

SRG/eROSITA–SDSS view on the relation between X-ray and UV emission for quasars

Sergey Prokhorenko^{a,b}, Sergey Sazonov^{a,b}, Marat Gilfanov^{a,c}, Sergey Balashev^d, Alexander Meshcheryakov^a, Alexander Ivanchik^d, Ilfan Bikmaev^{e,f}, Rashid Sunyaev^{a,c,g}

^aSpace Research Institute of the Russian Academy of Sciences, Profsoyuznaya Str. 84/32, Moscow, 117997, Russia

^bNational Research University Higher School of Economics, Pokrovsky Bulvar 11, Moscow, 101990, Russia

^cMax-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, Garching, D-85741, Germany

^dIoffe Institute, Politekhnicheskaya str. 26, St Petersburg, 194021, Russia

^eKazan Federal University, Kremlevskaya str. 18, Kazan, 420008, Russia

^fAcademy of Sciences of Tatarstan, Bauman Str. 20, Kazan, 420111, Russia

^gInstitute for Advanced Study, 1st Einstein Drive, Princeton, 08540, New Jersey, USA

Abstract

Motivated by the idea of using quasars as standardizable candles for cosmology, we examine the relation between X-ray (at 2 keV, $L_{2\text{keV}}$) and ultraviolet (at 2500 Å, L_{2500}) monochromatic luminosities of quasars using a sample of 2414 X-ray sources from the SRG/eROSITA all-sky survey cross-matched with the Sloan Digital Sky Survey data release 16 quasar catalogue (SDSS DR16Q), at redshifts between 0.5 and 2.5. These objects are bright both in X-rays and in the optical, so that the sample is characterized by nearly 100% statistical completeness. We have developed a new method for determining the $L_{2\text{keV}}-L_{2500}$ relation, which consistently takes into account (i) X-ray and UV flux limited object selection, (ii) X-ray and UV variability of quasars, and (iii) the decreasing space density of quasars with increasing luminosity. Assuming a linear relation between $\log L_X \equiv \log \frac{L_{2\text{keV}}}{\text{erg s}^{-1} \text{Hz}^{-1}}$ and $\log L_{\text{UV}} \equiv \log \frac{L_{2500}}{\text{erg s}^{-1} \text{Hz}^{-1}}$, we find the slope, $\gamma = 0.69 \pm 0.02$ (hereafter all uncertainties are quoted at the 68% confidence level), and normalization, $\log L_X = 26.45 \pm 0.02$ at $\log L_{\text{UV}} = 30.5$, of the $L_{2\text{keV}} (L_{2500})$ dependence. These values are not substantially different from the results of previous studies. A key novel aspect of our work is allowance for intrinsic scatter (which adds to the dispersion induced by quasar variability and flux measurement uncertainties) of the $L_{2\text{keV}}-L_{2500}$ relation in both variables, i.e. in X-ray and UV luminosity. The intrinsic X-ray scatter ($\sigma_{\text{intX}}^2 = 0.066 \pm 0.005$) strongly dominates over the UV one ($\sigma_{\text{intUV}}^2 = 0.001^{+0.003}_{-0.001}$). Further studies should seek to explain this behaviour in terms of accretion onto supermassive black holes and orientation of quasars with respect to the observer.

Keywords:

Quasars, Supermassive black holes, X-rays, Accretion discs

1. Introduction

Active galactic nuclei (AGN) are extremely diverse objects as regards their luminosity, spectral properties, radio-loudness, variability, polarization etc. (see Padovani et al. 2017 for a review). However, this diversity is predominantly driven by just a few physical parameters related to accretion of matter onto the central supermassive black hole (SMBH), namely its mass, spin and accretion rate, with a presumably lesser role played by the properties of the host galaxy. This should cause different AGN observable properties to correlate with each other, and a number of such correlations have indeed been found.

One of the most notable and reliably established relations is between the X-ray and ultraviolet (UV) luminosities, usually expressed in terms of monochromatic luminosities (per frequency) at rest-frame 2 keV and 2500 Å, which we will hereafter refer to as $L_{2\text{keV}}$ and L_{2500} , respectively. Its existence was first suggested in the 1970s–1980s (Tananbaum et al., 1979; Zamorani et al., 1981; Avni & Tananbaum, 1982, 1986) and has then been validated in many studies using progressively larger

AGN samples. Most authors conclude that the relation between $L_{2\text{keV}}$ and L_{2500} is nonlinear, namely that $L_{2\text{keV}}/L_{2500}$ decreases with increasing L_{2500} (e.g. Lusso et al. 2010; Bisogni et al. 2021; Dainotti et al. 2022; Signorini et al. 2024; Rankine et al. 2024; Sacchi et al. 2025). This implies a decreasing contribution of X-ray emission (presumably produced by Comptonization in a hot corona of the accretion disc) to the AGN bolometric luminosity as the latter increases. The 2500 Å luminosity, which can be conveniently measured by ground-based spectrographs for objects located at redshifts between approximately 0.5 and 3, is believed to be a good tracer of thermal emission from the inner regions of the accretion disc (e.g. Jin et al. 2023).

The $L_{2\text{keV}}-L_{2500}$ relation is not only important for exploring the internal structure of AGN and the physics of accretion onto black holes but also for cosmology, as it potentially provides an opportunity to use quasars (i.e. luminous and distant AGN), as standardizable candles for building the cosmic distance ladder, even at high redshifts (Risaliti & Lusso, 2015; Lusso et al., 2019, 2020; Khadka & Ratra, 2020; Li et al., 2022; Lenart et al., 2023). However, this can only be fulfilled if this relation is indeed tight and nonlinear. Otherwise it would be impossible to deduce the absolute distance to a quasar by measuring $L_{2\text{keV}}$ and

Email address: sprokhorenko@cosmos.ru (Sergey Prokhorenko)

L_{2500} . In reality, previous studies have reported significantly diverse values of the slope of the $L_{2\text{keV}}(L_{2500})$ dependence¹ (Vignali et al., 2003; Strateva et al., 2005; Steffen et al., 2006; Bisogni et al., 2021; Rankine et al., 2024) and found a substantial scatter in the correlation: $L_{2\text{keV}}/L_{2500}$ varies by a factor of ~ 2 for a given luminosity (Lusso & Risaliti, 2016, 2017). As a result, the practical usability of quasars for cosmology is still under question, urging further investigation. The precision of previous studies of the X-ray–UV luminosity relation was often limited by the size of the AGN samples used. In addition, it is important to carefully account for selection effects, both in the X-ray and optical–UV bands. Another important issue, as we demonstrate below, is to properly define what is actually meant by the X-ray–UV luminosity relation and its scatter, since different definitions will lead to different results.

Here, we use a new large and well defined sample of AGN (which we hereafter refer to as quasars) obtained by cross-matching the catalogue of X-ray sources detected during the all-sky survey by the eROSITA telescope (Predehl et al., 2021) aboard the *SRG* observatory (Sunyaev et al., 2021) in the Eastern Galactic half of the sky with the catalogue of quasars observed during the SDSS optical spectroscopic survey (Lyke et al., 2020). This sample comprises about two and a half thousand quasars and effectively spans two decades in luminosity. This allows us to conduct one of the most precise investigations of the $L_{2\text{keV}}-L_{2500}$ relation thus far. To this end, we have developed a novel method that self-consistently takes X-ray and optical/UV selection effects into account and, more importantly, for the first time allows for the presence of intrinsic scatter of the $L_{2\text{keV}}-L_{2500}$ relation in both variables, i.e. in X-ray and UV luminosity. Throughout the paper, we use a flat Λ CDM-cosmology model with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$.

2. SRG/eROSITA–SDSS sample

Our sample is composed of spectroscopically confirmed quasars from the 16th data release of the Sloan Digital Sky Survey (SDSS) Quasar Catalogue (SDSS DR16Q, Lyke et al. 2020) that were detected by eROSITA during the *SRG* all-sky survey in the $0^\circ < l < 180^\circ$ hemisphere. Specifically, we use the eROSITA catalogue of point sources detected on the summed map of the first five all-sky surveys² in the 0.3–2.3 keV energy band. The SDSS DR16Q catalogue covers 9376 square degrees in the sky, of which 7994 are at $0^\circ < l < 180^\circ$.

2.1. Raw sample

To minimize systematic uncertainties and biases associated with absorption in the interstellar medium of the Galaxy, we excluded sky regions with Galactic neutral hydrogen column

density (HI4PI Collaboration et al., 2016) $N_H > 4 \times 10^{20} \text{ cm}^{-2}$. The attenuation in the g band associated with Galactic dust (Schlegel et al., 1998; Schlafly & Finkbeiner, 2011) is less than 0.5 mag for the remaining sky area of 4965 square degrees. In addition, we selected for our analysis only those regions that had been covered by spectroscopy during all three of the SDSS-I/II, SDSS-III and SDSS-IV survey stages. Given the complex optical/infrared colour and variability based selection of quasar candidates for SDSS spectroscopy (Berk et al. 2005; Richards et al. 2006; Dawson et al. 2013; Myers et al. 2015), more than 98% of quasars brighter than $m_g = 19$ at redshifts between 0.5 and 2.5 (the range of our interest) are expected to have been covered by spectroscopy in these regions (as we explicitly demonstrate in Appendix A), whereas in regions that were covered by only one or two SDSS quasar programs (I/II, III or IV), there can be a deficit of either dim or bright quasars, which we would like to avoid. This further reduces the area of our study to 2858 square degrees, which finally defines our *SRG*/eROSITA–SDSS footprint.

We cross-matched the SDSS and eROSITA catalogues using the X-ray position uncertainties, r_{98} . This is the radius of the circle within which the source is located with a probability of 98%, according to eROSITA. We limited our sample of quasars in both the X-ray and optical flux, as described below. Upon application of the flux limits, the median value of r_{98} is $5''$, and 90% of objects have $r_{98} < 7''$. For 11 eROSITA sources, there are two SDSS quasars within r_{98} . We exclude these ambiguous cross-matches from our sample.

In the X-ray domain, our sample is limited by the Galactic absorption corrected flux in the observed 0.3–2.3 keV energy band (which corresponds to a range between ~ 0.5 and ~ 5 keV in the quasar rest-frame for a typical redshift $z \sim 1$ of our objects): $F_X > F_{X,\text{min}} = 6 \times 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$. At these X-ray fluxes, we have at least 20 eROSITA detector counts from each source (apart from two sources in the final sample, which have 19.9 and 17.0 background subtracted counts) and thus deal with nearly Gaussian statistics of measured fluxes.

We estimated source X-ray fluxes from the count rates measured by eROSITA, assuming an intrinsic power-law spectrum with a photon index $\Gamma = 2$, modified by absorption in the Galactic interstellar medium. To this end, we used the *XSPEC* software (version 12.12.0, Arnaud 1996), the Galactic interstellar absorption model of Wilms et al. (2000) (MODEL TBABS * POWERLAW command in *XSPEC*) with solar abundances of elements, and the hydrogen map from the HI4PI survey (HI4PI Collaboration et al., 2016). The adopted slope $\Gamma = 2$ may be regarded as a typical effective spectral slope for type I AGN in the ~ 0.5 to ~ 5 keV rest-frame range probed here, taking into account the presence of a primary X-ray continuum with $\Gamma \approx 1.8$ and a ubiquitous soft X-ray excess (i.e. steepening) at energies below 1–2 keV (e.g. Laor et al. 1997; Brightman et al. 2013; Ricci et al. 2017; Trakhtenbrot et al. 2017). In reality, the spectral slope may vary from one quasar to another. If we conservatively assume that individual Γ values in our sample are distributed according to a normal distribution around $\Gamma = 2$ with a standard deviation of 0.3, and take into account the eROSITA spectral response and the probed range of Galactic absorption,

¹ Virtually all authors have adopted, for simplicity, a power-law form for this relation.

² The fifth eROSITA survey was interrupted on February 26th, 2022 and has not resumed yet, so that for $\sim 60\%$ of the sky the data of only four surveys are available.

then this would result in an additional systematic uncertainty of less than 4% (8%) in the estimated absorption-corrected fluxes in the observed 0.3–2.3 keV band for 90% (99%) of the sample. This is much smaller than typical statistical uncertainties for our objects, which justifies our neglect of the uncertainty associated with X-ray spectral diversity. We also ignore any intrinsic X-ray absorption, since we are dealing with broad-line quasars only, and additionally exclude quasars with broad absorption lines (see Sect. 2.2).

We also imposed a limit on the optical flux, namely that magnitudes in the SDSS g -band $m_g < m_{g,\max} = 19$. To this end, we used the g -band photometric measurements from the SDSS catalogue, namely point-spread-function (PSF) magnitudes, as recommended for point sources, such as the SDSS DR16Q quasars (by construction of this catalogue). This removes 7919 quasars (out of 12969). In addition, we excluded 20 quasars with incorrect magnitude measurements (negative values) from the SDSS DR16Q catalogue table from Lyke et al. (2020). Note that we could increase our sample by going deeper in the optical flux, at least down to $m_g \sim 19.5$, while retaining nearly the same high level of completeness (see Appendix A). We decided not to do it, because the sample would then become strongly unbalanced in terms of X-ray and optical–UV depth, namely it would be heavily skewed towards one side of the X-ray–UV luminosity relation that we seek, which we would like to avoid.

For all quasars, we use the redshifts from Wu & Shen (2022) (ZSYS column in their catalogue), denoted hereafter as z . These authors have improved the redshift measurements reported in the original SDSS DR16Q catalogue by accurately accounting for velocity shifts of various lines with respect to the systemic velocity and correcting catastrophically wrong redshifts. Namely, the redshifts were corrected by more than 10000 km s^{-1} for 1943 objects and by more than 1500 km s^{-1} for ≈ 8000 objects out of the 750414 quasars in the DR16Q catalogue. We also note that if two or more SDSS spectra are available for a given object, the SDSS DR16Q catalogue uses the best available spectrum, as defined by the SDSS pipeline (Bolton et al., 2012).

We expect to have approximately four (less than seven with a probability of 90%) spurious eROSITA–SDSS matches (upon application of the X-ray and optical flux limits). We obtained this estimate by assigning small, but much larger than $r98$, angular offsets (the result is virtually independent of the offset in the tested range from 1' to 20') to the SDSS DR16Q quasars and then counting the number of cross-matches with the eROSITA catalogue.

Our choice of X-ray and optical flux thresholds has allowed us to obtain a large and statistically well defined sample, representative of both the bright and dim parts of the quasar luminosity function (LF). The resulting ‘raw’ sample comprises 5050 quasars.

It is worth noting that this quasar sample contains just a small ($\sim 15\%$) fraction of all AGN detected by eROSITA in the studied area of the sky above the (high) adopted X-ray flux limit (see also Kolodzig et al. 2013). The remaining X-ray bright AGN have not passed our selection criteria, namely (i) do not

Catalogue	Number of counterparts
Roma-BZCAT Multifrequency Catalogue of Blazars	147
4th Fermi/LAT catalogue of AGN DR3	50
Blazar Radio and Optical Survey	141
Combined Radio All-Sky Targeted Eight GHz Survey	141
All of the above	249

Table 1: Cross-correlation with blazar catalogues.

qualify for the SDSS DR16Q catalogue (i.e. not a point-like source or no broad emission lines in the spectrum) or (ii) fall below our optical flux threshold. The majority of such objects are expected to be relatively nearby ($z \lesssim 0.5$) Seyfert I and Seyfert II galaxies.

2.2. Cleaning the sample

We applied some additional cleaning to the raw sample in order to improve its purity and statistical completeness. This included the following steps.

- First, we excluded 161 X-ray sources that have multiple counterparts in the Gaia Data Release 3 (DR3) catalogue (Vallenari et al., 2023) within $r98$ of the eROSITA position, because in such cases we cannot be sure that the X-ray emission is associated with the SDSS DR16Q quasar rather than with another nearby optical object.
- We cleaned the sample from blazars (BL Lac objects and flat-spectrum radio quasars). These are AGN with a highly collimated emission component that is associated with a relativistic jet directed towards the observer, while the current study is focused on the relationship between the emission from the accretion disc and its corona. We used the same approach as in our previous work (Prokhorenko et al., 2024). Specifically, we cross-matched our sample with the 5th edition of the Roma-BZCAT Multifrequency Catalogue of Blazars (Massaro et al., 2015). This yielded 147 counterparts. We also cross-correlated our sample with the fourth catalogue of AGN detected by the Fermi Large Area Telescope, Data Release 3 (Ajello et al., 2022), namely with objects classified as ‘bli’ (BL Lac-type objects), ‘bcu’ (blazar candidates of unknown types) or ‘fsqr’ (Flat Spectrum Radio Quasars). This resulted in another five counterparts. We then cross-matched our sample with the Blazar Radio and Optical Survey (BROS) (Itoh et al., 2020), excluding Gigahertz peaked-spectrum sources and compact steep-spectrum radio sources (based on the corresponding flag in the catalogue) from the search. This provided an additional 51 sources. Finally, we cross-correlated our sample with the Combined Radio All-Sky Targeted Eight GHz Survey (Healey et al., 2007), which yielded another 46 counterparts. In all these cases, the

search was done within $10''$ of the SDSS DR16Q optical positions to take into account that some of the positions given in the blazar catalogues are not very precise (in particular, when only a radio position is available). The expected total number of spurious cross-matches with the blazar catalogues is less than one with a probability of 90%. In total, we found 249 blazars or blazar candidates associated with our objects (see Table 1 for details) and removed them from the sample.

- We excluded 72 broad-absorption-line (BAL) quasars, namely those that have a BAL probability greater than 0.2 and BALnicity index for CIV greater than 0, as quoted in Lyke et al. (2020). Such absorption lines indicate the presence of outflowing material, whose influence on the attenuation of the quasar's UV emission is difficult to assess. Note that information on the presence of BAL features is available only for quasars at $z > 1.5$, for which the potential CIV absorption systems are within the wavelength range of the SDSS spectrometer. Hence, we do not perform any BAL cleaning for the quasars at $z < 1.5$ in our sample.
- We leave only quasars at redshifts $0.5 < z < 2.5$ because their rest frame $2500 \text{ \AA}$ falls into the $3800\text{--}9100 \text{ \AA}$ waveband of the SDSS-I/II spectrograph, which was used for spectroscopy of most quasars in our sample (the BOSS spectrograph, which was employed in SDSS-III and SDSS-IV, has a wider spectral range of $3650\text{--}10400 \text{ \AA}$). This led to the removal of another 1453 quasars. Although we actually estimate the UV ($2500 \text{ \AA}$) luminosities of our quasars based on SDSS photometry rather than spectroscopy (see below), we nevertheless apply this redshift filter because we wish to account for the systematic uncertainty associated with the conversion of g -band photometric fluxes to UV luminosities (see Sect. 2.3 below).
- Some sources in our sample may be multiply imaged lensed quasars or pairs and groups of physically connected or closely projected quasars. To clean the sample from such sources, we first excluded two sources with the Renormalised Unit Weight Error (ruwe) in Gaia DR3 catalogue larger than 1.3, which is a strong indication that the source is multiple or extended. Then, we correlated our sample with several catalogues of lensed, dual and binary quasars using a $2''$ radius (Wu, Q. et al., 2024; Chen et al., 2025; Jing et al., 2025; Makarov & Secrest, 2022; Lemon et al., 2020, 2023), which led to the removal of another six sources.

The above selections leave 3109 quasars in the sample.

2.3. Estimation of the X-ray and UV luminosities and construction of the final sample

As was already mentioned in Section 1, we define $L_{2\text{keV}}$ as the monochromatic luminosity (per frequency) at rest-frame 2 keV, i.e. $L_{2\text{keV}} \equiv L_\nu(2 \text{ keV})$. It can be determined from the measured redshifts and unabsorbed X-ray fluxes in the observed

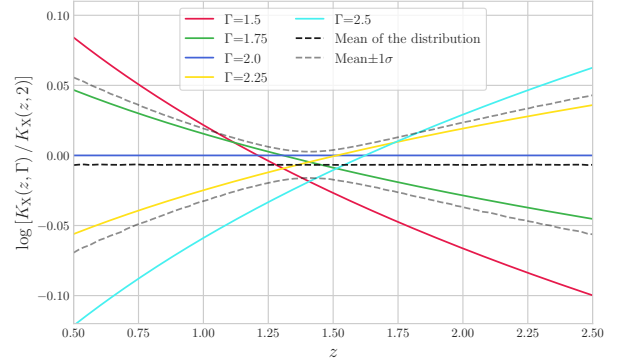


Figure 1: Solid lines show $\log K_X(z, \Gamma) / K_X(z, 2)$ as a function of z for different X-ray spectral indexes Γ , as indicated in the legend. The black dashed line marks the mean values of the $\log [K_X(z, \Gamma) / K_X(z, 2)]$ distribution as a function of z and assuming that Γ follows a normal distribution with $\mu = 2$ and $\sigma = 0.3$. The black dashed semitransparent lines mark the mean plus one standard deviation and the mean minus one standard deviation as a function of z . Hereafter, $\log$ means $\log_{10}$.

0.3–2.3 keV band, using the corresponding k -correction. For a power-law spectrum with slope Γ and redshift z , the latter is equal to

$$K_X(z, \Gamma) = \frac{4.136 \times 10^{-18} \text{ Hz}^{-1}}{2^{\Gamma-1}(1+z)^{1-\Gamma}} \times \begin{cases} \frac{1}{\ln(2.3/0.3)} & \text{if } \Gamma = 2, \\ \frac{2-\Gamma}{2.3^{2-\Gamma}-0.3^{2-\Gamma}} & \text{otherwise.} \end{cases} \quad (1)$$

Figure 1 shows $K_X(z, \Gamma)$ for several slopes.

As was discussed in Section 2.1, the intrinsic X-ray spectral slope may vary from one object to another. This has a negligible systematic effect on our estimation of unabsorbed fluxes in the observed 0.3–2.3 keV band from detector counts (see Sect. 2.1), but may significantly affect $K_X(z, \Gamma)$ and hence our $L_{2\text{keV}}$ estimates. To evaluate this effect, we assumed that the spectral slopes in our sample follow a normal distribution centered at $\Gamma = 2$ with a standard deviation of 0.3, as we already did in Section 2.1. We simulated the $\log \frac{K_X(z, \Gamma)}{K_X(z, 2)}$ distribution and calculated its mean value and standard deviation as a function of redshift (see Fig. 1). The mean equals -0.0067 (with the needed precision) independent of z , implying that the X-ray luminosity evaluated under our assumption of $\Gamma = 2$ will be slightly biased with respect to the expected value of $L_{2\text{keV}}$ for the adopted distribution of Γ . Hence, we define

$$L_{2\text{keV}} = 0.985 \frac{4\pi D_L^2(z) F_X}{1+z} K_X(z, 2), \quad (2)$$

where $D_L(z)$ is the luminosity distance.

The fact that the systematic error on F_X is small allows us to estimate the total uncertainty of $\log L_{2\text{keV}}$ as the square root of the sum of the squares of the statistical error of $\log F_X$ (due to detector count fluctuations) and the systematic uncertainty of $\log \frac{K_X(z, \Gamma)}{K_X(z, 2)}$. We define the latter as the standard deviation of the simulated distribution for each quasar's redshift. It is assumed

here (and simple simulations we have conducted provide support for that) that the $\log L_{2\text{keV}}$ distribution is close to a normal one even if $\log \frac{K_X(\Gamma)}{K_X(2)}$ is significantly skewed (as is the case for $z \in (1.1, 1.5)$), because it convolves with the nearly normally distributed $\log F_X$. As shown in Fig. 1, the systematic uncertainty in $\log L_{2\text{keV}}$ is especially important (reaching ≈ 0.05) for the nearest ($z \approx 0.5$) and most distant ($z \approx 2.5$) objects in our sample.

We next define L_{2500} as the monochromatic luminosity (per frequency) at rest-frame 2500 Å, i.e. $L_{2500} \equiv L_\nu(2500 \text{ Å})$, and this quantity can be determined from measured distance and flux via

$$L_{2500} = \frac{4\pi D_L^2(z) F_\nu(2500(1+z) \text{ Å})}{1+z}, \quad (3)$$

where in the numerator is the measured flux density per frequency at an observed wavelength of $2500(1+z) \text{ Å}$.

Importantly, we estimate L_{2500} based on SDSS photometry rather than SDSS spectroscopy. This allows us to reliably control the statistical completeness of the sample. Indeed, quasar candidates for SDSS spectroscopy had been selected based on their photometric magnitudes and colours and we know (see Appendix A) how the completeness of this selection depends on the apparent magnitude limit in the g band. Therefore, since we estimate L_{2500} at the epoch of target selection (SDSS photometry), we can specify a strict UV luminosity limit as a function of redshift, $L_{2500\text{min}}(z)$, above which our sample is statistically complete. This would be impossible to do if we instead used SDSS spectroscopy for estimating L_{2500} , since SDSS spectra were taken several years after the corresponding photometric observations, when the quasars had already changed their luminosity due to intrinsic variability.

Specifically, we estimated $F_\nu(2500(1+z) \text{ Å})$, for each quasar based on its SDSS g -band photometric measurement (m_g), SDSS spectroscopic redshift (z) and the corresponding k -correction, $K_{UV}(z)$:

$$F_\nu(2500(1+z) \text{ Å}) = F_{\text{mg}} K_{UV}(z), \quad (4)$$

where F_{mg} is the flux corresponding to m_g .

Assuming that the spectral shape did not change between the epochs of SDSS photometry and spectroscopy for each quasar, we estimated the K_{UV} values using the aforementioned catalogue of spectral properties by Wu & Shen (2022), which is based on SDSS spectroscopy. These authors used pyQSOFIT (see Shen et al. 2019) and fitted the SDSS spectra by a global continuum plus emission lines model. For the continuum they used a sum of a power-law, a third-order polynomial model and a component responsible for optical and UV FeII emission.

We used the parameters of the power-law and third-order polynomial quasi-continuum for each quasar to estimate $F_\nu^{\text{spec}}(2500(1+z) \text{ Å})$ (hereafter, the superscript ‘spec’ indicates that fluxes are estimated from spectra rather than from photometry) for the continuum emission, excluding the contribution of Fe II. Then we calculated K_{UV} as the ratio of the $F_\nu^{\text{spec}}(2500(1+z) \text{ Å})$ and the integrated flux (including the contribution of emission lines and FeII emission) over the observed

g band, F_g^{spec} :

$$K_{UV} = \frac{F_\nu^{\text{spec}}(2500(1+z) \text{ Å})}{F_g^{\text{spec}}}. \quad (5)$$

We use in our analysis the continuum rather than the total luminosity at 2500 Å, since the continuum emission presumably comes from the innermost region of the accretion disc near the corona and it varies roughly in concert with the coronal X-ray emission on the time scales we are interested in (see Sect. 3.3.2). However, the difference between the two quantities is modest: namely, the total flux density at rest-frame 2500 Å is greater than $F_\nu^{\text{spec}}(2500(1+z) \text{ Å})$ by just 15% on average and this fraction ranges from 8% to 27% for 90% for our quasars. We also note that the contribution of emission lines and FeII emission to the total flux in the g band is 13% on average and ranges from 6% to 27% for 90% of our sample. We neglect any intrinsic extinction in the UV, assuming that it is small for our quasars since they all have broad optical emission lines. It is thus clear that we underestimate the intrinsic UV luminosities, but the amplitude of this effect is difficult to assess. Note that a similar small systematic bias (due to intrinsic absorption) should affect our estimates of the X-ray luminosities, as was already mentioned in Section 2.1.

We then used the $F_\nu(2500(1+z) \text{ Å})$ values to estimate L_{2500} . The uncertainty of L_{2500} is composed of the photometry uncertainty $\sigma(m_g)$ and the uncertainty in the k -correction, $\sigma(\log K_{UV})$. The latter is derived from the uncertainties in $F_\nu^{\text{spec}}(2500(1+z) \text{ Å})$ and F_g^{spec} by propagation of error in equation (5), neglecting the uncertainty in spectroscopic redshift. We adopted the $F_\nu^{\text{spec}}(2500(1+z) \text{ Å})$ uncertainty directly from the relevant LOGL2500_ERR column in Wu & Shen (2022), and estimated the F_g^{spec} uncertainty by averaging the uncertainty in the spectral bins across the g -band. Hence,

$$\sigma(\log F_\nu(2500(1+z) \text{ Å})) = \sqrt{[0.4\sigma(m_g)]^2 + \sigma(\log K_{UV})^2}. \quad (6)$$

During this procedure, we removed from our sample 50 quasars with missing LOGL2500_ERR values, which indicates that the spectrum breaks off around 2500 Å, so the uncertainty cannot be evaluated. It turns out that the flux uncertainty $\sigma(\log F_\nu^{\text{spec}}(2500(1+z) \text{ Å}))$ is dominated by photometry. The median value of $0.4\sigma(m_g)/\sigma(\log K_{UV})$ is 3.5.

We finally intend to impose a strict luminosity limit $L_{2500\text{min}}(z)$ on our sample, in order to minimize the impact of any k -correction (i.e. optical–UV spectral shape) related selection bias on our analysis of the X-ray–UV luminosity correlation. To this end, we divided our quasar sample into ten redshift bins containing nearly the same number of objects (from 313 to 316 objects per bin). As is shown in the top panel of Fig. 2, the scatter in the k -correction within a given redshift bin is not large. Namely, 90% (5–95%) of quasars have k -corrections differing by less than 0.25 (less than 0.18 for all bins except the highest redshift one). For each bin, we determined a limiting value K_{99} such that 99% of quasars in that bin have $K_{UV} < K_{99}$ ³.

³This threshold is sufficiently conservative for our purposes as it is com-

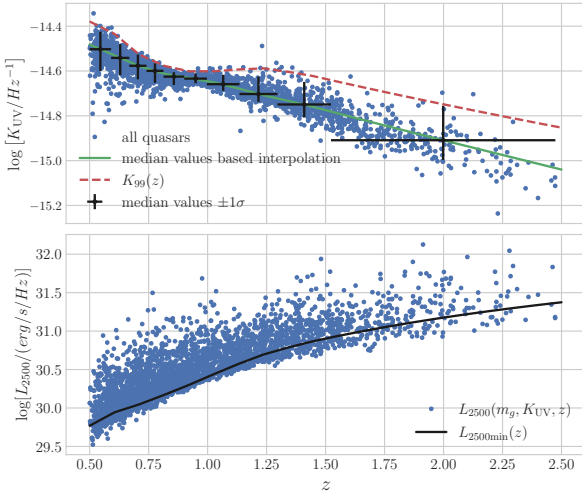


Figure 2: Top panel: Scatter-plot of k -correction vs redshift for quasars from our sample before the additional UV luminosity selection. The red dashed line bounds the region containing 99% of objects in each redshift bin. Bottom panel: Corresponding scatter-plot of 2500Å luminosity vs redshift. The black line depicts the luminosity limit $L_{2500min}$ as a function of z .

We then derived a smooth dependence $K_{99}(z)$ by interpolation between these discrete K_{99} values (see the top panel of Fig. 2). Finally, we defined a UV luminosity limit,

$$L_{2500min}(z) = \frac{4\pi D_L^2(z) F_{mg=19} K_{99}(z)}{1+z}, \quad (7)$$

where $F_{mg=19}$ is the flux in the g band corresponding to $m_{g,max} = 19$, and excluded all sources (645) with $L_{2500} < L_{2500min}(z)$ from our sample (see the bottom panel of Fig. 2). This ensures that quasars with $L_{2500} > L_{2500min}(z)$, selected for the subsequent analysis, will fairly represent the overall quasar population with all its spectral diversity.

The resulting sample consists of 2414 quasars, and it is used in our analysis below. We will call it the *SRG/eROSITA*–SDSS quasar sample.

2.4. Properties of the quasar sample

Figure 3 shows the distribution of the logarithms of the X-ray luminosities and the corresponding uncertainties, $\sigma(\log L_{2keV})$ (due to photon count fluctuations and the systematic uncertainty in the X-ray spectral shape), based on *SRG/eROSITA* data. The median value of $[\sigma(\log L_{2keV})]^2 = 4.5 \times 10^{-3}$. Note that it would be 3.0×10^{-3} if we neglected the systematic uncertainty. Figure 4 shows the distribution of the logarithms of the UV luminosities and their uncertainties, $\sigma(\log L_{2500})$. The latter is

parable to the aforementioned completeness ($\gtrsim 98\%$) of SDSS spectroscopic coverage of quasars.

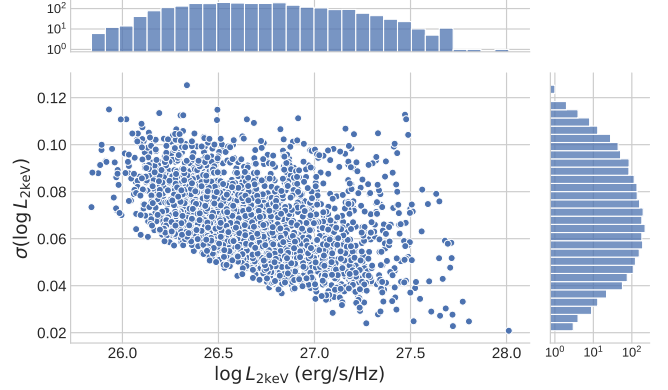


Figure 3: Scatter-plot of the logarithms of X-ray luminosities vs. their uncertainties for the *SRG/eROSITA*–SDSS quasar sample. The histograms shown on the sides reflect projections of the two-dimensional distribution on the corresponding axes (note the logarithmic scale for the number of objects).

equal to the uncertainty in $\log F_{2500}$, defined in equation (6). The median $[\sigma(\log L_{2500})]^2 = 7.4 \times 10^{-5}$, much smaller than for the X-ray luminosity. Note that we neglect uncertainties in the spectroscopic z measurements and in the cosmological model.

Figure 5 shows the g -band photometric magnitudes measured by SDSS. We have checked that the apparent drop of the m_g distribution near our adopted optical limit of $m_{g,max} = 19$ is not due to paucity of such quasars in the DR16Q catalogue but rather a consequence of the limits that we imposed on UV luminosity, $L_{2500min}$, and X-ray flux, $F_{X,min}$ (via the intrinsic correlation between X-ray and UV luminosity of quasars).

Figure 6 shows the distributions of: (i) redshifts, (ii) X-ray luminosities and (iii) UV luminosities. The median redshift of our sample is 0.897. Importantly, we cover a broad dynamic range, ~ 2 dex, in both X-ray and UV luminosity. The smooth boundary of the distribution in the middle panel reflects the fact that by defining L_{2500} as a function of m_g and z we have achieved a strict cut in L_{2500} for each z , just as for L_{2keV} , which is important for the following statistical analysis.

3. Method

The inferred parameters of the L_{2keV} – L_{2500} correlation can be affected by selection effects and the specifics of methods applied to the data. Although such biases have been addressed before, we see some potentially significant shortcomings in the previous analyses.

In particular, Dainotti et al. (2022) took into account flux selection in treating the cosmological evolution of quasars but not in determining the L_{2keV} – L_{2500} correlation parameters (based on de-evolved luminosities). Other authors attempted to take flux selection into account by considering quasars undetected in X-rays. In particular, Lusso & Risaliti (2016) and Rankine et al. (2024) took into account X-ray flux upper limits for optically selected quasars, but did not discuss how the correlation might

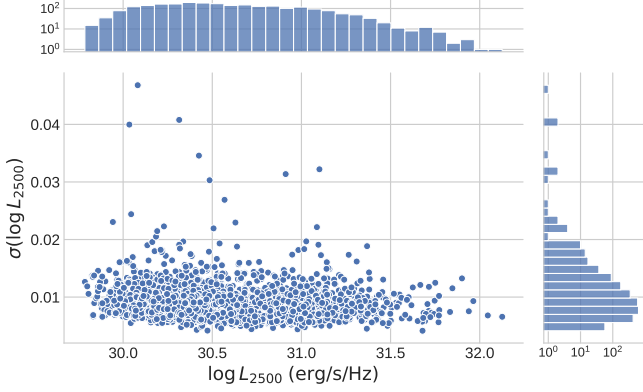


Figure 4: Scatter-plot of the logarithms of UV (2500 Å) luminosities vs. their uncertainties for the *SRG/eROSITA*–SDSS quasar sample. The histograms shown on the sides reflect projections of the two-dimensional distribution on the corresponding axes (note the logarithmic scale for the number of objects).

be affected by X-ray bright quasars that are undetected in the optical.

We propose a novel method, which uses quasars detected both in the X-ray and optical–UV bands and fully accounts for biases associated with objects undetected in each of these bands. A similar approach was used, in a different context, by Takeuchi et al. (2013), who constructed a far-ultraviolet and far-infrared bivariate luminosity function of galaxies using a copula approach, which accounts for selection effects in a similar way to ours (see equation 25 in that article).

Another novelty of our method consists in a more general definition of the scatter of the correlation between X-ray and UV luminosities compared to previous works. All preceding studies (e.g., Lusso et al. 2010; Bisogni et al. 2021; Dainotti et al. 2022; Signorini et al. 2024; Rankine et al. 2024) attributed this scatter to either $L_{2\text{keV}}$ or L_{2500} , i.e. to only one of the two physical quantities involved in the correlation. In doing so, it is assumed that either $L_{2\text{keV}}$ depends on L_{2500} or vice versa, and results will depend on this assumption. For example, in the ordinary least squares (OLS) analysis conducted by Steffen et al. (2006), significantly different results were obtained when the scatter was ascribed to X-ray or to UV luminosity: the authors found slopes of 0.642 ± 0.021 and 0.815 ± 0.023 for OLS(X|UV) and OLS(UV|X), respectively. However, there is no obvious reason to regard either X-ray or UV luminosity as the dependent variable. Thus, we do not give any preference to either $L_{2\text{keV}}$ and L_{2500} and take the scatter in both $L_{2\text{keV}}$ and L_{2500} into account.

3.1. Mathematical formulation

For convenience and to facilitate comparison with previous studies, we will work in logarithmic luminosity space and thus introduce $l_X = \log \frac{L_X}{\text{erg s}^{-1} \text{Hz}^{-1}}$, $l_{UV} = \log \frac{L_{2500}}{\text{erg s}^{-1} \text{Hz}^{-1}}$. Suppose that these ‘observed’ luminosities are related to each other with some scatter. Then we can introduce ‘primary’ luminosities

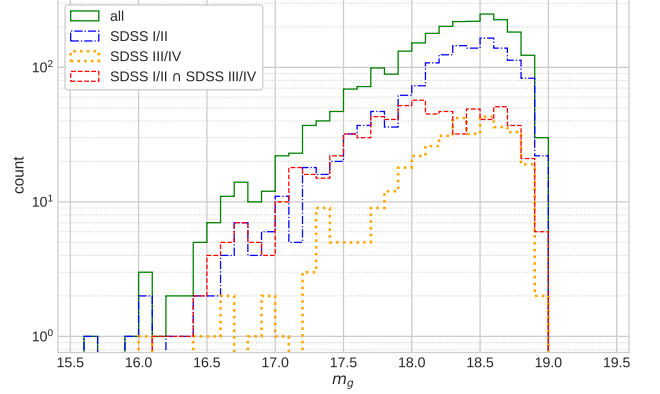


Figure 5: Distribution of SDSS *g*-band photometric magnitudes for the *SRG/eROSITA*–SDSS quasar sample. The full sample is shown in green. The other histograms show subsamples according to SDSS spectroscopy: blue, quasars observed in SDSS I/II only; orange, those observed in SDSS III/IV only; red, those observed both in SDSS I/II and SDSS III/IV.

$\tilde{l}_X = \log \frac{\tilde{L}_X}{\text{erg s}^{-1} \text{Hz}^{-1}}$ and $\tilde{l}_{UV} = \log \frac{\tilde{L}_{UV}}{\text{erg s}^{-1} \text{Hz}^{-1}}$ such that one is related to the other via a continuous dependence, and l_X and l_{UV} are scattered around this ‘primary’ relation between $\tilde{l}_X$ and $\tilde{l}_{UV}$. The scatter arises due to (i) ‘intrinsic scatter’, (ii) variability and (iii) flux measurement uncertainties. Therefore, by ‘intrinsic’ we mean all sources of scatter that are not related to variability and flux measurement uncertainties. Specifically, we seek a correlation between X-ray and UV luminosities measured simultaneously and averaged over a certain time scale. The exact meaning of ‘simultaneity’ and the ‘averaging time scale’ will be revealed in Section 3.3.

We further assume, for simplicity and computational efficiency, that all three sources of scatter, both for X-ray and UV luminosity, are mutually independent and normally distributed, that is the observed X-ray and UV luminosities obey lognormal distributions relative to the primary ones. Then the distribution of l_X for a given $\tilde{l}_X$ is a normal one:

$$\mathcal{N}(l_X|\tilde{l}_X, \sigma_X) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{(\tilde{l}_X - l_X)^2}{2\sigma_X^2}\right), \quad (8)$$

and, similarly, l_{UV} for a given $\tilde{l}_{UV}$:

$$\mathcal{N}(l_{UV}|\tilde{l}_{UV}, \sigma_{UV}) = \frac{1}{\sqrt{2\pi}\sigma_{UV}} \exp\left(-\frac{(\tilde{l}_{UV} - l_{UV})^2}{2\sigma_{UV}^2}\right), \quad (9)$$

where

$$\begin{aligned} \sigma_X^2 &= \sigma_{mX}^2 + \sigma_{\text{varX}}^2 + \sigma_{\text{intX}}^2, \\ \sigma_{UV}^2 &= \sigma_{mUV}^2 + \sigma_{\text{varUV}}^2 + \sigma_{\text{intUV}}^2. \end{aligned} \quad (10)$$

Here σ_{mX} and σ_{mUV} are the X-ray and UV measurement uncertainties, σ_{varX} and σ_{varUV} the contributions associated with

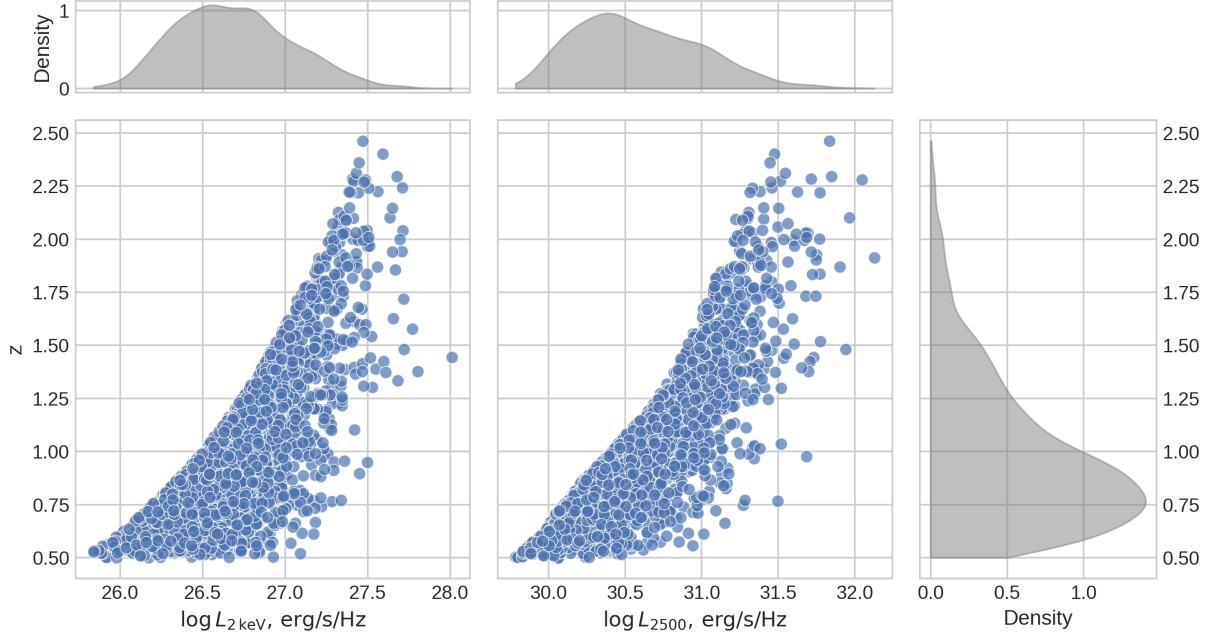


Figure 6: Scatter-plots of redshift vs. X-ray luminosity (left) and UV luminosity (right), with the corresponding marginal distributions for the *SRG/eROSITA–SDSS* sample.

X-ray and UV variability, and σ_{intX} and σ_{intUV} account for the intrinsic scatter.

It should be emphasized that we do not know the nature of the intrinsic scatter and thus assume, for simplicity, that its contributions along the X-ray and UV luminosity axes are independent of each other. The real situation may be more complicated, with some cross-talk between X-ray and UV intrinsic scatter. Nevertheless, the current formulation is more general than if we had assumed the intrinsic scatter to be associated with either X-ray or UV luminosity. The contributions of X-ray and UV measurement uncertainties obviously do not correlate with each other or with other contributions to the observed scatter. The exact meaning of the X-ray and UV variability contributions to the scatter will be clarified in Section 3.3.

Following previous studies (e.g. Lusso & Risaliti 2016), we seek the primary dependence in a linear form (in logarithmic space):

$$\tilde{l}_X = \beta + \gamma \tilde{l}_{UV}, \quad (11)$$

where we assume that γ and β are independent of redshift (at least in the interval $0.5 < z < 2.5$).

We also define a ‘primary quasar luminosity function (LF)’:

$$\widetilde{LF}(\tilde{l}_X, z) \equiv \frac{dN(z)}{d\tilde{l}_X}, \quad (12)$$

which is the number of quasars per comoving volume of the Universe per $\tilde{l}_X$ as a function of z , and correspondingly the nor-

malized primary LF:

$$\widetilde{NLF}(\tilde{l}_X|z) \equiv \frac{\widetilde{LF}(\tilde{l}_X, z)}{\int \widetilde{LF}(\tilde{l}_X, z) d\tilde{l}_X}. \quad (13)$$

Here, the integration is to be carried out over $\tilde{l}_X$ from $-\infty$ to ∞ . However, for computational convenience we set a lower limit on luminosity, l_{cut} , below which there are no quasars in the Universe. That is, we additionally postulate that $\widetilde{NLF}(\tilde{l}_X|z) = 0$ for $\tilde{l}_X < l_{\text{cut}}$. The specific choice of this threshold does not affect our results if l_{cut} is less than the luminosity corresponding to the X-ray and optical/UV flux limits at a given redshift.

With these definitions, the probability density for a quasar at redshift z to have observed luminosities l_X and l_{UV} (without yet taking selection by luminosity thresholds into account) is

$$\begin{aligned} f(l_X, l_{UV}|z, \sigma_{UV}, \sigma_X) &= \int \int \mathcal{N}(l_X|\tilde{l}_X, \sigma_X) \mathcal{N}(l_{UV}|\tilde{l}_{UV}, \sigma_{UV}) \\ &\quad \times \delta(\tilde{l}_X - (\beta + \gamma \tilde{l}_{UV})) |\gamma| \widetilde{NLF}(\tilde{l}_X|z) d\tilde{l}_X d\tilde{l}_{UV} \\ &= \int \mathcal{N}(l_X|\tilde{l}_X, \sigma_X) \mathcal{N}(l_{UV}|\tilde{l}_X - \beta)/\gamma, \sigma_{UV}) \widetilde{NLF}(\tilde{l}_X|z) d\tilde{l}_X, \end{aligned} \quad (14)$$

where δ is the Dirac delta function and the integration is again carried out from $-\infty$ to ∞ .

Alternatively, we could consider the UV (at 2500 Å) LF, $\widetilde{LF}_{UV}(\tilde{l}_{UV}, z)$. The model given by equation (14) has the following meaning: the delta function determines the primary dependence, $\mathcal{N}$ describes the scatter around it along the X-ray and

UV luminosity axes, and $\widetilde{NLF}$ describes the changing (normalized) number density of quasars along the primary dependence.

3.1.1. Account for selection effects

Next, we wish to take into account the selection effects associated with the X-ray and UV luminosity thresholds. Suppose that: (i) $P(A_i)$ is the probability of having a quasar with X-ray luminosity $l_{X,i}$ and UV luminosity $l_{UV,i}$ in the Universe at redshift z , and (ii) $P(B)$ is the probability to pass the selection criterion. Then, the probability that a quasar with $l_{X,i}$ and $l_{UV,i}$ will enter the resulting sample is $P(B|A_i)$. When analyzing a sample of thus selected quasars, we are actually interested in another probability, $P(A_i|B)$, which can be found using Bayes' theorem:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)}. \quad (15)$$

Here, the summation in the denominator on the right-hand side is carried out over all possible pairs of $l_{X,j}$ and $l_{UV,j}$.

We can rewrite this equation in terms of probability density by replacing $P(A_i)$ and $P(B|A_i)$ by $f(l_{X,i}, l_{UV,i}|z_i, \sigma_{UV,i}, \sigma_{X,i})dl_{UV}dl_X$ and $\rho(l_{X,i}, l_{UV,i}, z_i)$, respectively, where ρ is a 'selection function' for given $l_{X,i}$, $l_{UV,i}$ and z_i . Then the observed probability density f_{obs} , i.e. the probability density of finding a quasar with particular $l_{UV,i}$ and $l_{X,i}$ at given z_i , provided it has been selected (based on its $l_{X,i}$, $l_{UV,i}$ and z_i), is:

$$f_{obs,i} \equiv f_{obs}(l_{X,i}, l_{UV,i}|z_i, \sigma_{mX,i}^2, \sigma_{mUV,i}^2, \sigma_{varX,i}^2, \sigma_{varUV,i}^2) = \frac{\rho(l_{X,i}, l_{UV,i}, z_i)f(l_{X,i}, l_{UV,i}|z_i, \sigma_{UV,i}, \sigma_{X,i})}{\int \int \rho(l_X, l_{UV}, z_i)f(l_X, l_{UV}|z_i, \sigma_{UV,i}, \sigma_{X,i})dl_{UV}dl_X}, \quad (16)$$

where, in our case, $\rho(l_X, l_{UV}, z) = \Theta(l_X - l_{Xmin}(z))\Theta(l_{UV} - l_{UVmin}(z))$ is a two-dimensional step function. In fact, there should be a multiplier responsible for the z limits we applied ($0.5 < z < 2.5$), but it does not depend on the luminosities, so it cancels out in the numerator and denominator of equation (16).

In our case, the limiting X-ray luminosity at z_i is determined by the *SRG/eROSITA* flux limit as $l_{Xmin}(z_i) = \log \frac{4\pi D_L^2(z_i)F_{Xmin}}{(1+z_i)(\text{erg s}^{-1})} + K_X(z_i)$, where $K_X(z)$ is the k -correction from the observed-frame 0.3–2.3 keV flux to the flux density per frequency at observed $2/(1+z)$ keV (using our adopted power-law model with photon index $\Gamma = 2$). The limiting UV luminosity $l_{UVmin}(z_i) \equiv L_{2500min}(z_i)$ was specified in Section 2.3.

As a result, equation (16) simplifies to:

$$f_{obs,i} = \frac{f(l_{X,i}, l_{UV,i}|z_i, \sigma_{UV,i}, \sigma_{X,i})}{\int_{l_{UVmin}(z_i)}^{\infty} \int_{l_{Xmin}(z_i)}^{\infty} f(l_X, l_{UV}|z_i, \sigma_{UV,i}, \sigma_{X,i})dl_Xdl_{UV}}, \quad (17)$$

where we have omitted the theta functions in the numerator because we are sampling quasars that by definition are above the imposed luminosity thresholds. Thus, the observed probability density $f_{obs,i}$ for quasar i from our sample is $f(l_{X,i}, l_{UV,i}|z_i, \sigma_{UV,i}, \sigma_{X,i})$ normalized by the integral over the $[l_{UV}, l_X]$ parameter area above the luminosity thresholds at a given z_i (see also Takeuchi et al. 2013).

Crucially for the purposes of this study, the integration in equation (14) properly accounts for flux limit related selection biases. Objects with given $\tilde{l}_X$ or $\tilde{l}_{UV}$ can have higher or lower observed l_X or l_{UV} due to (i) intrinsic scatter around the primary dependence between $\tilde{l}_X$ or $\tilde{l}_{UV}$, (ii) variability and (iii) photon count fluctuations at the detector (the latter, combined with the higher space density of less luminous objects leads to the well-known Eddington bias). In our method, we self-consistently account for all these effects.

Furthermore, the absence of a redshift integral in equation (16) renders our method insensitive to redshift selection. Essentially, we adjust each quasar to the probability density corresponding to its z_i , and the relative number density of quasars at different redshifts does not play a role. Thus, the lack of objects at some redshift in the sample would merely lead to less information about quasars at this z , so that such objects would contribute less statistical weight to the final result. Therefore, we could also, for example, split our sample into bins by z , and apply our method to each bin separately, without applying a redshift selection to the algorithm. We finally note that the SDSS quasar catalogue is actually not significantly affected by redshift selection in our adopted parameter range of $m_g < 19$ and $0.5 < z < 2.5$ (see Appendix A).

3.1.2. Final algorithm

Eventually, we use the following likelihood function for the *SRG/eROSITA*–SDSS quasar sample:

$$\mathcal{L}(\beta, \gamma, \sigma_{intUV}^2, \sigma_{intX}^2, \mathbf{p}_{LF}) = \prod_i f_{obs}(l_{X,i}, l_{UV,i}|z_i, \sigma_{mX,i}^2, \sigma_{mUV,i}^2, \sigma_{varX,i}^2, \sigma_{varUV,i}^2; \beta, \gamma, \sigma_{intUV}^2, \sigma_{intX}^2, \mathbf{p}_{LF}), \quad (18)$$

where the summation is carried out over all quasars in the *SRG/eROSITA*–SDSS sample, $\mathbf{p}_{LF}$ are parameters of the LF, which we define later.

In practice, we could fix the quasar $\widetilde{LF}$ to some model from the literature and consider only the four parameters of the correlation, namely, β , γ , σ_{intUV}^2 and σ_{intX}^2 as free ones. Alternatively, we could also simultaneously derive the parameters of $\widetilde{LF}$. We adopted the latter approach and used a Markov chain Monte Carlo (MCMC) method for Bayesian inference of parameter distribution, unless otherwise specified. Specifically, we utilize the Python package *emcee* (Foreman-Mackey et al., 2013), an implementation of the affine-invariant ensemble sampler for MCMC proposed by Goodman and Weare.

3.2. Luminosity function

Our method relies on assumptions about the AGN LF, which limits the precision of determination of the X-ray–UV luminosity correlation parameters. On the other hand, this allows us to test various hypotheses about the LF based on the available *SRG/eROSITA* and SDSS data, i.e. we can control the impact of uncertainties in the LF on the inferred correlation parameters.

As introduced in Section 3.1, we based our analysis on the primary LF, $\widetilde{LF}$, which gives the number density of quasars

per $\tilde{l}_X$. However, observational studies always report on LFs that are defined as number density per observed X-ray (or optical/UV) luminosity, l_X . This observed X-ray luminosity function, LF , is different from our $\widetilde{LF}$ due to the intrinsic scatter of the X-ray–UV luminosity correlation and due to the broadening caused by variability. Namely, NLF (normalized LF) results from convolution of $\widetilde{NLF}$ with a Gaussian of width $\sqrt{\sigma_{\text{intX}}^2 + \sigma_{\text{varX}}^2}$:

$$NLF(l_X|z) = \int \mathcal{N}(l_X|\tilde{l}_X, \sqrt{\sigma_{\text{intX}}^2 + \sigma_{\text{varX}}^2}) \widetilde{NLF}(\tilde{l}_X|z) d\tilde{l}_X, \quad (19)$$

where $\mathcal{N}(l_X|\tilde{l}_X, \sqrt{\sigma_{\text{intX}}^2 + \sigma_{\text{varX}}^2})$ is the normal distribution defined in equation (8). We can thus expect NLF to be different from $\widetilde{NLF}$, especially in the (steep) high-luminosity part. This is demonstrated explicitly in Appendix B.

In what follows, we use a luminosity-and-density-evolution (LADE) model, following Aird et al. (2015), as our baseline $\widetilde{LF}$ model (in the X-ray band). Since, as explained above, we are not sensitive to the redshift evolution of quasar number density, we cannot limit the parameters describing the evolution of the LF normalization. Thus, we parameterize the model as follows:

$$\widetilde{LF}(\tilde{l}_X, z) \propto \left[\left(\frac{\tilde{l}_X}{e(z)L^*} \right)^{g_1} + \left(\frac{\tilde{l}_X}{e(z)L^*} \right)^{g_2} \right]^{-1}, \quad (20)$$

where g_1 and g_2 are the slopes of the faint and bright parts of the luminosity dependence, and $e(z)$ is a function describing the cosmological evolution of the characteristic luminosity:

$$e(z) = \left[\left(\frac{1+z_c}{1+z} \right)^{p_1} + \left(\frac{1+z_c}{1+z} \right)^{p_2} \right]^{-1}, \quad (21)$$

where z_c is a transition redshift, $L^*/2$ is the bend luminosity at $z = z_c$, and p_1 and p_2 are the slopes of the redshift dependence.

A LADE model of this kind is known to be a good approximation of the observed X-ray LF of AGN at redshifts $z \lesssim 3.5$ over the broad range of luminosities, $10^{42} < L_{2-10} < 10^{46} \text{ erg s}^{-1}$ (e.g. Aird et al. 2015), where L_{2-10} is the luminosity in the 2–10 keV rest-frame energy band. Moreover, similar models have also been successfully used to describe optical/UV LFs of AGN, namely at 2500 Å and in the b_j band, (Boyle et al., 2000; Palanque-Delabrouille et al., 2013; Ross et al., 2013). Therefore, the above model is suitable for our purposes, since our *SRG*/eROSITA–SDSS sample does not contain high-redshift ($z > 2.5$) quasars and low-luminosity ($L_{2-10} < 10^{43} \text{ erg s}^{-1}$) AGN. In what follows, we regard all parameters of the LADE model (we do not consider the normalization and its evolution, as mentioned above) as free ones, albeit with priors, expect for the high-redshift slope p_2 in the evolution function, which is difficult to constrain with our data and so we fix it at the best-fit value of -3.58 from table 5 in Aird et al. (2015). If we take the corresponding prior from table 5 in Aird et al. (2015), we get the same results.

In reality, as has been shown in many X-ray and optically based studies (e.g. Ueda et al. 2014; Aird et al. 2015), the cosmological evolution of AGN is more complicated than LADE,

but this only becomes important at large redshifts, which are not probed by this study. Therefore, the use of more complicated existing LF models, with their many additional parameters related to cosmological evolution, would unnecessarily complicate our treatment.

3.3. Variability

One of the key elements of our analysis (see Section 3.1) consists in estimating the contribution of quasar variability to the observed scatter of the X-ray–UV luminosity correlation and separating it from the intrinsic scatter. As is well known, AGN are variable both in the X-ray and optical–UV bands on time scales ranging from hours (or even minutes in some Seyfert galaxies) up to many years (essentially up to the Hubble time). We thus need to combine the general knowledge of AGN variability with a proper account of the timing properties of the *SRG*/eROSITA and SDSS observations on which the current study is based.

3.3.1. Timing properties of the data

Let us first consider how the data used in this study were obtained. As regards the optical band, typically a single photometric measurement was obtained for each quasar during the SDSS imaging, with an exposure time of about 55 seconds per band. Therefore, the UV luminosity estimates (l_{UV}) that we use are single-epoch measurements obtained on an observed time scale of ~ 1 minute, which corresponds to a time scale of ~ 30 seconds in the quasar rest frame at typical redshifts $z \sim 0.5$ – 1.5 .

As regards the eROSITA data, each quasar was observed during four or five (depending on the celestial position) *SRG* all-sky scans with intervals of six months. Each of such visits lasted between one day (for sources located near the ecliptic plane) and approximately one week (for sources with high ecliptic latitude; note that yet longer exposures are not present in this study because the SDSS footprint does not cover the north ecliptic pole), and in fact was a series of 40-s long exposures separated by four hours (see Sunyaev et al. 2021 for more details about the *SRG* all-sky survey). The X-ray luminosity estimates (l_X) that we use are based on source fluxes measured on the map of the Eastern Galactic hemisphere obtained by combining all eROSITA data accumulated during the all-sky survey between Dec. 2019 and Feb. 2022. Therefore, these measurements are affected by quasar variability on several time scales, namely: ~ 40 seconds (single exposure), ~ 4 hours (interval between exposures), $\gtrsim 1$ day (single visit), ~ 6 months (interval between consecutive scans), and ~ 2 years (duration of the *SRG*/eROSITA all-sky survey). These times pertain to the observer’s frame; in the quasar rest frame, they are shorter by a factor of $(1+z) \sim 2$.

The next important point to consider is that the X-ray data (*SRG*/eROSITA) were obtained with a significant delay with respect to the optical photometry data (SDSS). Figure 7 (second left panel) shows the distribution of lags in the quasar rest frame, ΔT_2 , between the SDSS and eROSITA observations for our sample. To this end, the X-ray measurement instant for a

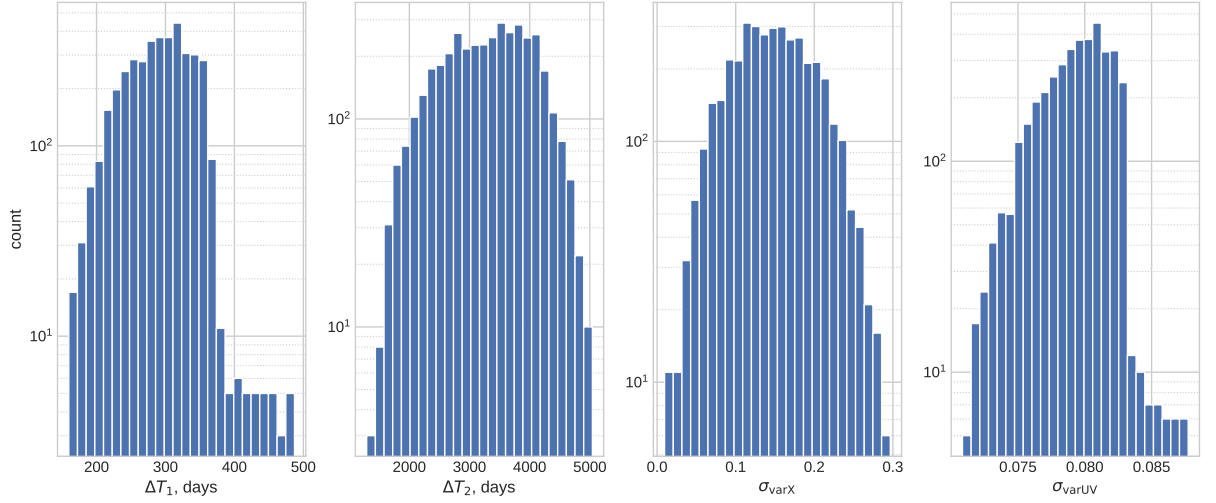


Figure 7: From left to right: distributions of the rest-frame X-ray (eROSITA) observation periods, rest-frame lags between X-ray and optical photometric (SDSS) measurements and the corresponding distributions of the expected long-term X-ray and short-term UV variability contributions (see main text) to the X-ray–UV luminosity scatter.

given quasar is defined as the average between the start time of the first scan and the end time of the last scan (4th or 5th). The rest-frame X-ray–UV lags range between 3.5 and 13.8 years, and the median $\Delta T_2 = 8.6$ years.

3.3.2. Basic considerations and clarification of the definition of the X-ray–UV luminosity correlation

Despite the non-trivial timing properties of the data, a number of well-established observational facts and reasonable assumptions can help us proceed with the treatment of variability. First, we assume that for a given quasar with luminosity $L(t)$ in some energy band at time t , the distribution of the logarithms of possible luminosities $l(t + \Delta t)$ after a period of time Δt is a Gaussian one centered at $l(t) = \log L(t)$, with variance

$$\sigma_{\text{var}}^2 = \langle [l(t + \Delta t) - l(t)]^2 \rangle = \langle [\log L(t + \Delta t) - \log L(t)]^2 \rangle \equiv S F^2(\Delta t) \quad (22)$$

Here, we have introduced a structure function (SF) (or first-order structure function if we follow Simonetti et al. (1985) definition), which is often used in variability studies of AGN and other astrophysical objects. In what follows, we will use a ‘two-argument structure function’, $\widetilde{S F}^2(\Delta t_1, \Delta t_2)$, which characterizes the variability on rest-frame time scale Δt_2 of the luminosity averaged over $\Delta t_1 < \Delta t_2$, i.e.

$$\widetilde{S F}^2(\Delta t_1, \Delta t_2) \equiv \left\langle \left[\overline{l(t + \Delta t_2)} - \overline{l(t)} \right]^2 \right\rangle, \quad (23)$$

where $\overline{l}$ means averaging over Δt_1 .

As known from many studies of Seyfert galaxies and quasars, X-ray and optical variability (e.g. expressed in terms of SF) in

a given AGN grows with increasing time scale, and this dependence can to a first approximation be described by a power-law on time scales from days to a couple of decades both in X-rays (Vagnetti et al., 2011, 2016; Middei et al., 2017; Prokhorenko et al., 2024) and in the optical (see e.g. MacLeod et al. 2010; Morganson et al. 2014). Therefore, to a first approximation our l_X estimates may be regarded as luminosities averaged over a $\sim 2/(1+z)$ year period (which corresponds to the full time span of the *SRG* all-sky survey). In contrast, as already mentioned, the adopted l_{UV} values represent much shorter (less than a minute), snapshot estimates of the UV luminosities.

Recent long-term (~ 1 year and longer) monitoring (so-called continuum reverberation mapping) campaigns of several AGN have demonstrated that variability in the X-ray and optical–UV bands is usually strongly correlated on time scales of months to years, but much less so on shorter (days to weeks) scales (Edelson et al., 2019; Hernández Santisteban et al., 2020; Cackett et al., 2023; Lawther et al., 2023; Edelson et al., 2024). Hagen et al. (2024) suggest that long-term X-ray–UV variability correlation can result from accretion-rate fluctuations that propagate inwards through the accretion disc (as in the model of Lyubarskii 1997) and modulate not only the optical–UV emission produced in the disc but also the X-rays generated in the central hot corona, whereas the complex short-term behaviour is due to additional X-ray variability induced in the corona and due to reprocessing of the X-rays into lower frequency radiation in the disc and, perhaps, in its wind. Although continuum reverberation mapping programs have so far been mostly targeted at nearby, moderate luminosity ($\sim 10^{43}$ – 10^{44} erg s $^{-1}$ in the X-ray band) Seyfert galaxies, a few similar campaigns have been conducted for higher luminosity AGN. In particular,

Arévalo et al. (2008) reported on the results of 2.5 year-long monitoring of quasar MR 2251–178, whose X-ray luminosity ($\sim 5 \times 10^{44} \text{ erg s}^{-1}$, 2–10 keV) is typical of our *SRG/eROSITA*–SDSS quasar sample, and found its X-ray and optical light curves to be strongly correlated on times scales of months to years.

It is thus reasonable to assume that for each quasar there is a ‘correlation time scale’, ΔT_c , shorter than a few months, on which its UV and X-ray emission begin to strongly correlate with each other. More precisely, we assume that intrinsic variability on time scales longer than ΔT_c ‘moves’ quasars along the primary dependence given by equation (11) on the UV–X-ray luminosity diagram, i.e. if we could average the X-ray and UV luminosities of a given quasar over the same time interval of duration longer than ΔT_c , then the following relation between the mean X-ray and UV luminosities would hold:

$$\langle l_X \rangle = \beta + \gamma \langle l_{UV} \rangle. \quad (24)$$

This allows us to narrow down the definition of the X-ray–UV luminosity interdependence that we wish to determine. Namely, we seek a relation between UV and X-ray luminosities averaged over a period of time (the same for the X-ray and UV measurements, i.e. they have to be simultaneous) of length $\Delta T_1 \geq \Delta T_c$. The particular value of ΔT_1 is not important as long as this condition is satisfied. In practice, we define ΔT_1 individually for each quasar as the rest-frame time span of its monitoring during the *eROSITA* all-sky survey, i.e. the time passed between the first and last X-ray observations. This is justified in our case, as $\Delta T_1 \sim 1$ year for the *SRG/eROSITA*–SDSS quasar sample, so that the condition $\Delta T_1 > \Delta T_c$ is expected to hold. The distribution of ΔT_1 for the *SRG/eROSITA*–SDSS quasar sample is shown in the leftmost panel on Fig. 7. Accordingly, σ_{intUV} and σ_{intX} denote the intrinsic scatter of UV and X-ray luminosities with respect to the so-defined primary dependence.

Given this definition, the parameters of the primary dependence and its scatter could be measured directly if the UV and X-ray observations were conducted simultaneously, on the same time scale (ΔT_1). Unfortunately, neither condition is met in our case. Therefore, we need to take X-ray and UV variability into account, as described below.

We illustrate our approach and the key concepts and assumptions of this study in Fig. 8. Note that it is made for illustrative purposes only and may be inaccurate in subtleties. The upper panel shows how the X-ray and optical light curves of a typical quasar might look like, as well as the time intervals in which we measure the X-ray and optical fluxes and then infer the corresponding X-ray and UV luminosities using the *SRG/eROSITA* and SDSS data. We generated the X-ray light curve using the algorithm of Timmer & König (1995), assuming a broken power-law power density spectrum with slopes of -1 and -2 and a break time scale of 50 years. It has a length of 7 years and a sampling interval of half a day. The UV light curve was obtained by smoothing the X-ray one using a moving mean with uniform weights and a window size of 1 month (all the quoted time scales being in the quasar’s rest frame).

The lower panel of Fig. 8 demonstrates the impact of variability and intrinsic scatter on the location of quasars on the

X-ray vs. UV luminosity plane. We adopt the moment of optical flux measurement (which then translates into an estimate of the UV luminosity), t_{opt} , as the instant at which we investigate the primary dependence. Note that we could just as well have chosen the time of measurement of the X-ray flux, t_{Xray} , for the same purpose. We then assume that different quasars have somewhat different intrinsic properties, so that their X-ray luminosities (l_X and l_{UV}) averaged over the same ΔT_1 interval around t_{opt} are scattered around the primary dependence, which is illustrated by the green point lying away from the corresponding black point on the primary dependence. However, the X-ray and UV luminosities that we actually infer from observation (here, we ignore the measurement uncertainties for clarity) deviate even more. First, the UV luminosity (the green diamond on the plot) is shifted due to the short-term variability, since the optical flux measured in snapshot SDSS observations fluctuates with variance σ_{varUV}^2 with respect to the ΔT_1 -averaged value at t_{opt} . Second, the X-ray luminosity, which we estimate from the *eROSITA* data (the blue point on the plot), shifts due to the long-term variability, as the X-ray flux changes with variance σ_{varX}^2 over the $\Delta T_2 > \Delta T_1$ time passed between the optical and X-ray observations. As a result, the observed (l_X , l_{UV}) pair (the red cross on the plot) lies significantly away from the primary dependence. We have also shown in the bottom panel of Fig. 8 a trajectory that a quasar could follow due to its long-term variability, which we here assumed (for simplicity) to be parallel to the primary dependence.

3.3.3. Long-term X-ray variability

Let us, for convenience, define the l_X – l_{UV} correlation at the moment of optical measurement. We then first need to estimate how the X-ray luminosity of a given quasar changed over the period of time between its observations by SDSS and *SRG/eROSITA*. Then in equation (10),

$$\sigma_{\text{varX}}^2 = \widetilde{SF}_X^2(\Delta T_1, \Delta T_2), \quad (25)$$

(using the definition of the $\widetilde{SF}^2$ in eq. 23). Note that $\widetilde{SF}_X^2$ defined here for the monochromatic luminosity at the rest-frame energy of 2 keV ($L_{2\text{keV}}$) is the same as $\widetilde{SF}_X^2$ for the measured flux in the 0.3–2.3 keV band, because these quantities are related by a fixed spectral model (power-law with $\Gamma = 2$), so that the difference of the corresponding logarithms is the same.

We have recently conducted a systematic analysis of X-ray variability of quasars (Prokhorenko et al., 2024) using a similar, but smaller sample of SDSS quasars detected by *SRG/eROSITA* (and also by *XMM-Newton* for a small part of the sample). The results were presented in terms of $\widetilde{SF}_X^2(\sim 1 \text{ day}, \Delta T_2)$, with ΔT_2 ranging up to ~ 10 years in the rest frame. We can apply these results to the current analysis using the following relation:

$$\widetilde{SF}_X^2(\Delta T_1, \Delta T_2) = \widetilde{SF}_X^2(\sim 1 \text{ day}, \Delta T_2) - 2 \text{Var}_X(\sim 1 \text{ day}, \Delta T_1), \quad (26)$$

which is derived in Appendix C. Here, $\text{Var}(\Delta t_1, \Delta t_2)$ denotes the variance of luminosity evaluated in Δt_1 intervals with respect to the average luminosity over a longer period $\Delta t_2 > \Delta t_1$.

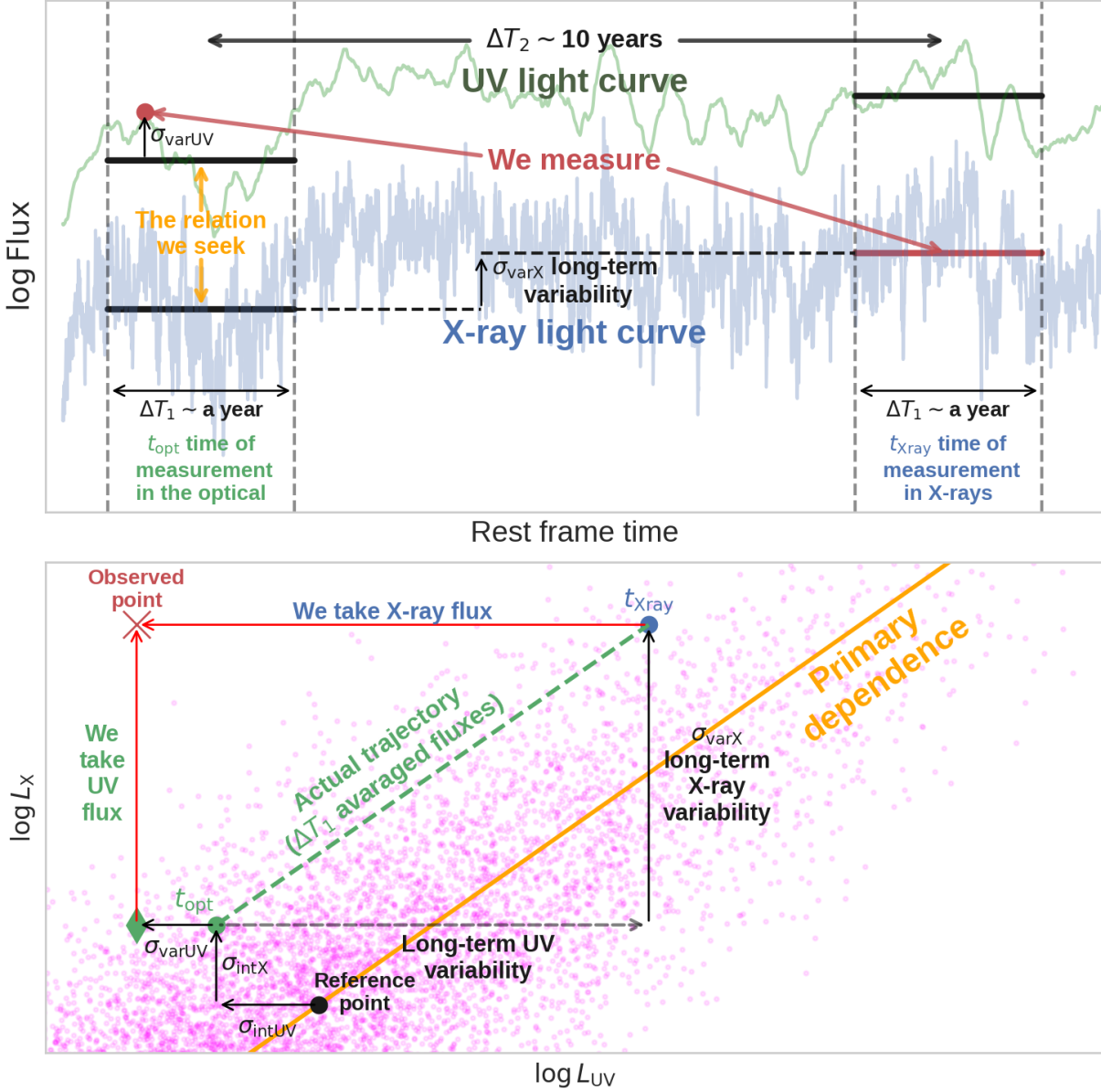


Figure 8: Illustration of the key aspects of this study. The top panel depicts X-ray (blue) and optical-UV (green) simulated light curves of a quasar and demonstrates what information we can extract from such data. The bottom panel demonstrates the impact of variability and intrinsic scatter on the location of quasars on the X-ray vs. UV luminosity plane. The orange solid line indicates the primary dependence, whose parameters we wish to evaluate. Intrinsic scatter within the quasar population both in X-ray and UV luminosity shifts individual objects away from this dependence, as shown by the black arrows, while variability imposes additional shifts. If X-ray and UV measurements could be done simultaneously and were averaged over sufficiently long intervals (of length ΔT_1), a quasar would travel with time, as a result of variability, along the trajectory shown with the green dashed line. In reality, the X-ray and UV measurements are taken at different instants and the UV (snapshot) measurement is additionally affected by short-term (within ΔT_1) variability. The numerous translucent magenta dots show the intrinsic distribution of the quasar population on this plane at some certain moment in time (taking into account the luminosity function).

Prokhorenko et al. (2024) demonstrated that the amplitude of X-ray variability on a given time scale is anti-correlated with the black hole mass and accretion rate (Eddington ratio), and with luminosity. Similar trends have been found for optical variability (see e.g. Arévalo et al. 2023). We adopt from Prokhorenko et al. (2024) the following power-law approximation for the dependence of the SF on time lag: $\widetilde{SF}_X^2(\sim 1 \text{ day}, \Delta T_2) = A^2(\Delta T_2/\text{year})^\beta$, with the normalization and slope depending on luminosity: $A = -0.111 \log \frac{L_{2-10}}{\text{erg s}^{-1}} + 5.145$, $\beta =$

$-0.127 \log \frac{L_{2-10}}{\text{erg s}^{-1}} + 5.828$, where L_{2-10} is the luminosity in the 2–10 keV rest-frame energy band. Following Prokhorenko et al. (2024), we estimate L_{2-10} for our quasars from their unabsorbed fluxes in the observed 0.3–2.3 keV band assuming a power-law spectrum with $\Gamma = 1.8$ (which is a typical spectral slope for AGN in the 2–10 keV energy band).

Therefore, we use for each quasar its specific values of $\Delta T_{1,i}$ and $\Delta T_{2,i}$ and apply the above SF model to calculate $\widetilde{SF}_{X,i}^2(\sim 1 \text{ day}, \Delta T_{2,i})$ and $\text{Var}_{X,i}(\sim 1 \text{ day}, \Delta T_{1,i})$ in equation (26) (see Ap-

pendix C and, in particular, equation C.10). The second right panel of Fig. 7 shows the resulting distribution of σ_{varX}^2 for our sample. The median value of σ_{varX}^2 is 0.020.

3.3.4. Short-term UV variability

We should now recall that our UV luminosity estimates (l_{UV}) are based on snapshot SDSS flux measurements, which are affected by short-term UV variability that is uncorrelated with that in X-rays. Therefore, in accordance with our definition of the X-ray–UV luminosity correlation, we must take into account this short-term UV variability. Namely, we need to know the expected scatter of ~ 0.5 -minute UV luminosity measurements within the ΔT_1 (~ 1 year) interval, that is, in equation (10),

$$\sigma_{\text{varUV}}^2 = \text{Var}_{\text{UV}}(\sim 0.5 \text{ minute}, \Delta T_1). \quad (27)$$

The variance $\text{Var}(\Delta t_1, \Delta t_2)$ can be unambiguously calculated from $\widetilde{SF}^2(\Delta t_1, \Delta t_2)$ (see Appendix C for details).

To estimate the UV variability, we use the best-fit parameters of the multidimensional power-law model from the work by Li et al. (2018), who explored the relationship between the SF and z , bolometric luminosity, rest-frame wavelength and black hole mass based on a sample of 119,305 SDSS and Dark Energy Camera Legacy Survey (DECaLS) quasars. For simplicity, we ignore all dependencies except the one on wavelength (we adopted a value of 2500 Å). Specifically (see eq. (8) in Li et al. 2018), we assume that $\log SF_{\text{UV}}^2(\tau) = \log 0.076 - 0.551 \log(2500 \text{ Å}/10000 \text{ Å}) + [0.487 + 0.172 \log(2500 \text{ Å}/10000 \text{ Å})] \log \tau - \log 2.5^2$, where the time lag τ is measured in years and we have added the last term to convert the SF from Li et al. (2018) defined in terms of mag^2 to our SF defined in terms of dex^2 . We ignore the differences between the exposures of SDSS spectroscopy and SDSS and DECaLS photometry. Given the power-law increase of SF^2 with time scale and the fact that we consider SF^2 on the scale of a year, whereas the characteristic exposure times are days or hours, this will be a minor correction.

The resulting distribution of σ_{varUV}^2 for the *SRG*/eROSITA–SDSS sample is shown in the rightmost panel of Fig. 7. The median $\sigma_{\text{varUV}}^2 = 0.006$. The contribution of short-term UV variability is small compared to that of long-term X-ray variability (σ_{varX}^2). Hence, short-term (on time scales less than ~ 1 year) optical–UV variability of quasars plays a minor role in shaping the observed X-ray–UV luminosity relation.

4. Results

We applied our method to the *SRG*/eROSITA–SDSS quasar sample. We adopted the mean values and standard deviations of the LADE model of Aird et al. (2015) (their Table 5) based on their ‘soft band’ (0.5–2 keV) selected AGN sample⁴ as Gaussian priors on the parameters of $\widetilde{NLF}$, except for g_2 and l^* ,

Parameter	Prior type	Prior specification
γ	Uniform	from 0 to 3
β	Uniform	from $-\infty$ to $+\infty$
$\log \sigma_{\text{intUV}}^2$	Uniform	from $-\infty$ to $+\infty$
$\log \sigma_{\text{intX}}^2$	Uniform	from $-\infty$ to $+\infty$
g_1	Normal	mean 0.44, std 0.02
g_2	Normal	mean 3.00, std 2.00
l^*	Normal	mean 44.93, std 1.00
p_1	Normal	mean 3.39, std 0.08
z_c	Normal	mean 2.31, std 0.07

Table 2: Priors for the model parameters.

for which we adopted much larger standard deviations. This is done because, as was already discussed in Section 3.2, the primary LF of our interest is expected to have somewhat different high-luminosity slope and bend luminosity compared to the observed LF. We used broad uniform priors for the parameters of the the X-ray–UV luminosity correlation (uniform in logarithm scale for the intrinsic scatter parameters). All the priors are listed in Table 2. To facilitate comparison of our results for the LF with those of previous studies, in particular with Aird et al. (2015), we define the bend luminosity (see eq. 20) in the rest-frame 2–10 keV energy band, namely use the definition

$$l^* = \log \frac{L^*}{\text{erg s}^{-1} \text{Hz}^{-1}} + 17.90, \quad (28)$$

where L^* is the corresponding monochromatic luminosity (per frequency) at 2 keV and the constant term corresponds to a power-law spectrum with $\Gamma = 2$.

4.1. Baseline model

We first consider our ‘baseline’ model, in which there is independent intrinsic scatter in UV and X-ray luminosities (σ_{intUV} , σ_{intX}). Figure 9 shows the scatter plot of these luminosities for our sample and the inferred correlation between them. The posterior distributions of the model parameters are depicted in Fig. 10, and their key characteristics (mean and median values, standard deviations and percentiles) are presented in Table 3. The primary dependence shown in Fig. 9 corresponds to the median γ and β values of the posterior distributions of these parameters, whereas the green and yellow regions around the primary dependence comprise, respectively, $[l_{\text{UV}}, l_{\text{X}}]$ points with $\tilde{l}_{\text{X}}$ lying within the 1σ and 2σ credible intervals (i.e. 16–84% and 2.5–97.5%) of the $\gamma \tilde{l}_{\text{UV}} + \beta$ distribution for a given $\tilde{l}_{\text{UV}}$. Namely, for each $\tilde{l}_{\text{UV}}$, we drew a sample of $[\gamma, \beta]$ pairs from their joint posterior distribution, calculated $\tilde{l}_{\text{X}} = \gamma \tilde{l}_{\text{UV}} + \beta$ and determined the 1σ and 2σ credible intervals for $\tilde{l}_{\text{X}}$ based on the resulting $\tilde{l}_{\text{X}}$ distribution.

We tightly constrain the slope of the $l_{\text{X}}-l_{\text{UV}}$ relation: $\gamma = 0.69 \pm 0.02$ (hereafter, we quote for our results the median values of parameters of interest and the corresponding 1σ uncertainties based on the 16–84% percentile region of the posterior distribution). This value is slightly higher than the slope $\gamma = 0.60-0.65$ reported by Lusso & Risaliti (2016) and significantly steeper

⁴Although Aird et al. (2015) did not explicitly exclude type II AGN from their study, as we do, their soft-band sample is expected to be strongly dominated by type I AGN.

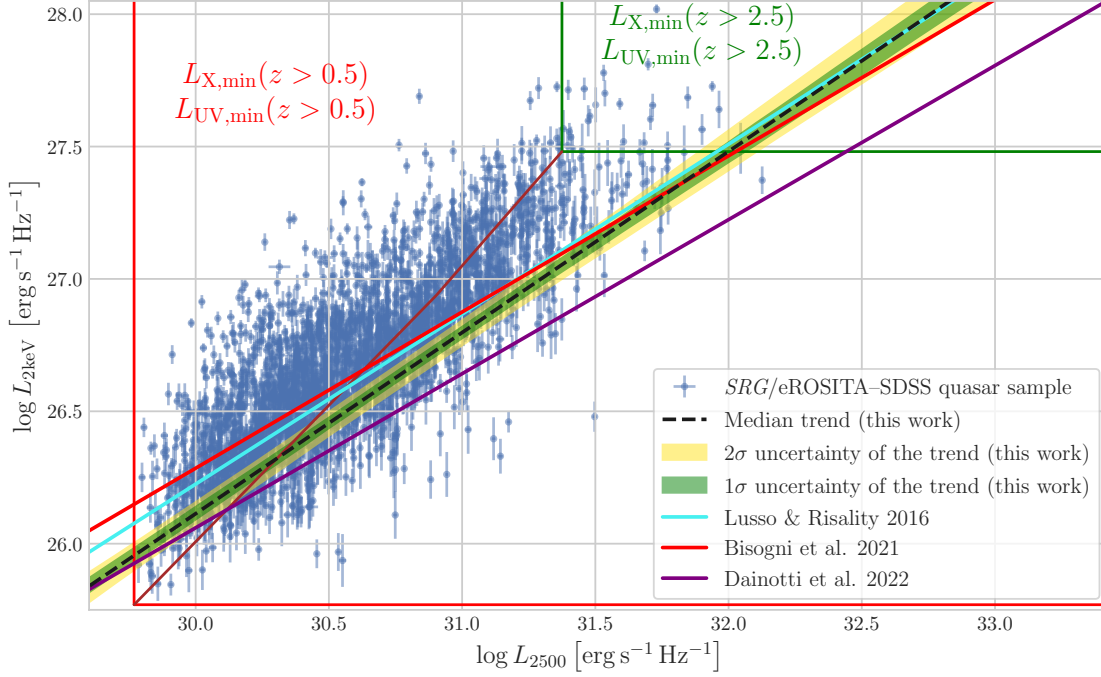


Figure 9: Scatter plot of the X-ray and UV luminosities with their measurement uncertainties for the *SRG*/eROSITA–SDSS quasar sample (points with error bars) and the inferred correlation between these quantities for the baseline model: the dashed line corresponds to the median γ and β values from the posterior distributions and the green and yellow regions show the corresponding 1σ and 2σ uncertainty. For comparison, we show (without uncertainties) the l_X – l_{UV} relations obtained in previous studies, as indicated in the legend. We also show the regions in the X-ray–UV luminosity space where the current study probes quasars at $z > 0.5$ and at $z > 2.5$ (the selection region moves along the brown solid line with increasing redshift).

than the slopes $\gamma = 0.59 \pm 0.04$ and $\gamma = 0.583 \pm 0.011$ determined by Bisogni et al. (2021) and Dainotti et al. (2022), respectively. As regards the normalization of the l_X – l_{UV} relation, we find $\tilde{l}_X(\tilde{l}_{UV} = 30.5) = 26.45 \pm 0.02$, which is somewhat lower than the corresponding values of 26.56 ± 0.02 in Lusso & Risaliti (2016) and 26.58 ± 0.04 in Bisogni et al. (2021) but higher than 26.35 ± 0.03 in Dainotti et al. (2022). This can also be clearly seen from a direct comparison of the X-ray–UV luminosity dependencies in Fig. 9. Hereafter, we utilize the normalization at $\tilde{l}_{UV} = 30.5$, because it is a typical luminosity probed in the current study (the median l_{UV} value in our sample is 30.537) and in the quoted works.

As a byproduct of our analysis, we have also determined the parameters of the quasar primary luminosity function $\tilde{LF}$. As can be seen from Fig. 10 and Table 3, the derived $\tilde{LF}$ parameters are in reasonable agreement with previous determinations of the AGN X-ray LF (Ueda et al., 2003; Hasinger et al., 2005; Aird et al., 2015; Ueda et al., 2014), which adds credibility to our results for the l_X – l_{UV} correlation (note that all these studies used the same cosmological parameters as ours). The only exception is the slope of the LF bright end, which we find to be significantly steeper ($g_2 = 4.4 \pm 0.2$) than in previous works: 2.57 ± 0.16 (Hasinger et al., 2005), 2.71 ± 0.09 (Ueda et al., 2014) and 2.55 ± 0.05 (Aird et al., 2015). (Here, we quote

the values corresponding to the luminosity-dependent density evolution models in these studies, since the corresponding samples reached into high redshifts, where the LADE model becomes inadequate.) However, this difference is actually expected, since, as discussed above, the observed LF reported by these authors results from convolution of the $\tilde{LF}$ determined here with the intrinsic and variability related scatter of the l_X – l_{UV} correlation, which increases the number of quasars with high observed luminosities. We prove this explicitly in Appendix B.

Importantly, we find that σ_{intX}^2 provides a substantial contribution [more than (long-term) X-ray variability and much more than X-ray flux measurement uncertainties] to the total observed scatter. In contrast, σ_{intUV}^2 just marginally exceeds zero and its contribution to the total observed scatter is small [as are the contributions of (short-term) UV variability and UV flux measurement uncertainties]. Note, however, that these intrinsic scatter parameters significantly anti-correlate with each other and correlate or anti-correlate with the slope and normalizing constant of the l_X – l_{UV} relation.

4.2. Intrinsic scatter models

Essentially all previous studies (Lusso & Risaliti, 2016; Bisogni et al., 2021; Dainotti et al., 2022) ascribed the en-

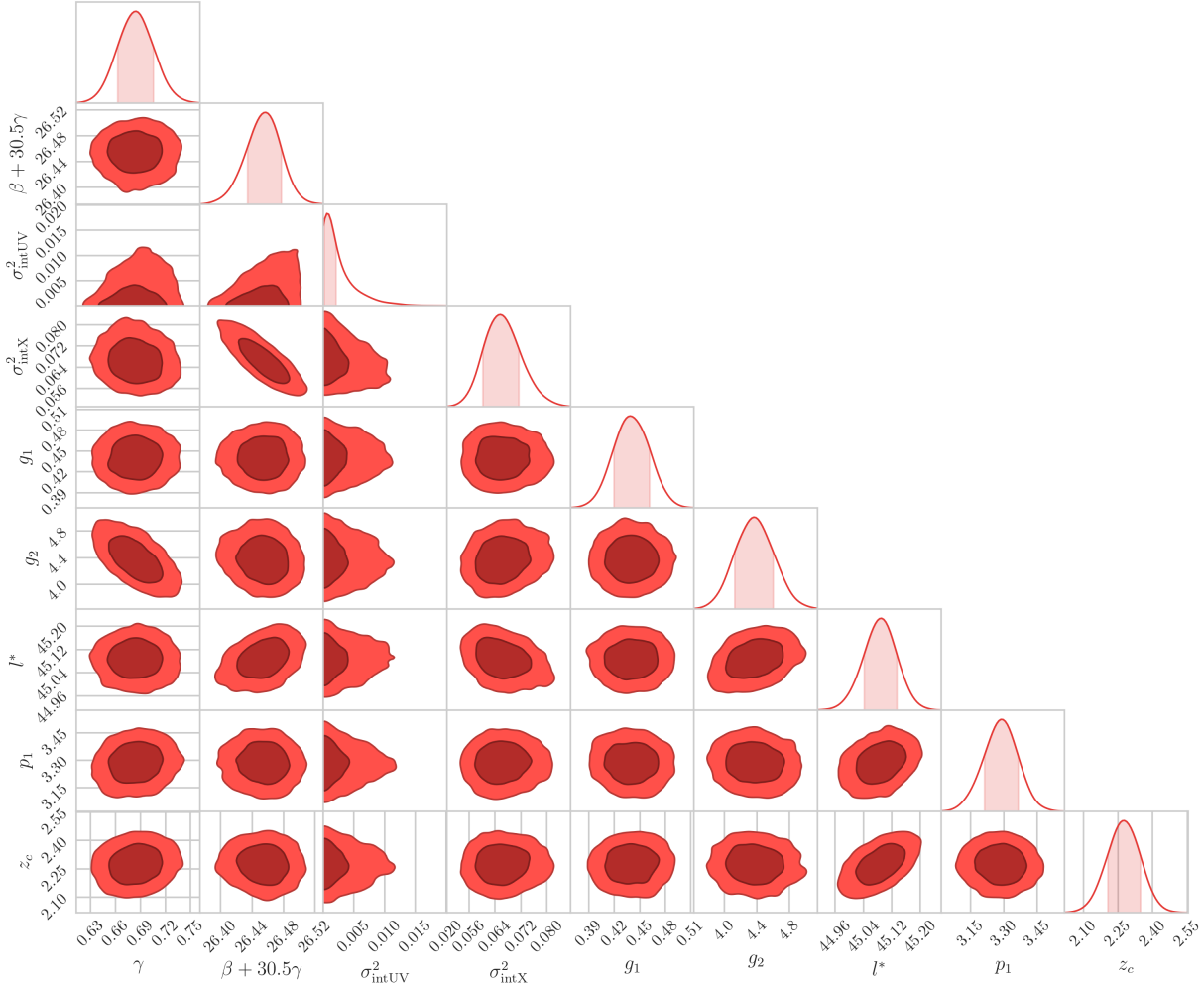


Figure 10: Posterior distributions of model parameters for the baseline model, including the parameters of the l_X – l_{UV} relation and of the LF. The shaded areas of the 1D distributions mark the 68% smallest credible intervals. The dark and light red areas of the 2D distributions correspond to the 68% and 95% smallest credible regions, respectively.

ture intrinsic scatter to the X-ray luminosity, i.e. assumed that $\sigma_{\text{intUV}} = 0$. If we make the same assumption and repeat our analysis, we obtain a slope, $\gamma = 0.69 \pm 0.02$, and a normalization, $\beta + 30.5\gamma = 26.45 \pm 0.02$, that are nearly the same as for the baseline model. This is expected, since in that case we obtained $\sigma_{\text{intUV}} \approx 0$. The results of this test are presented in the left panel of Fig. 11 and in the middle section of Table 3.

Alternatively, we can ascribe the entire intrinsic scatter to the UV axis, i.e. assume $\sigma_{\text{intX}} = 0$. In this case (see the right panel of Fig. 11 and the bottom section of Table 3), we find nearly the same slope, $\gamma = 0.71 \pm 0.03$, but a much higher normalization, $\beta + 30.5\gamma = 26.95 \pm 0.03$, compared to the baseline model. Therefore, the trend line in the $\sigma_{\text{intX}} = 0$ case is nearly parallel to that in the $\sigma_{\text{intUV}} = 0$ case but shifted upwards. Moreover, the ratio of $\sigma_{\text{intX}}(\sigma_{\text{intUV}} = 0)$ to $\sigma_{\text{intUV}}(\sigma_{\text{intX}} = 0)$ is equal to γ within the uncertainties.

This behaviour may look surprising at first glance but finds a natural explanation. As demonstrated in Appendix D, if

the quasars in the Universe obeyed a power-law primary LF, $NLF(\tilde{l}_X) \propto \tilde{l}_X^{-\alpha}$, with scatter either only along the X-ray or only along the UV axis, the probability density distribution of quasars on the l_X vs. l_{UV} plane, $f(l_X, l_{UV} | \sigma_{UV}, \sigma_X, \gamma, \beta)$, could be equivalently described by two solutions: one with $\sigma_{UV} = 0$ and another with $\sigma_X = 0$, with the trend line of the latter being translated by $\alpha\sigma_X^2(\sigma_{\text{intUV}} = 0) \ln 10$ along the X-ray axis from the trend line of the former and with $\sigma_X(\sigma_{UV} = 0)/\sigma_{UV}(\sigma_X = 0) = \gamma$. This behaviour, i.e. the existence of two alternative solutions, could be revealed directly, namely by linear regression $l_X(l_{UV})$ and $l_{UV}(l_X)$, if one had a volume limited sample of quasars; or, for a flux limited sample, using the maximum likelihood or MCMC based approach developed in our work, which properly takes into account the biases associated with the sample's X-ray and UV flux limits.

In reality, of course, we are dealing with a more complicated LF, namely with a double power-law model for $\tilde{LF}$, hence the $\sigma_{\text{intUV}} = 0$ and $\sigma_{\text{intX}} = 0$ solutions are not expected to be par-

Parameter	Mean value and standard deviation	2.5% percentile	16% percentile	median	84% percentile	97.5% percentile
Baseline model						
γ	0.686, 0.021	0.64	0.66	0.68	0.71	0.73
$\beta + 30.5\gamma$	26.453, 0.023	26.41	26.43	26.45	26.47	26.49
σ_{intUV}^2	0.002, 0.003	0.000	0.000	0.001	0.004	0.010
σ_{intX}^2	0.067, 0.006	0.056	0.061	0.066	0.072	0.079
g_1	0.438, 0.019	0.40	0.42	0.44	0.46	0.48
g_2	4.37, 0.23	3.9	4.1	4.4	4.6	4.8
l^*	45.08, 0.04	44.99	45.04	45.09	45.13	45.17
p_1	3.28, 0.07	3.14	3.21	3.28	3.34	3.43
z_c	2.27, 0.07	2.14	2.21	2.27	2.34	2.41
Assuming $\sigma_{\text{intUV}} = 0$						
γ	0.687, 0.021	0.65	0.67	0.69	0.71	0.73
$\beta + 30.5\gamma$	26.453, 0.019	26.41	26.43	26.45	26.47	26.49
σ_{intX}^2	0.067, 0.005	0.058	0.062	0.067	0.073	0.078
g_1	0.439, 0.021	0.40	0.42	0.44	0.46	0.48
g_2	4.32, 0.21	4.0	4.1	4.3	4.5	4.7
l^*	45.08, 0.04	45.00	45.04	45.09	45.14	45.18
p_1	3.28, 0.07	3.14	3.21	3.29	3.37	3.44
z_c	2.28, 0.07	2.14	2.20	2.28	2.35	2.42
Assuming $\sigma_{\text{intX}} = 0$						
γ	0.708, 0.027	0.66	0.68	0.71	0.74	0.76
$\beta + 30.5\gamma$	26.95, 0.03	26.89	26.92	26.95	26.98	27.02
σ_{intUV}^2	0.136, 0.014	0.11	0.12	0.14	0.15	0.17
g_1	0.439, 0.021	0.40	0.42	0.44	0.46	0.48
g_2	3.272, 0.12	3.1	3.2	3.3	3.4	3.5
l^*	44.0, 0.6	42.6	43.4	44.1	44.7	44.9
p_1	3.38, 0.08	3.21	3.30	3.39	3.47	3.54
z_c	2.31, 0.07	2.17	2.24	2.31	2.37	2.44

Table 3: Characteristics of the posterior distributions of the model parameters.

Model	k	$-2 \ln \mathcal{L}_{\text{max}}$	AIC
Both σ_{intX} and σ_{intUV} are free	9	-4231.01	-4213.01
Assuming $\sigma_{\text{intUV}} = 0$	8	-4229.98	-4213.98
Assuming $\sigma_{\text{intX}} = 0$	8	-4159.44	-4143.44

Table 4: Comparison of the different intrinsic scatter models.

allel to each other anymore. However, to a first approximation (also ignoring the influence of an additional scatter induced by flux measurement uncertainties and variability) we may still expect the $\sigma_{\text{intUV}} = 0$ and $\sigma_{\text{intX}} = 0$ solutions to be shifted from each other by an amount that is determined by an ‘effective’ slope of the LF. The measured shift $\beta(\sigma_{\text{intX}} = 0) - \beta(\sigma_{\text{intUV}} = 0) \approx 0.5$ falls between the values of 0.07 and 0.66 expected in the single power-law LF scenario for our inferred $\widetilde{LF}$ slopes $g_1 = 0.44$ and $g_2 = 4.3$ below and above l^* , respectively. We thus see a reasonable agreement.

We can formally compare the results obtained for the baseline model and for the $\sigma_{\text{intUV}} = 0$ and $\sigma_{\text{intX}} = 0$ models using

the Akaike information criterion (AIC):

$$AIC = 2k - 2 \ln \mathcal{L}_{\text{max}}, \quad (29)$$

where k is the number of estimated parameters in a given model and $\mathcal{L}_{\text{max}}$ is the maximum likelihood value, as defined in equation (18). Thus, here we use the standard maximum likelihood estimation method instead of the Bayesian approach. The $\sigma_{\text{intX}} = 0$ model provides a significantly worse description of the data compared to the baseline model (see Table 4). The $\sigma_{\text{intUV}} = 0$ model is as good as the baseline model in terms of the AIC.

Our derived value for intrinsic dispersion $\sigma_{\text{intX}}^2 = 0.066 \pm 0.005$ corresponds to a standard deviation of ~ 0.25 dex. This is somewhat larger than ~ 0.2 dex estimated by some previous authors (Lusso & Risaliti, 2016; Chiaraluce et al., 2018; Bisogni et al., 2021; Dainotti et al., 2022). The main reason for this difference is likely the following: σ_{intX}^2 in our model is defined as the *true* dispersion of the two-dimensional luminosity function [σ_X^2 from equation (14) minus the contribution of variability and measurement uncertainties], rather than as the *observed* dispersion (minus the contribution of variability and measurement uncertainties) of a given sample as in the previ-

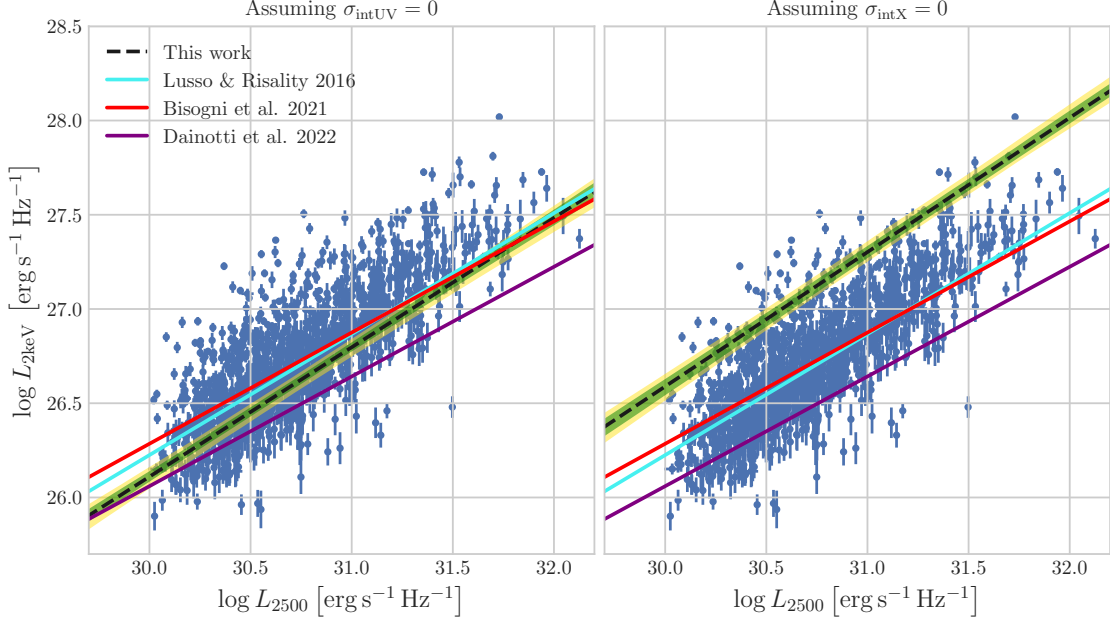


Figure 11: Same as Fig. 9, but for the $\sigma_{\text{intUV}} = 0$ and $\sigma_{\text{intX}} = 0$ models.

ous studies. As a result, our σ_{intX}^2 is expected to be larger because part of the intrinsic distribution of objects is cut off by flux selection in our and other quasar samples. Note that the *observed* total dispersion along the X-ray axis, σ_{obsX}^2 , minus the mean σ_{varX}^2 and mean σ_{mX}^2 for the *SRG/eROSITA*-SDSS quasar sample is equal to 0.045 (equivalent to a standard deviation of ~ 0.2 dex), similar to the previous studies based on other quasar samples. Another, perhaps less important, reason for the difference may be related to our sample containing a larger fraction of rare high-luminosity quasars for a given redshift (due to the larger sky coverage by *SRG/eROSITA*) or exclusion of quasars with inferred significant intrinsic absorption from the previous analyses mentioned above (unlike in our study).

4.3. Comparison with other methods

As we noted before, in most previous studies (e.g. Bisogni et al. 2021; Salvestrini et al. 2019; Dainotti et al. 2022, 2024), the authors regarded l_X as the “dependent” variable, i.e. examined the dependence $l_X(l_{\text{UV}})$. One then needs to minimize a likelihood of the form

$$\mathcal{L} = \prod_i \exp \left\{ -[l_X - \mathcal{F}(l_{\text{UV}})]^2 / 2\delta_X^2 \right\}, \quad (30)$$

where $\mathcal{F}$ is an arbitrary (usually assumed to be linear) function and δ_X^2 accounts for the scatter along the X-ray luminosity axis (it can include intrinsic scatter, variability and flux measurement uncertainties). One could also consider the opposite case of UV luminosity being the dependent variable, i.e. $l_{\text{UV}}(l_X)$,

and then

$$\mathcal{L} = \prod_i \exp \left\{ -[l_{\text{UV}} - \mathcal{F}(l_X)]^2 / 2\delta_{\text{UV}}^2 \right\}, \quad (31)$$

where δ_{UV}^2 defines the scatter along the UV luminosity axis. Hence, the scatter in this approach, first described by D’Agostini (2005) (we thus refer to it as the D’Agostini method below), pertains fully to either X-ray or UV luminosity.

Other authors, in particular Lusso & Risaliti (2016); Bisogni et al. (2021), used the LINMIXERR method following Kelly (2007), regarding l_X as the “dependent” variable. This is a linear regression analysis based on the assumptions that (i) the measurement uncertainties are Gaussian with zero mean and known variance, (ii) the intrinsic scatter in the dependent variable about the regression line is Gaussian, and (iii) the intrinsic distribution of the independent variable can be approximated as a mixture of Gaussian functions. The method allows for heteroscedastic and possibly correlated measurement uncertainties.

We have tried to apply these two methods to our *SRG/eROSITA*-SDSS quasar sample. Specifically, for the D’Agostini method, we used its variant described by Dainotti et al. (2022). Figure 12 compares the results of these analyses, considering $l_X(l_{\text{UV}})$ and alternatively $l_{\text{UV}}(l_X)$, with the results obtained above by our own method (for the baseline model, Sect. 4.1). It turns out that if the X-ray luminosity is assumed to be the dependent variable, then both the D’Agostini and LINMIXERR methods yield nearly the same slope 0.694 ± 0.010 and 0.695 ± 0.010 , but significantly higher normalization of

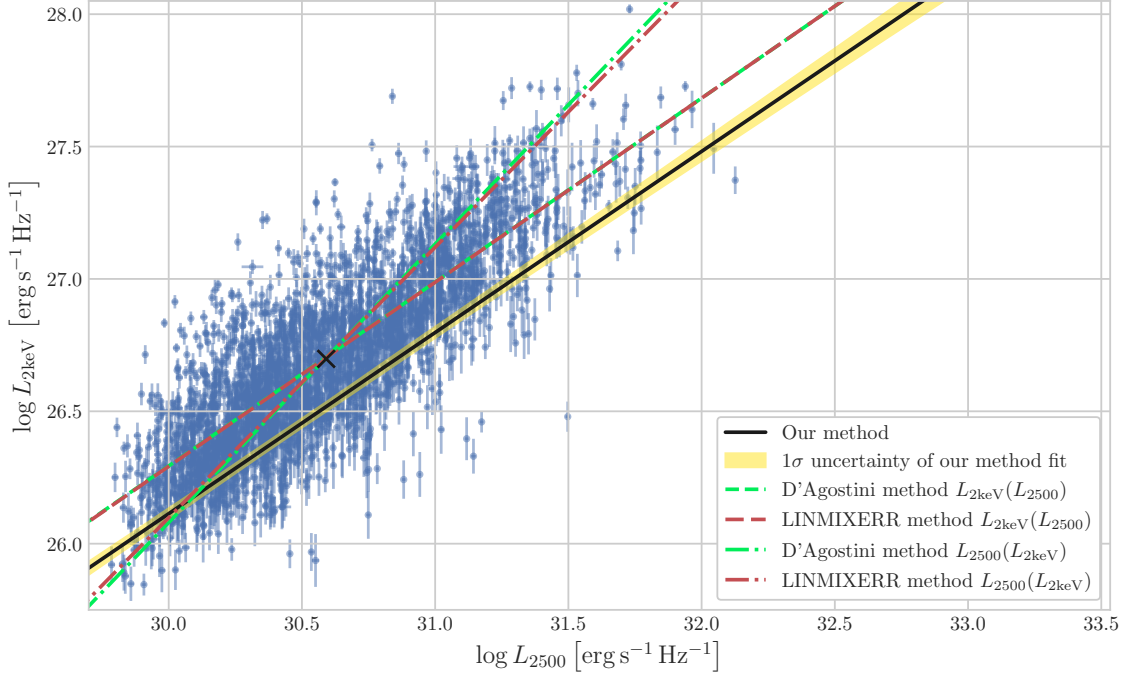


Figure 12: Comparison of our method with methods from previous works applied to our sample. The solid line corresponds to the median γ and β values from the posterior distributions and the yellow region shows the corresponding 1σ uncertainty for our baseline model. We also show (without uncertainties) the l_X – l_{UV} relations for our sample obtained using the D’Agostini and LINMIXERR methods assuming $L_{2keV}(L_{2500})$ or $L_{2500}(L_{2keV})$ dependencies, as indicated in the legend. The black cross marks the mean values of l_X and l_{UV} for *SRG/eROSITA*–SDSS quasar sample.

the correlation compared to our method. The relations obtained by these methods pass through the center of the cloud of data points, as expected. In contrast, our method is not tied to the observed two-dimensional luminosity distribution (which strongly depends on the flux limits adopted for a given quasar sample). However, if the opposite assumption is made, i.e. the dependence $l_{UV}(l_X)$ is considered, then both the D’Agostini and LINMIXERR methods result in a correlation that is very different from both the result obtained assuming the $l_{UV}(l_X)$ dependence and that obtained by our own method: namely it has a slope of 0.99–1.04. This clearly demonstrates the problem of choosing the ‘right’ dependent variable, which is inherent in the D’Agostini and LINMIXERR methods but is irrelevant for our method.

Why do then our results show such a good agreement with those of Lusso & Risaliti (2016), although their analysis was affected by the choice of the dependent luminosity (namely, they adopted l_X to depend on l_{UV})? Apparently, this happened because they used a sample with a deep X-ray flux limit, $\sim 10^{-15}$ erg s $^{-1}$ cm $^{-2}$ in the 0.5–2 keV band. As a result, their sample contained nearly the same number of quasars below and above the true $l_X(l_{UV})$ dependence, so that the linear regression analysis naturally led to a nearly unbiased result (see also a relevant discussion in Sazonov et al. 2012). The fact that we have

obtained a similar result using a sample that is characterized by a much higher X-ray flux threshold demonstrates the insensitivity of other method to the choice of sample as regards its X-ray and optical/UV flux limits.

To further demonstrate this point, we selected from our *SRG/eROSITA*–SDSS quasar sample a subsample of optically bright objects, namely quasars with $m_g < 18.5$ instead of $m_g < 19$ (retaining the same X-ray flux limit) and repeated the entire analysis. As demonstrated in Appendix E, the slope and normalization of the $l_X(l_{UV})$ dependence have remained the same within the 1σ uncertainties (which have of course increased for the smaller sample).

5. Discussion

Below, we discuss several aspects of the results of this study and what could be improved in future work. Let us recall that we have divided the observed dispersion of the l_X – l_{UV} correlation into three components: luminosity measurement uncertainties, variability and intrinsic scatter. We rely on the results of previous statistical studies of quasars in estimating the contribution of (long-term) X-ray and (short-term) optical–UV variability, whereas the intrinsic dispersion is described by a pair of free parameters, σ_{intUV} and σ_{intX} .

Since the current *SRG*/eROSITA–SDSS sample is composed of fairly bright quasars (in terms of their X-ray and optical–UV fluxes), the X-ray and UV luminosity measurement uncertainties provide minor contributions to the total observed dispersion. We can formally define the latter with respect to the primary dependence and along the X-ray axis, σ_{totX}^2 , as a sum of σ_{intX}^2 , the mean value of σ_{varX}^2 and the mean value of σ_{mX}^2 over the sample. We can similarly define the total dispersion relative to the UV axis, σ_{totUV}^2 .

We thus find that $\sigma_{\text{totX}}^2 = 0.092 \pm 0.006$ is composed of $73.4 \pm 1.8\%$ intrinsic scatter, $22.7 \pm 1.6\%$ variability (namely, long-term variability) and $3.83 \pm 0.25\%$ luminosity measurement uncertainties. For $\sigma_{\text{totUV}}^2 = 0.009 \pm 0.003$, the corresponding fractions are $32^{+16}_{-30}\%$, $67^{+31}_{-16}\%$ (namely, short-term variability) and $1.0^{+0.4}_{-0.2}\%$. Therefore, long-term (on time scales of years) variability provides a significant but subdominant contribution to the total dispersion of the l_X – l_{UV} correlation for the current sample, allowing us to obtain robust constraints on the intrinsic scatter.

As was discussed at length in Section 3.3, AGN variability (e.g., characterized by the structure function) grows with increasing time scale. The X-ray flux measurements (*SRG*/eROSITA) used in our analysis were typically performed with fairly long rest-frame lags of ~ 10 years with respect to the corresponding optical–UV observations (SDSS). If better synchronized X-ray and optical–UV observations were available for the same quasar sample, the contribution of variability to the total scatter of the l_X – l_{UV} correlation would presumably become smaller. It will be interesting to check this prediction using the data of the Dark Energy Spectroscopic Instrument (DESI) survey, which have already provided (DESI Collaboration et al., 2024; Dawson & Palanque-Delabrouille, 2025) new high-quality optical spectra for many *SRG*/eROSITA quasars, taken nearly at the same time (May 2021–June 2022) as our X-ray observations (Dec. 2019–Feb. 2022).

At the present state of research, we can only speculate about the origin of the intrinsic scatter of the l_X – l_{UV} correlation, and in particular why it is aligned with the X-ray luminosity axis. It is natural to expect that at least such key physical parameters as black hole mass and accretion rate (Eddington ratio), as well as, perhaps, spin, play important roles in shaping this correlation. It should be possible to address this question quantitatively using the same *SRG*/eROSITA–SDSS sample, as there are available single-epoch spectroscopy mass estimates for all these quasars.

Another possible source of the scatter is intrinsic absorption of X-ray and UV emission. We neglected it because our current sample is composed of quasars with broad optical emission lines, however even such objects may be affected by intrinsic absorption, originating near the SMBH and/or in the host galaxy. This, in fact, might be one of the causes of the observed dispersion of the optical–UV k -corrections (which reflects the diversity of observed spectral shapes, see the top panel of Fig. 2). In future work, we plan to address this issue in detail via a careful analysis of the optical–UV spectra of quasars using SDSS and DESI data, which will allow us to clean the sample from strongly reddened spectra and/or correct all spectra for intrinsic extinction. Unfortunately, a similar effort in the

X-ray domain is more problematic, as even our current, relatively X-ray bright quasar sample mostly consists of sources from which just a few tens of photons have been detected by eROSITA, which precludes any sensible spectral analysis.

In addition, as has been discussed in the literature (e.g., Risaliti et al. 2011; Bisogni et al. 2021; Signorini et al. 2024), inclination of the accretion disc (usually assumed to be geometrically thin and optically thick, Shakura & Sunyaev 1973) with respect to our line of sight could cause a significant scatter in the observed UV luminosity. Moreover, orientation can similarly affect the observed X-ray luminosity, since the hot corona of the accretion disc is inferred to be moderately optically thick with respect to its X-ray radiation (Thomson optical depth $\tau \sim 1$, e.g. Malizia et al. 2014; Lubiński et al. 2016; Ricci et al. 2018), and near 2 keV there can also be a significant contribution of emission from an optically thick ($\tau \sim 20$) ‘warm’ corona (e.g. Magdziarz et al. 1998; Kubota & Done 2018). It is difficult to investigate these viewing-angle driven effects, as one usually cannot reliably infer the orientation of AGN from observations other than discerning open and obscured directions (type I vs. type II AGN). In order to make progress in this direction, it might be worth using additional information, in particular in the infrared band, which in combination with the optical data could help constrain the opening angle of the obscuring ‘torus’ and the inclination of the accretion disc in a given object.

6. Summary

We have developed a statistical model of the joint X-ray–UV luminosity distribution of quasars that is based on: (i) intrinsic linear relation (‘primary dependence’) between the logarithms of these luminosities (l_X and l_{UV}), (ii) dispersion along both axes, and (iii) changing density of quasars along the primary dependence (‘primary luminosity function’). We also estimated the expected contributions of intrinsic variability (resulting from the non-simultaneity of X-ray and optical/UV observations) and flux measurement uncertainties to the total observed dispersion of the l_X – l_{UV} dependence, which allowed us to evaluate the residual (‘intrinsic’) scatter of the latter. Our method properly accounts for all relevant observational biases and for the evolution of quasar space density with redshift.

We started with a sample of quasars from the SDSS DR16Q catalogue cross-matched with the *SRG*/eROSITA all-sky survey X-ray source catalogue at spectroscopic redshifts between 0.5 and 2.5, with optical and X-ray flux limits of $m_g > 19$ and $F_X > 6 \times 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$, respectively. With this bright sample, we achieved nearly 100% statistical completeness. After application of several additional filters (removal of blazars etc.), we obtained a ‘clean’ *SRG*/eROSITA–SDSS quasar sample, containing 2414 objects. We then applied our method to this sample and obtained the following main results:

- The slope of the primary $l_X(l_{\text{UV}})$ dependence $\gamma = 0.69 \pm 0.02$ and its normalization at $l_{\text{UV}} = 30.5$ is $l_X = 26.45 \pm 0.02$.
- The X-ray intrinsic scatter ($\sigma_{\text{intX}}^2 = 0.066 \pm 0.005$) dominates over the UV intrinsic scatter ($\sigma_{\text{intUV}}^2 = 0.001^{+0.003}_{-0.001}$).

We emphasize that this is the *true* dispersion of the entire population of quasars rather than the *observed* (smaller) dispersion of a given flux limited quasar sample, which was estimated by previous authors.

The derived parameters of the X-ray–UV luminosity relation are in satisfactory agreement with the results of some (but not all) previous studies. However, we emphasize that none of these works accounted for observational selection biases as fully as has been done here, nor did they consider the dispersion of the l_X (l_{UV}) dependence for both variables at the same time. As a byproduct of our analysis, we also determined the shape of the X-ray LF of quasars, which is consistent with the results of previous studies.

In future work, we would like to apply the method developed here to other samples, in particular to DESI optical quasars cross-correlated with *SRG/eROSITA* X-ray sources. These two surveys were conducted almost simultaneously, which makes it possible to reduce the influence of the variability factor, which impairs the accuracy of our method. We also plan to investigate the influence of SMBH mass and accretion rate, as well as redshift and intrinsic absorption on the X-ray–UV luminosity relation and its dispersion.

Acknowledgements

This work is based on observations with *eROSITA* telescope onboard SRG observatory. The SRG observatory was created by Roskosmos in the interests of the Russian Academy of Sciences represented by its Space Research Institute (IKI) in the framework of the Russian Federal Space Program, with the participation of Germany. The *SRG/eROSITA* X-ray telescope was built by a consortium of German Institutes led by MPE, and supported by Deutsches Zentrum für Luft- und Raumfahrt (DLR). The SRG spacecraft was designed, built, launched and is operated by the Lavochkin Association and its subcontractors. The science data are downlinked via the Deep Space Network Antennae in Bear Lakes, Ussurijsk, and Baykonur, funded by Roskosmos.

This work also made use of data supplied by Sloan Digital Sky Survey IV, provided by the Alfred P. Sloan Foundation.

SP, SS and MG acknowledge support by grant 075-15-2024-647 from the Ministry of Science and Higher Education.

Data Availability

The SDSS data used in this article are accessible through the corresponding web pages. The X-ray data analysed in this article were used by permission of the Russian *SRG/eROSITA* consortium. The data will become publicly available as part of the corresponding *SRG/eROSITA* data release along with the appropriate calibration information.

References

Abdurro’uf et al., 2022, *The Astrophysical Journal Supplement Series*, 259, 35

Aird J., Coil A. L., Georgakakis A., Nandra K., Barro G., Pérez-González P. G., 2015, *Monthly Notices of the Royal Astronomical Society*, 451, 1892

Ajello M., et al., 2022, *The Astrophysical Journal Supplement Series*, 263, 24

Arévalo P., Uttley P., Kaspi S., Breedt E., Lira P., McHardy I. M., 2008, *MNRAS*, 389, 1479

Arévalo P., Lira P., Sánchez-Sáez P., Patel P., López-Navas E., Churazov E., Hernández-García L., 2023, *Monthly Notices of the Royal Astronomical Society*, 526, 6078

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

Avni Y., Tananbaum H., 1982, *Astrophys. J. Lett.*, 262, L17

Avni Y., Tananbaum H., 1986, *Astrophys. J.*, 305, 83

Berk D. E. V., et al., 2005, *The Astronomical Journal*, 129, 2047

Bisogni S., Lusso E., Civano F., Nardini E., Risaliti G., Elvis M., Fabbiano G., 2021, *Astronomy & Astrophysics*, 655, A109

Bolton A. S., et al., 2012, *Astron. J.*, 144, 144

Boyle B. J., Shanks T., Croom S., Smith R., Miller L., Loaring N., Heymans C., 2000, *Monthly Notices of the Royal Astronomical Society*, 317, 1014

Brightman M., et al., 2013, *MNRAS*, 433, 2485

Cackett E. M., et al., 2023, *Astrophys. J.*, 958, 195

Chen Q., Jing L., Zhu X., Fang Y., He Z., Deng Z., Xiang C., Wu J., 2025, *Search for Quasar Pairs with Gaia Astrometric Data. I. Method and Candidates* (arXiv:2504.17777), <https://arxiv.org/abs/2504.17777>

Chiaraluce E., Vagnetti F., Tombesi F., Paolillo M., 2018, *Astron Astrophys.*, 619, A95

D’Agostini G., 2005, arXiv preprint physics/0511182

DESI Collaboration et al., 2024, *Astron. J.*, 168, 58

Dainotti M. G., Bargiacchi G., Lenart A. Ł., Capozziello S., Colgáin E. Ó., Solomon R., Stojkovic D., Sheikh-Jabbari M., 2022, *The Astrophysical Journal*, 931, 106

Dainotti M. G., Lenart A., Yengejeh M. G., Chakraborty S., Fraija N., Di Valentino E., Montani G., 2024, *Physics of the Dark Universe*, 44, 101428

Dawson K., Palanque-Delabrouille N., 2025, *Journal of Cosmology and Astroparticle Physics*, 2025, 123

Dawson K. S., et al., 2013, *Astron. J.*, 145, 10

Edelson R., et al., 2019, *Astrophys. J.*, 870, 123

- Edelson R., Peterson B. M., Gelbord J., Horne K., Goad M., McHardy I., Vaughan S., Vestergaard M., 2024, arXiv e-prints, p. arXiv:2407.09445
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, Publications of the Astronomical Society of the Pacific, 125, 306
- HI4PI Collaboration et al., 2016, Astron Astrophys., 594, A116
- Hagen S., Done C., Edelson R., 2024, Monthly Notices of the Royal Astronomical Society, 530, 4850
- Hasinger G., Miyaji T., Schmidt M., 2005, Astronomy & Astrophysics, 441, 417
- Healey S. E., Romani R. W., Taylor G. B., Sadler E. M., Ricci R., Murphy T., Ulvestad J. S., Winn J. N., 2007, The Astrophysical Journal Supplement Series, 171, 61
- Hernández Santisteban J. V., et al., 2020, MNRAS, 498, 5399
- Itoh R., Utsumi Y., Inoue Y., Ohta K., Doi A., Morokuma T., Kawabata K. S., Tanaka Y. T., 2020, The Astrophysical Journal, 901, 3
- Jin C., Lusso E., Ward M., Done C., Middei R., 2023, Monthly Notices of the Royal Astronomical Society, 527, 356
- Jing L., Chen Q., Deng Z., Zhu X., Zou H., Wu J., 2025, A Quasar Pair Catalog Compiled from DESI DR1 (arXiv:2505.03103), <https://arxiv.org/abs/2505.03103>
- Kelly B. C., 2007, The Astrophysical Journal, 665, 1489
- Khadka N., Ratra B., 2020, Monthly Notices of the Royal Astronomical Society, 497, 263
- Kolodzig A., Gilfanov M., Sunyaev R., Sazonov S., Brusa M., 2013, Astron Astrophys., 558, A89
- Kubota A., Done C., 2018, MNRAS, 480, 1247
- Laor A., Fiore F., Elvis M., Wilkes B. J., McDowell J. C., 1997, Astrophys. J., 477, 93
- Lawther D., Vestergaard M., Raimundo S., Koay J. Y., Peterson B. M., Fan X., Grupe D., Mathur S., 2023, MNRAS, 519, 3903
- Lemon C., et al., 2020, Monthly Notices of the Royal Astronomical Society, 494, 3491
- Lemon C., et al., 2023, Monthly Notices of the Royal Astronomical Society, 520, 3305
- Lenart A. L., Bargiacchi G., Dainotti M. G., Nagataki S., Capozziello S., 2023, The Astrophysical Journal Supplement Series, 264, 46
- Li Z., McGreer I. D., Wu X.-B., Fan X., Yang Q., 2018, The Astrophysical Journal, 861, 6
- Li Z., Huang L., Wang J., 2022, Monthly Notices of the Royal Astronomical Society, 517, 1901
- Lubiński P., et al., 2016, MNRAS, 458, 2454
- Lusso E., Risaliti G., 2016, The Astrophysical Journal, 819, 154
- Lusso E., Risaliti G., 2017, Astron Astrophys., 602, A79
- Lusso E., et al., 2010, Astronomy & Astrophysics, 512, A34
- Lusso E., Piedipalumbo E., Risaliti G., Paolillo M., Bisogni S., Nardini E., Amati L., 2019, Astron Astrophys., 628
- Lusso E., et al., 2020, Astronomy & Astrophysics, 642, A150
- Lyke B. W., et al., 2020, The Astrophysical Journal Supplement Series, 250, 8
- Lyubarskii Y. E., 1997, MNRAS, 292, 679
- MacLeod C. L., et al., 2010, The Astrophysical Journal, 721, 1014
- Magdziarz P., Blaes O. M., Zdziarski A. A., Johnson W. N., Smith D. A., 1998, MNRAS, 301, 179
- Makarov V. V., Secrest N. J., 2022, The Astrophysical Journal Supplement Series, 264, 4
- Malizia A., Molina M., Bassani L., Stephen J. B., Bazzano A., Ubertini P., Bird A. J., 2014, Astrophys. J. Lett., 782, L25
- Massaro E., Maselli A., Leto C., Marchegiani P., Perri M., Giommi P., Piranomonte S., 2015, Astrophysics and Space Science, 357, 1
- Middei R., Vagnetti F., Bianchi S., La Franca F., Paolillo M., Ursini F., 2017, Astronomy & Astrophysics, 599, A82
- Morganson E., et al., 2014, The Astrophysical Journal, 784, 92
- Myers A. D., et al., 2015, The Astrophysical Journal Supplement Series, 221, 27
- Padovani P., et al., 2017, Astron. Astrophys. Rev., 25, 2
- Palanque-Delabrouille N., et al., 2013, Astronomy & Astrophysics, 551, A29
- Predehl P., et al., 2021, Astron Astrophys., 647, A1
- Prokhorenko S., et al., 2024, Monthly Notices of the Royal Astronomical Society, 528, 5972
- Rankine A. L., Aird J., Ruiz A., Georgakakis A., 2024, Monthly Notices of the Royal Astronomical Society, 527, 9004
- Ricci C., et al., 2017, Astrophys. J. Suppl. Ser., 233, 17
- Ricci C., et al., 2018, MNRAS, 480, 1819
- Richards G. T., et al., 2006, The Astronomical Journal, 131, 2766

- Risaliti G., Lusso E., 2015, *The Astrophysical Journal*, 815, 33
- Risaliti G., Salvati M., Marconi A., 2011, *Monthly Notices of the Royal Astronomical Society*, 411, 2223
- Ross N. P., et al., 2013, *The Astrophysical Journal*, 773, 14
- Sacchi A., Risaliti G., Signorini M., Nardini E., Lusso E., Trefoloni B., 2025, *Astron Astrophys.*, 703, A273
- Salvestrini F., Risaliti G., Bisogni S., Lusso E., Vignali C., 2019, *Astron Astrophys.*, 631, A120
- Sazonov S., et al., 2012, *Astrophys. J.*, 757, 181
- Schlafly E. F., Finkbeiner D. P., 2011, *The Astrophysical Journal*, 737, 103
- Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *The Astrophysical Journal*, 500, 525
- Shakura N. I., Sunyaev R. A., 1973, *Astron Astrophys.*, 24, 337
- Shen Y., et al., 2019, *The Astrophysical Journal Supplement Series*, 241, 34
- Signorini M., Risaliti G., Lusso E., Nardini E., Bargiacchi G., Sacchi A., Trefoloni B., 2024, *Astron Astrophys.*, 687, A32
- Simonetti J., Cordes J., Heeschen D., 1985, *The Astrophysical Journal*, 296, 46
- Steffen A. T., Strateva I., Brandt W., Alexander D., Koekemoer A., Lehmer B., Schneider D., Vignali C., 2006, *The Astronomical Journal*, 131, 2826
- Strateva I. V., Brandt W. N., Schneider D. P., Vanden Berk D. G., Vignali C., 2005, *Astron. J.*, 130, 387
- Sunyaev R., et al., 2021, *Astron Astrophys.*, 656, A132
- Takeuchi T. T., Sakurai A., Yuan F.-T., Buat V., Burgarella D., 2013, *Earth, Planets and Space*, 65, 281
- Tananbaum H., et al., 1979, *Astrophys. J. Lett.*, 234, L9
- Timmer J., König M., 1995, *Astron Astrophys.*, 300, 707
- Trakhtenbrot B., et al., 2017, *MNRAS*, 470, 800
- Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, *The Astrophysical Journal*, 598, 886
- Ueda Y., Akiyama M., Hasinger G., Miyaji T., Watson M. G., 2014, *The Astrophysical Journal*, 786, 104
- Vagnetti F., Turriziani S., Trevese D., 2011, *Astronomy & Astrophysics*, 536, A84
- Vagnetti F., Middei R., Antonucci M., Paolillo M., Serafinelli R., 2016, *Astronomy & Astrophysics*, 593, A55
- Vallenari A., et al., 2023, *Astronomy & Astrophysics*, 674, A1
- Vaughan S., Edelson R., Warwick R., Uttley P., 2003, *Monthly Notices of the Royal Astronomical Society*, 345, 1271
- Vignali C., Brandt W. N., Schneider D. P., 2003, *The Astronomical Journal*, 125, 433
- Wilms J., Allen A., McCray R., 2000, *Astrophys. J.*, 542, 914
- Wu Q., Shen Y., 2022, *The Astrophysical Journal Supplement Series*, 263, 42
- Wu, Q. Scialpi, M. Liao, S. Mannucci, F. Qi, Z. 2024, *Astron Astrophys.*, 692, A154
- Zamorani G., et al., 1981, *Astrophys. J.*, 245, 357

Appendix A. Completeness of the SDSS spectroscopic quasar sample

Here, we empirically investigate the spectroscopic completeness of the *SRG/eROSITA*–SDSS sample. Specifically, we are interested in knowing the completeness of SDSS spectroscopy for optical quasars with SDSS dereddened photometric magnitudes $m_g < 19$ and $m_g < 19.5$ in the redshift interval from 0.5 to 2.5. It is important to note that the SDSS photometric catalogue is highly completed down to much fainter fluxes, namely at $m_g < 23.13$ (Abdurro’uf et al., 2022).

We considered the DESI One-Percent spectroscopic survey (DESI Collaboration et al., 2024) as a highly complete reference sample of optical quasars (specifically, we use objects with SPECTYPE=QSO and ZWARN=0 from the DESI EDR SV3 spectroscopic catalogue). The One-Percent survey consists of 20 round areas (‘rosettes’) of 8.48 square degrees each, where highest spectroscopic completeness has been achieved for targets with $m_g < 23$, i.e. much deeper than our adopted brightness limits for the *SRG/eROSITA*–SDSS sample. Five of these rosettes are fully (more than 98%) covered by the *SRG/eROSITA*–SDSS footprint, and we use these regions to evaluate the spectroscopic completeness of the SDSS DR16Q catalogue at the photometric limits quoted above.

Figure A.13 shows the the number (bottom panels) and fraction (top panels) of One-Percent quasars that have also been covered by SDSS spectroscopy, in individual One-Percent rosettes. Quasars with $m_g < 19$ and $m_g < 19.5$, as defined by SDSS photometry, are presented in the left and right panels, respectively. Also shown is the SDSS spectroscopic completeness based on the sum over the five rosettes. In Fig. A.14, the same statistics is summed over all rosettes and shown as a function of redshift.

The SDSS DR16 catalogue proves to be highly complete (95% confidence intervals are (94.5, 97.8)% and (95.6, 98.6)% for the $m_g < 19$ and $m_g < 19.5$ SDSS photometric cuts, respectively) with respect to the much deeper DESI One-Percent spectroscopic survey across the redshift range of our interest.

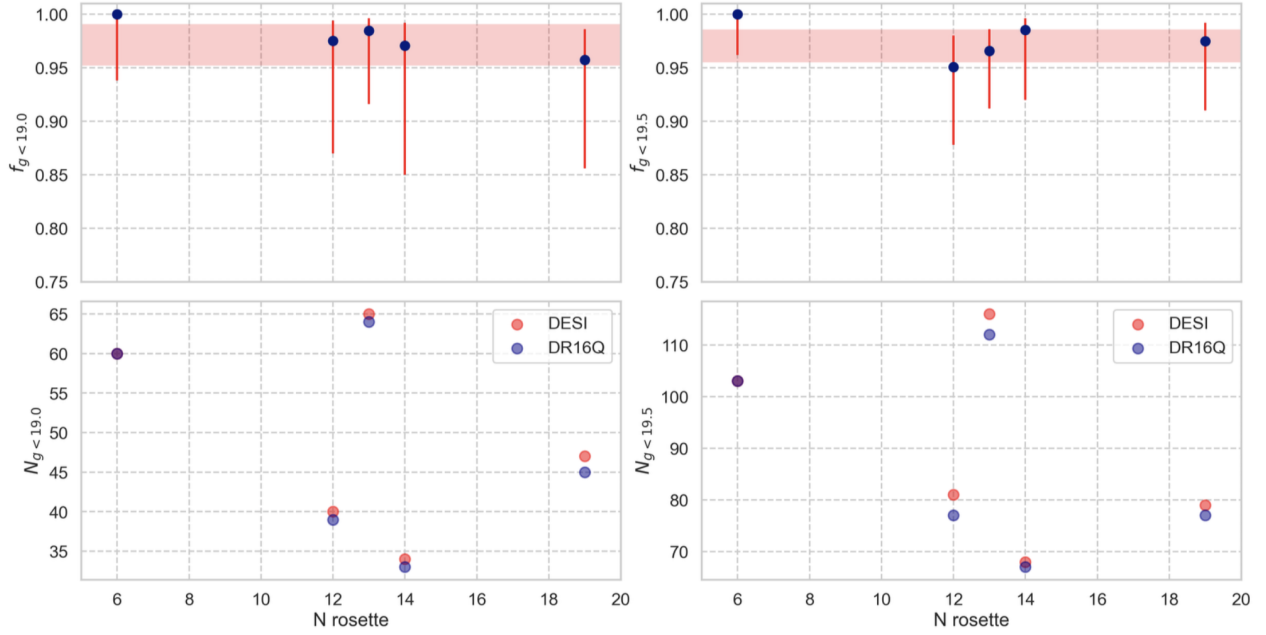


Figure A.13: SDSS DR16Q spectroscopic completeness (top panels) and number of quasars (bottom panels) in different DESI One-Percent survey rosettes within the *SRG*/eROSITA–SDSS footprint. Quasars with $m_g < 19$ and $m_g < 19.5$ are presented in the left and right panels, respectively. The error bars for the completeness measurements correspond to the 95% confidence level. The completeness 95% confidence interval based on all five rosettes is shown by the shaded area for each optical magnitude limit.

Appendix B. Relationship between the primary and observed luminosity functions

Let us demonstrate explicitly how the primary LF ($\widetilde{LF}$, or its normalized variant $\widetilde{NLF}$, which differs from the former by a constant factor, according to eq. 13) corresponds to the observed LF (NLF , which differs from LF by the same constant). For $\widetilde{NLF}$, we adopt the double-power law model (eq. 20) with parameters taken from the range allowed by our baseline l_X – l_{UV} correlation model.

For simplicity, we fix $z = 1$, which is a typical redshift for our *SRG*/eROSITA–SDSS quasar sample. Using equation (19), we calculate NLF from the assumed $\widetilde{NLF}$ performing the convolution with $\mathcal{N}(l_X|\widetilde{l}_X, \sqrt{\sigma_{\text{intX}}^2 + \sigma_{\text{varX}}^2})$. Here, σ_{varX}^2 depends on l_X as described in Section 3.3.3, particularly by equation (26), and we use the median ΔT_1 and ΔT_2 values for our sample. Then we fit the resulting dependence by the double-power law model, with its parameters (g_1 , g_2 and l_b) being free, by minimizing deviations between the model and LF (using the least-squares method). Here, $l_b = l^* + \log e(1)$, with $e(z)$ and l^* having been defined in equation (21) and equation (28), respectively. To this end, we use 50 X-ray luminosity (l_X) bins spaced uniformly over the interval 24.5–28, which is broader than the range effectively probed by our data at $z \lesssim 1$ (see Fig. 6).

First, we fixed the parameters of $\widetilde{LF}$ at the 0.5-quantile (median) values of our baseline model (listed in the top section of Table 3). The resulting $\widetilde{LF}$, LF and the corresponding fit of the latter along with the inferred parameter values are presented in

the left panel of Fig. B.15. Both $\widetilde{LF}$ and LF have been normalized to the best a posteriori value (mind that it has uncertainty) of the soft-band LADE model from Aird et al. (2015) (see their Table 5) at $z = 1$ and $l_X = 26.1$ (which corresponds to a luminosity of 10^{44} erg s $^{-1}$ in the 2–10 keV energy band) and thus have the meaning of quasar space density per l_X . We see a fairly good but not perfect correspondence between the results of Aird et al. (2015) and ours. Namely, the g_1 value inferred from our fit to LF is smaller by $\sim 2\sigma_A$, where σ_A is the uncertainty of this parameter in Aird et al. (2015), while our best-fit g_2 and l^* values deviate from theirs by $\sim 2\sigma_A$ and $\sim 4\sigma_A$ respectively. However, it is possible to bring our modeled LF to better agreement with the result of Aird et al. (2015) by slightly changing the parameters of the adopted $\widetilde{LF}$. Namely, if we increase g_1 by 1σ , decrease g_2 by 2σ , increase l^* by 2σ and increase σ_{intX}^2 by 2σ (where σ 's are the corresponding uncertainties in our baseline model, see again the top section of Table 3), i.e. within the ranges marginally allowed by our baseline model posterior distributions, then we can reproduce the LF of Aird et al. (2015) very well, as demonstrated in the right panel in Fig. B.15.

We conclude that the results of the current study are in satisfactory agreement with the work of Aird et al. (2015) as regards the AGN LF, given that we have used many assumptions in our analysis that were irrelevant to theirs (in particular, we assumed σ_{intX}^2 to be independent of z and l_X and a constant σ_{varX}^2 , both of which might not be true). Furthermore, we see that LF can be fairly well described by a double-power law. However, the observed deviations at high luminosities indicate that this kind

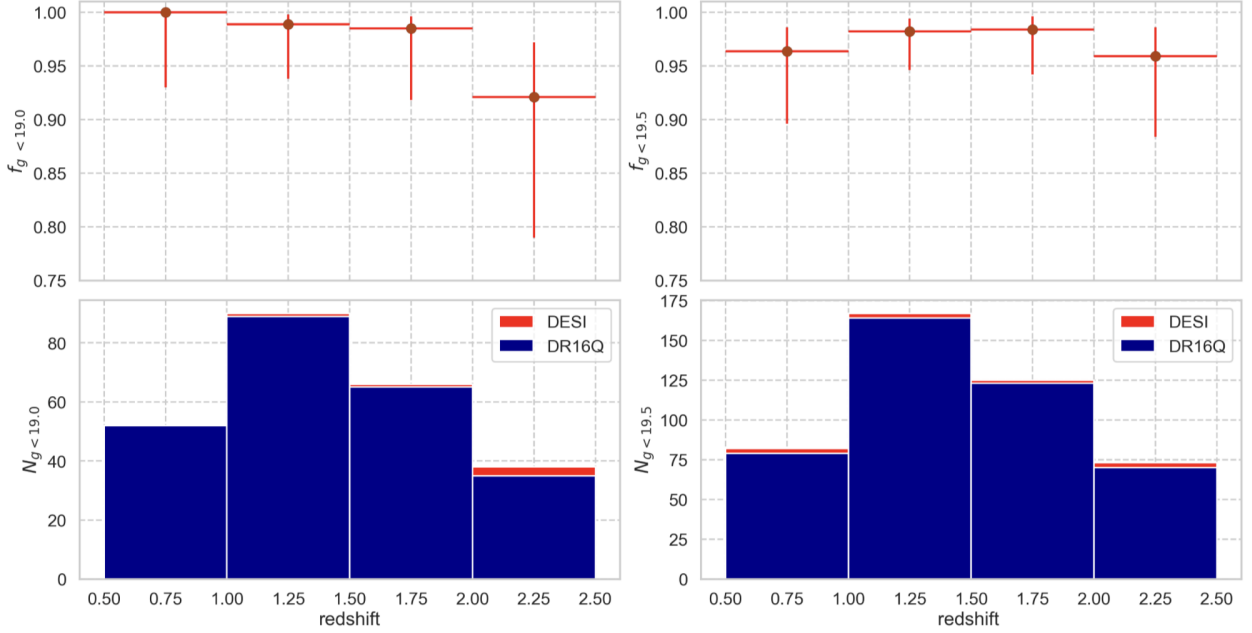


Figure A.14: SDSS DR16Q spectroscopic completeness (top panels) and number of quasars (bottom panels) in different redshift bins. Quasars with $m_g < 19$ and $m_g < 19.5$ are presented in the left and right panels, respectively. The error bars for the completeness measurements correspond to the 95% confidence level.

of model may not be a perfect choice for the primary LF.

Appendix C. Relationship between variance and structure function

Consider a weakly non-stationary variability process (see Vaughan et al. 2003) and its individual realization – a light curve of length Δt_2 that consists of N_1 consecutive measurements f_i of length $\Delta t_1 = \Delta t_2/N_1$ (without gaps between the measurements, i.e. we collect entire fluence in Δt_2). We can calculate the squared deviation of the measurements from the mean value $\bar{f}$ for this light curve. Let us define the variance on timescales from Δt_1 to Δt_2 to be the arithmetic mean of the squared deviation: $\hat{\text{Var}} = \frac{1}{N_1} \sum_{i=1}^{N_1} (f_i - \bar{f})^2$. We are interested in the expected value of $\hat{\text{Var}}$ for the stochastic process, $\text{Var}(\Delta t_1, \Delta t_2) = \langle \hat{\text{Var}} \rangle_{\text{lc}}$, i.e. the average over all possible realizations (light curves) of this process.

Let us also define a two-argument structure function $\hat{S}F^2$ for the light curve as the mean squared difference between f_i separated by time lag τ , such that $\Delta t_1 \leq \tau \leq (\Delta t_2 - \Delta t_1)$ and $\tau/\Delta t_1$ is a natural number: $\hat{S}F^2(\Delta t_1, \tau) = \frac{1}{N_p} \sum_i (f_i - f_{i+\tau/\Delta t_1})^2$, where the summation is performed over all measurement pairs separated by τ and $N_p = (\Delta t_2 - \tau)/\Delta t_1 = N_1 - \tau/\Delta t_1$ is the number of such pairs. Accordingly, we define $\bar{S}F^2 = \langle \hat{S}F^2 \rangle_{\text{lc}}$ as the expected value of $\hat{S}F^2$ over all possible realizations of the stochastic process.

Now, a long light curve of length $\Delta t_3 \gg \Delta t_2$, consisting of consecutive measurements f_i of length Δt_1 , can be divided

into consecutive segments of length Δt_2 . Then, the total number of measurement pairs separated by τ , such that $\Delta t_2 \leq \tau \leq (\Delta t_3 - \Delta t_2)$ and $\tau/\Delta t_2$ is a natural number, in the long light curve is $N_p = (\Delta t_3 - \tau)/\Delta t_1 = (\Delta t_2/\Delta t_1)(\Delta t_3 - \tau)/\Delta t_2 = N_1 N_{p2}$, where $N_{p2} \equiv (\Delta t_3 - \tau)/\Delta t_2$ is the total number of pairs of these segments. Let us also define $n = \tau/\Delta t_1$, $m = \tau/\Delta t_2$ and for k -th light curve of length Δt_2 we denote its mean and variance as $\bar{f}_k$ and $\hat{\text{Var}}_k = \frac{1}{N_1} \sum_j (f_j - \bar{f}_k)^2$. Then, we have

$$\begin{aligned} \hat{S}F^2(\Delta t_1, \tau) &= \frac{1}{N_p} \sum_{i=1}^{N_p} (f_i - f_{i+n})^2 \\ &= \frac{1}{N_p} \sum_{k=1}^{N_{p2}} \sum_{j=(k-1)N_1+1}^{kN_1} \left[(f_j - \bar{f}_k) + (\bar{f}_{k+m} - f_{j+n}) + (\bar{f}_k - \bar{f}_{k+m}) \right]^2 \\ &= \frac{1}{N_{p2}} \sum_{k=1}^{N_{p2}} (\hat{\text{Var}}_k + \hat{\text{Var}}_{k+m}) + \hat{S}F^2(\Delta t_2, \tau) \end{aligned} \quad (\text{C.1})$$

$$+ \frac{2}{N_{p2}} \sum_{k=1}^{N_{p2}} \sum_{j=(k-1)N_1+1}^{kN_1} \frac{(f_j - \bar{f}_k)(\bar{f}_{k+m} - f_{j+n})}{N_1} \quad (\text{C.2})$$

$$+ \frac{2}{N_{p2}} \sum_{k=1}^{N_{p2}} \left[(\bar{f}_k - \bar{f}_{k+m}) \sum_{j=(k-1)N_1+1}^{kN_1} \left(\frac{f_j - \bar{f}_k}{N_1} + \frac{\bar{f}_{k+m} - f_{j+n}}{N_1} \right) \right]. \quad (\text{C.3})$$

The term (C.3) obviously vanishes due to the definitions of arithmetic mean $\bar{f}_k$ and $\bar{f}_{k+m}$. Let us average $\hat{S}F^2(\Delta t_1, \tau)$ over

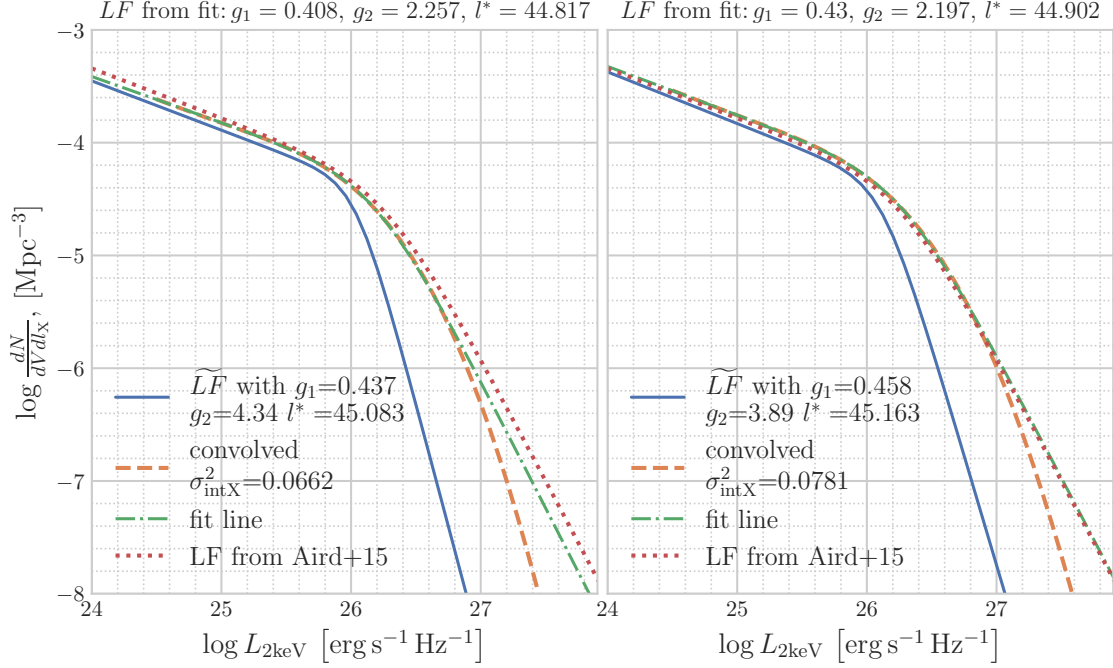


Figure B.15: In each panel, the blue line shows $\tilde{LF}$ (see eq. 12) at $z = 1$ with the parameters (of our baseline model) given in the legend; the orange line shows LF at $z = 1$ obtained from this $\tilde{LF}$ by applying equation (19) with intrinsic and variability related scatter (see the main text); and the green line shows the result of fitting LF by the double power-law model (eq. 20), with the inferred parameters indicated in the header. The left panel shows the results obtained for the median values of g_1, g_2, l^* and σ_{intX}^2 of our baseline model; the right panel for g_1 increased by 1σ , g_2 decreased by 2σ , l^* increased by 2σ and σ_{intX}^2 increased by 2σ . The red line in each panel shows the soft-band LADE model of Aird et al. (2015) (their Table 5) at $z = 1$, i.e. with $g_1 = 0.44 \pm 0.02$, $g_2 = 2.18 \pm 0.04$, $l^* = 44.93 \pm 0.03$, $p_1 = 3.39 \pm 0.08$, $p_2 = -3.58 \pm 0.26$ and $z_c = 2.31 \pm 0.07$.

many long light curves. Due to the linearity of the expected value, it will be the sum of the average values of (C.1) and (C.2). Since we have assumed that the light curve has no ‘memory’ (the i -th value of the light curve is not affected by the $(i + n)$ -th value), the term (C.2) vanishes upon averaging, and we finally get:

$$\widetilde{SF}^2(\Delta t_1, \tau) = 2\text{Var}(\Delta t_1, \Delta t_2) + \widetilde{SF}^2(\Delta t_2, \tau). \quad (\text{C.4})$$

Let us now establish the relation between $\widetilde{SF}^2(\Delta t_1, \tau)$ (defined for $\Delta t_1 \leq \tau \leq (\Delta t_2 - \Delta t_1)$) and $\text{Var}(\Delta t_1, \Delta t_2)$. We can compose a total of $N = N_1(N_1 - 1)/2$ measurement pairs from our light curve of length Δt_2 . On the one hand, we can consider the mean squared difference between two arbitrary measurements

from the light curve f_i and f_k , and get:

$$\begin{aligned} \sum_{k>i} \frac{(f_i - f_k)^2}{N} &= \sum_{k \neq i} \frac{(f_i - f_k)^2}{2N} = \sum_{i=1}^{N_1} \sum_{k=1}^{N_1} \frac{(f_i - \bar{f} + \bar{f} - f_k)^2}{N_1(N_1 - 1)} \\ &= \sum_{i=1}^{N_1} \frac{(f_i - \bar{f})^2}{N_1 - 1} + \sum_{k=1}^{N_1} \frac{(f_k - \bar{f})^2}{N_1 - 1} - 2 \sum_{i=1}^{N_1} \sum_{k=1}^{N_1} \frac{(f_i - \bar{f})(f_k - \bar{f})}{N_1(N_1 - 1)} \\ &= 2 \frac{N_1}{N_1 - 1} \text{Var}(\Delta t_1, \Delta t_2) - 2 \sum_{i=1}^{N_1} \frac{f_i - \bar{f}}{N_1 - 1} \sum_{k=1}^{N_1} \frac{f_k - \bar{f}}{N_1} \\ &= 2 \frac{N_1}{N_1 - 1} \text{Var}(\Delta t_1, \Delta t_2), \end{aligned} \quad (\text{C.5})$$

where the definition of $\bar{f}$ was used to cancel the last term in the expression preceding the final one. On the other hand, each pair of measurements is characterized by the time interval $\tau_j = j\Delta t_1$, where j can take values from 1 to $N_1 - 1$. Then, for a given τ_j there are $p_j = N_1 - j$ pairs, and we can write

$$\frac{\sum_{k>i} (f_i - f_k)^2}{N} = \frac{\sum_{j=1}^{N_1-1} \sum_{i=1}^{p_j} (f_i - f_{i+j})^2}{\sum_{j=1}^{N_1-1} p_j} = \frac{\sum_{j=1}^{N_1-1} p_j \hat{SF}_j^2}{\sum_{j=1}^{N_1-1} p_j}. \quad (\text{C.6})$$

Here, $\hat{S}F_j^2 \equiv \frac{\sum_{i=1}^{p_j} (f_i - f_{i+j})^2}{p_j}$, and it is related to $\widetilde{SF}^2(\Delta t_1, \tau_j)$ as $\widetilde{SF}^2(\Delta t_1, \tau_j) = \left\langle \hat{S}F_j^2 \right\rangle_{lc}$. Therefore, equations (C.5) and (C.6) give:

$$\frac{\Delta t_2}{\Delta t_2 - \Delta t_1} \hat{\text{Var}}(\Delta t_1, \Delta t_2) = \frac{\sum_{j=1}^{\Delta t_2/\Delta t_1 - 1} (\Delta t_2 - j\Delta t_1) \hat{S}F_j^2}{(\Delta t_2 - \Delta t_1) \Delta t_2/\Delta t_1}, \quad (\text{C.7})$$

which, due to the linearity of the expected value over all possible realizations of the process yields

$$\text{Var}(\Delta t_1, \Delta t_2) = \frac{\sum_{j=1}^{\Delta t_2/\Delta t_1 - 1} (\Delta t_2 - j\Delta t_1) \widetilde{SF}^2(\Delta t_1, j\Delta t_1)}{\Delta t_2^2/\Delta t_1}. \quad (\text{C.8})$$

This proof reflects the fact that the structure function determines the contribution of measurement pairs with a given time lag to the variance. Note that we would come to the same result if we randomly chose N' times first f_i and then f_k from the set of measurements. That is, the proportion between the total number of measurement pairs N' and the number of measurement pairs with particular τ_j , for all j , does not change when $N' \rightarrow \infty$.

In the limit $\Delta t_2 \gg \Delta t_1$,

$$\begin{aligned} \text{Var}(\Delta t_1, \Delta t_2) &\approx \frac{1}{2} \int_0^{\infty} T(0, 0, \Delta t_2, \tau) \widetilde{SF}^2(\Delta t_1, \tau) d\tau \\ &= \int_0^{\Delta t_2} \frac{(\Delta t_2 - \tau)}{\Delta t_2^2} \widetilde{SF}^2(\Delta t_1, \tau) d\tau, \end{aligned} \quad (\text{C.9})$$

where $T(a, c, b, x)$ is the probability density function of the triangular distribution.

For example, in the case of a power-law structure function, $\widetilde{SF}^2(\Delta t_1, \tau) = A\tau^\beta$ ($\beta \geq 0$), equation (C.9) gives

$$\text{Var}(\Delta t_1, \Delta t_2) \approx \frac{A\Delta t_2^\beta}{(\beta+1)(\beta+2)}. \quad (\text{C.10})$$

Figure C.16 compares $\text{Var}(\Delta t_1, \Delta t_2)$ calculated via equation (C.8) with its approximation given by equation (C.10) for different values of $\Delta t_2/\Delta t_1$ and different slopes β of the power-law structure function. For $\Delta t_2/\Delta t_1 \geq 10$, the approximate expression is accurate within 11% for all probed slopes.

Thus, knowing the structure function on some time scales, one can readily estimate the variance on these time scales. Moreover, using equation (C.4) it is also possible to evaluate the structure function on a different time scale.

Appendix D. The case of a power-law luminosity function and scatter along either the UV or X-ray axis.

It is instructive to consider the special case for the probability density $f(l_X, l_{UV}|\sigma_{UV}, \sigma_X, \gamma, \beta)$ (defined in Section 3.1) where the normalized primary LF has a power-law shape: $\widetilde{NLF}(\tilde{l}_X) \propto \tilde{l}_X^{-\alpha} \propto 10^{-\alpha l_X}$. If we also assume that there is no scatter in the

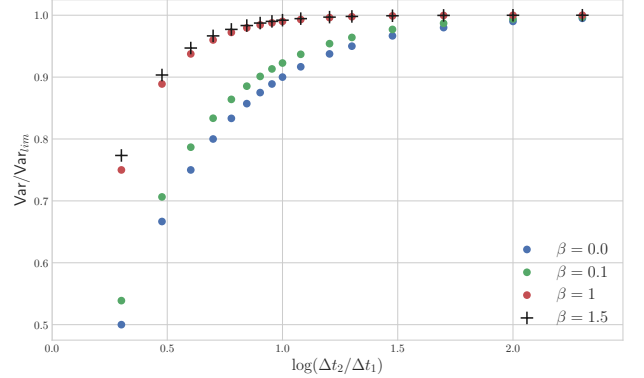


Figure C.16: Ratio of Var given by equation (C.8) and its approximation Var_{lim} given by equation (C.10) for a power-law structure function $\widetilde{SF}^2(\Delta t_1, \tau) = A\tau^\beta$ as a function of $\Delta t_2/\Delta t_1$, for different slopes (indicated by different colours).

primary X-ray–UV luminosity relation along the X-ray axis, i.e. $\sigma_X = 0$, then equation (14) simplifies to:

$$\begin{aligned} f(l_X, l_{UV}|\sigma_{UV}, \sigma_X = 0, \gamma, \beta) &\propto \mathcal{N}(l_{UV}|(l_X - \beta)/\gamma, \sigma_{UV}) 10^{-\alpha l_X} \\ &\propto \exp\left(-\alpha l_X \ln 10 - \frac{[\gamma l_{UV} - l_X + \beta]^2}{2\sigma_{UV}^2 \gamma^2}\right) \end{aligned} \quad (\text{D.1})$$

If, instead, we assume $\sigma_{UV} = 0$, we get:

$$\begin{aligned} f(l_X, l_{UV}|\sigma_{UV} = 0, \sigma_X, \gamma, \beta) &\propto \mathcal{N}(l_X|\gamma l_{UV} + \beta, \sigma_X) 10^{-\alpha(\gamma l_{UV} + \beta)} \\ &\propto \exp\left(-\alpha \gamma l_{UV} \ln 10 - \frac{[l_X - \gamma l_{UV} - \beta]^2}{2\sigma_X^2}\right) \\ &\propto \exp\left(-\alpha l_X \ln 10 - \frac{[\gamma l_{UV} - l_X + \beta + \alpha \sigma_X^2 \ln 10]^2}{2\sigma_X^2}\right) \end{aligned} \quad (\text{D.2})$$

We thus see that the probability density $f(l_X, l_{UV}|\sigma_{UV}, \sigma_X, \gamma, \beta)$ for parameter values (1) $\gamma = \gamma_0$, $\beta = \beta_0$, $\sigma_{UV} = 0$, $\sigma_X = \sigma_{X0}$ is the same as for (2) $\gamma = \gamma_0$, $\beta = \beta_0 + \alpha \sigma_{X0}^2 \ln 10$, $\sigma_{UV} = \sigma_{X0}/\gamma_0$, $\sigma_X = 0$ [the right-hand sides of equations (D.1) and (D.2) become identical]. This means that given a $f(l_X, l_{UV}|\sigma_{UV}, \sigma_X, \gamma, \beta)$ distribution or a sample from it, it is impossible to give preference to one of these two sets of parameter values. In other words, in the special case of a power-law LF and under assumption that the X-ray–UV luminosity relation is scattered in only one direction (l_X or l_{UV}), there arises degeneracy between the $\sigma_X = 0$ and $\sigma_{UV} = 0$ solutions. On the l_X vs. l_{UV} plane, the primary dependence line of the $\sigma_X = 0$ solution is translated by $\alpha \sigma_{X0}^2 \ln 10$ along the X-ray axis from the primary dependence line for the $\sigma_{UV} = 0$ solution.

However, this symmetry breaks down for any realistic LF, such as a double power-law LF used in this work. This allows us to give preference to either the $\sigma_X = 0$ or the $\sigma_{UV} = 0$ solution using e.g. the maximum likelihood method. Moreover, with the approach taken in this work we can consider the more general case where both σ_X and σ_{UV} are not equal to zero.

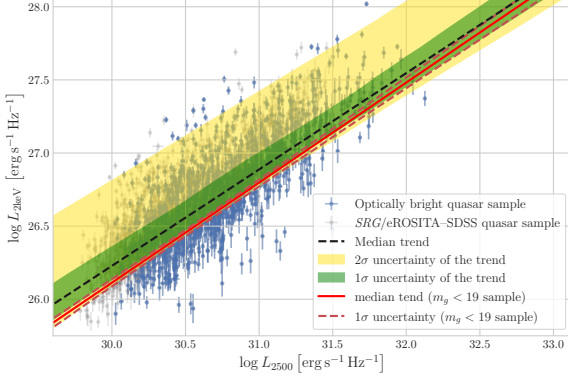


Figure E.17: Scatter plot of the X-ray and UV luminosities with their measurement uncertainties for the optically bright ($m_g < 18.5$) quasar sample (points with error bars) and the inferred dependence between these quantities for the baseline model (the median trend and 1σ and 2σ uncertainty regions). For comparison, the red solid and red dashed lines show the median trend and 1σ uncertainty region for our main *SRG/eROSITA*–SDSS quasar sample ($m_g < 19$).

Appendix E. Result for a higher optical flux threshold

To additionally test the robustness of the $L_{2\text{keV}}\text{--}L_{2500}$ relation derived in this work, we repeated the entire analysis using a quasar sample with a higher optical flux threshold compared to our main *SRG/eROSITA*–SDSS quasar sample. Namely, we adopted $m_g < 18.5$ instead of $m_g < 19$, retaining the same X-ray flux limit. The resulting sample contains 1242 quasars.

Figure E.17 shows the $L_{2\text{keV}}$ vs. L_{2500} scatter-plot and the $L_{2\text{keV}}(L_{2500})$ primary dependence obtained by applying our baseline model (see Section 4.1) to the optically bright sample, in comparison with the result for the main sample. Table E.5 presents the posterior distribution characteristics of the model parameters for the optically bright sample. In addition, in Fig. E.18 we compare the slope and normalization of the $L_{2\text{keV}}(L_{2500})$ dependencies inferred for the optically bright and main samples. Namely, we plot the joint posterior distribution contours corresponding to the 68% and 95% smallest credible regions. Note that the latter do not coincide with the quantile-based credible intervals listed in Table E.5, because the posterior distributions (especially for $\beta + 30.5\gamma$) are highly skewed.

The results obtained for the optically bright and main samples are in good agreement with each other. This strengthens the robustness of the conclusions of this work. It is important to note the $L_{2\text{keV}}(L_{2500})$ trend line passes approximately through the centre of the optically bright sample in Fig. E.17, which was not the case for our main sample. This explicitly demonstrates that our approach is insensitive to adopted X-ray and optical flux limits.

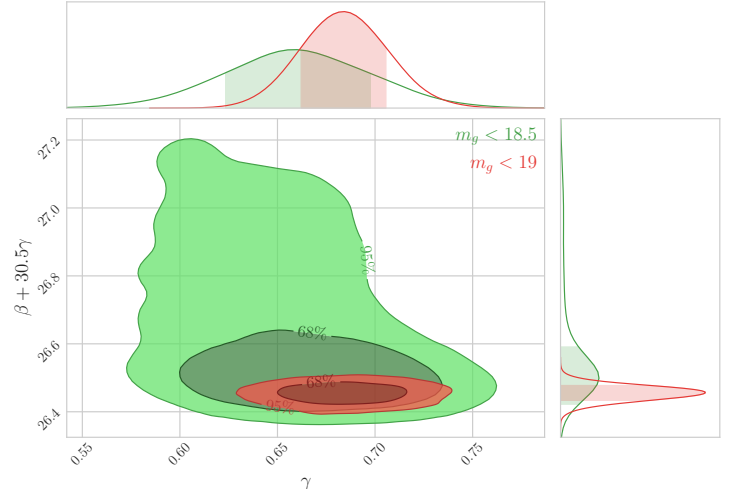


Figure E.18: Main panel: contour plots in the slope–normalization plane of the 68% and 95% smallest credible regions of the joint posterior distribution for the *SRG/eROSITA*–SDSS quasar sample ($m_g < 19$, red color) and the optically bright quasar sample ($m_g < 18.5$, green color). Side panels: posterior distributions of the corresponding parameter for both samples and their 68% smallest credible intervals (shaded regions).

Parameter	Mean value and standard deviation	2.5% percentile	16% percentile	median	84% percentile	97.5% percentile
Baseline model						
γ	0.66, 0.04	0.59	0.62	0.66	0.70	0.73
$\beta + 30.5\gamma$	26.58, 0.17	26.45	26.47	26.51	26.72	27.07
σ_{intUV}^2	0.03, 0.04	0.000	0.001	0.007	0.062	0.160
σ_{intX}^2	0.054, 0.018	0.002	0.040	0.060	0.068	0.076
g_1	0.440, 0.021	0.40	0.42	0.44	0.46	0.48
g_2	4.2, 0.3	3.7	3.9	4.2	4.5	4.8
l^*	44.91, 0.29	44.25	44.53	45.05	45.15	45.21
p_1	3.34, 0.08	3.18	3.26	3.34	3.43	3.51
z_c	2.29, 0.07	2.15	2.22	2.29	2.36	2.44

Table E.5: Characteristics of the posterior distributions of the model parameters based on the optically bright quasar sample.