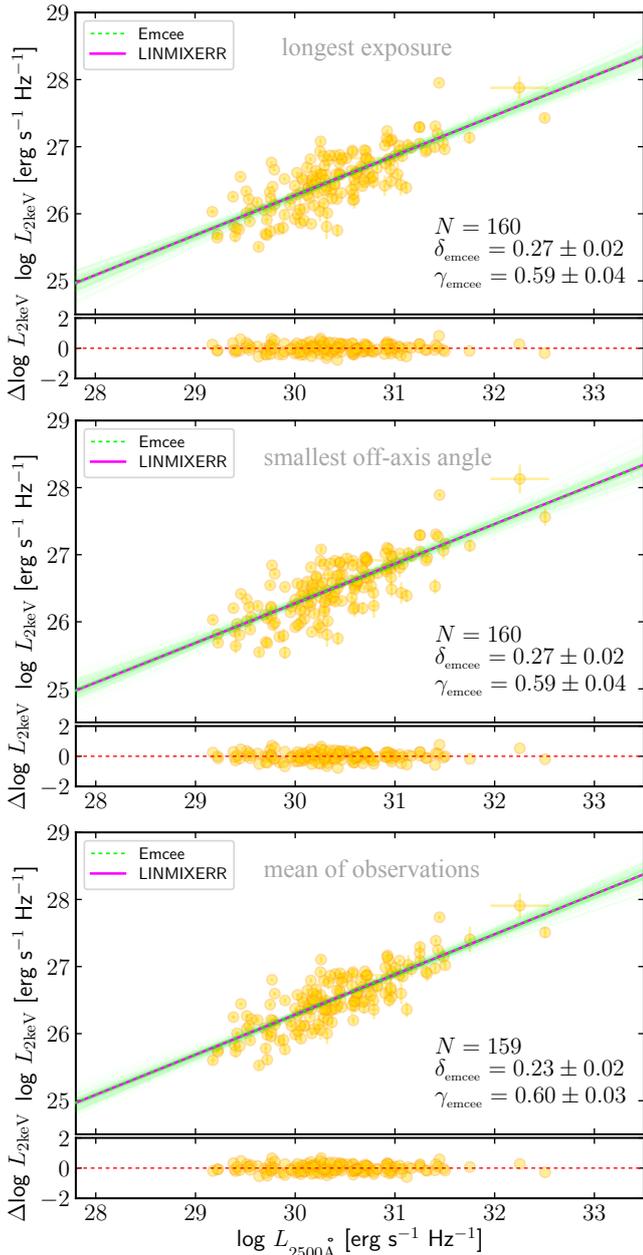
Another contribution to the dispersion is associated with the variability in the UV band, even if to a much lesser extent ( $\sim 0.05$  mag; Sesar et al. 2007; Kozłowski et al. 2010; Ai et al. 2010). Light-curve fitting is the ideal tool to quantify variability. Multiple photometric observations spanning several years and calibrated to better than 0.02 mag are required to produce well-sampled light curves for each object. The relatively low number

of optical/UV observations (two on average, performed during the SDSS and BOSS campaigns) for a small number of objects in our final cleaned sample (<20% have more than one observation, SDSS or BOSS, and only 4% have two sets of *ugriz* magnitudes from SDSS and BOSS) and the sparse sampling do not allow a reliable quantification of variability using this methodology in our sample. Low-order statistics (e.g. rms scatter) is a viable option but the number of objects in our final sample for which this analysis can be performed is so small that any number would be affected by high uncertainties, much higher than the observed variability amplitude. Some additional artificial variability introduced by the non-simultaneity of the X-ray and UV data could be present, although it does not represent a significant contribution to the dispersion (see Sect. 5 in Lusso & Risaliti 2016).

An additional important element to be taken into account is the orientation of the accretion disc, which causes its intrinsic emission to be scaled by a function of the inclination angle. We are planning to apply a correction for the disc inclination with respect to the line of sight by making use of the rest-frame equivalent width of the [O III] emission line,  $EW[\text{O III}]$ , for the sources for which this line is available in SDSS ( $z < 0.8$ ). This quantity has been proposed by our group as an orientation indicator of the accretion disc (Risaliti et al. 2011; Bisogni et al. 2017b, 2019). We will further analyse this aspect in a future work.

### 5.3. Simulating the effect of observational contaminants to the dispersion

To verify whether the decrease in the dispersion examined in this study is reasonable, we performed the following exercise: starting from a simulated sample, we progressively added the observational effects that we have removed from the real one. We started by bootstrapping a distribution of UV fluxes from the data-sets composed by 2332 SDSS DR14 and 288 COSMOS sources (2620 quasars in total) with a measured flux limit at



**Fig. 11.**  $L_X$ – $L_{UV}$  relation for the sub-sample of 160 sources with multiple observations. The X-ray luminosity has been inferred for each source from the longest observation (*top panel*), from the one with the smallest off-axis angle (*middle panel*) and from an arithmetic mean of all the fluxes available, that is of all the observations surviving the quality filters that pertain to the single source.

X-rays. As discussed in Sect. 2, the above sample was selected to be homogeneous in the optical/UV, where observational biases and contaminants are minimised. To a reasonable extent, the sample contains only sources free of – or very poorly affected by – dust reddening and host galaxy contamination, which can be used as a suitable pool for the simulated fluxes. Fainter objects are preferentially mapped, in the X-ray band, by the deeper *Chandra* COSMOS Legacy survey, with a flux limit per observation and a flux error that are smaller than those of CSC 2.0 observations. In order to shape a realistic sample, we performed 2000 extractions of UV fluxes (similar in size to the original sample) and, contextually, extracted errors for the UV and X-ray measurements, and a flux limit value in the X-ray band. Assuming

a perfect, that is dispersion-free,  $L_X$ – $L_{UV}$  relation, we inferred fluxes at 2 keV for all the sources.

We first simulated the contribution of X-ray variability to the dispersion in the  $L_X$ – $L_{UV}$  as follows. We assumed X-ray variability for a source to be represented by a log-normal distribution centred on its average flux value, corresponding, in our case, to the ‘expected’  $L_X$  inferred directly from the relation with a standard deviation  $\sigma$  representing the characteristic fractional variability of the sources (Maughan & Reiprich 2019). The effect of variability was simulated by extracting a random value from the log-normal distribution for each source to reproduce the observed  $L_X$ – $L_{UV}$  relation.

AGN variability is found to increase with rest-frame timescale (Paolillo et al. 2004; Vagnetti et al. 2011), reaching up to 50% on timescales of decades (Middei et al. 2017), but it is also found to depend on the source luminosity (Vagnetti et al. 2016). The characteristic fractional variability  $\sigma$  was then chosen depending on the X-ray luminosity of the source. Following Vagnetti et al. (2016), we assigned a range of fractional variability increasing with 0.5–4.5 keV luminosity<sup>10</sup> of  $L < 10^{44}$ ,  $10^{44} < L < 10^{45}$ ,  $L > 10^{45}$  erg s<sup>–1</sup>. The  $\sigma$  to be associated with a source, was computed as the mean of 1000 extractions from a uniform distribution of a fractional variability ranging from 0 to a maximum value that depended on the luminosity of the source (0.45 for the faintest, 0.26 for the intermediate, and 0.14 for the brightest sources). An extraction from the log-normal distribution for each source provided the ‘observed’  $L_X$  value.

A fit to the simulated data, when a single value was extracted from the log-normal distribution, matched the dispersion of 0.14 dex due to variability in the real sample and gave a slope  $\gamma = (0.65 \pm 0.01)$ . When the mean of 1000 extractions was used instead, the same dispersion dropped to  $\sim 0.12$  dex.

The inclusion of sources on an emission level higher than their average is preferentially located in – even if not limited to – the low-luminosity end of the relation, and causes a variation in the overall slope of the relation. The flux limits associated to the simulated sample were used to exclude the sources with flux at 2 keV close to the observation flux limit (Eddington bias). This step did not significantly modify the slope observed in the relation and reduced only slightly the amount of dispersion ( $\delta = 0.13$  dex vs 0.14 dex).

Finally, to account for the unlikely presence of heavy obscuration ( $N_H > 10^{22}$  cm<sup>–2</sup>) in the sample, we performed a simulation to roughly quantify this effect. It is important to note that, in the redshift range covered by our sample, we are observing the rest-frame hard band which is less affected by the presence of absorption. Specifically, we multiplied the luminosities for the exponential term  $e^{-N_H \sigma_E}$ , where the photoionisation cross-section  $\sigma_E$  has been computed at 2 keV following Verner et al. (1996). We then considered  $N_H$  ranging from  $0.5 \times 10^{21}$  to  $0.5 \times 10^{23}$  cm<sup>–2</sup>, finding a slope of the relation steadily around  $\sim 0.6$  and an intercept that decreases progressively from  $6.8 \pm 0.2$  ( $N_H = 0.5 \times 10^{21}$  cm<sup>–2</sup>) to  $6.7 \pm 0.2$  for ( $N_H = 10^{22}$  cm<sup>–2</sup>) and  $6.4 \pm 0.2$  for the most heavily obscured cases ( $N_H = 0.5 \times 10^{23}$  cm<sup>–2</sup>), while the intercept for the real sample is  $8.5 \pm 0.4$ . For column densities larger than that,  $\gamma$  starts decreasing sensibly. Even if some obscuration in the X-ray band is left in the sample after the selection in terms of the photon index  $\Gamma$ , a comparison between slope and intercept in the case of the real data and those for the simulated sample sets  $N_H \sim 10^{22}$  cm<sup>–2</sup> as an

<sup>10</sup> The luminosity in the 0.5–4.5 keV band was computed from the luminosity at 2 keV by adopting a photon index of  $\Gamma = 1.7$  (Vagnetti et al. 2016).

upper-limit estimate of the column density for the sources in the original sample. This simulation is meant to be highly conservative and overestimates the number of absorbed sources. A similar effect is found if we assume absorption in the optical band: lower optical luminosities for larger extinction result in an increasing intercept of the relation.

## 6. Conclusions

We presented an analysis of the  $L_X-L_{UV}$  relation on a sample of 3430 observations (2332 sources) from the latest release of the *Chandra* Source Catalog (CSC 2.0), and of 273 sources from the *Chandra* COSMOS Legacy survey. The aim of this work was to employ the largest spectroscopic X-ray sample available up to now with *Chandra* to decrease the dispersion in the relation, verifying its non-evolution with redshift and examining the X-ray variability contribution thanks to the multiple observations available. We pre-selected the parent sample to include only the sources representative of the intrinsic relation, that is type-1, radio-quiet, non-BAL, non-dust-reddened quasars. We then filtered out all the observations that did not meet the required quality criteria in the X-ray band: we excluded possibly absorbed sources characterised by a photon index  $\Gamma < 1.7$ , and observations affected by the Eddington bias. The analysis of the relation delivers the following results:

1. The X-ray-to-UV flux relation analysed in small redshift bins – small enough to make the difference in the luminosity distance for sources in the same bin negligible – does not show any statistically significant evolution with redshift. The slope stays around the expected value of 0.6 within the entire redshift range probed by our sample ( $z \sim 0.5-4.5$ ).
2. The dispersion in the relation between fluxes is in agreement with what found in previous works ( $\langle\delta\rangle \sim 0.24$  dex) up to  $z \sim 3$ . It strongly decreases in the last redshift bins ( $\langle\delta_{\text{obs}}\rangle \sim 0.15$  dex). This behaviour is explained by two facts: (a) at the highest redshifts, the majority of the objects have dedicated, pointed X-ray observations for which calibration issues are minimised; (b) the decrease is also ascribable to the use of spectroscopic data in the X-ray band and to the very low level of background of *Chandra* observations,
3. Over the entire clean sample, the relation between the X-ray and UV/optical luminosities has a dispersion of  $\sim 0.24$  dex, comparable to the one found by previous works with samples of similar size.
4. The analysis of the sub-sample of sources with multiple observations available shows that the mean 2-keV flux, obtained from all the observations that have survived the quality filters, minimises the dispersion in the  $L_X-L_{UV}$  relation with respect to the choice of the single ‘best’ observation, in terms of both longest exposure time and smallest off-axis angle. Performing a mean of the observations corresponds to removing part of the dispersion, mostly due to X-ray variability, but also to possible issues in the flux calibration related to the off-axis angle of the source. We estimated a contribution of the X-ray variability or calibration issues of the order of  $\sim 0.14$  dex, in agreement with previous results.
5. The analysis of observational contaminants on a simulated quasar sample confirms the amount of dispersion in the relation ascribable to X-ray variability ( $\delta \sim 0.14$  dex) and allows an estimate of the upper limit for the average column density in the sample ( $N_{\text{H}} \leq 10^{22} \text{ cm}^{-2}$ ).

Points 1. and 2. have major implications. First of all the interplay between accretion disc and hot corona over the entire redshift

range examined so far has to be universal. The relation is very tight once we exclude any possible contribution to the dispersion that is not intrinsic. Secondly, we can use the relation to robustly infer cosmological distances at every redshift. These two results further justify the employment of quasars as standardisable candles in cosmology.

*Acknowledgements.* SB acknowledges funding from the INAF Ob.Fun. 1.05.03.01.09 Supporto Arizona – LBT Italia. SB was also supported by NASA through the *Chandra* award no. AR7-18013X issued by the *Chandra* X-ray Observatory Center, operated by the Smithsonian Astrophysical Observatory for and on behalf of NASA under contract NAS8-03060, and partially by grant HST-AR-13240.009. EL acknowledges the support of grant ID: 45780 Fondazione Cassa di Risparmio Firenze. This research has made use of data obtained from the *Chandra* Source Catalog, provided by the *Chandra* X-ray Center (CXC) as part of the *Chandra* Data Archive. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the US Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS website is [www.sdss.org](http://www.sdss.org). SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University. This work made use of matplotlib, a Python library for publication quality graphics (Hunter 2007), and of the software for the analysis and manipulation of catalogues and tables TOPCAT (Taylor 2005).

## References

- Ai, Y. L., Yuan, W., Zhou, H. Y., et al. 2010, *ApJ*, 716, L31  
 Allen, J. T., Hewett, P. C., Maddox, N., Richards, G. T., & Belokurov, V. 2011, *MNRAS*, 410, 860  
 Arnaud, K. A. 1996, in *Astronomical Data Analysis Software and Systems V*, eds. G. H. Jacoby, & J. Barnes, *ASP Conf. Ser.*, 101, 17  
 Avni, Y., & Tananbaum, H. 1986, *ApJ*, 305, 83  
 Bañados, E., Venemans, B. P., Mazzucchelli, C., et al. 2018, *Nature*, 553, 473  
 Bargiacchi, G., Risaliti, G., Benetti, M., et al. 2021, *A&A*, 649, A65  
 Bisogni, S., Risaliti, G., & Lusso, E. 2017a, *Front. Astron. Space Sci.*, 4, 68  
 Bisogni, S., Marconi, A., & Risaliti, G. 2017b, *MNRAS*, 464, 385  
 Bisogni, S., Lusso, E., Marconi, A., & Risaliti, G. 2019, *MNRAS*, 485, 1405  
 Civano, F., Elvis, M., Brusa, M., et al. 2012, *ApJS*, 201, 30  
 Civano, F., Marchesi, S., Comastri, A., et al. 2016, *ApJ*, 819, 62  
 Di Matteo, T. 1998, *MNRAS*, 299, L15  
 Elvis, M. 2000, *ApJ*, 545, 63  
 Elvis, M., Civano, F., Vignali, C., et al. 2009, *ApJS*, 184, 158  
 Evans, I. N., Primini, F. A., Glotfelty, K. J., et al. 2010, *ApJS*, 189, 37  
 Fitzpatrick, E. L. 1999, *PASP*, 111, 63  
 Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306  
 Gibson, R. R., Jiang, L., Brandt, W. N., et al. 2009, *ApJ*, 692, 758  
 Haardt, F., & Maraschi, L. 1991, *ApJ*, 380, L51  
 Haardt, F., & Maraschi, L. 1993, *ApJ*, 413, 507  
 Hao, H., Elvis, M., Civano, F., et al. 2014, *MNRAS*, 438, 1288  
 Hunter, J. D. 2007, *Comput. Sci. Eng.*, 9, 90  
 Just, D. W., Brandt, W. N., Shemmer, O., et al. 2007, *ApJ*, 665, 1004  
 Kellermann, K. I., Sramek, R., Schmidt, M., Shaffer, D. B., & Green, R. 1989, *AJ*, 98, 1195  
 Kelly, B. C. 2007, *ApJ*, 665, 1489  
 Kozłowski, S., Kochanek, C. S., Udalski, A., et al. 2010, *ApJ*, 708, 927

- Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, *ApJS*, **224**, 24
- Liu, B. F., Mineshige, S., & Shibata, K. 2002, *ApJ*, **572**, L173
- Lusso, E., & Risaliti, G. 2016, *ApJ*, **819**, 154
- Lusso, E., & Risaliti, G. 2017, *A&A*, **602**, A79
- Lusso, E., Comastri, A., Vignali, C., et al. 2010, *A&A*, **512**, A34
- Lusso, E., Piedipalumbo, E., Risaliti, G., et al. 2019, *A&A*, **628**, L4
- Lusso, E., Risaliti, G., Nardini, E., et al. 2020, *A&A*, **642**, A150
- Marchesi, S., Civano, F., Elvis, M., et al. 2016, *ApJ*, **817**, 34
- Maughan, B. J., & Reiprich, T. H. 2019, *Open J. Astrophys.*, **2**, 9
- McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, *A&A*, **544**, A156
- Merloni, A., & Fabian, A. C. 2001, *MNRAS*, **321**, 549
- Middei, R., Vagnetti, F., Bianchi, S., et al. 2017, *A&A*, **599**, A82
- Mingo, B., Watson, M. G., Rosen, S. R., et al. 2016, *MNRAS*, **462**, 2631
- Murray, N., Chiang, J., Grossman, S. A., & Voit, G. M. 1995, *ApJ*, **451**, 498
- Nardini, E., Lusso, E., Risaliti, G., et al. 2019, *A&A*, **632**, A109
- Paolillo, M., Schreier, E. J., Giacconi, R., Koekemoer, A. M., & Grogin, N. A. 2004, *ApJ*, **611**, 93
- Pâris, I., Petitjean, P., Ross, N. P., et al. 2017, *A&A*, **597**, A79
- Pâris, I., Petitjean, P., Aubourg, É., et al. 2018, *A&A*, **613**, A51
- Planck Collaboration VI. 2020, *A&A*, **641**, A6
- Prevot, M. L., Lequeux, J., Maurice, E., Prevot, L., & Rocca-Volmerange, B. 1984, *A&A*, **132**, 389
- Puccetti, S., Vignali, C., Cappelluti, N., et al. 2009, *ApJS*, **185**, 586
- Richards, G. T., Lacy, M., Storrie-Lombardi, L. J., et al. 2006, *ApJS*, **166**, 470
- Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. 2019, *ApJ*, **876**, 85
- Risaliti, G., & Lusso, E. 2015, *ApJ*, **815**, 33
- Risaliti, G., & Lusso, E. 2019, *Nat. Astron.*, **3**, 272
- Risaliti, G., Salvati, M., & Marconi, A. 2011, *MNRAS*, **411**, 2223
- Salvestrini, F., Risaliti, G., Bisogni, S., Lusso, E., & Vignali, C. 2019, *A&A*, **631**, A120
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, **500**, 525
- Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, *ApJ*, **859**, 101
- Sesar, B., Ivezić, Ž., Lupton, R. H., et al. 2007, *AJ*, **134**, 2236
- Shen, Y., Richards, G. T., Strauss, M. A., et al. 2011, *ApJS*, **194**, 45
- Smolčić, V., Novak, M., Bondi, M., et al. 2017, *A&A*, **602**, A1
- Steffen, A. T., Strateva, I., Brandt, W. N., et al. 2006, *AJ*, **131**, 2826
- Tananbaum, H., Avni, Y., Branduardi, G., et al. 1979, *ApJ*, **234**, L9
- Taylor, M. B. 2005, in *Astronomical Data Analysis Software and Systems XIV*, eds. P. Shopbell, M. Britton, & R. Ebert, *ASP Conf. Ser.*, **347**, 29
- Vagnetti, F., Turriziani, S., Trevese, D., & Antonucci, M. 2010, *A&A*, **519**, A17
- Vagnetti, F., Turriziani, S., & Trevese, D. 2011, *A&A*, **536**, A84
- Vagnetti, F., Middei, R., Antonucci, M., Paolillo, M., & Serafinelli, R. 2016, *A&A*, **593**, A55
- Verner, D. A., Ferland, G. J., Korista, K. T., & Yakovlev, D. G. 1996, *ApJ*, **465**, 487
- Vignali, C., Brandt, W. N., & Schneider, D. P. 2003, *AJ*, **125**, 433
- Weymann, R. J., Morris, S. L., Foltz, C. B., & Hewett, P. C. 1991, *ApJ*, **373**, 23
- Wilkes, B. J., & Elvis, M. 1987, *ApJ*, **323**, 243
- Wilkes, B. J., Tananbaum, H., Worrall, D. M., et al. 1994, *ApJS*, **92**, 53
- Young, M., Elvis, M., & Risaliti, G. 2010, *ApJ*, **708**, 1388
- Zamorani, G., Henry, J. P., Maccacaro, T., et al. 1981, *ApJ*, **245**, 357

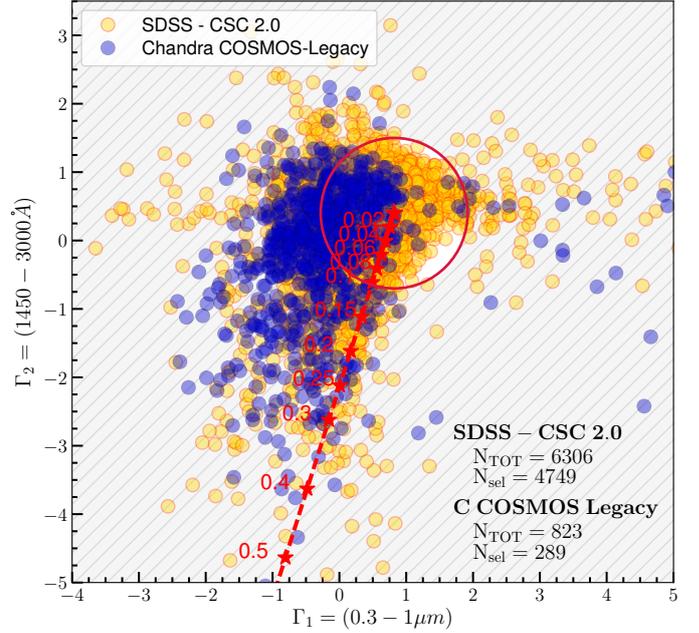
## Appendix A: UV/optical/NIR SED analysis

**Table A.1.** Instruments and filters for UV/optical/NIR photometry for the SDSS-CSC2.0 and the *Chandra* COSMOS Legacy samples.

Band	Instrument	$\lambda$ ( $\text{\AA}$ )
<b>SDSS – CSC 2.0</b>		
<i>FUV</i>	<i>GALEX</i>	1539
<i>NUV</i>	-	2316
<i>u</i>	SDSS	3557
<i>g</i>	-	4825
<i>r</i>	-	6261
<i>i</i>	-	7672
<i>z</i>	-	9097
<i>Y</i>	UKIDSS	10305
<i>J</i>	-	12483
<i>H</i>	-	16313
<i>K</i>	-	22010
<i>J</i>	2MASS	12350
<i>H</i>	-	16620
<i>K</i>	-	21590
<i>W<sub>1</sub></i>	<i>WISE</i>	34000
<i>W<sub>2</sub></i>	-	46000
<i>W<sub>3</sub></i>	-	120000
<i>W<sub>4</sub></i>	-	220000
ch1	<i>Spitzer</i> (IRAC)	36000
ch2	-	45000
ch3	-	58000
ch4	-	80000
<b>Chandra COSMOS Legacy</b>		
<i>NUV</i>	<i>GALEX</i>	2316
<i>u</i>	CFHT(MegaCam)	3881.58
<i>B</i>	SUBARU (Suprime-Cam)	4478
<i>V</i>	-	5493
<i>r</i>	-	6315.
<i>i</i>	-	7683.9
	CFHT (WIRCam)	7617.66
	SDSS	7672.0
<i>z<sup>+</sup></i>	SUBARU (Suprime-Cam)	9021.6
<i>Y-HSC</i>	SUBARU (HyperSuprimeCam)	9779.93
<i>Y</i>	VISTA (VIRCAM)	10214.2
<i>J</i>	-	12534.6
<i>H</i>	-	16453.4
<i>K<sub>S</sub></i>	-	21539.9
	CFHT (WIRCam)	21480.2
ch1	<i>Spitzer</i> (IRAC)	35550
	-	35634.3
ch2	-	45110.1
ch3	-	57593.4
ch4	-	79594.9

In this appendix we give a detailed description of the analysis in the UV/optical/NIR bands for the two sub-samples.

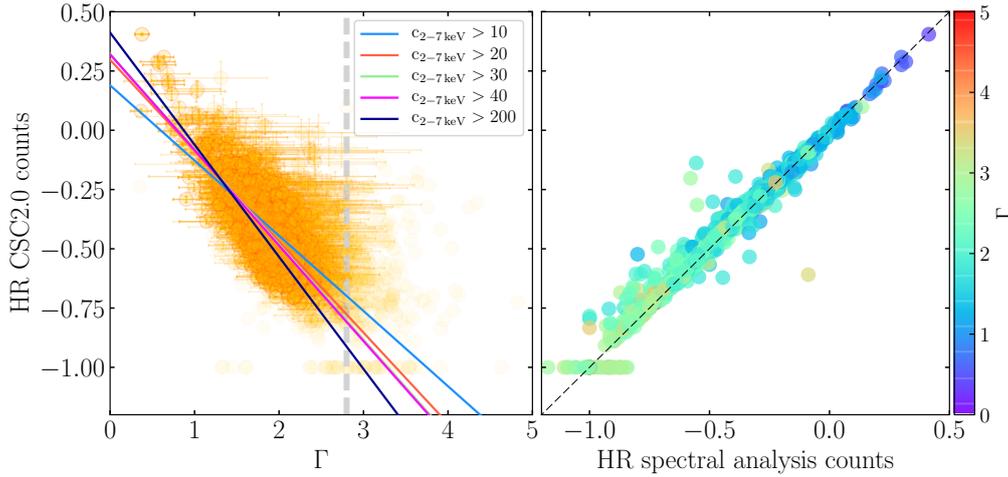
**SDSS-CSC2.0:** The photometry available allowed us to build an interpolated SED for each individual object, from which we inferred the rest-frame flux at 2500  $\text{\AA}$ , proxy for the accretion disc emission. As discussed below, we are interested in estimating the continuum up to 1  $\mu\text{m}$  rest-frame. For sources at  $z \sim 2$  and above, this wavelength falls into the mid-infrared and near-infrared bands. We therefore used all the magnitudes provided



**Fig. A.1.** Distribution of the quasar sample (excluding radio-loud and BAL sources) in a  $\Gamma_1$ – $\Gamma_2$  plane, where  $\Gamma_1$  and  $\Gamma_2$  are the slopes of a power law in the  $\log \nu$ –( $\log \nu L_\nu$ ) plane, at 0.3–1  $\mu\text{m}$  and 1450–3000  $\text{\AA}$ , respectively. The red stars represent  $\Gamma_1$ – $\Gamma_2$  values for the intrinsic mean SED for quasars of Richards et al. (2006) reddened by the presence of dust (extinction law of Prevot et al. 1984) with increasing  $E(B-V)$ , ranging from 0 to 0.5. We selected the sources inside the red circle of centre  $E(B-V) = 0$  and radius  $\sim 1$  (corresponding to  $E(B-V) = 0.1$ ), characterised by the minimum extinction. Other possible explanations to the deviation of points from the locus of blue quasars are, for example, the presence of host-galaxy contamination and, trivially, wrong photometric measurements.

by Pâris et al. (2018), that is *GALEX*, SDSS, UKIDSS, 2MASS, *WISE* and *Spitzer*, whose wavelengths are listed in Tab. 3. We first corrected the fluxes densities in each band for Galactic extinction, using the  $E(B-V)$  values estimated from the Schlegel et al. (1998) maps, as listed in the NASA/IPAC Infrared Science Archive<sup>11</sup> (IRSA), and the reddening law by Fitzpatrick (1999) with  $R_V = 3.1$ . We then computed the rest-frame fluxes according to the redshift listed in Pâris et al. (2018) and performed a linear interpolation to retrieve the flux (and luminosity) value at 2500  $\text{\AA}$  (a higher-order polynomial could introduce spurious features and lead to higher inaccuracy in the UV flux estimation). When the combination of source redshift and data available did not provide the wavelength range to cover the rest-frame 2500  $\text{\AA}$ , the flux was extrapolated from the last two photometric points available. Continuum fluxes estimated in this way can be contaminated by the presence of the broad emission lines characterising type-1 quasars spectra. Depending on the source redshift, broad emission lines are found in different bands and therefore affect different observed magnitudes. In the case of SDSS-CSC2.0 data, we were able to compute a correction taking advantage of the spectral fitting performed by Shen et al. (2011) on DR7 quasars. Shen et al. (2011) take into account continuum, broad and narrow emission lines and Fe II emission, hence providing an accurate measure of the 2500- $\text{\AA}$  continuum flux. Comparing our photometric measures with the spectral ones by Shen et al. (2011), we inferred a correction as a function of redshift,

<sup>11</sup> <https://irsa.ipac.caltech.edu/applications/DUST/>



**Fig. A.2.** Left. Distribution of the HR computed from the catalogued counts as a function of the photon index output of our spectral analysis. Right. Comparison between the HR computed from the catalogued counts and the one resulting from our spectral analysis. Points are coloured as a function of the source photon index (right bar).

that could be applied to all the SDSS-CSC2.0 photometric 2500-Å flux estimates.

Rest-frame flux (and luminosity) at 1450 Å, 3000 Å and 1 μm were also computed. Through a simple linear least-square regression the spectral slopes of the power law in  $\log(\nu)$ – $\log(\nu L_\nu)$  describing the SED ( $\Gamma_1$ , between 0.3 and 1 μm, and  $\Gamma_2$ , between 1450 Å and 3000 Å) were estimated, and used to estimate the effect by dust reddening and host-galaxy contamination, following the same criteria as in [Lusso & Risaliti \(2016\)](#). Specifically, we excluded from the sample the sources in the  $\Gamma_1$ – $\Gamma_2$  plane not included in the circle of radius  $\sim 1$  (Fig. A.1), corresponding to  $E(B - V) \sim 0.1$ , and centred in  $E(B - V) = 0$ , that is the value of the slopes for which the UV spectrum is completely unaffected by dust reddening or host contamination ( $\Gamma_1 = 0.82$  and  $\Gamma_2 = 0.4$ ), based on the intrinsic mean SED for quasars of [Richards et al. \(2006\)](#). For DR7 sources, we used the 2500-Å flux made available by [Shen et al. \(2011\)](#) through spectral fitting. Since [Shen et al. \(2011\)](#) do not provide errors on this quantity, we assumed a 2%, considering that they list a 3% average error on bolometric luminosity, where also the uncertainty on the bolometric correction is included.

**COSMOS:** The same analysis was performed, in this case with the photometry by [Laigle et al. \(2016\)](#), making available deep NIR observations for the COSMOS field from the Ultraviolet and SPLASH surveys. Ranging from the UV to the infrared bands, as listed in the lower part of Tab. 3, the photometry allowed the determination of the rest-frame 2500-Å flux and the estimation of the SED spectral slopes to exclude sources affected by dust absorption and host-galaxy contamination (Fig. A.1), following the same criteria described above for the SDSS-CSC2.0 sample. The 4400-Å flux and luminosity were also computed and used, along with the radio luminosity at 5 GHz, to infer the radio loudness parameter  $R$  and get rid of radio-loud sources, as described in Section 2.1.

## Appendix B: Check on hardness ratios

We further checked all sources with extreme photon index values derived from the spectral analysis - namely, either very low (flat,  $\Gamma < 1.4$ ) or unexpectedly high (steep,  $\Gamma > 3.0$ ). Fig. A.2, right panel, shows the comparison between the hardness ratios listed in the CSC 2.0 and those computed from the spectral

counts retrieved using XSPEC. In both cases, the hardness ratio is defined as

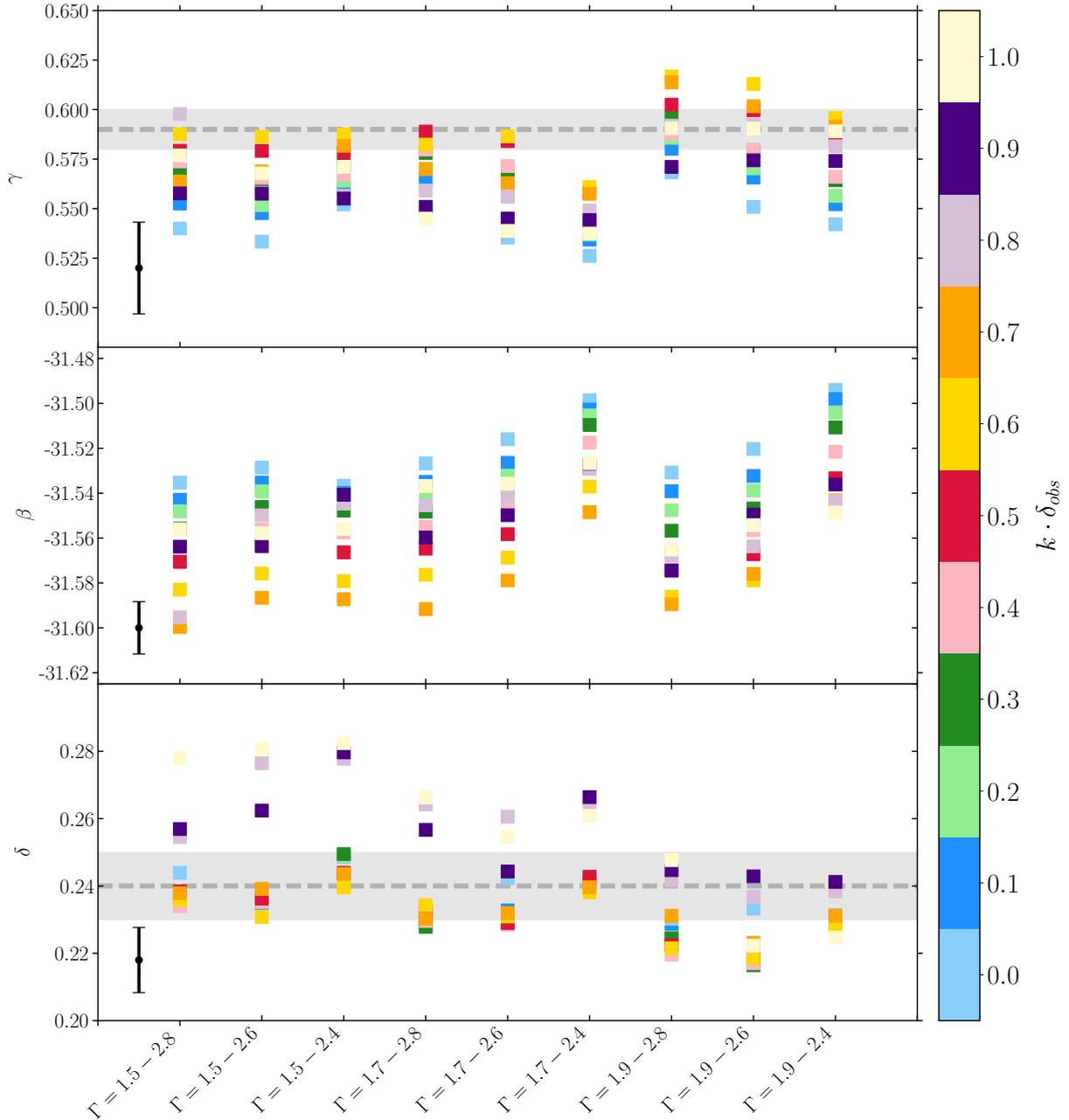
$$HR = \frac{H - S}{H + S}, \quad (\text{B.1})$$

where  $H$  and  $S$  are the net counts in the hard (2–7 keV) and soft (0.5–2 keV) band respectively. The agreement between the two measurements, the first one based on aperture photometry and the second one on spectroscopic data, is excellent, with the exception of a few outliers. This represents an independent validation of the spectral analysis we carried out. The observations with a non detection in the hard band, namely  $HR = -1.0$  (50 for the CSC 2.0, while only 20 have no counts in the hard band from the spectral analysis) are located in the lower left corner of the panel.

The HR is a coarse measure of the spectral slope, nonetheless we expect to observe a correlation between HR and the photon index. The left panel of Fig. A.2 shows the hardness ratio from the CSC 2.0 as a function of the photon index from the spectral analysis for an increasing number of net counts in the hard band. For sources with a photon index  $\Gamma < 2.8$ , the value chosen as an upper threshold, the CSC 2.0 HR is a good proxy for the photon index, improving with the number of counts in the hard and, consequently, in the soft band. This is an obvious consequence of the fact that points deviating from the bestfit relation in the HR vs  $\Gamma$  plane, populating the locus of  $\Gamma > 2.8$  and  $HR < -0.25$ , are generally associated with observations with a small number of counts ( $< 10$  in the hard band).

## Appendix C: Choice of selection criteria in X-ray band

In this work, the parent sample has been pre-selected for the properties concerning the UV band, that is broad-line absorption and dust reddening. Here we discuss our choices for the selection in the X-ray band, which led to the final sample of 1385 CSC 2.0 observations, corresponding to 958 sources, and the 140 COSMOS sources. To decide which combination of selection in photon index  $\Gamma$ , that is X-ray absorption, and Eddington bias, that is X-ray flux threshold, is a good trade-off between a small observed dispersion and a large enough number of sources, we



**Fig. C.1.** Results of the analysis on the  $f_X - f_{UV}$  relation. Different choices for the photon index  $\Gamma$  are on the x axis, while those for  $k \delta_{obs}$  are shown as different colours. The squares represent the mean slope (top panel), intercept (mid panel) and dispersion (bottom panel) for all the redshift bins ( $\Delta \log(z) = 0.06$ ) with more than five objects after the chosen cuts in  $\Gamma$  and  $k \delta_{obs}$  are applied. In black, we show the typical error on the estimates for the three quantities.

have analysed the behaviour of the  $f_X - f_{UV}$  relation for different choices of these quantities. Tab. C.1 and Fig. C.1 summarise the results of the analysis. Fig. C.1 shows the slope, intercept and dispersion for the  $f_X - f_{UV}$  relation analysed as follows: as explained in Section 4, when dealing with fluxes, we have to divide the sample in redshift bins. The fit to the data in each redshift bin ( $\Delta \log(z) = 0.06$ ), gave us a value for slope, intercept and dispersion. Then, for each quantity, we computed a mean of the values for all the redshift bins ( $\Delta \log(z) = 0.06$ ) with more than five objects. Here we are comparing these values for different choices of  $\Gamma$  (x axis) and  $k \delta_{obs}$  (colour bar)<sup>12</sup>. Values obtained

<sup>12</sup> In this appendix we are showing the analysis for nine different ranges of  $\Gamma$  and eleven different values of  $k \delta_{obs}$ .

for the  $L_X - L_{UV}$  relation are plotted in grey as a reference. In Tab. C.1 we also list, for each choice of  $\Gamma$  and  $k \delta_{obs}$  the number of sources that outlived the selection.

We looked at the dispersion and at the slope of the  $f_X - f_{UV}$  relation: the bottom panel of Fig. C.1 shows that, for all the choices of  $\Gamma$ , the observed dispersion has a minimum for  $k \delta_{obs} = 0.6$ , while it increases for both lower and higher choices of this quantity, meaning not only that the inclusion of sources affected by the Eddington bias (i.e. light blue and pale blue squares) can modify the dispersion in the relation, but also that a too severe selection (i.e. purple, light yellow) implies a drastic decrease in the number of sources and increasing dispersion with respect to best fit. Moreover, while for a choice of  $k \delta_{obs} \sim 0.6$  the slope of the relation stays steadily around  $\gamma \sim 0.59$  for every

**Table C.1.** Results of the analysis on the  $f_X - f_{UV}$  relation.

		$\Gamma$								
		1.5 – 2.8	1.5 – 2.6	1.5 – 2.4	1.7 – 2.8	1.7 – 2.6	1.7 – 2.4	1.9 – 2.8	1.7 – 2.6	1.7 – 2.4
0.0	$N$	1841	1727	1573	1630	1515	1357	1277	1157	991
	$\gamma$	$0.54 \pm 0.02$	$0.53 \pm 0.02$	$0.55 \pm 0.03$	$0.55 \pm 0.02$	$0.54 \pm 0.02$	$0.53 \pm 0.02$	$0.57 \pm 0.02$	$0.55 \pm 0.02$	$0.54 \pm 0.02$
	$\beta$	$-31.54 \pm 0.01$	$-31.53 \pm 0.01$	$-31.54 \pm 0.01$	$-31.53 \pm 0.01$	$-31.52 \pm 0.01$	$-31.50 \pm 0.01$	$-31.53 \pm 0.01$	$-31.52 \pm 0.01$	$-31.49 \pm 0.01$
	$\delta$	$0.24 \pm 0.01$	$0.24 \pm 0.01$	$0.25 \pm 0.01$	$0.23 \pm 0.01$	$0.24 \pm 0.01$	$0.24 \pm 0.01$	$0.23 \pm 0.01$	$0.23 \pm 0.01$	$0.23 \pm 0.01$
0.1	$N$	1778	1672	1533	1571	1464	1322	1226	1114	964
	$\gamma$	$0.55 \pm 0.02$	$0.55 \pm 0.02$	$0.56 \pm 0.03$	$0.56 \pm 0.02$	$0.56 \pm 0.02$	$0.53 \pm 0.02$	$0.58 \pm 0.02$	$0.57 \pm 0.02$	$0.55 \pm 0.02$
	$\beta$	$-31.54 \pm 0.01$	$-31.54 \pm 0.01$	$-31.54 \pm 0.01$	$-31.53 \pm 0.01$	$-31.53 \pm 0.01$	$-31.50 \pm 0.01$	$-31.54 \pm 0.01$	$-31.53 \pm 0.01$	$-31.50 \pm 0.01$
	$\delta$	$0.24 \pm 0.01$	$0.23 \pm 0.01$	$0.25 \pm 0.01$	$0.23 \pm 0.01$	$0.23 \pm 0.01$	$0.24 \pm 0.01$	$0.23 \pm 0.01$	$0.22 \pm 0.01$	$0.23 \pm 0.01$
0.2	$N$	1690	1596	1477	1489	1394	1272	1151	1053	922
	$\gamma$	$0.56 \pm 0.02$	$0.55 \pm 0.02$	$0.56 \pm 0.03$	$0.57 \pm 0.02$	$0.56 \pm 0.02$	$0.54 \pm 0.02$	$0.59 \pm 0.02$	$0.57 \pm 0.02$	$0.56 \pm 0.02$
	$\beta$	$-31.55 \pm 0.01$	$-31.54 \pm 0.01$	$-31.54 \pm 0.01$	$-31.54 \pm 0.01$	$-31.53 \pm 0.01$	$-31.51 \pm 0.01$	$-31.55 \pm 0.01$	$-31.54 \pm 0.01$	$-31.50 \pm 0.01$
	$\delta$	$0.24 \pm 0.01$	$0.24 \pm 0.01$	$0.25 \pm 0.01$	$0.23 \pm 0.01$	$0.23 \pm 0.01$	$0.24 \pm 0.01$	$0.22 \pm 0.01$	$0.22 \pm 0.01$	$0.23 \pm 0.01$
0.3	$N$	1579	1500	1395	1393	1313	1205	1063	982	867
	$\gamma$	$0.57 \pm 0.02$	$0.56 \pm 0.02$	$0.57 \pm 0.03$	$0.58 \pm 0.02$	$0.57 \pm 0.02$	$0.54 \pm 0.02$	$0.60 \pm 0.02$	$0.58 \pm 0.02$	$0.56 \pm 0.02$
	$\beta$	$-31.56 \pm 0.01$	$-31.55 \pm 0.01$	$-31.55 \pm 0.01$	$-31.55 \pm 0.01$	$-31.54 \pm 0.01$	$-31.51 \pm 0.01$	$-31.56 \pm 0.01$	$-31.55 \pm 0.01$	$-31.51 \pm 0.01$
	$\delta$	$0.24 \pm 0.01$	$0.24 \pm 0.01$	$0.25 \pm 0.01$	$0.23 \pm 0.01$	$0.23 \pm 0.01$	$0.24 \pm 0.01$	$0.22 \pm 0.01$	$0.22 \pm 0.01$	$0.23 \pm 0.01$
0.4	$N$	1440	1377	1293	1260	1197	1110	943	879	786
	$\gamma$	$0.57 \pm 0.02$	$0.57 \pm 0.02$	$0.57 \pm 0.03$	$0.58 \pm 0.02$	$0.57 \pm 0.02$	$0.55 \pm 0.02$	$0.59 \pm 0.02$	$0.58 \pm 0.02$	$0.57 \pm 0.02$
	$\beta$	$-31.56 \pm 0.01$	$-31.56 \pm 0.01$	$-31.56 \pm 0.01$	$-31.55 \pm 0.01$	$-31.55 \pm 0.01$	$-31.52 \pm 0.01$	$-31.57 \pm 0.01$	$-31.56 \pm 0.01$	$-31.52 \pm 0.01$
	$\delta$	$0.23 \pm 0.01$	$0.24 \pm 0.01$	$0.24 \pm 0.01$	$0.23 \pm 0.01$	$0.23 \pm 0.01$	$0.24 \pm 0.01$	$0.22 \pm 0.01$	$0.22 \pm 0.01$	$0.23 \pm 0.01$
$k\delta$ 0.5	$N$	1289	1237	1168	1120	1068	996	827	775	701
	$\gamma$	$0.58 \pm 0.02$	$0.58 \pm 0.02$	$0.58 \pm 0.03$	$0.59 \pm 0.02$	$0.58 \pm 0.02$	$0.56 \pm 0.02$	$0.60 \pm 0.02$	$0.60 \pm 0.02$	$0.58 \pm 0.02$
	$\beta$	$-31.57 \pm 0.01$	$-31.56 \pm 0.01$	$-31.57 \pm 0.01$	$-31.56 \pm 0.01$	$-31.56 \pm 0.01$	$-31.53 \pm 0.01$	$-31.57 \pm 0.01$	$-31.57 \pm 0.01$	$-31.53 \pm 0.01$
	$\delta$	$0.24 \pm 0.01$	$0.24 \pm 0.01$	$0.24 \pm 0.01$	$0.23 \pm 0.01$	$0.23 \pm 0.01$	$0.24 \pm 0.01$	$0.22 \pm 0.01$	$0.22 \pm 0.01$	$0.23 \pm 0.01$
0.6	$N$	1133	1088	1025	1098	937	872	720	675	608
	$\gamma$	$0.59 \pm 0.02$	$0.59 \pm 0.02$	$0.59 \pm 0.03$	$0.58 \pm 0.02$	$0.59 \pm 0.02$	$0.56 \pm 0.02$	$0.62 \pm 0.02$	$0.61 \pm 0.02$	$0.60 \pm 0.03$
	$\beta$	$-31.58 \pm 0.01$	$-31.58 \pm 0.01$	$-31.58 \pm 0.01$	$-31.58 \pm 0.01$	$-31.57 \pm 0.01$	$-31.54 \pm 0.01$	$-31.59 \pm 0.01$	$-31.58 \pm 0.01$	$-31.55 \pm 0.01$
	$\delta$	$0.24 \pm 0.01$	$0.23 \pm 0.01$	$0.24 \pm 0.01$	$0.23 \pm 0.01$	$0.23 \pm 0.01$	$0.24 \pm 0.01$	$0.22 \pm 0.01$	$0.22 \pm 0.01$	$0.23 \pm 0.01$
0.7	$N$	937	904	855	814	781	730	588	555	502
	$\gamma$	$0.56 \pm 0.03$	$0.57 \pm 0.03$	$0.58 \pm 0.03$	$0.57 \pm 0.03$	$0.56 \pm 0.03$	$0.56 \pm 0.03$	$0.61 \pm 0.03$	$0.60 \pm 0.03$	$0.59 \pm 0.03$
	$\beta$	$-31.60 \pm 0.01$	$-31.59 \pm 0.01$	$-31.59 \pm 0.01$	$-31.59 \pm 0.01$	$-31.58 \pm 0.01$	$-31.55 \pm 0.01$	$-31.59 \pm 0.01$	$-31.58 \pm 0.01$	$-31.54 \pm 0.01$
	$\delta$	$0.24 \pm 0.01$	$0.24 \pm 0.01$	$0.24 \pm 0.01$	$0.23 \pm 0.01$	$0.23 \pm 0.01$	$0.24 \pm 0.01$	$0.23 \pm 0.01$	$0.22 \pm 0.01$	$0.23 \pm 0.01$
0.8	$N$	792	772	734	684	664	624	487	467	425
	$\gamma$	$0.60 \pm 0.03$	$0.56 \pm 0.03$	$0.56 \pm 0.03$	$0.56 \pm 0.03$	$0.56 \pm 0.03$	$0.55 \pm 0.03$	$0.59 \pm 0.03$	$0.59 \pm 0.03$	$0.58 \pm 0.03$
	$\beta$	$-31.60 \pm 0.02$	$-31.55 \pm 0.01$	$-31.54 \pm 0.01$	$-31.55 \pm 0.01$	$-31.54 \pm 0.01$	$-31.53 \pm 0.02$	$-31.57 \pm 0.02$	$-31.56 \pm 0.02$	$-31.54 \pm 0.02$
	$\delta$	$0.25 \pm 0.01$	$0.28 \pm 0.01$	$0.28 \pm 0.01$	$0.26 \pm 0.01$	$0.26 \pm 0.01$	$0.26 \pm 0.01$	$0.24 \pm 0.01$	$0.24 \pm 0.01$	$0.24 \pm 0.01$
0.9	$N$	617	602	576	536	521	495	392	377	351
	$\gamma$	$0.56 \pm 0.03$	$0.56 \pm 0.03$	$0.56 \pm 0.03$	$0.55 \pm 0.03$	$0.54 \pm 0.03$	$0.54 \pm 0.03$	$0.57 \pm 0.03$	$0.57 \pm 0.03$	$0.57 \pm 0.03$
	$\beta$	$-31.56 \pm 0.02$	$-31.56 \pm 0.02$	$-31.54 \pm 0.02$	$-31.56 \pm 0.02$	$-31.55 \pm 0.02$	$-31.53 \pm 0.02$	$-31.57 \pm 0.02$	$-31.55 \pm 0.02$	$-31.54 \pm 0.02$
	$\delta$	$0.26 \pm 0.01$	$0.26 \pm 0.01$	$0.28 \pm 0.01$	$0.26 \pm 0.01$	$0.24 \pm 0.01$	$0.27 \pm 0.01$	$0.24 \pm 0.02$	$0.24 \pm 0.01$	$0.24 \pm 0.01$
1.0	$N$	487	476	458	426	415	397	305	294	276
	$\gamma$	$0.58 \pm 0.04$	$0.57 \pm 0.04$	$0.57 \pm 0.04$	$0.54 \pm 0.04$	$0.54 \pm 0.04$	$0.54 \pm 0.04$	$0.59 \pm 0.04$	$0.59 \pm 0.04$	$0.59 \pm 0.04$
	$\beta$	$-31.56 \pm 0.02$	$-31.56 \pm 0.02$	$-31.56 \pm 0.02$	$-31.54 \pm 0.02$	$-31.54 \pm 0.02$	$-31.53 \pm 0.02$	$-31.57 \pm 0.02$	$-31.55 \pm 0.02$	$-31.55 \pm 0.02$
	$\delta$	$0.28 \pm 0.01$	$0.28 \pm 0.01$	$0.28 \pm 0.01$	$0.27 \pm 0.01$	$0.25 \pm 0.01$	$0.26 \pm 0.01$	$0.25 \pm 0.02$	$0.22 \pm 0.01$	$0.22 \pm 0.02$

**Notes.** Results of the analysis on the  $f_X - f_{UV}$  relation. We list, for every combination of cuts in  $\Gamma$  and  $k\delta_{obs}$ , the slope, intercept and dispersion obtained as a mean of all those in the redshift bins ( $\Delta\log(z) = 0.06$ ) with more than five objects, as long as the number of sources that survived the selection.

choice of  $\Gamma$ , for both low and high  $k\delta_{obs}$  values the slope is significantly lower than the reference value of  $\sim 0.6$  (upper panel). Finally, exploring the possible choices of the photon index  $\Gamma$ , we realised that a smaller dispersion is achieved by a more conservative selection at lower values than at higher values; in fact, setting the lower boundary progressively from  $\Gamma_{min} = 1.5$  to  $\Gamma_{min} = 1.9$  produces a decrease in the dispersion, while this is not generally true going from  $\Gamma_{max} = 2.8$  to  $\Gamma_{max} = 2.4$ . Even if the slope of the relation is closer to the reference value

and the dispersion is slightly smaller in the case of the most conservative selection  $\Gamma = 1.9$ – $2.8$ , we decided for the compromise  $\Gamma = 1.7$ – $2.8$  because of the larger number of complying sources, and also for consistency with the cut adopted in the analysis of the SDSS-XMM sample of [Risaliti & Lusso \(2019\)](#). For the sake of completeness, we also plot the intercepts for the different choices of the parameters (middle panel), even if this quantity is not taken into account for the purpose of selection.