

X-ray selected broad absorption line quasars in SDSS-V: BALs and non-BALs span the same range of X-ray properties

Pranavi Hiremath¹, Amy L. Rankine¹★, James Aird¹, W. N. Brandt^{2,3,4}, Paola Rodríguez Hidalgo⁵, Scott F. Anderson⁶, Catarina Aydar^{7,8}, Claudio Ricci^{9,10}, Donald P. Schneider^{11,12}, M. Vivek^{11,12}, Zsofi Igo^{7,8}, Sean Morrison¹³ and Mara Salvato⁷

¹Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

²Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA

³Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA

⁴Department of Physics, 104 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA

⁵Physical Sciences Division, School of STEM, University of Washington Bothell, Bothell, WA 98011, USA

⁶Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195, USA

⁷Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse 1, D-85748 Garching, Germany

⁸Excellence Cluster ORIGINS, Boltzmannstrasse 2, D-85748 Garching, Germany

⁹Instituto de Estudios Astrofísicos, Facultad de Ingeniería y Ciencias, Universidad Diego Portales, Av. Ejército Libertador 441, Santiago, Chile

¹⁰Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, China

¹¹Indian Institute of Astrophysics, Koramangala II block, Bangalore 560034, Karnataka, India

¹²Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Monchhofstr. 12-14, D-69120 Heidelberg, Germany

¹³Department of Astronomy, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Accepted 2025 August 4. Received 2025 July 23; in original form 2025 March 31

ABSTRACT

Broad absorption line (BAL) quasars are often considered X-ray weak relative to their optical/UV luminosity, whether intrinsically (i.e. the coronal emission is fainter) or due to large column densities of absorbing material. The SDSS-V is providing optical spectroscopy for samples of quasar candidates identified by *eROSITA* as well as *Chandra*, *XMM*, or *Swift*, making the resulting data sets ideal for characterizing the BAL quasar population within an X-ray selected sample. We use the Balnicity Index (BI) to identify the BAL quasars based on absorption of the C IV λ 1549 emission line in the optical spectra, finding 143 BAL quasars in our sample of 2317 X-ray selected quasars within $1.5 \leq z \leq 3.5$. This observed BAL fraction of ≈ 6 per cent is comparable to that found in optically selected samples. We also identify absorption systems via the Absorption Index (AI) which includes mini-BALs and NALs, finding 954 quasars with $AI > 0$. We consider the C IV emission space (equivalent width versus blueshift) to study the BAL outflows within the context of the radiatively driven accretion disc–wind model. X-ray selection excludes the highest outflow velocities in emission but includes the full range of absorption velocities which we suggest is consistent with the BAL gas being located further from the X-ray corona than the emitting gas. We observe both X-ray weak and X-ray strong BALs (via the optical-to-X-ray spectral slope, α_{ox}) and detect little evidence for differing column densities between the BAL and non-BAL quasars, suggesting the BALs and non-BALs have the same shielding gas and intrinsic X-ray emission.

Key words: quasars: absorption lines – quasars: emission lines – X-rays: galaxies.

1 INTRODUCTION

At the centre of every massive galaxy, a common consensus is the existence of a supermassive black hole (SMBH) that goes through active phases during its lifetime, accreting material from the inner regions of the galaxy and is therefore known as an active galactic nucleus (AGN; Lynden-Bell 1969). The varying scales in the model of an AGN (see Heckman & Best 2014), along with the observational evidence found for the important role AGN play in the formation and evolution of galaxies (e.g. Magorrian et al. 1998; Aird et al. 2010;

Brandt & Yang 2022), suggest the presence of feedback. The spectral energy distribution (SED) of an AGN covers the full electromagnetic spectrum meaning that AGN populations can be identified at X-ray, ultra-violet (UV), optical, as well as other wavelengths.

Quasars are the brightest of all AGNs and represent the most radiatively efficient phase of SMBH growth, releasing energy as kinetic luminosity through sub-relativistic wide-angle winds (e.g. Silk & Rees 1998). Broad absorption line (BAL) quasars are a subset of quasars with rest-frame UV spectra that show evidence for outflows via the presence of blueshifted absorption lines, usually identified based on absorption bluewards of the C IV λ 1549 broad emission line. The definition most frequently used to define a BAL quasar requires absorption troughs with widths $\geq 2000 \text{ km s}^{-1}$

★ E-mail: amy.rankine@ed.ac.uk

(Weymann et al. 1991). Other definitions can include much narrower absorption features ($\geq 450 \text{ km s}^{-1}$; Hall et al. 2002) often called mini-BALs; however, this can result in contamination by strong narrow absorption lines (NALs: $\text{FWHM} < 500 \text{ km s}^{-1}$; e.g. Hamann & Ferland 1999; Bowler et al. 2014). Quasars can additionally have absorption in lower ionization species such as $\text{Al III } \lambda 1857$ and $\text{Mg II } \lambda 2800$ – the so-called LoBALs – however we will focus on BAL quasars with absorption only in the high-ionization lines – the HiBALs, hereafter referred to as BALs. While the driver of BAL quasar outflows may be radiation, thermal, or magnetic pressure, BAL features are strong evidence for significant momentum transfer from a powerful radiation field to the gas, which results in the propulsion of the gas to high velocities (e.g. Weymann et al. 1991; Allen et al. 2011). These observations support a radiatively driven hydrodynamical scenario for quasar outflows whereby outflowing winds are propelled by radiation pressure caused by the accretion disc’s UV emissions (Proga 2007). Correlations between outflow velocity, Eddington ratio, and luminosity lead Vivek & Wylezalek (2025) to conclude that radiation pressure is important for driving outflows, but suggest that additional mechanisms (i.e. greater outflow distances, thicker discs, softer SEDs) are required to explain the presence of BALs in low-Eddington ratio sources.

Richards et al. (2011) demonstrated that the distribution of radio-quiet/loud quasars in the C IV emission space [C IV equivalent width (EW) versus blueshift] can be well understood via a ‘disc’ versus ‘wind’ model. In detail, a disc-dominated quasar would be expected to have symmetric line profiles (low blueshift) with the line dominated by the disc emission, and a wind-dominated system would have lower C IV EW and high blueshift. Rivera et al. (2020) and Richards et al. (2021) defined the ‘C IV distance’ metric in order to group quasars along the best-fitting line through C IV space with higher C IV distance values associated with stronger winds. He II EW (an SED indicator; Leighly 2004) is known to correlate with the C IV emission space (Baskin, Laor & Hamann 2013, 2015; Rankine et al. 2020 [hereafter R20; Timlin, Brandt & Laor 2021] and Rivera et al. (2022) showed that He II EW is correlated most strongly with C IV distance over blueshift or EW. Thus the C IV distance is a convenient way to explore the impact of changes in the SED under a disc–wind model. Such an accretion disc–wind model is often invoked to explain observed BAL features (see fig. 1 of Luo et al. 2013) and according to this model, radiatively driven winds launched from the accretion disc result in BALs when inclination angles are large, implying the wind is in the direct line of sight (Leighly et al. 2015). Here, these accretion disc winds may be equatorial (e.g. Murray et al. 1995) and/or polar (e.g. Punnsly 1999), and can be continuous or consist of knots/blobs of matter (Ghisellini, Haardt & Matt 2004; Matthews et al. 2016). Therefore, BAL quasars can be considered generally to be wind-dominated quasars which were investigated by R20 within the C IV space. They reported that the optically selected BAL quasars have moderate EW and high blueshift but crucially, however, they can be found in all regions of the space. Additionally, Ahmed et al. (2025) found that BALs and non-BALs have the same SEDs, from mid-infrared to X-ray.

Typically 10–20 per cent of optically selected quasars show broad absorption features; however, the intrinsic fraction is expected to be higher (up to 40 per cent; Allen et al. 2011; Dai, Shankar & Sivakoff 2012) due to considerations of geometrical and evolutionary aspects of BAL quasar outflows. Specifically, in a geometrical scenario, the fraction of BALs observed depends on the solid angle the BAL winds subtend (e.g. Ghosh & Punnsly 2007) while an evolutionary scenario considers the variability in the appearance/disappearance of the BAL winds (e.g. Mishra et al. 2021). Regardless of the scenario, it implies

that a classified non-BAL quasar could be harbouring BAL winds that are not currently within the line of sight and therefore making the intrinsic fraction of BAL quasars difficult to ascertain. Moreover, the observed fraction of BAL quasars also depends on the signal-to-noise (S/N) ratio of the sample studied (Gibson et al. 2009a, hereafter G09; Allen et al. 2011).

Additionally, to ensure the outflowing gas is not overionized by the X-ray and extreme-UV radiation from the corona/accretion disc i.e. to explain the observed high velocities, the disc–wind model also invokes X-ray shielding gas with high column densities (Murray et al. 1995). This gas allows for efficient line driving of the UV disc winds and is a natural consequence in the hydrodynamical 2D simulations of Proga & Kallman (2004). Luo et al. (2013) considered the X-ray shielding material to originate from ‘failed winds’ i.e. gas that was unable to reach the local escape velocity due to overionization. Baskin, Laor & Stern (2014) provide an alternate theory via Radiation Pressure Compression or Confinement (RPC) whereby radiation pressure compresses the gas along the radial direction, resulting in a high electron density and low ionization parameter to allow absorption in the UV and efficient driving of the outflows to high velocities. This theory does not invoke X-ray shielding gas but rather suggests the radiation force itself can lead to high column densities through compression of the absorbing gas, preventing it from becoming overionized. Many studies have indeed shown that BAL quasars often appear to be X-ray weak relative to non-BAL quasars (e.g. Green et al. 1995; Laor et al. 1997; Brandt, Laor & Wills 2000; Gallagher et al. 2006, G09; Luo et al. 2014; Saccheo et al. 2023).

The X-ray weakness of BAL quasars can be studied by quantifying the relationship between X-ray and UV luminosities. Studies such as Avni & Tananbaum (1982) parametrized this relationship between X-ray (L_X) and UV luminosity (L_{UV}) as $L_X \propto L_{UV}^\gamma$ with $\gamma \sim 0.6$. The relationship between L_X and L_{UV} is often quantified using α_{ox} , which is the energy index or slope that connects the optical with the X-ray band for a power-law spectrum ($F_\nu \propto \nu^\alpha$). In detail, Tananbaum et al. (1979) introduced this correlation specifically between monochromatic luminosities at 2500 Å and 2 keV, although the physical mechanism responsible for this relation remains unclear, i.e. the relationship between the UV disc and the X-ray emissions from the hot corona. Moreover, the relationship between α_{ox} and L_{2500} is often observed to be an anticorrelation (e.g. Steffen et al. 2006; Lusso & Risaliti 2016; Rankine et al. 2024), which implies that as L_{2500} increases the corona becomes (relatively) weaker and thus the AGN becomes more disc-dominated. However, there is intrinsic scatter in this relation and objects with greater than average X-ray luminosities given their optical luminosity are expected to drive winds less efficiently (Giustini & Proga 2019; Timlin et al. 2021; Rivera et al. 2022; Temple et al. 2023). While the choice of L_{2500} is somewhat arbitrary, recent studies such as Timlin et al. (2021) and Jin et al. (2024) have shown that the L_{2500} alone is sufficient to describe the UV/optical emissions of a quasar over a broader wavelength range, making L_{2500} suitable for investigating the relationship between UV and X-ray emissions.

Laor et al. (1997) demonstrated that the X-ray flux of BAL quasars was 10–30 times lower than expected for their UV flux, classifying them as ‘soft-X-ray weak’ objects and Laor & Brandt (2002) found that these quasars had the strongest absorption features. G09 (further studied in Wu et al. 2010, W10 hereafter) also found that observed X-ray weakness was correlated to maximum velocity and absorption strengths of the C IV BAL troughs, suggesting the impact of X-ray absorption on UV disc winds and further supporting the accretion disc wind model. However, Giustini, Cappi & Vignali (2008) studied

a sample of 54 X-ray-selected quasars to find that one-third of the sample showed little to no absorption and the distributions of the optical-to-X-ray spectral slope, α_{ox} , of the BAL and non-BAL populations were similar, suggesting that these BAL quasars were not X-ray weaker relative to the non-BALs. They attributed their results to the X-ray selected sample preferentially being viewed from smaller angles with respect to the accretion disc rotation axis than optically selected BAL quasars. While they conducted X-ray spectral analysis considering a neutral absorber model, Streblyanska et al. (2010) considered both neutral and ionized absorption models for a sample of 39 X-ray-selected BAL quasars, finding that while 36 per cent of BALs did not show absorption in a neutral absorption model, 90 per cent of them were absorbed in a physically motivated ionized model, although even this model is likely to be too simplistic.

This paper aims to study the multiwavelength properties of X-ray selected (including *eROSITA*, *Chandra*, *XMM*, and *Swift*) SDSS-V BAL quasars relative to the non-BAL quasars to further investigate the X-ray weakness of ‘classical’ BAL quasars previously predominantly observed in optically selected samples. This observed X-ray weakness of BAL quasars makes them interesting candidates to study within the X-ray-selected SDSS-V sample, so the preliminary aim of this paper is to find the fraction of BAL quasars present in this sample. The *eROSITA*/SDSS-V collaboration marks the first time we have a large (soft) X-ray-selected sample with optical and corresponding X-ray data readily available. Therefore, we can build on the previous investigations of X-ray-selected BAL quasars through statistical analyses of a larger sample. Additionally, the work in this paper can be considered to be an extension of R20; while they focused on the study of emission and absorption properties of optically selected BAL quasars relative to non-BAL quasars, we aim to extend this investigation to X-ray-selected BAL quasars to obtain a more complete picture for the BAL population. The sample used and the results found by R20 form a basis of comparison to better understand the X-ray-selected sample studied in this paper.

The outline of this paper is as follows: we first describe in further detail our sample in Section 2. Section 3 presents the BAL identification methods employed to achieve the first aim of this paper, i.e. identify X-ray-selected BAL quasars. Section 4 investigates the identified subpopulations by placing them in the C IV emission space and further explores the identified BAL quasars in detail. Section 5 studies the X-ray properties of the BAL and non-BAL quasars by determining their X-ray luminosities, α_{ox} and $\Delta\alpha_{\text{ox}}$ values. Section 6 discusses and Section 7 summarizes the obtained results.

A Λ CDM cosmology with $h_0 = 0.71$, $\Omega_M = 0.27$, and $\Omega_\Lambda = 0.73$ is adopted for determining quantities such as quasar luminosities.

2 DATA

The *eROSITA* (extended ROentgen Survey with an Imaging Telescope Array) instrument aboard the *SRG* (Russian–German Spektrum Roentgen Gamma; Sunyaev et al. 2021) space mission, launched in July 2019, presents a step forward for the exploration of X-ray AGN properties (Predehl et al. 2021). *eROSITA* is an X-ray-sensitive telescope capable of discovering new X-ray sources within the energy range of ~ 0.2 –8 keV (Merloni et al. 2024). SDSS-V sets out to be the first of its kind to provide multi-epoch, all-sky surveys in the optical and IR wavelengths (Kollmeier et al. 2025). The black hole mapper programme (Anderson et al., in preparation) of SDSS-V forms one of the three top-level goals aimed to provide spectroscopic measurements such as redshifts for $\sim 300\,000$ X-ray sources identified by *eROSITA*.

Optical spectra have been obtained using the BOSS multi-fibre spectrographs (Smee et al. 2013) on the Apache Point Observatory 2.5 m SDSS telescope (Gunn et al. 2006) and the 2.54 m Irénée du Pont Telescope at Las Campanas Observatory (Bowen & Vaughan 1973). We select sources with redshifts in the range 1.5–3.5 and $S/N > 5$ per SDSS spectrum pixel ($\Delta v = 69 \text{ km s}^{-1}$) within the wavelength range of investigation (1260–3000 Å), and spectroscopically classified as ‘QSO’ by the SDSS pipeline. Our sample consists of 2680 objects targeted by SDSS-V (Kollmeier et al. 2025) as part of the SPIDERS programme (SPectroscopic IDentification of EROSITA Sources; Dwelly et al. 2017; Comparat et al. 2020; Aydar et al. 2025) which aims to obtain optical/near-infrared spectra of sources detected by *eROSITA*. The SPIDERS targets in our sample were detected in either the full survey depth eFEDS region (*eROSITA* Final Equatorial Depth Survey; Brunner et al. 2022; Salvato et al. 2022, 43 per cent of our sample) or the first all-sky survey eRASS:1 (32 per cent; Merloni et al. 2024). We supplement the *eROSITA*-selected sample with 695 sources from the *Chandra* Source Catalogue (Evans et al. 2024) and another 55 from the combination of *XMM-Newton* and *Swift*-XRT serendipitous catalogues (Webb et al. 2020; Delaney et al. 2023),¹ all in the 1.5–3.5 redshift range and with $S/N \geq 5$ per SDSS pixel. The *Chandra* and *XMM/Swift* catalogues have been used to define additional targets for SDSS-V, probing a larger sky area and deeper X-ray observations. All spectra were reduced with the v6.1.3 version of the SDSS BOSS pipeline, idlspec2d (Bolton et al. 2012; Morrison et al., in preparation) as part of the SDSS DR19 (SDSS Collaboration 2025).

Our sample is X-ray selected by design and so will be affected by the sensitivity and depth of the X-ray instruments and surveys. The *XMM* and *Swift* detections (4.5–12 and 2–10 keV, respectively) are at harder X-ray energies than *eROSITA* where most of the detections are below 2 keV. Meanwhile the *Chandra* sources were mostly detected at 0.5–7 keV but observations are not uniform in depth. The largest subset of our sample is from eFEDS and so we consider the eFEDS flux limit ($10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$) throughout the paper.

The redshift range 1.5–3.5 was chosen for examination because, at $z < 1.5$, the wavelength range of interest for C IV BAL detection begins to shift out of the spectrograph’s range, while above $z > 3.5$ the C III] emission line is absent which is necessary for spectral reconstructions (see Section 3). The detectability of BAL-troughs decreases significantly as the S/N of a given spectrum decreases (G09; Allen et al. 2011). As we aim to investigate the BAL quasars relative to the non-BAL quasars in this X-ray-selected sample, a minimum spectrum S/N threshold is adopted to only include spectra with $S/N > 5$.

The C IV emission-line blueshifts that we are interested in cover velocities of order a few thousand km s^{-1} and so rely on accurate systemic redshifts. The BAL classifications are also dependent upon the exact redshift estimate for a quasar where the troughs lie close to the edges of the velocity window considered (see Section 3). We produced updated redshifts using the cross-correlation algorithm outlined in section 4.2 of Hewett & Wild (2010) and updated in Stepney et al. (2023). More information regarding the redshifts for the *eROSITA*-selected subsample can be found in Rankine et al. (in preparation) and we adopt the same procedure for the additional *Chandra*- and *XMM/Swift*-selected sources. In practice,

¹The *XMM-Newton* and *Swift*-XRT catalogues were combined to form a single target catalogue for SDSS-V to assign spare fibres (i.e. as an ‘open fibre programme’), preferentially adopting the *XMM-Newton* positions and X-ray fluxes, when available.

> 85 per cent of the sample has redshift estimates that differ from the SDSS pipeline by $|\Delta z| < 500 \text{ km s}^{-1}$ and < 5 per cent have $|\Delta z| > 1000 \text{ km s}^{-1}$.

A table containing the optical and X-ray properties underpinning this paper is available as supplementary data. Appendix A describes the data available.

3 BAL IDENTIFICATION

The first aim of this paper is to identify the BAL quasars within the X-ray-selected SDSS-V sample, and we have chosen to make this categorization based on the C IV emission line, while excluding quasars defined as LoBALs due to Al III and/or Mg II absorption. Two methods are commonly used to classify quasars as BAL or non-BAL: Weymann et al. (1991) defined the Balnicity Index (BI) while Hall et al. (2002) introduced the Absorption Index (AI) to incorporate absorption close to systemic velocity and narrower troughs, often called mini-BALs. However, the AI metric can also contaminate BAL samples with strong narrow absorption lines (NALs) and blended NALs (see Knigge et al. 2008 for a full discussion and Appendix B for example spectra). BI and AI are calculated as follows:

$$\text{BI} = \int_{3000}^{25000} \left(1 - \frac{f(V)}{0.9}\right) C dV, \quad (1)$$

and

$$\text{AI} = \int_0^{25000} \left(1 - \frac{f(V)}{0.9}\right) C dV. \quad (2)$$

$f(V)$ is the spectrum flux normalized by the continuum and V is the velocity in km s^{-1} defined with respect to the quasar systemic velocity. In equation (1), $C = 1$ where $f(V) < 0.9$ contiguously for at least 2000 km s^{-1} and zero otherwise. The integration occurs between 3000 and 25000 km s^{-1} from systemic velocity. Quasars with $\text{BI} > 0$ are considered BAL quasars ($\text{BI} = 0$ are non-BI BALs). In equation (2), $C = 1$ where $f(V) < 0.9$ contiguously for at least 450 km s^{-1} and zero otherwise, integrated over 0 – 25000 km s^{-1} . Quasars with $\text{AI} > 0$ could either be mini-BALs or have strong NALs in their spectra; therefore we do not refer to this population as BALs in this paper and instead denote them as ‘AI quasars’ ($\text{AI} = 0$ are non-AI quasars). Quasars with BI and/or AI = 0 are collectively referred to as non-BALs in this paper. Given the narrower width requirement, we apply $\chi^2_{\text{trough}} \geq 10$ in order to account for spurious troughs (Trump et al. 2006) using:

$$\chi^2_{\text{trough}} = \sum \frac{1}{N} \left(\frac{1 - f(V)}{\sigma} \right)^2, \quad (3)$$

where N is the number of pixels in the potential false-positive AI troughs (due to noise) and σ is the estimated noise for a given pixel.

To obtain $f(V)$, we reconstruct the quasar spectra using the same procedure outlined in R20 using the same spectral components generated via mean-field Independent Component Analysis (ICA). The resulting reconstructions essentially fill in the BAL troughs to produce a smooth unabsorbed continuum, and this continuum is used to normalize the spectrum. The normalized spectrum, $f(V)$, is used for distinguishing it as a BAL or non-BAL using the BAL-identification code by Rogerson et al. (2018), where the code uses equations (1), (2), and (3).

We made use of the same ICA components as R20 that were generated using the SDSS-IV quasar sample of Pâris et al. (2018) (see section 4.2 of R20). The SDSS-V target selection differs greatly from that of previous SDSS surveys (for this paper, mainly due to targeting an X-ray selected quasar sample); thus it is not a foregone conclusion

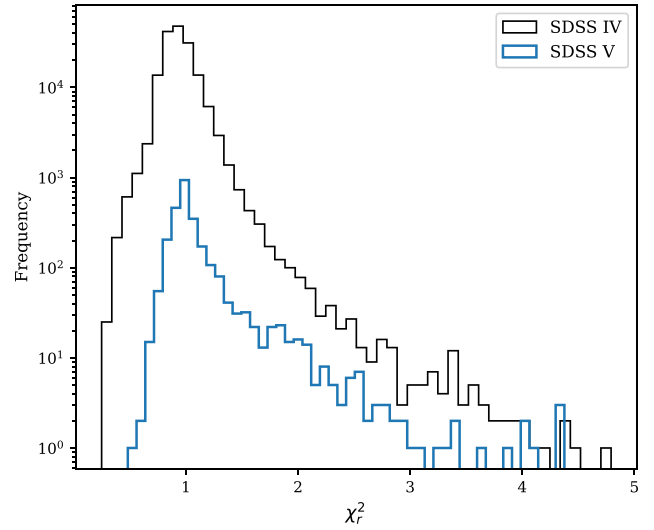


Figure 1. Distribution of the reduced χ^2 values of the SDSS-IV and SDSS-V samples consisting of $\sim 165\,000$ and 2680 quasars respectively, show that the two samples are comparable with a common peak occurring at $\chi_r^2 \approx 1$. This suggests that the MFICA components generated using SDSS-IV are suitable for constructing a smooth, high S/N continuum for the spectra in the SDSS-V sample.

that the components would perform adequately as the SDSS-V quasar population may have significantly different optical/UV spectroscopic properties. We compare the quality of the reconstructions of the SDSS-IV and SDSS-V samples using the reduced χ^2 of the fits. The use of the same components in reconstructing the SDSS-V spectra is justified due to the similarities in the distribution of the reduced χ^2 values of the two samples as shown in Fig. 1.

We limit our final sample to quasars that have spectral reconstructions with $\chi_r^2 < 2$ and place them in the C IV emission line EW versus blueshift space (see Section 4) to visually inspect the objects lying in the extremes of C IV emission space. Following R20, we removed quasars with large negative C IV blueshifts ($< -1200 \text{ km s}^{-1}$), small C IV EW ($< 10 \text{ \AA}$) and within the bottom left corner of the space ($\log_{10} \text{CIV(EW)} < -2.3077 \times 10^{-4} \times \text{CIV (blueshift)} + 1.321$) due to suboptimal reconstructions of quasar spectra in this region. These considerations produced a final sample of 2317 quasars.

Fig. 2 shows that the fraction of BI BAL and AI quasars relative to non-BAL quasars in general increases with increasing S/N. This increase is steeper for the AI quasars due to the additional χ^2_{trough} requirement which introduces an additional bias against low-S/N spectra. Although the majority of the BI BAL quasars in this sample have low S/N, each spectrum has been visually inspected to confirm the presence of a broad absorption trough in the original spectrum. While there are relatively more BI BAL and AI quasars than non-BAL quasars in the high S/N bins, the first panel indicates that the overall number of quasar spectra in this high-S/N range is small.

The observed fraction of BAL quasars in Fig. 2 will be different from the intrinsic fraction due to selection effects resulting from the magnitude-limited nature of SDSS and S/N of the spectra. Allen et al. (2011) quantified the effect of BAL quasars being redder (thus more likely to not satisfy the colour selection of previous SDSS surveys) and the effect of S/N (noise spikes halting the BI integration). SDSS-V target selection does not employ any colour constraints but does, for our sample, have an X-ray flux cut in its place. However, determining the intrinsic BAL fraction from this sample is beyond the scope of this paper (and is likely not feasible due to the complex way in

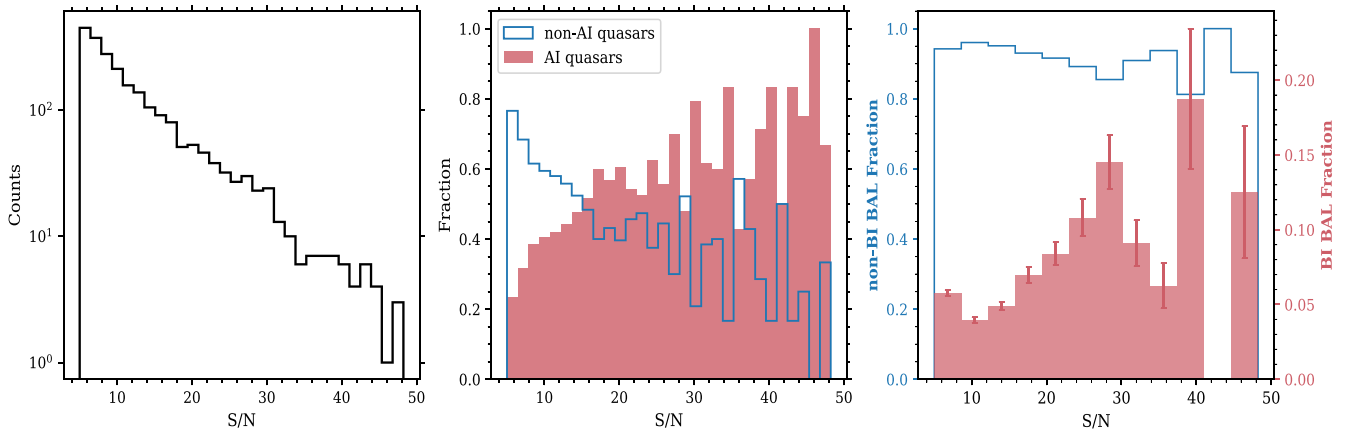


Figure 2. The *first* panel shows the median S/N distribution in the SDSS spectra of our final X-ray selected sample numbering 2317 quasars, reduced from 2680 due to the removal of sources with suboptimal spectral reconstructions (see Section 3). The *second* and *third* panels present the distribution of the fraction of AI quasars, BI BALs, and non-BAL quasars as a function of median S/N. The fraction of BALs and non-BALs adds to unity in each bin (note the different axis for the BI BAL and non-BAL fraction in the *third* panel). The fraction of BI BAL and AI quasars decreases at low S/N due to noisy spectra that make it harder to identify absorption features.

which X-ray highly obscured systems are lost from the sample); we instead aim to understand the BAL *properties* of this new and large X-ray-selected BAL quasar sample.

Previously studied optically selected samples have observed BAL quasars to be X-ray weaker relative to non-BAL quasars, so the initial aim of this paper is to identify BAL quasars within our SDSS-V sample to investigate their extent in the X-ray selected sample. Using the methods outlined in this section, we found our sample to consist of 143 or 6.2 per cent of BI BAL quasars (excluding 9 LoBALs) and 954 or 41.2 per cent of AI quasars. Several studies find the observed BAL fraction to be in the range of 10 per cent to 15 per cent within optically selected quasar samples (e.g. Tolea, Krolik & Tsvetanov 2002, Hewett & Foltz 2003, Knigge et al. 2008, G09). Trump et al. (2006) quote BI and AI fractions of 10.4 and 60.6 per cent, respectively. Note that their definition of AI extends to $29\,000\text{ km s}^{-1}$ (cf. our $25\,000\text{ km s}^{-1}$) and they also obtain an AI fraction of 26.0 per cent if they increase the minimum trough width from 450 to 1000 km s^{-1} . Thus, the fraction of BAL or AI quasars identified within our X-ray-selected quasar sample is comparable to – albeit slightly lower than – the fractions found in optically selected samples.

In the following sections, we proceed to examine where the X-ray selected BAL quasars identified in our paper lie in the C IV emission space compared to the non-BALs, and to explore their UV/optical and X-ray properties.

4 C IV EMISSION SPACE

As mentioned in Section 1, Richards et al. (2011) demonstrated that the distribution of quasars within the C IV emission space is useful for understanding which component – disc (high C IV EW and low blueshift) or wind (low C IV EW and high blueshift) – is dominant in the context of an accretion disc–wind model. R20 showed that optically selected BAL quasars exist with a range of C IV emission properties (and were similar to the non-BAL quasars across various UV-derived properties). Hence, we employ this space to study the X-ray selected SDSS-V BALs relative to non-BAL quasars identified in this paper.

Both panels of Fig. 3 show the optically selected SDSS-IV sample in grey dots (from R20, and limited to the CORE SDSS-IV targets)

relative to the SDSS-V sample in coloured dots, these samples are placed in the C IV emission space where the EWs and blueshifts of the C IV emission line have been measured non-parametrically using the ICA reconstructions, consistently to the methods used by R20. The SDSS-IV sample is a significantly larger sample that extends to higher C IV blueshifts relative to the X-ray-selected SDSS-V sample studied in this paper. While a more detailed comparison of the SDSS-IV and SDSS-V samples in the C IV emission space will be presented by Rankine et al. (in preparation), the difference in the extent of the C IV blueshifts could be leading to the differences in their distributions, which can be seen via the 1D blueshift distributions in the same figure. Furthermore, a 1D Kolmogorov–Smirnov (KS) test indicates an extremely low probability that the SDSS-IV and SDSS-V samples are drawn from the same underlying distribution of EW ($p = 3.2 \times 10^{-62}$) or blueshift ($p = 1.2 \times 10^{-17}$) and thus we can conclude that they are significantly different.

R20 reported that while the optically selected BAL quasars in their SDSS-IV sample are distributed across the entire C IV space, they are most likely to have moderate to low EW and high blueshifts. In detail, the optically selected BAL quasars in the SDSS-IV sample are mildly skewed towards low EW and high blueshifts relative to the non-BAL quasars (see their fig. 8). Fig. 3 demonstrates that the SDSS-V BI BAL and AI quasars are distributed across the C IV space similarly to the non-BAL quasars. The small p -values ($< 10^{-5}$) resulting from KS tests, where we require $p < 0.05$ to reject the null hypothesis of two populations being drawn from the same underlying distribution, suggest that the BAL and non-BAL populations are unlikely to be from the same underlying distribution.

In an earlier study of BALs in the SDSS DR6 spectroscopic survey, Allen et al. (2011) found an intrinsic redshift-evolving BAL fraction which results in an observed optical luminosity dependence when accounting for various selection effects. R20 described an optical luminosity dependence on C IV location. Rankine et al. (in preparation) will perform a more detailed analysis of the different redshift distributions; however, our sample does cover the same range of redshifts as R20 and here we compile a redshift-matched BAL and non-BAL sample to check if the differences in the *observed* redshift distribution between the BALs and non-BALs in our sample are leading to the differences in the C IV emission space. For a given BI BAL quasar we chose three non-BI BAL quasars with the closest z

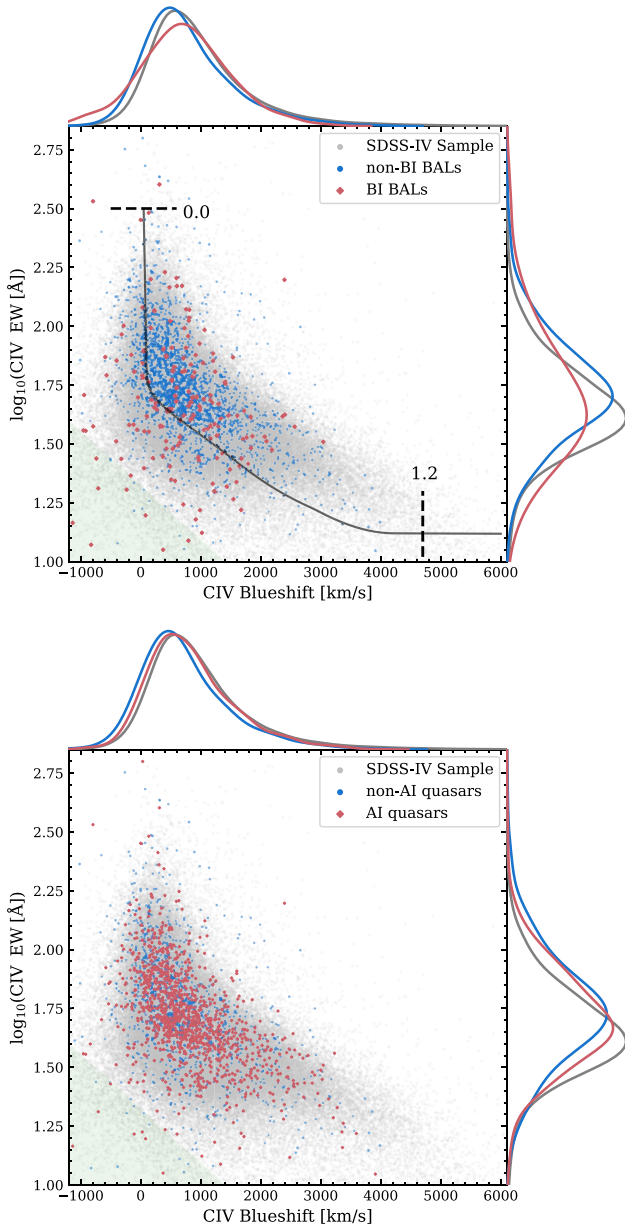


Figure 3. The C IV space with BI- (*top panel*) and AI- (*bottom panel*) defined quasars in red and non-BAL quasars in blue. These X-ray-selected BI BALs and AI quasars overlap with their non-BAL equivalents in the C IV space. The optically selected SDSS-IV sample shown in grey in both panels (BALs and non-BALs) extends to higher C IV blueshifts relative to the X-ray-selected sample in coloured dots. Additionally, the *top panel* presents the best-fitting relation based on the SDSS-RM sample that is used to define C IV distance that increases with decreasing EW and increasing blueshift. The green-shaded region in both panels represents the sources excluded from our sample due to suboptimal reconstruction of their spectra (see Section 3).

values, while for every AI quasar we chose one non-AI quasar with the closest z value. We performed KS tests to compare the BAL and the AI sample to their redshift-matched non-BAL equivalents in the C IV space and report the resulting p -values in Table 1.

The blueshifts of the BI sample are statistically consistent with being drawn from the same distribution as the non-BI BALs ($p = 0.5$), whereas the lower p -value when comparing the blueshifts of AI quasars to the equivalent non-AI quasars provide statistical

Table 1. KS test p -values for the redshift matched BI and AI populations. The null hypothesis of the populations being drawn from the same distribution is rejected when $p < 0.05$. X-ray properties of the BI/AI populations suggest that the BAL and non-BAL quasars are drawn from the same underlying distribution.

KS test	BI populations	AI populations
1D (EW)	0.03	0.01
1D (Blueshift)	0.5	0.03
2D (EW-blueshift)	0.02	0.01
1D (L_X)	0.2	0.4
2D (L_X - z)	0.3	0.4
1D (N_H)	0.6	0.5
1D ($\Delta\alpha_{ox}$)	0.4	0.8

significance ($p < 0.05$) for rejecting the null hypothesis of being drawn from the same population; we therefore conclude that the underlying populations differ. We also find statistically significant p values allowing rejection of the null hypothesis of the same underlying distributions when comparing the EWs of both the BI and AI samples to their non-BAL equivalents, as well as statistically significant p values when performing a 2D KS test² to compare the distributions of both BI BALs and AI quasars to their non-BAL equivalents in the 2D EW–blueshift space. We conclude overall that the BAL samples have a distinct distribution in the C IV EW–blueshift space compared to non-BALs that cannot be accounted for by differences in the observed redshift distributions. However, it is important to note that – while statistically significant – the differences are *small* and appear mainly to be driven by the BALs being preferentially found in quasars with slightly lower C IV EWs, i.e. BI BALs have median $\log_{10}(\text{EW } [\text{\AA}]) = 1.66 \pm 0.03$ compared to a median $\log_{10}(\text{EW } [\text{\AA}]) = 1.719 \pm 0.006$ for non-BI BALs; similarly, AI quasars have median $\log_{10}(\text{EW } [\text{\AA}]) = 1.697 \pm 0.008$ compared to $\log_{10}(\text{EW } [\text{\AA}]) = 1.738 \pm 0.008$ for non-AI quasars. None the less, BALs are distributed across the C IV space occupied by the SDSS-V X-ray selected quasars. We note that using the redshift-matched sample did not change the rest of our results so we retained the whole sample for our subsequent analysis for better number statistics.

Next, we investigate how the *strength* of the BAL features (rather than simply their presence) varies across the C IV emission space. We divide the BI BAL sample into different regions of the C IV emission space based on their ‘C IV distance’, which is defined by Richards et al. (2021) as the distance along the best-fitting relation between C IV EW and blueshift for the SDSS-RM sample from Rivera et al. (2020) and is shown by the black curve in the top panel of Fig. 3. In detail, each object in our sample is assigned a C IV distance based on its closest position to this best-fitting relation to obtain its distance along the curve, using the code by McCaffrey & Richards (2021). As noted in the figure, large C IV distances imply low EWs and high blueshifts. Our sample typically lies to the right or above the Richards et al. (2021) curve due to slight differences in the way that the blueshift and EW are measured, but this does not affect the measurements of the C IV distance.

Fig. 4 displays the median composites of the X-ray-selected BI BAL quasars in different C IV distance bins, with C IV distance increasing from left to right. To generate the composites the original and the reconstructed spectra are shifted to the quasar rest frame

²The 2D KS tests were performed using the public code NDTEST written by Zhaozhou Li, <https://github.com/syrte/ndtest>.

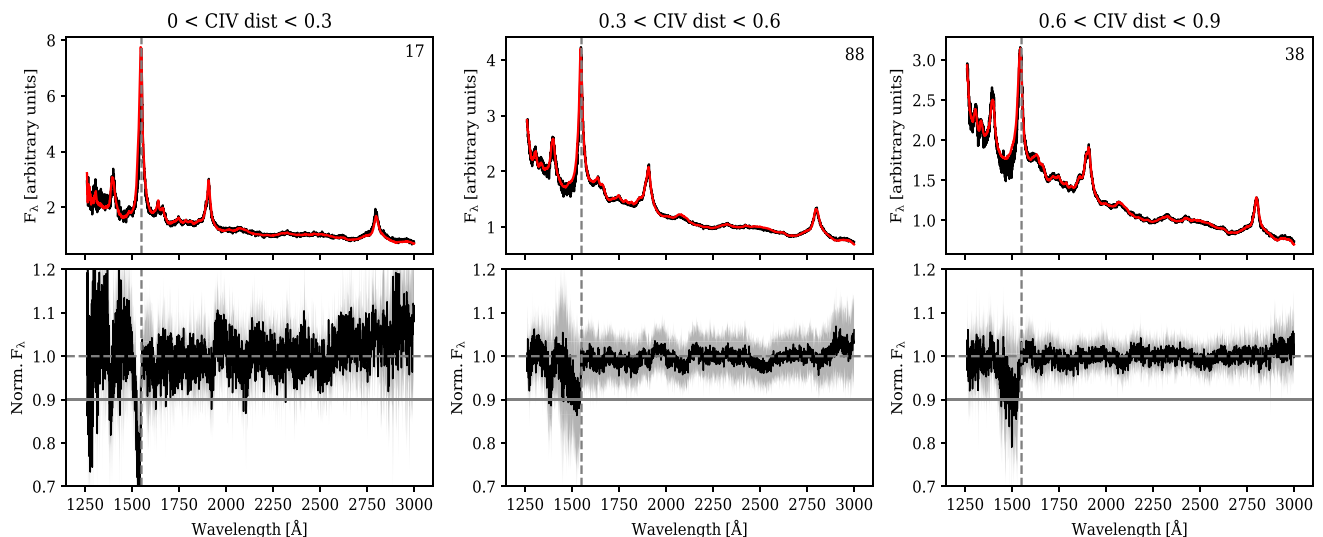


Figure 4. Median composite spectra of the BI BAL quasars in different CIV distance bins with the *top panels* showing the composite flux (black) and reconstruction (red) along with the number of sources that went into making the composites in the top right corner, the *bottom panels* show composite spectra normalized by the reconstructions along with its Median Absolute Deviation (MAD) in grey. The solid horizontal grey line marks the 0.9 threshold for the BI calculation. Considering that our X-ray-selected BAL quasars do not extend to as high CIV blueshifts as the SDSS-IV sample, expected absorption features are observed in our BI BAL quasars with the troughs becoming broader with increasing CIV distance.

and the median flux is calculated at each pixel for the original, reconstructed, and normalized (original/reconstruction) spectra. The composite spectra reveal that as the CIV distance increases, the absorption features become broader. We remind the reader that the BI BAL sample does not extend to as high CIV blueshifts/distances as the SDSS-IV sample from R20, which results in no BI quasars in the highest CIV distance bin (CIV distance > 0.9), and thus we do not observe the strongest BAL features, as per the correlation between emission and absorption signatures discussed in R20. Moreover, it was also observed that the absorption troughs for these BAL quasars were found to exist at a wide range of velocities (as observed in the median absolute deviation at each pixel – the grey-shaded region in the bottom panels) such that the depth and width of many of the troughs are washed out in the composites. Fig. 5 reveals that our X-ray selected sample is indeed missing the quasars with the strongest absorption troughs (largest BI values) compared to the optically selected population. Given the known correlation between BI and the CIV emission properties, we limit the optically selected sample to $-1000 \leq \text{CIV blueshift} \leq 3000 \text{ km s}^{-1}$. Removing the highest CIV blueshifts objects is insufficient to account for the lack of objects with $\text{BI} > 10^4 \text{ km s}^{-1}$ in the X-ray selected sample. Our X-ray sample also lacks $\text{BI} < 10^2 \text{ km s}^{-1}$ objects. To test if the different distribution is purely due to our smaller sample size, we perform bootstrap sampling by randomly sampling 143 BI BALs from the SDSS-IV sample 100 times to find the average distribution shown as the dashed grey histogram in Fig. 5. While the SDSS-V and the bootstrapped SDSS-IV BI distributions are similar between $1.5 < \log_{10}(\text{BI} [\text{km s}^{-1}]) < 4$, the differences between the distributions persist. Our sample is also missing the lowest BI objects, likely due to the lower average S/N of our spectra compared to the SDSS-IV sample.

Fig. 6 uses the same CIV distance bins used in Fig. 4, to examine how the median BI values (from equation 1) change as the CIV distance increases, where the BI value provides a measure similar to an EW measure (in km s^{-1}) of the broad absorption features that are $\geq 2000 \text{ km s}^{-1}$ wide bluewards of the CIV emission line. These median BI values, shown as the blue dots in the figure, suggest an

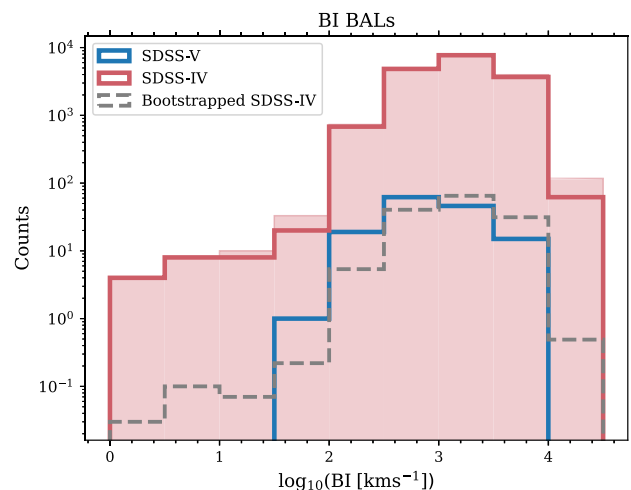


Figure 5. Comparing the BI values (in km s^{-1} , quantifying the strength of absorption in BALs, similar to an EW measure) of the BI-defined BAL quasars in the X-ray selected SDSS-V sample (in blue) and the optically selected SDSS-IV sample (in red). The filled red histogram shows the full SDSS-IV sample, whereas the open histograms are restricted to samples with $-1000 \text{ km s}^{-1} \leq \text{CIV blueshift} \leq 3000 \text{ km s}^{-1}$, i.e. they are CIV blueshift matched. Our relatively smaller X-ray selected sample picks up the most common BI BALs, however, with a lack of objects with $\text{BI} > 10^4 \text{ km s}^{-1}$. The grey histogram shows a bootstrapped subsample from SDSS-IV, indicating that the differences in the shapes of the BI distributions of the SDSS-IV and SDSS-V BALs persists, most notably in the low and high BI tails. The absence of BALs at these extremes in SDSS-V is likely due to its relatively lower S/N and X-ray selection.

increase in absorption as the CIV distance increases for the X-ray-selected BI BAL quasars in our sample, as previously noted via the composite spectra. While this trend agrees with the trend observed in the optically selected SDSS-IV sample studied by R20 (median

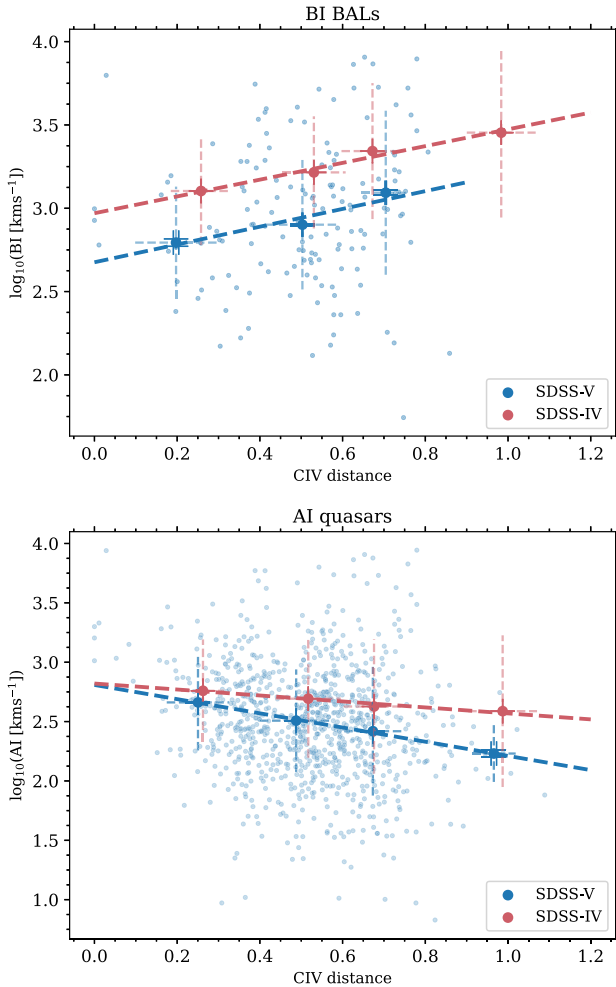


Figure 6. The relations of BI (*top panel*) and AI (*bottom panel*) against C IV distance with the blue scatter points displaying all of the BI- and AI-defined quasars. The blue dots show the median BI and AI values in C IV distance bins for the quasars identified in the SDSS-V sample (the same bins used in Fig. 4 for BI and Fig. 7 for AI). Similarly, the red dots indicate the BI- and AI-defined quasars in the optically selected SDSS-IV sample. The dashed lines show the best fit for the medians, the dashed error bars indicate the standard deviation, and the capped error bars are the standard errors in the medians. The median behaviour in the X-ray selected BI BAL (AI) quasars in our sample is such that the absorption increases (decreases) with increasing C IV distance. The SDSS-IV sample has systematically higher BI and AI values relative to the SDSS-V sample, although both samples show a high level of overlap due to the large scatter.

values shown as the red dots), we note a substantial scatter, i.e. the dashed error bars show the standard deviation of the BI values in each C IV distance bin which suggests the presence of a broad scatter. However, the median BI values are well constrained as indicated by the standard errors in the medians shown using the solid error bars, i.e. the median values increase from 2.79 ± 0.02 at 0.2 C IV distance to 3.01 ± 0.01 at 0.7 C IV distance, indicating a significant increase in the average strength of absorption within BI BAL quasars associated with strong C IV outflows. While the SDSS-IV sample has systematically higher BI values relative to the SDSS-V sample, they are consistent within the errors.

The AI metric, however, is seen to decrease with increasing C IV distance (bottom panel of Fig. 6), in contrast to the increasing BI

values with this distance. The AI definition allows the inclusion of mini-BALs (velocity widths of ~ 1000 – 2000 km s^{-1} ; Trump et al. 2006) as well as strong narrow absorption lines (NALs: FWHM $< 500 \text{ km s}^{-1}$; e.g. Hamann & Ferland 1999; Bowler et al. 2014 – see Appendix B for representative rest frame spectra). Composite spectra of the AI quasars in the quasar rest frame (similar to Fig. 4) contained no evidence of absorption due to the narrow width and broad velocity range of the AI-defined troughs. We instead generate composite spectra in the absorber rest frame, z_{abs} as defined in equation (1) of Hall et al. (2002) [introduced in Weymann et al. (1979) & Foltz et al. (1986)]:

$$\beta = \frac{v}{c} = \frac{R^2 - 1}{R^2 + 1}; R = \frac{1 + z}{1 + z_{\text{abs}}}, \quad (4)$$

where z is the redshift of the quasar, c is the speed of light, and v is the velocity of the deepest absorption relative to the C IV line. The composites are presented in Fig. 7 and show a clear C IV NAL doublet. The absorbers producing these features have very different internal kinematics compared to absorbers that produce BI-defined absorption troughs. In addition to outflowing absorbers, some fraction are likely to be a consequence of intervening material along the line of sight to the quasar. The decreasing strength of the absorption with increasing C IV distance is also evident in the normalized spectra (bottom panels). This differing behaviour of the BI- and AI-defined troughs in C IV space was also observed for the optically selected sample in R20.

Fig. 8 shows the distribution of the maximum velocity at which the broad absorption blueward of the C IV emission line, referred to as V_{max} (in km s^{-1}), occurs in the BI BAL SDSS-IV and SDSS-V quasars. The open histograms compare the C IV blueshift-matched optically selected and X-ray selected BI BAL quasars revealing that the absorption velocities span the full range independent of the selection. The pile-up at 25000 km s^{-1} in this figure is explained by considering the BI definition in equation (1), where the upper limit on the integral causes the troughs that extend beyond this limit to be truncated. Moreover, the full SDSS-IV sample in red (shaded) matches the range of V_{max} of the SDSS-V sample, indicating that while the C IV emission velocities of the samples show dependence on selection (Fig. 3), the range of absorption velocities does not. This is further explored in Fig. 9 that shows the dependence of V_{max} on C IV distance, with both samples sharing a common trend of increasing V_{max} with distance, consistent with the trend in composite spectra as previously shown in Fig. 4. Therefore, this implies that while the extent of the C IV blueshifts differs in the optical and X-ray selected samples, i.e. the emission velocities differ, the absorption velocities appear to match with each other. The similar range of V_{max} combined with the changing median V_{max} with C IV distance suggest that the *shape* of the V_{max} distribution is changing with location in C IV space. This is a new result and applies to both the optically and X-ray-selected populations.

5 X-RAY PROPERTIES OF BAL QUASARS

In this section, we study the X-ray properties of the BAL and non-BAL quasars in our X-ray-selected sample to investigate the X-ray weakness of BAL quasars. The rest-frame 0.2 – 2.3 keV X-ray luminosities (L_X) of our sample were calculated assuming a photon index of $\Gamma = 1.9$ and using fluxes from eFEDS (43 per cent of the sample), eRASS:1 (32 per cent), *XMM/Swift* (~ 1 per cent) or *Chandra* (23 per cent) in descending order of preference for objects with matches in multiple X-ray catalogues. We adopted the fluxes measured in the 0.2 – 2.3 keV band for the *eROSITA* (eFEDS and

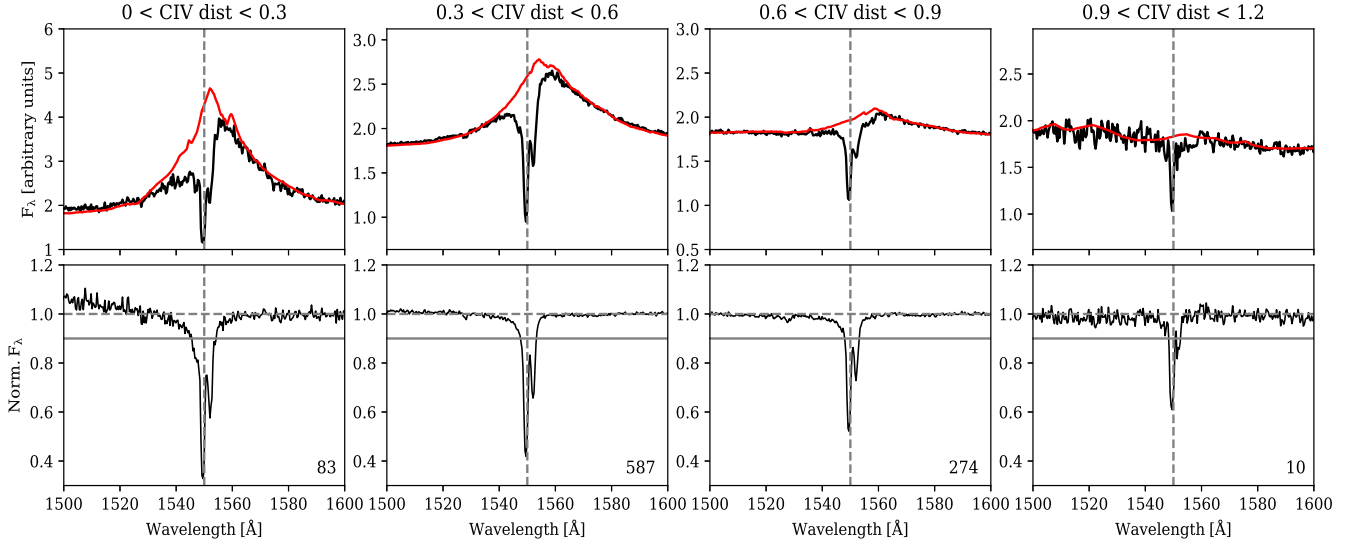


Figure 7. Same as Fig. 4 but for the AI quasars and in the rest frame of the absorber (the MAD is also omitted for clarity). The X-ray-selected AI quasars show the presence of NALs that decrease in strength at higher CIV distance (see also bottom panel of Fig. 6).

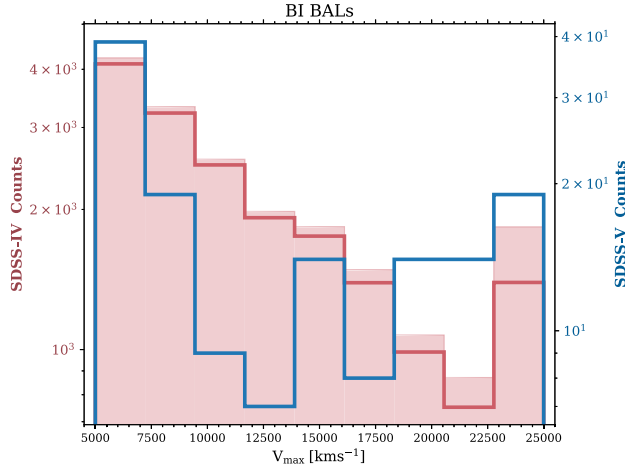


Figure 8. Comparing the maximum velocity at which the C IV absorption or the $V_{\max}$ occurs for the BI BAL quasars in the X-ray selected SDSS-V sample (in blue) and the optically selected SDSS-IV sample (in red). The red histogram shows the full SDSS-IV sample. The step-style histograms compare $V_{\max}$ of the CIV blueshift matched SDSS-IV and SDSS-V samples, i.e. BI BAL quasars between $-1000 \text{ km s}^{-1} \leq \text{CIV blueshift} \leq 3000 \text{ km s}^{-1}$. These BI BALs show that regardless of optical/X-ray selection BAL absorptions occur at a full range of velocities, in contrast to the BAL C IV emission line velocities or blueshifts that show dependence on the wavelength selection.

eRASS:1) sources, the 4.5–12 keV band for the sources identified by *XMM*, the 2–10 keV band for the sources from *Swift*, while for *Chandra* sources we use fluxes measured in the ACIS broad band (0.5–7 keV), medium band (1.2–2 keV), soft band (0.5–1.2 keV), and HRC broad band (0.1–10 keV) as provided in the *Chandra* Source Catalogue (Evans et al. 2024). See Table 2 for a summary. The catalogue fluxes were calculated assuming Galactic absorption values of $N_{\text{H}} = 3 \times 10^{20} \text{ cm}^{-2}$ for *eROSITA* sources (Brunner et al. 2022; Merloni et al. 2024) and *XMM* (Rosen et al. 2016), and $4 \times 10^{20} \text{ cm}^{-2}$ for *Swift* sources (Delaney et al. 2023). Galactic N_{H} was estimated for each *Chandra* source individually (Evans et al. 2024).

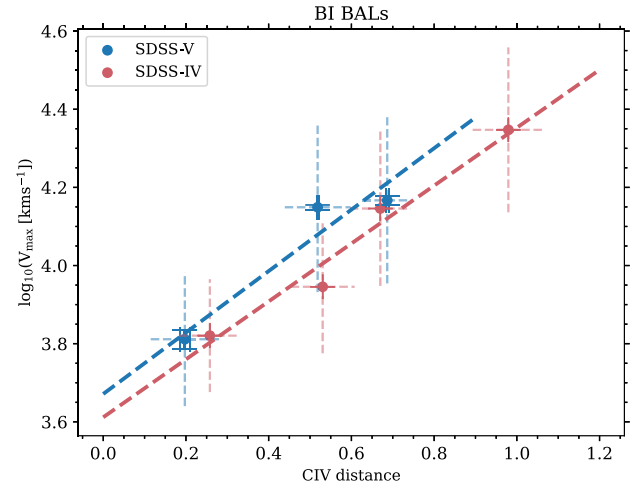


Figure 9. The median $V_{\max}$ is observed to increase with increasing CIV distance for both the BI BAL SDSS-V quasars (in blue) and SDSS-IV (in red). Where the error bars are the same as in Fig. 6 and the median $V_{\max}$ is determined in CIV distance bins of size $\Delta \text{CIV distance} = 0.3$ between $0 \leq \text{CIV distance} \leq 1.2$. The dashed lines show the best fit for the medians. This increase in $V_{\max}$ with CIV distance suggests that while the extent of the C IV blueshifts, i.e. the emission velocities of the optically and X-ray samples differ (the BI BALs in the SDSS-V sample do not have sources in the last CIV distance bin, see Fig. 3), the absorption velocities are relatively consistent at a given CIV distance.

Table 2. Origin of the X-ray data in our sample.

Instrument	Observed energy band keV	Number of sources
<i>eROSITA</i> (eFEDS)	0.2–2.3	959
<i>eROSITA</i> (eRASS:1)	0.2–2.3	716
<i>XMM</i>	4.5–12	10
<i>SWIFT</i>	2–10	12
<i>Chandra</i>	0.5–7 (ACIS broad)	508
	1.2–2 (medium)	14
	0.5–1.2 (soft)	1
	0.1–10 (HRC broad)	2

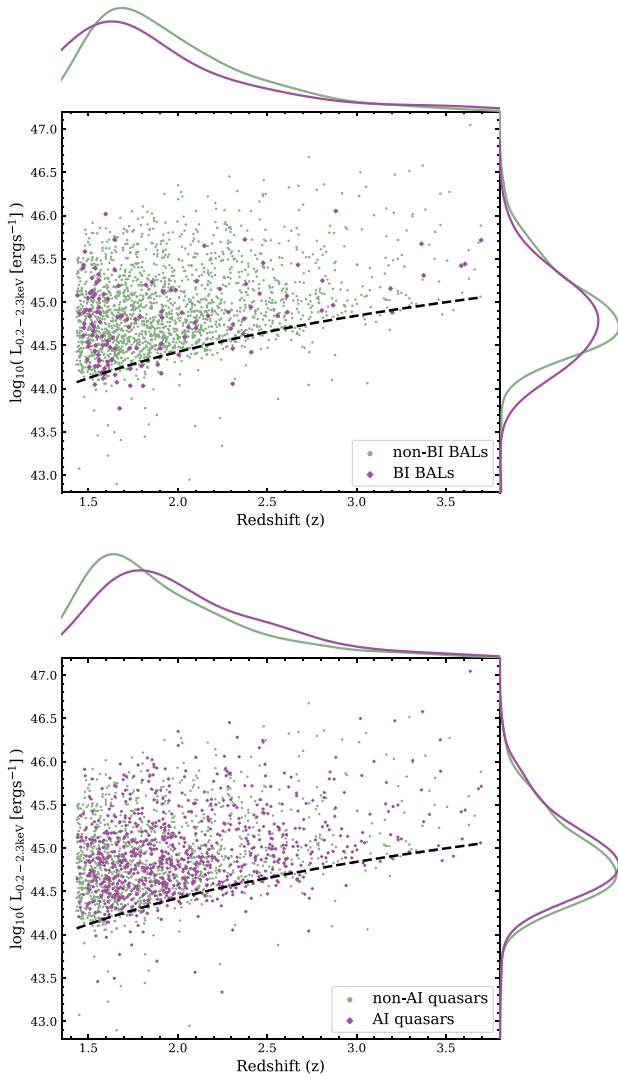


Figure 10. The observed X-ray luminosities of the BI and AI populations. The purple diamonds represent the BI BAL and AI quasars and the green dots represent the non-BAL quasars, along with the black dashed line showing the luminosity at the limiting flux (adopting the eFEDS flux limit of 10^{-14} erg s $^{-1}$ cm $^{-2}$; Brunner et al. 2022). Note that the size of BI BAL points has been increased for visibility. The distribution of the X-ray luminosities of all the subpopulations is similar over redshift, consistent with the previous result of finding these populations to be overlapping in the C IV space (see Fig. 3). The X-ray selected BI BAL quasars do not appear X-ray weaker relative to the non-BALs, i.e. by construction, our sample is probing the X-ray bright BAL quasars.

Fig. 10 shows that the L_X distribution of the BI- and AI-defined quasars over redshift is similar to the non-BAL quasars, suggesting that the X-ray selected BI BAL and AI quasars do not appear to be X-ray weaker relative to the non-BAL quasars in the sample studied in this paper (cf. Green et al. 1995; Laor et al. 1997; Brandt et al. 2000; Gallagher et al. 2006, G09; Luo et al. 2014; Saccheo et al. 2023). This similarity in the L_X values may be due to the fact that we are working with an X-ray-selected sample and thus our selection is biased towards the X-ray brightest BAL quasars. Thus, our sample – by construction – will only probe the X-ray bright tail of the underlying BAL quasar population. The majority of the objects

below the eFEDS flux limit are from the deeper *Chandra* sample.³ The similar X-ray luminosities of the X-ray selected BAL and non-BAL quasars are consistent with our finding that the distribution of these two populations in the C IV space are also similar (see Fig. 3 and Section 4), suggesting that – for an X-ray-selected sample – the BAL and non-BAL quasars are indistinguishable in their characteristics except for the broad (or narrow) absorption troughs that identify the BI BAL and AI quasars. Additionally, the KS test p -values presented in Table 1 for L_X and $L_X - z$ space of the BI/AI populations, also imply that the BALs and non-BALs in our sample are drawn from the same underlying distribution.

This result motivates us to explore in further detail the observed X-ray properties of the BAL quasars relative to non-BALs. Hence, we employ α_{ox} to investigate the observed lack of difference in L_X values of these two sub-populations. As discussed in Section 1, the UV/optical and X-ray luminosities of AGN are related (e.g. Avni & Tananbaum 1982), and this relationship is often quantified using the spectral slope, α_{ox} , between the optical and X-ray as

$$\alpha_{ox} = \frac{\log_{10} \left(L_{2 \text{ keV}} / L_{2500 \text{ \AA}} \right)}{\log_{10} \left(\nu_{2 \text{ keV}} / \nu_{2500 \text{ \AA}} \right)}, \quad (5)$$

where $L_{2500 \text{ \AA}}$ is the monochromatic optical luminosity at 2500 Å and $\nu_{2500 \text{ \AA}}$ is the frequency corresponding to this wavelength, while $L_{2 \text{ keV}}$ and $\nu_{2 \text{ keV}}$ are, equivalently, the monochromatic X-ray luminosity at 2 keV and corresponding frequency. The monochromatic X-ray luminosities were calculated by converting full-band luminosities assuming a photon index of $\Gamma = 1.9$ (e.g. Nandra et al. 2007). Fig. 11 shows the calculated α_{ox} as a function of $L_{2500 \text{ \AA}}$ for the BI- and AI-defined sub-populations in our X-ray-selected sample. All quasars in our sample show the expected anticorrelation between $L_{2500 \text{ \AA}}$ and α_{ox} , which suggests that while optically brighter quasars are more X-ray luminous, overall, the *relative* strength of the X-ray emission (compared to the optical emission) decreases at higher optical luminosities. While the contours for the BI BAL and AI quasars show that these quasars defined in our X-ray selected sample overlap with each other, they also overlap with the non-BAL quasars consistent with our previous result finding these subpopulations having similarly distributed L_X . This result further suggests that the X-ray selected BI BAL quasars show a lack of X-ray weakness relative to the non-BAL quasars with all of the subpopulations showing similar UV/optical and X-ray properties. This conclusion is further motivated by the p values presented in Table 1 for a KS test comparing the α_{ox} distributions.

α_{ox} is a function of $L_{2500 \text{ \AA}}$ for quasars and Just et al. (2007) found this correlation to be

$$\alpha_{ox}(L_{2500 \text{ \AA}}) = (-0.140 \pm 0.007) \log(L_{2500 \text{ \AA}}) + (2.705 \pm 0.212), \quad (6)$$

for a sample of optically bright SDSS (York et al. 2000) quasars within the redshift range of $1.5 \leq z \leq 4.5$ with additional complementary high luminosity quasars with $z \geq 4$, and this correlation is shown as the black dashed line in Fig. 11. This figure also shows a more recent linear anticorrelation equation found by Timlin et al. (2021) that made use of the same sample as Just et al. (2007) with additional restrictions of requiring the quasars to be within $1.6 \leq z \leq$

³We adopt the eFEDS flux limit as the bulk of our X-ray selected sample were identified with *eROSITA* (from eRASS:1 or eFEDS catalogues) and the eFEDS flux limit corresponds to the deeper portion of this imaging.

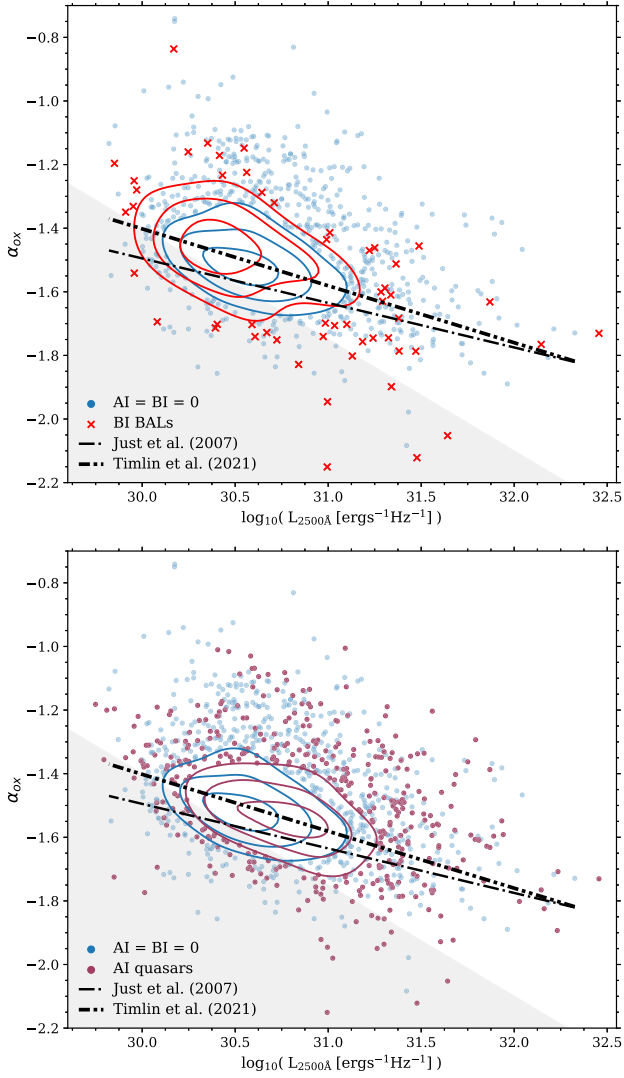


Figure 11. α_{ox} versus L_{2500} for our sample of X-ray selected SDSS-V quasars with the non-BALs (AI = BI = 0) in blue, the BI BALs in red (*top panel*), and the AI quasars in dark red (*bottom panel*). The contours are defined at 10 per cent, 30 per cent, and 50 per cent of each of the samples. The BI BAL sub-population overlaps with the AI = BI = 0, suggesting that the X-ray weakness of a quasar is independent of its categorization as a BAL or non-BAL. Similarly, the AI quasars overlap with the AI = BI = 0 defined non-BAL population. The grey shaded region shows the eFEDS flux limit, with the Just et al. (2007) and Timlin et al. (2021) relations as the black dotted/dashed lines. The relations have slopes shallower than the sample studied in this paper, with the flux limit of our sample introducing a tilt in its distribution and missing the X-ray weaker (i.e. lower α_{ox}) population at lower optical luminosities.

3.5, be radio-quiet, and devoid of BAL features:

$$\alpha_{\text{ox}}(L_{2500\text{\AA}}) = -0.179\log(L_{2500\text{\AA}}) + 3.968. \quad (7)$$

Fig. 11 demonstrates that the Just et al. (2007) and Timlin et al. (2021) slopes are much shallower relative to the sample studied in this paper due to their differing sample selection. In this paper, however, we focus on comparing the properties of BALs and non-BALs in our new X-ray-selected sample. The grey-shaded region in Fig. 11 indicates the impact of the eFEDS flux limit in this space, evaluated at the redshift of our full quasar sample ($1.5 \leq z \leq 3.5$) and thus how

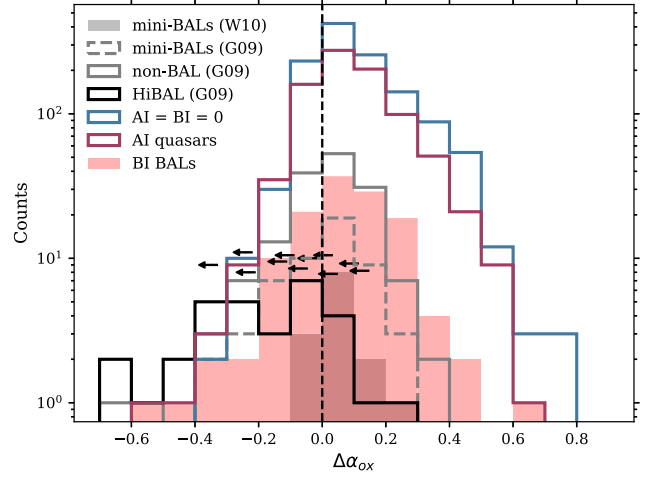


Figure 12. The distribution of $\Delta\alpha_{\text{ox}}$ for the BI BALs, AI quasars, and non-BAL quasars (following the same colour scheme as in Fig. 11), with the vertical dashed line indicating where $\Delta\alpha_{\text{ox}} = 0$ occurs. The grey histogram shows the optically bright mini-BALs studied by W10. The optically selected HiBALs, mini-BALs, and non-BAL samples studied by G09 are shown as the black, dashed grey, and grey step-style histograms, respectively. The arrows indicate the 1σ upper limits of the non-detections in the G09 sample (the y-coordinates of these arrows are arbitrary). The relatively high $\Delta\alpha_{\text{ox}}$ values of our sample are expected since X-ray selection biases us towards X-ray brighter sources at the optical luminosities probed by SDSS-V. The BI BAL quasars peak around $\Delta\alpha_{\text{ox}} = 0.2$ and have a broad distribution extending between $-0.5 \lesssim \Delta\alpha_{\text{ox}} \lesssim 0.6$. These findings indicate that we probe the high $\Delta\alpha_{\text{ox}}$ region more fully and identify a relatively X-ray bright BAL population.

the X-ray selection will miss the X-ray weaker (i.e. lower α_{ox}) quasar population at lower optical luminosities. Again, the majority of the sources below the flux limit are from the *Chandra* sample. The flux limit also introduces a tilt in the distribution of our samples in this space that leads to a deviation in the *apparent* slope of the $\alpha_{\text{ox}}-L_{2500\text{\AA}}$ sample compared to the underlying relations measured by Just et al. (2007) and Timlin et al. (2021). By measuring the offset from these correlations,

$$\Delta\alpha_{\text{ox}} = \alpha_{\text{ox}} - \alpha_{\text{ox}}(L_{2500\text{\AA}}), \quad (8)$$

where $\alpha_{\text{ox}}(L_{2500})$ is the predicted α_{ox} at a given L_{2500} based on the Just et al. (2007) best-fitting relation, we can compare the relative X-ray brightness across the full optical luminosity range between the X-ray selected BAL quasars and the non-BAL quasars. A lower $\Delta\alpha_{\text{ox}}$ value implies X-ray weakness, i.e. if $\Delta\alpha_{\text{ox}} = -0.4$, then the quasar is X-ray weaker by a factor of ≈ 11 with respect to quasars with similar UV luminosities (G09). We define X-ray strong as $\Delta\alpha_{\text{ox}} > 0.2$ and X-ray weak as $\Delta\alpha_{\text{ox}} < -0.2$. Fig. 12 shows the distribution of $\Delta\alpha_{\text{ox}}$ for all the defined subpopulations relative to the Just et al. (2007) relation. The relative skew of the $\Delta\alpha_{\text{ox}}$ distribution towards more positive values is expected as the X-ray selection introduces an X-ray flux limit that biases us towards the X-ray brighter population at the optical luminosities probed with SDSS-V. However, our X-ray selection does also identify X-ray weak populations as ~ 7 per cent of the sample has $\Delta\alpha_{\text{ox}} < -0.2$. This figure demonstrates that we find both X-ray weak, X-ray normal,⁴ and X-ray strong populations within our X-ray selected sample. 21/25 of the $\Delta\alpha_{\text{ox}} < -0.2$ sources are from *Chandra*.

⁴We define ‘X-ray normal’ here as $-0.2 < \Delta\alpha_{\text{ox}} < 0.2$.

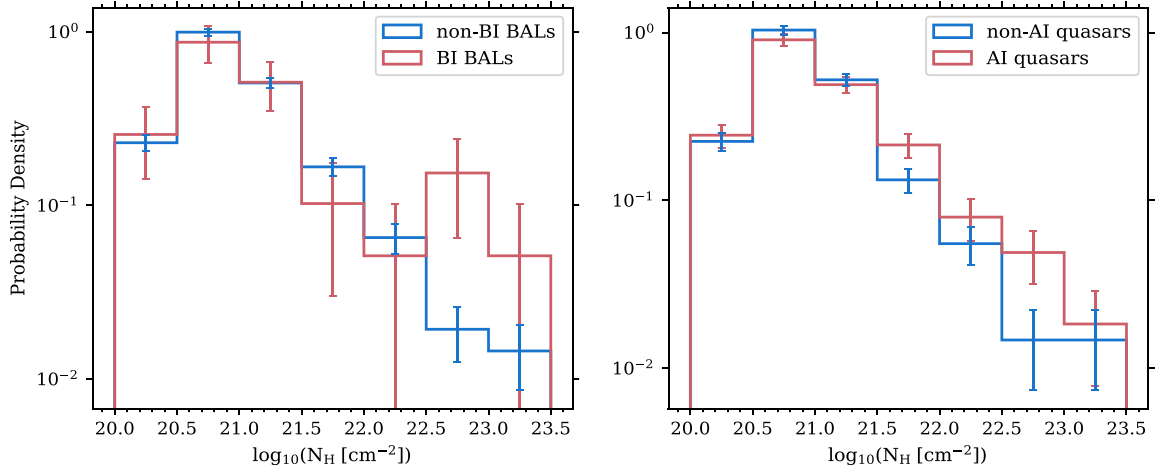


Figure 13. The distribution of column densities or N_H (obtained from Liu et al. 2022) of the four subpopulations in the eFEDS sample where we have X-ray spectral information, with the lowest bin in each panel representing the unobscured fraction. While noting the overall small counts of the populations and the simplistic X-ray spectral model considered, the BAL and non-BAL quasars appear to be similarly distributed, i.e. the BAL quasars do not have higher column densities relative to the non-BAL quasars.

Fig. 12 also presents the optically selected HiBALs studied by G09 in black, where they have also calculated their $\Delta\alpha_{\text{ox}}$ values relative to the Just et al. (2007) relation. The majority of the G09 HiBAL sample has $\Delta\alpha_{\text{ox}} < 0$ although it does include a small number of objects with $\Delta\alpha_{\text{ox}} > 0$. However, the $\Delta\alpha_{\text{ox}}$ distribution of the G09 HiBALs does not align with the BI BAL quasar population isolated from our sample due to the differences in selection. Additionally, we observe that isolating the BAL quasars from the rest of the sample introduces a slight skew in our $\Delta\alpha_{\text{ox}}$ distribution towards lower values relative to the non-BALs, although the BALs still show a large fraction of $\Delta\alpha_{\text{ox}} > 0$ objects and many are X-ray stronger than those in the G09 sample. In Table 1, the p -value for a KS test comparing the $\Delta\alpha_{\text{ox}}$ indicates that we should accept the hypothesis that the BI or AI-defined quasars and non-BALs are drawn from the same distribution. Gibson et al. (2009b) and W10 found that the absorption systems they labelled as mini-BAL quasars had X-ray properties more similar to the non-BALs than to the BALs. Our BI BALs and AI quasars do not appear to be X-ray weaker than the non-BALs and there is insufficient evidence to suggest that the X-ray-selected BALs preferentially show substantial X-ray weakness relative to non-BALs. While we do find X-ray weak objects, the lack of evidence of relative BAL X-ray weakness is likely due to the selection bias of our X-ray quasar sample that prefers X-ray strong quasars. The $\Delta\alpha_{\text{ox}}$ values for our sample were also computed relative to the Timlin et al. (2021) relation, obtained for an optically selected sample, finding the distributions of all the subpopulations agreed with Fig. 12.

Liu et al. (2022) performed X-ray spectral fitting of the eFEDS sample in order to derive various properties including the intrinsic column density, N_H , of the sources; however, note that the absorption model is likely too simplistic for BAL X-ray absorption. Fig. 13 presents these column densities for our eFEDS subsample of quasars. We include in the figure those that are classified as ‘unobscured’, ‘mildly-measured’, and ‘well-measured’ via the neutral absorption power-law model.⁵ We do not use the tabulated N_H values for the unobscured population; all of these objects are placed in the lowest

N_H bin (see section 4.3 of Liu et al. 2022 for details). As shown in the figure, the N_H values for the BI BAL and non-BAL quasars as well as the AI and non-AI quasars are similar while noting the marginally higher number of BAL quasars with higher N_H values. The four BI BAL quasars with $N_H > 10^{22.5} \text{ cm}^{-2}$ have a median (mean) $\Delta\alpha_{\text{ox}} \simeq -0.091$ (-0.119) placing them in the weaker half of the X-ray normal population. Future larger samples with detailed X-ray spectral analysis are required to investigate further, but the $N_H > 10^{22.5} \text{ cm}^{-2}$ sample could be indicative of the absorbed X-ray weak BAL population observed in previous optically selected studies.

6 DISCUSSION

The spectroscopic follow-up of *eROSITA*, *Chandra*, *XMM*, and *Swift* sources with SDSS-V provides a large sample of X-ray selected quasars to investigate BAL quasars relative to non-BALs. This synergy has enabled, for the first time, to study a large sample of X-ray selected BAL quasars, i.e. 143 BI BALs and 954 AI quasars identified in this paper. Previous studies that investigated the X-ray-selected BAL quasars include Giustini et al. (2008) who studied 54 sources, Streblyanska et al. (2010) who studied 88 sources, and Page et al. (2017) who studied six sources. Our increase in sample size allows us to probe a broader dynamic range of BAL physical properties and explore trends with other signatures of outflowing gas (e.g. C IV blueshift). We have also consistently selected a large sample of non-BAL quasars that provides a robust comparison to the BALs and their properties. In the following sections, we discuss the implications of the results obtained in this paper and contribute to the discussions of these previous studies.

6.1 X-ray selection excludes the highest outflow velocities in emission but spans the full range of absorption velocities

As discussed in Section 1, previous studies that investigated optically selected quasars have found BAL quasars to be X-ray weaker relative to non-BAL quasars. Before comparing the BAL quasars with the non-BAL quasars in our study, let us first consider the impact of the

⁵Model 1 of Liu et al. (2022).

X-ray selection on the properties of our sample relative to those of the optically selected comparison sample from SDSS-IV and what it could mean for the nature of the winds.

The C IV emission blueshift is often considered as evidence for disc winds [e.g. Gaskell 1982; Sulentic et al. 2000; Richards et al. 2002, 2011; Leighly 2004; but see Gaskell & Goosmann (2013, 2016) for an alternative view]. The distribution of quasars within the C IV space alongside its correlation with the He II EW (a tracer of the strength of the far-UV SED; Leighly 2004) point towards radiative line driving as being an important mechanism for driving strong winds (Baskin et al. 2013, 2015; Rankine et al. 2020; Temple et al. 2023). In detail, radiatively efficient winds occur when the outflowing gas can be propelled to high velocities by radiation pressure caused by the accretion disc's UV emission; both physical arguments and numerical simulations support the notion that radiation-driven outflows are feasible only when the gas surrounding the disc is not highly ionized since highly ionized gas would have a low concentration of ions that are capable of producing the observed UV line opacities [Proga 2007; Matthews et al. 2023; but see Baskin et al. (2014) for a different view]. Through X-ray selection, we are missing the quasars with the highest C IV emission blueshifts ($\gtrsim 3500 \text{ km s}^{-1}$), which is consistent with the scenario where X-rays can overionize the gas resulting in lower velocity outflows. Therefore, through X-ray selection, we preferentially probe the high X-ray luminosity and low UV emission velocity part of the quasar population. R20 reported that the observed BAL fraction increases with increasing C IV blueshift for a given C IV EW (without correcting for selection effects) thus by consequence we may be missing a large fraction of the BAL population (in agreement with Giustini et al. 2008). Since we are not probing the whole BAL population through our X-ray sample, i.e. there are no high C IV blueshift quasars in this sample, any conclusions drawn can only be applied to the low-to-moderate C IV blueshift BAL/non-BAL quasar populations.

Despite the effect of X-ray selection on the occupation of the C IV emission space, in Section 3 we found the overall fraction of BAL quasars (we find 6 per cent for our sample) is comparable to the observed fractions in optically selected samples, that typically find 10 per cent to 15 per cent (e.g. Tolea et al. 2002; Hewett & Foltz 2003; Knigge et al. 2008, G09). While our observed BAL fraction is slightly lower than found in optically selected samples, it remains a surprising result given that BAL quasars are generally thought to be X-ray weak relative to non-BALs (Green et al. 1995; Laor et al. 1997; Brandt et al. 2000; Gallagher et al. 2006, G09; Luo et al. 2014; Saccheo et al. 2023). However, while the BAL fraction has been observed to increase with C IV blueshifts, most quasars (and thus most BALs) are at lower C IV blueshifts, and excluding the highest C IV blueshifts via X-ray selection does not lead to a large change in the overall BAL fraction.

We also find that while X-ray selection preferentially identifies sources with relatively low-velocity disc winds (as traced by the C IV emission line), the velocities of the *absorbing* gas span the full range of BAL velocities seen in the optically selected BALs (see Fig. 8). Furthermore, at a given C IV distance (i.e. emission line velocity), the velocities of the absorbing gas for both X-ray and optically selected BALs are consistent (see Fig. 9). Thus, the presence of absorbing gas and its properties appear to be consistent, regardless of the selection wavelength.

The increase in BAL velocity with increasing C IV distance (Fig. 9) is expected within a radiatively driven wind scenario since more line transitions within the gas lead to both stronger absorption features as well as the propulsion of the gas to higher velocities. However, the similar absorption velocities between the X-ray and optically

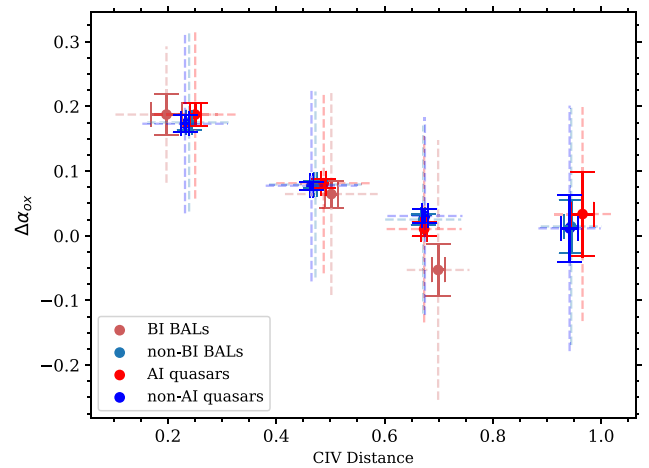


Figure 14. $\Delta\alpha_{\text{ox}}$ versus C IV distance for all the subpopulations in our X-ray selected sample, with each point showing the median $\Delta\alpha_{\text{ox}}$ value in different C IV distance bins. While the rest of the subpopulations have four bins, the BI BAL quasars are not found in the last bin with the highest C IV distance. The error bars are similar to Fig. 6. The trend suggests a decrease in median $\Delta\alpha_{\text{ox}}$ value with increasing C IV distance for all but the highest C IV distance bin, although the error bars are large – quasars appear to be X-ray weaker at higher C IV distances. This trend is observed for all the subpopulations independent of their categorization as a BAL or non-BAL. The trend suggests that, for a given quasar, low C IV EWs and high C IV blueshifts are associated with relatively X-ray-weak systems.

selected samples indicate that the X-rays may have a relatively larger impact on the C IV emission line compared to the absorption. This behaviour is seen again in Fig. 14 which shows the trend observed between $\Delta\alpha_{\text{ox}}$ and C IV distance. The median $\Delta\alpha_{\text{ox}}$ value decreases with increasing C IV distance for all the subpopulations, implying that systems with higher velocity C IV-emitting outflows are X-ray weaker relative to those with lower velocity outflows. A similar anticorrelation between $\Delta\alpha_{\text{ox}}$ and C IV distance was observed by Timlin et al. (2020b) and Rivera et al. (2022). It was also an expected trend considering the relationship between He II emission and C IV blueshift (see Section 1 for references). Timlin et al. (2021) have also reported a strong correlation between α_{ox} and He II across broad redshift and luminosity ranges. Importantly for this paper, R20 found no difference between the BALs and non-BALs in regards to the He II distributions in the C IV space which aligns with the similar trends observed here between the X-ray selected BALs and non-BALs: there is little difference in the $\Delta\alpha_{\text{ox}}$ trends with C IV distance between the BAL and non-BAL quasars.

W10 found correlations between various trough parameters and $\Delta\alpha_{\text{ox}}$ (see their figs 9–11) with slower and narrower absorption features present in objects with stronger X-ray emission. In Fig. 15, we investigate how the absorption properties relate to the X-ray strength for our sample and compare with some of the results from fig. 9 of W10. Unlike W10 we do not see any correlations for our X-ray selected sample. We caution, however, that the X-ray flux limit may be affecting any correlation by removing sources at low $\Delta\alpha_{\text{ox}}$. Many of the W10 sample fall below our $\Delta\alpha_{\text{ox}}$ limit (see the vertical black dashed lines based on the median flux limit at the sample's median L_{2500}). However, our new sample illustrates the broad range of values that these measures of absorption properties span for our sample of sources at $\Delta\alpha_{\text{ox}} \gtrsim -0.1$, and that these properties do not appear to depend on the X-ray properties. This does not preclude a mild dependence for samples that span a broader range of $\Delta\alpha_{\text{ox}}$ but does suggest there must be a large scatter in any such relation. While

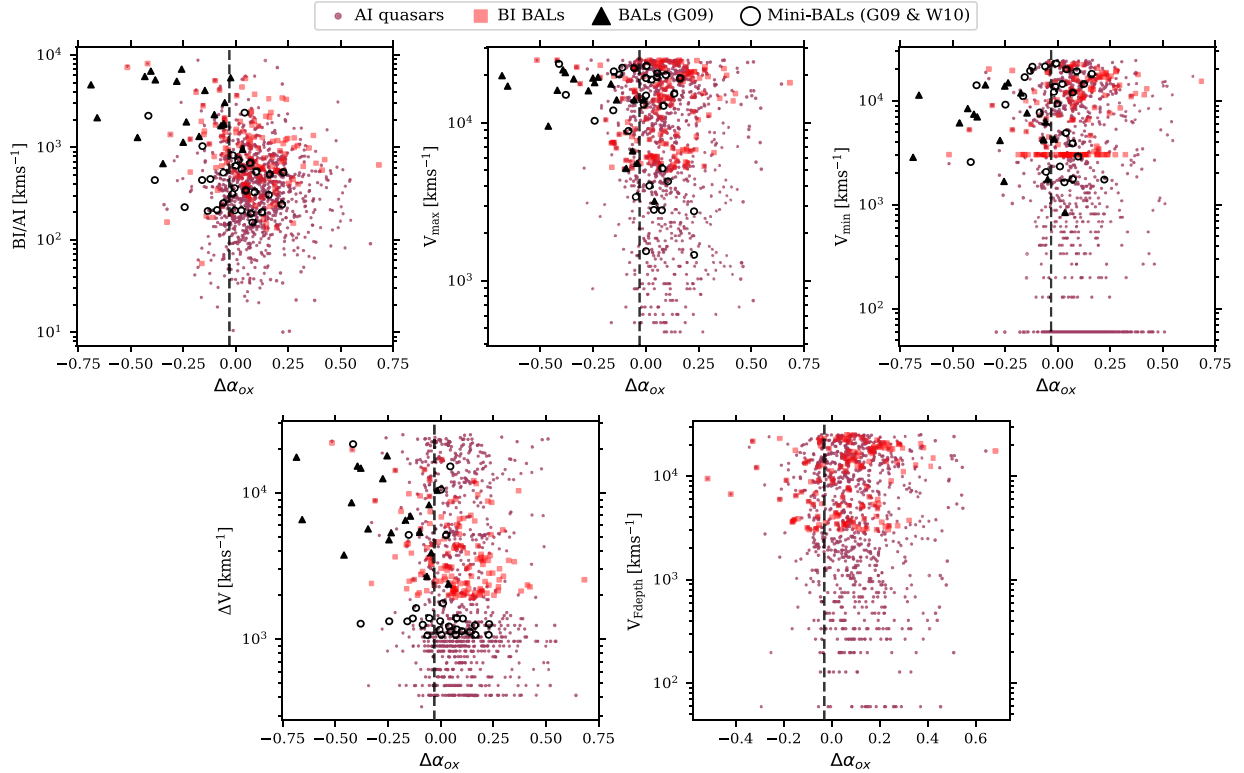


Figure 15. Various absorption properties as a function of $\Delta\alpha_{\text{ox}}$ for the BI and AI absorption systems. The coloured dots (AI quasars) and squares (BI BALs) are from our sample. The black triangles are HiBALs from G09 and open circles are mini-BALs from G09 & W10, see fig. 9 of W10 for further details. From top left: BI or AI value [as defined by equations (1) and (2)], maximum trough velocity, minimum trough velocity, velocity extent of absorption ((maximum – minimum) trough velocity which is the width of the trough for objects with one absorption feature), and velocity of deepest absorption. The velocity measurements show discretization due to the finite width of the SDSS pixels. The build-up of objects at the limits of the BI and AI integration is most evident in the V_{min} panel (top right). The vertical dashed line in each panel shows the typical $\Delta\alpha_{\text{ox}}$ limit at the median X-ray sensitivity and median optical luminosity. The samples studied by W10 fall below our $\Delta\alpha_{\text{ox}}$ limit and we do not find any clear trends within our X-ray flux limited sample, suggesting that these absorption properties do not depend on the X-ray properties for sources in our sample.

we do not see a correlation between BI and $\Delta\alpha_{\text{ox}}$ (top-left panel), in Fig. 5 we see that the X-ray selected nature of our sample is likely to have some effect on the types of BALs we are observing, since the BI distribution of the X-ray selected sample is missing the strongest BALs. However, apart from at the highest and lowest BI values, the BI distributions are very similar. The lack of a significant connection between the X-ray properties and the presence and properties of any absorption (either BI or AI troughs) furthers the conclusion that the wind signatures in emission are affected more than the wind signatures in absorption. These results could be a manifestation of the different physical locations of the C IV-emitting gas and the C IV-absorbing gas. The emission line is produced from gas much closer to the X-ray corona (at sub-parsec scales) compared to the absorption that is thought to occur further from the central engine (at hundreds to thousands of parsecs; Arav et al. 2018, 2020; Xu et al. 2019; Miller et al. 2020). However, Extremely High Velocity Objects (EHVOs) are being found at distances closer to the broad emission line region (e.g. Rodríguez Hidalgo, Hamann & Hall 2011; Hamann et al. 2013; Rodríguez Hidalgo et al. 2025). Via the study of NALs, Bowler et al. (2014) and Perrotta et al. (2018) have shown that the ionization parameter (i.e. the number of ionizing photons per atom) of the outflowing gas decreases, and therefore distance from the quasar increases, with increasing absorber velocity (in BALs and non-BALs alike; Bowler et al. 2014). By this logic, the faster

the BALs, the further out they may be where the impact of the X-ray emission is less pronounced, thereby reducing any correlation between the velocity of the wind and $\Delta\alpha_{\text{ox}}$. The AI population is likely also to be contaminated by intervening absorbers that will not be influenced by the X-ray properties of the background quasar.

To explore further any relationship (or lack thereof) between the velocity of the absorption systems and the X-ray properties, we could extend this study to systems with even larger absorption speeds. EHVOs (Rodríguez Hidalgo et al. 2011) extend the speed baseline as they are defined with minimum absorption speeds of $30\,000\text{ km s}^{-1}$. Most of these systems are found, due to the convenience of avoiding the Ly α forest, with maximum speeds of $60\,000\text{ km s}^{-1}$, although some exceptions with larger speeds have been found (Hamann et al. 2018; Hall et al., in preparation). EHVOs also show, on average, larger speeds of C IV-emitting outflows than BALs (Rodríguez Hidalgo & Rankine 2022). Thus, studying the X-ray properties of EHVOs would allow us to test whether these systems are X-ray weaker relative to those with lower velocity outflows, both in emission and absorption. Testing the X-ray properties of EHVOs as a function of C IV blueshifts would allow us to test whether X-rays have, as appears to be in the case of BALs, a larger impact on emission than absorption properties.

6.2 BAL and non-BAL quasars span the same range of X-ray properties

With X-ray information available for the BAL and non-BAL quasars in our sample, we have investigated the similarity in the properties derived from their X-ray data. Previously studied optically selected samples have observed BAL quasars to have lower observed X-ray luminosities relative to non-BAL quasars, which is well-explained within the accretion disc–wind model that invokes X-ray shielding (Murray et al. 1995; Proga & Kallman 2004; Luo et al. 2013) to explain the X-ray weakness observed in BAL quasars.

One key result of our paper is that we find BAL quasars across the full range of relative X-ray strength probed by our sample (i.e. $\Delta\alpha_{\text{ox}}$; Fig. 12), and we do still identify moderately X-ray weak systems despite the bias due to the X-ray flux limits. Previous studies have also shown that HiBALs can be X-ray weaker than found in our sample (e.g. reaching $\Delta\alpha_{\text{ox}} \sim -0.7$; Gallagher et al. 2006) and thus BALs exist across the full range of relative X-ray strength of quasars. X-ray strong BALs are relatively rare and strong BAL features rare also. In an optically selected flux-limited sample one is biased towards optically brighter sources where X-ray strong objects are rarer still (see the lack of objects towards the top-right of Fig. 11). In our X-ray selected sample, we are biased towards X-ray strong objects, particularly at low optical luminosities (see where the flux limit and α_{ox} relations intersect). Given the previous sample sizes of BALs with X-ray observations and the expected rarity of X-ray strong BALs, our sample selection is more likely to include these objects.

We raise two questions that others have also posed: (1) Is the X-ray weakness in the BAL quasars that *do* have $\Delta\alpha_{\text{ox}} < 0$ a consequence of intrinsic X-ray weakness caused by a dim X-ray corona (relative to the disc) or due to large column densities of absorbing material along the line-of-sight? (2) Why in an X-ray strong source do we still see BALs if the X-ray emission is likely to overionize the wind? We discuss these questions with the help of Fig. 16 where we split the BI and AI quasars into X-ray weak ($\Delta\alpha_{\text{ox}} < -0.2$) and strong ($\Delta\alpha_{\text{ox}} > 0.2$) and plot their N_{H} distributions.⁶ While the sample sizes are small and the error bars large, all of the X-ray strong BI BALs have $N_{\text{H}} < 10^{21.5} \text{ cm}^{-2}$ and there is tentative evidence (one object) that the X-ray weak BI BALs are more likely to have larger column densities than the X-ray strong BI BALs. It should be noted that the X-ray normal BI BAL population contains objects up to $N_{\text{H}} \sim 10^{23.5} \text{ cm}^{-2}$. This result suggests that the X-ray weak BALs in our sample are X-ray weak due to large column densities along the line of sight and are not intrinsically X-ray weak. This absorption scenario is favoured in most cases (e.g. Gallagher et al. 2006).

Meanwhile, the majority of the X-ray weak AI quasars have $N_{\text{H}} < 10^{21.5} \text{ cm}^{-2}$, similar to the X-ray strong AI quasars. This result suggests that an enhanced absorbing column is not the cause of the X-ray weakness in the AI quasars.

When combining the X-ray weak, normal, and strong eFEDS samples, the p values for the column densities and the redshift-matched observed (not corrected for obscuration) X-ray luminosities (see Table 1) suggest that the intrinsic X-ray properties of our BAL and non-BAL quasar samples are drawn from the same distribution. This analysis is presented for a spectral model with a neutral absorber; however, Streblyanska et al. (2010) find a more significant fraction of absorbed BAL quasars when X-ray spectral fitting was performed with an ionized absorber model. Nevertheless, we determined the intrinsic luminosity p values for the eFEDS subsamples (similarly

as in Section 4 and using luminosities from Liu et al. 2022), finding the BI BALs/non-BALs are consistent with being drawn from the same underlying distribution ($p = 0.6$) while the p -value ($= 0.002$) rejects this hypothesis for the quasars with and without AI-defined absorption systems. This is in contrast to Liu et al. (2018) who found a higher percentage of HiBAL quasars that are *intrinsically* X-ray weak (7–10 per cent) as compared to the percentage of non-BAL quasars (< 2 per cent; Gibson, Brandt & Schneider 2008). We note, however, that eFEDS is biased towards less obscured objects and so the X-ray selection may be selecting against the obscured BAL quasars resulting in observed similarities between the X-ray properties of the BAL and non-BAL quasars that are not intrinsic to the whole population.

Of course, our line of sight may not be the same as the line of sight between the BAL gas and the inner corona, so geometry may play a role. Giustini et al. (2008) also observed a lack of X-ray weakness in a population of X-ray-selected BAL quasars, albeit a sample numbering only 54. They argued that the lack of X-ray weakness in their BALs can be accommodated within the theoretical model of the accretion disc wind by considering that the subset of BAL quasars that are X-ray strong (X-ray selected in their case) may be being viewed from smaller angles with respect to the accretion disc than the traditional optically selected BALs. We are therefore missing sources where our line-of-sight intersects with a large amount of inner shielding gas that absorbs X-rays (resulting in the observed X-ray weakness), suppresses the ionizing X-ray radiation and results in highest velocity UV outflows (seen in C IV emission). However, as discussed in Section 1, Baskin et al. (2014) suggest RPC to address the problem of over-ionization without invoking an additional X-ray shielding gas: the radial compression of the absorbing gas by disc radiation prevents the gas from being overionized and thus BAL quasars are not required to be X-ray weak, but they can still exist. Separately, Hamann et al. (2013) observed that mini-BALs approaching speeds of $0.1\text{--}0.2c$ do not have significantly higher column densities than their lower velocity counterparts. We do not find a difference in the X-ray luminosities (observed or intrinsic) of the BI BAL and non-BAL quasars or the AI and non-AI quasars identified in our paper, and the RPC framework could explain the efficient outflows we observe without requiring an additional shielding component.

Jets can also contribute to the observed X-ray luminosity (see reviews by Worrall 2009; Hardcastle & Croston 2020). Since we are studying the X-ray selected population of SDSS-V, we do not remove quasars from our sample on the basis of their radio properties. Furthermore, our sample does not have uniform or complete radio coverage. Fourteen per cent (332 objects) of our sample is within the footprint of the second data release of the LOFAR Two-Metre Sky Survey (LoTSS; Shimwell et al. 2022) and 53 per cent (1239 objects) is within the Rapid ASKAP Continuum Survey footprint (RACS; McConnell et al. 2020). Regardless, X-ray emission from jets could be responsible for some of the X-ray strong objects in our sample. Investigating the radio properties of the SDSS-V X-ray selected BAL quasars is beyond the scope of this paper; however, we state that out of the objects with radio coverage, 125 are radio-loud, all of which have $\Delta\alpha_{\text{ox}} > 0$.⁷ To our interest 1/16 (~ 6 per cent) BI

⁶The sample is limited to the eFEDS sources as before.

⁷We match to the optical counterpart catalogue of LOFAR by Hardcastle et al. (2023), and utilize a cross-match between Legacy Survey DR10 and the RACS catalogue (Hale et al. 2021) performed using NWay (Salvato et al. 2022). We employ a radio-loud threshold of $\log R > 2$ for LOFAR and $\log R > 1.5$ for RACS sources where $R = L_{\text{radio}}/L_{2500}$. The thresholds

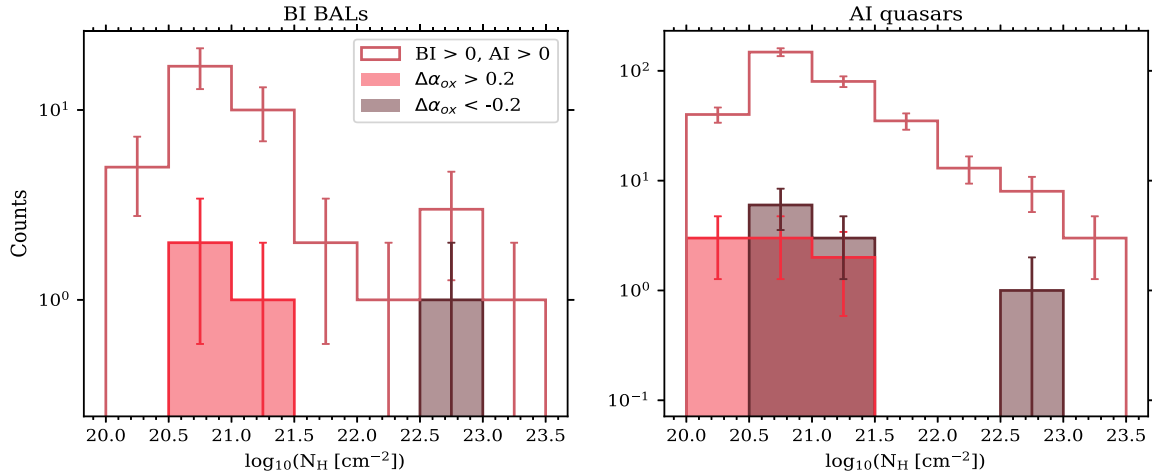


Figure 16. The distribution of column densities or N_{H} of the BI BALs (left) and AI quasars (right) where we have X-ray spectral information, similarly as in Fig. 13. The open histograms contain the whole BI/AI samples (the same red histograms as in Fig. 13). The X-ray weak subpopulations are in brown and the X-ray strong in filled red. Noting the small sample sizes and the large error bars, one X-ray weak BI BAL has a large column density whilst the X-ray strong BI BALs have lower column densities. The majority of the X-ray strong and weak AI quasars, on the other hand, have $N_{\text{H}} \leq 21.5$.

BALs, and 14/135 (~ 10 per cent) AI quasars with $\Delta\alpha_{\text{ox}} > 0.2$ and within the radio coverage are radio-loud. However, radio-loudness criteria overpredict the fraction of AGN-dominated radio sources (Yue et al. 2025) and so these fractions may be inflated. None the less, by our criteria, jets are not sufficient to explain all of the X-ray strong BI BALs and AI quasars.

Intrinsic X-ray variability may also be important to consider as the inner X-ray corona is likely to vary fairly rapidly (hours to days; Mushotzky, Done & Pounds 1993); perhaps the X-ray strong phases ($\Delta\alpha_{\text{ox}} > 0$) are too short-lived to have sufficient impact on the (further out) gas associated with the BALs. In addition to intrinsic X-ray variability, studies by Saez et al. (2012) and Timlin, Brandt & Zhu (2020a) revealed that the long-term X-ray variability (rest frame 0.1–30 yr) likely due to movement of any shielding gas is independent of the classification as a BAL or non-BAL (despite these BALs being X-ray weak) which could fit with the picture emerging here where the presence of a BAL is somewhat independent of the shielding gas. Regardless of time-scale, any X-ray variability may be responsible for the broad $\Delta\alpha_{\text{ox}}$ distribution in the first place.

Moreover, we find that the median column densities of both BALs and non-BALs are constant with increasing C IV distance suggesting the intrinsic X-ray properties do not change with C IV distance, although it is important to note that smaller numbers available for the investigation of N_{H} , the simplistic X-ray spectral model, and the selection bias of the sample used by Liu et al. (2022) may limit our ability to identify any changes. We also calculated Eddington ratios for our sample by determining black hole masses from the Mg II line (see Vestergaard & Osmer 2009) to calculate $L_{\text{bol}}/L_{\text{Edd}}$, finding this to increase with increasing C IV distance for all of the quasar populations as expected from optically selected studies (e.g. R20; Temple et al. 2023) and weak emission-line quasar studies (e.g. Luo et al. 2015). See Appendix C for figures showing these measurements of N_{H} and Eddington ratio with C IV distance.

are based on extrapolating the traditional threshold of $\log R > 1$ at 5 GHz to 144 MHz and 888 MHz, respectively. Radio luminosities are calculated assuming synchrotron emission with $S_{\nu} \sim \nu^{-0.7}$.

R20 showed that the BALs and non-BALs in the SDSS-IV (and precursors) optically selected sample are co-located in C IV emission space with a changing BAL fraction across the space. Here we have investigated the X-ray selected sample from SDSS-V and find similar results with the KS tests not allowing us to rule out the possibility that the BALs and non-BALs have different C IV blueshift-EW distributions. It is clear, however, that the X-ray properties of the BALs and non-BALs in our X-ray-selected sample are not statistically distinct from one another.

7 CONCLUSIONS

This paper aimed to study the X-ray-selected SDSS-V BAL quasars within the redshift range of $1.5 \leq z \leq 3.5$. A sample of 2317 quasars with $S/N > 5$ was selected based on best-fitting ICA spectral reconstructions and BAL quasars were identified using the BI metric. Other absorption systems were also identified via the AI metric. We have studied the UV outflow signatures in the BAL quasars and investigated their X-ray properties relative to non-BAL quasars. Our study of a large X-ray-selected sample builds on prior studies of BAL and non-BAL quasars that predominantly used optically selected samples, hence we have drawn comparisons with previous results from G09 and R20. The main results of this paper are as follows:

(i) This X-ray-selected quasar sample consists of 143 or 6.2 per cent BI BAL quasars and 954 or 41.2 per cent quasars with AI absorption systems, i.e. the observed fractions are slightly lower than but comparable to those previously found in optically selected samples. We note that these observed BAL fractions are different from the intrinsic fraction due to selection effects at play, including the X-ray flux limits, the optical-magnitude limit for SDSS-V spectroscopy, and the S/N required to identify BALs in the optical spectra.

(ii) The X-ray selected quasars lie at lower C IV blueshifts relative to optically selected quasars (see Fig. 3). However, within the X-ray selected sample, the BI BALs, AI quasars, and non-BAL quasars overlap in C IV space. While BALs span the same range of C IV emission properties as X-ray selected non-BALs, we found

small – but statistically significant – differences such that BALs are preferentially found in quasars with slightly lower C IV EWs.

(iii) While the extent of the C IV emission-line blueshifts differ for the optically and X-ray selected samples, the *absorption* velocities are independent of the target selection (Figs 8 and 9), suggesting that X-rays may have a larger impact on the gas associated with C IV emission compared to the (likely distinct, physically further out) gas responsible for the absorption in BALs (see Section 6.1).

This picture is supported by the result that X-ray weaker sources are at higher C IV distances (Fig. 14) and the lack of connection between the absorption properties of the X-ray selected BI BAL and AI quasars relative to X-ray strength or $\Delta\alpha_{\text{ox}}$ (Fig. 15).

(iv) For the eFEDS subset of our sample with X-ray spectral fitting results, the observed X-ray luminosities (i.e. not corrected for obscuration) are the same for the BAL and non-BAL quasars (see Fig. 10). Additionally, we found a similar distribution of X-ray absorption column densities, N_{H} , (see Fig. 13) suggesting that the *intrinsic* X-ray properties of the BAL and non-BAL quasars studied in this paper are also similar – validated for the eFEDS detected BI populations with a KS test p -value of 0.6 – although we note that the soft X-ray *eROSITA* selection likely misses heavily obscured objects.

(v) We found that our non-BAL quasars and the BI BAL and AI quasars overlap in the α_{ox} versus L_{2500} space (see Fig. 11), and that BALs exist across the entire range of relative X-ray strength (see Fig. 12) probed with our X-ray selected sample. Due to this X-ray selection, our BAL and non-BAL subsamples are very similar, unlike the optically selected HiBALs studied by G09 that show X-ray weakness.

(vi) X-ray selection enabled the identification of examples of (relatively) X-ray bright BAL quasars, i.e. the $\Delta\alpha_{\text{ox}} > 0$ range is more fully sampled through X-ray selection compared to the samples in G09 (see Fig. 12). Our sample spans a much wider range than G09, reaching $\Delta\alpha_{\text{ox}} > 0.4$, which includes sources that were not found by G09.

Overall, the main result of this investigation is that the X-ray-selected BAL and non-BAL quasars studied in this paper appear to have similar X-ray properties, i.e. the X-ray properties of a given quasar are independent of its categorization as a BAL or non-BAL. Further investigation of the intrinsic versus observed X-ray properties of the X-ray selected BALs relative to the non-BAL quasars is required to understand the nature of X-ray weakness and brightness of these BAL quasars. X-ray spectral analysis using a larger sample and deeper X-ray data will be possible with future *eROSITA* data releases and will allow for further exploration of the column densities present in BAL quasars and the investigation into the presence (or absence) of any X-ray absorbing gas/shielding mechanism with greater statistics.

ACKNOWLEDGEMENTS

PH, ALR, and JA acknowledge support from a UK Research and Innovation Future Leaders Fellowship (grant code: MR/T020989/1). ALR also acknowledges support from a Leverhulme Early Career Fellowship. WNB acknowledges support from National Science Foundation grant AST-2407089, and the Penn State Eberly Endowment. PRH acknowledges support from the National Science Foundation AAG Award AST-2107960 and the SDSS FAST-III programme. CA and ZI acknowledge the support of the Excellence Cluster ORIGINS, which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2094-390783311. CR acknowledges support from FONDECYT Regular grant 1230345, ANID BASAL

project FB210003 and the China-Chile joint research fund. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

Funding for the Sloan Digital Sky Survey V has been provided by the Alfred P. Sloan Foundation, the Heising-Simons Foundation, the National Science Foundation, and the Participating Institutions. SDSS acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. SDSS telescopes are located at Apache Point Observatory, funded by the Astrophysical Research Consortium and operated by New Mexico State University, and at Las Campanas Observatory, operated by the Carnegie Institution for Science. The SDSS web site is www.sdss.org.

SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration, including Caltech, The Carnegie Institution for Science, Chilean National Time Allocation Committee (CNTAC) ratified researchers, The Flatiron Institute, the Gotham Participation Group, Harvard University, Heidelberg University, The Johns Hopkins University, L'Ecole polytechnique fédérale de Lausanne (EPFL), Leibniz-Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Extraterrestrische Physik (MPE), Nanjing University, National Astronomical Observatories of China (NAOC), New Mexico State University, The Ohio State University, Pennsylvania State University, Smithsonian Astrophysical Observatory, Space Telescope Science Institute (STScI), the Stellar Astrophysics Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Illinois at Urbana-Champaign, University of Toronto, University of Utah, University of Virginia, Yale University, and Yunnan University.

This work is based on data from *eROSITA*, the soft X-ray instrument aboard SRG, a joint Russian-German science mission supported by the Russian Space Agency (Roskosmos), in the interests of the Russian Academy of Sciences represented by its Space Research Institute (IKI), and the Deutsches Zentrum für Luft- und Raumfahrt (DLR). The SRG spacecraft was built by Lavochkin Association (NPOL) and its subcontractors, and is operated by NPOL with support from the Max Planck Institute for Extraterrestrial Physics (MPE). The development and construction of the *eROSITA* X-ray instrument was led by MPE, with contributions from the Dr Karl Remeis Observatory Bamberg & ECAP (FAU Erlangen-Nürnberg), the University of Hamburg Observatory, the Leibniz Institute for Astrophysics Potsdam (AIP), and the Institute for Astronomy and Astrophysics of the University of Tübingen, with the support of DLR and the Max Planck Society. The Argelander Institute for Astronomy of the University of Bonn and the Ludwig Maximilians Universität Munich also participated in the science preparation for *eROSITA*.

This research has also made use of data obtained from the Chandra Source Catalogue, provided by the Chandra X-ray Center (CXC).

This research has made use of NASA's Astrophysics Data System and adstex (<https://github.com/yymao/adstex>).

DATA AVAILABILITY

The SDSS-V data, including optical counterpart identifications to the X-ray sources as part of targeting and the optical spectra that underlie this article, are publicly available at <https://www.sdss.org/>. Data products including C IV emission line measurements, BAL properties, and X-ray properties (excluding eRASS:1-derived properties) are available online (see Appendix A). The eRASS:1 optical counterpart identification will be made public in Salvato et al.

(in preparation). After which, the eRASS:1 data will be made available upon reasonable request. *eROSITA* data is available at <https://erossita.mpe.mpg.de>. The parent *Chandra* catalogue used for targeting is available at <https://doi.org/10.25574/csc2>. The parent *Swift*-XRT catalogue is available at <https://www.swift.ac.uk/2SXP/S/excess/> and the parent *XMM-Newton* catalogue is provided at <http://xmm-catalog.irap.omp.eu/>.

REFERENCES

- Ahmed H., Gallagher S. C., Shemmer O., Brotherton M. S., Dix C., Parrott L., Richards G. T., 2025, *ApJ*, 985, 207
- Aird J. et al., 2010, *MNRAS*, 401, 2531
- Allen J. T., Hewett P. C., Maddox N., Richards G. T., Belokurov V., 2011, *MNRAS*, 410, 860
- Arav N., Liu G., Xu X., Stidham J., Benn C., Chamberlain C., 2018, *ApJ*, 857, 60
- Arav N., Xu X., Miller T., Kriss G. A., Plesha R., 2020, *ApJS*, 247, 37
- Avni Y., Tananbaum H., 1982, *ApJ*, 262, L17
- Aydar C. et al., 2025, *A&A*, 698, A132
- Baskin A., Laor A., Hamann F., 2013, *MNRAS*, 432, 1525
- Baskin A., Laor A., Hamann F., 2015, *MNRAS*, 449, 1593
- Baskin A., Laor A., Stern J., 2014, *MNRAS*, 445, 3025
- Bolton A. S. et al., 2012, *AJ*, 144, 144
- Bowen I. S., Vaughan A. H. J., 1973, *Appl. Opt.*, 12, 1430
- Bowler R. A. A., Hewett P. C., Allen J. T., Ferland G. J., 2014, *MNRAS*, 445, 359
- Brandt W. N., Laor A., Wills B. J., 2000, *ApJ*, 528, 637
- Brandt W. N., Yang G., 2022, in Bambi C., Sanganello A., eds, *Handbook of X-ray and Gamma-ray Astrophysics*. Springer Nature, Singapore, p. 78
- Brunner H. et al., 2022, *A&A*, 661, A1
- Comparat J. et al., 2020, *A&A*, 636, A97
- Dai X., Shankar F., Sivakoff G. R., 2012, *ApJ*, 757, 180
- Delaney J. N., Aird J., Evans P. A., Barlow-Hall C., Osborne J. P., Watson M. G., 2023, *MNRAS*, 521, 1620
- Dwelly T. et al., 2017, *MNRAS*, 469, 1065
- Evans I. N. et al., 2024, *ApJS*, 274, 22
- Foltz C. B., Weymann R. J., Peterson B. M., Sun L., Malkan M. A., Chaffee F. H., Jr., 1986, *ApJ*, 307, 504
- Gallagher S. C., Brandt W. N., Chartas G., Priddey R., Garmire G. P., Sambruna R. M., 2006, *ApJ*, 644, 709
- Gaskell C. M., 1982, *ApJ*, 263, 79
- Gaskell C. M., Goosmann R. W., 2013, *ApJ*, 769, 30
- Gaskell C. M., Goosmann R. W., 2016, *Ap&SS*, 361, 67
- Ghisellini G., Haardt F., Matt G., 2004, *A&A*, 413, 535
- Ghosh K. K., Punsly B., 2007, *ApJ*, 661, L139
- Gibson R. R. et al., 2009a, *ApJ*, 692, 758 (G09)
- Gibson R. R., Brandt W. N., Gallagher S. C., Schneider D. P., 2009b, *ApJ*, 696, 924
- Gibson R. R., Brandt W. N., Schneider D. P., 2008, *ApJ*, 685, 773
- Giustini M., Cappi M., Vignali C., 2008, *A&A*, 491, 425
- Giustini M., Proga D., 2019, *A&A*, 630, A94
- Green P. J. et al., 1995, *ApJ*, 450, 51
- Gunn J. E. et al., 2006, *AJ*, 131, 2332
- Hale C. L. et al., 2021, *Publ. Astron. Soc. Aust.*, 38, e058
- Hall P. B. et al., 2002, *ApJS*, 141, 267
- Hamann F., Chartas G., McGraw S., Rodríguez Hidalgo P., Shields J., Capellupo D., Charlton J., Eracleous M., 2013, *MNRAS*, 435, 133
- Hamann F., Chartas G., Reeves J., Nardini E., 2018, *MNRAS*, 476, 943
- Hamann F., Ferland G., 1999, *ARA&A*, 37, 487
- Hardcastle M. J. et al., 2023, *A&A*, 678, A151
- Hardcastle M. J., Croston J. H., 2020, *New Astron. Rev.*, 88, 101539
- Heckman T. M., Best P. N., 2014, *ARA&A*, 52, 589
- Hewett P. C., Foltz C. B., 2003, *AJ*, 125, 1784
- Hewett P. C., Wild V., 2010, *MNRAS*, 405, 2302
- Jin C., Lusso E., Ward M., Done C., Middei R., 2024, *MNRAS*, 527, 356
- Just D. W., Brandt W. N., Shemmer O., Steffen A. T., Schneider D. P., Chartas G., Garmire G. P., 2007, *ApJ*, 665, 1004
- Knigge C., Scaringi S., Goad M. R., Cottis C. E., 2008, *MNRAS*, 386, 1426
- Kollmeier J. A. et al., 2025, preprint (arXiv:2507.06989)
- Laor A., Brandt W. N., 2002, *ApJ*, 569, 641
- Laor A., Fiore F., Elvis M., Wilkes B. J., McDowell J. C., 1997, *ApJ*, 477, 93
- Leighly K. M., 2004, *ApJ*, 611, 125
- Leighly K. M., Cooper E., Grupe D., Terndrup D. M., Komossa S., 2015, *ApJ*, 809, L13
- Liu H., Luo B., Brandt W. N., Gallagher S. C., Garmire G. P., 2018, *ApJ*, 859, 113
- Liu T. et al., 2022, *A&A*, 661, A5
- Luo B. et al., 2013, *ApJ*, 772, 153
- Luo B. et al., 2014, *ApJ*, 794, 70
- Luo B. et al., 2015, *ApJ*, 805, 122
- Lusso E., Risaliti G., 2016, *ApJ*, 819, 154
- Lynden-Bell D., 1969, *Nature*, 223, 690
- Magorrian J. et al., 1998, *AJ*, 115, 2285
- Matthews J. H. et al., 2023, *MNRAS*, 526, 3967
- Matthews J. H., Knigge C., Long K. S., Sim S. A., Higginbottom N., Mangham S. W., 2016, *MNRAS*, 458, 293
- McCaffrey T. V., Richards G. T., 2021
- McConnell D. et al., 2020, *Publ. Astron. Soc. Aust.*, 37, e048
- Merloni A. et al., 2024, *A&A*, 682, A34
- Miller T. R., Arav N., Xu X., Kriss G. A., 2020, *MNRAS*, 499, 1522
- Mishra S., Vivek M., Chand H., Joshi R., 2021, *MNRAS*, 504, 3187
- Murray N., Chiang J., Grossman S. A., Voit G. M., 1995, *ApJ*, 451, 498
- Mushotzky R. F., Done C., Pounds K. A., 1993, *ARA&A*, 31, 717
- Nandra K., O'Neill P. M., George I. M., Reeves J. N., 2007, *MNRAS*, 382, 194
- Page M. J. et al., 2017, *MNRAS*, 464, 4586
- Pâris I. et al., 2018, *A&A*, 613, A51
- Perrotta S. et al., 2018, *MNRAS*, 481, 105
- Predehl P. et al., 2021, *A&A*, 647, A1
- Proga D., 2007, in Ho L. C., Wang J. W., eds, *ASP Conf. Ser. Vol. 373, The Central Engine of Active Galactic Nuclei*. Astron. Soc. Pac., San Francisco, p. 267
- Proga D., Kallman T. R., 2004, *ApJ*, 616, 688
- Punsly B., 1999, *ApJ*, 527, 624
- Rankine A. L., Aird J., Ruiz A., Georgakakis A., 2024, *MNRAS*, 527, 9004
- Rankine A. L., Hewett P. C., Banerji M., Richards G. T., 2020, *MNRAS*, 492, 4553 (R20)
- Richards G. T. et al., 2011, *AJ*, 141, 167
- Richards G. T., McCaffrey T. V., Kimball A., Rankine A. L., Matthews J. H., Hewett P. C., Rivera A. B., 2021, *AJ*, 162, 270
- Richards G. T., Vanden Berk D. E., Reichard T. A., Hall P. B., Schneider D. P., Subba Rao M., Thakar A. R., York D. G., 2002, *AJ*, 124, 1
- Rivera A. B., Richards G. T., Gallagher S. C., McCaffrey T. V., Rankine A. L., Hewett P. C., Shemmer O., 2022, *ApJ*, 931, 154
- Rivera A. B., Richards G. T., Hewett P. C., Rankine A. L., 2020, *ApJ*, 899, 96
- Rodríguez Hidalgo P. et al., 2025, preprint (arXiv:2508.14221)
- Rodríguez Hidalgo P., Hamann F., Hall P., 2011, *MNRAS*, 411, 247
- Rodríguez Hidalgo P., Rankine A. L., 2022, *ApJ*, 939, L24
- Rogerson J. A., Hall P. B., Ahmed N. S., Rodríguez Hidalgo P., Brandt W. N., Filiz Ak N., 2018, *ApJ*, 862, 22
- Rosen S. R. et al., 2016, *A&A*, 590, A1
- Saccheo I. et al., 2023, *A&A*, 671, A34
- Saez C., Brandt W. N., Gallagher S. C., Bauer F. E., Garmire G. P., 2012, *ApJ*, 759, 42
- Salvato M. et al., 2022, *A&A*, 661, A3
- SDSS Collaboration, 2025, preprint (arXiv:2507.07093)
- Shimwell T. W. et al., 2022, *A&A*, 659, A1
- Silk J., Rees M. J., 1998, *A&A*, 331, L1
- Smee S. A. et al., 2013, *AJ*, 146, 32
- Steffen A. T., Strateva I., Brandt W. N., Alexander D. M., Koekemoer A. M., Lehmer B. D., Schneider D. P., Vignali C., 2006, *AJ*, 131, 2826

- Stepney M., Banerji M., Hewett P. C., Temple M. J., Rankine A. L., Matthews J. H., Richards G. T., 2023, *MNRAS*, 524, 5497
- Streblyanska A., Barcons X., Carrera F. J., Gil-Merino R., 2010, *A&A*, 515, A2
- Sulentic J. W., Zwitter T., Marziani P., Dultzin-Hacyan D., 2000, *ApJ*, 536, L5
- Sunyaev R. et al., 2021, *A&A*, 656, A132
- Tananbaum H. et al., 1979, *ApJ*, 234, L9
- Temple M. J. et al., 2023, *MNRAS*, 523, 646
- Timlin III J. D., Brandt W. N., Laor A., 2021, *MNRAS*, 504, 5556
- Timlin J. D., Brandt W. N., Ni Q., Luo B., Pu X., Schneider D. P., Vivek M., Yi W., 2020b, *MNRAS*, 492, 719
- Timlin J., Brandt W. N., Zhu S., 2020a, *Res. Notes Am. Astron. Soc.*, 4, 168
- Tolea A., Krolik J. H., Tsvetanov Z., 2002, *ApJ*, 578, L31
- Trump J. R. et al., 2006, *ApJS*, 165, 1
- Vestergaard M., Osmer P. S., 2009, *ApJ*, 699, 800
- Vivek M., Wylezalek D., 2025, *A&A*, 695, L22
- Webb N. A. et al., 2020, *A&A*, 641, A136
- Weymann R. J., Morris S. L., Foltz C. B., Hewett P. C., 1991, *ApJ*, 373, 23
- Weymann R. J., Williams R. E., Peterson B. M., Turnshek D. A., 1979, *ApJ*, 234, 33
- Worrall D. M., 2009, *A&AR*, 17, 1
- Wu J., Brandt W. N., Comins M. L., Gibson R. R., Shemmer O., Garmire G. P., Schneider D. P., 2010, *ApJ*, 724, 762

- Xu X., Arav N., Miller T., Benn C., 2019, *ApJ*, 876, 105
- York D. G. et al., 2000, *AJ*, 120, 1579
- Yue B. H. et al., 2025, *MNRAS*, 537, 858

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: DATA

Table A1 contains the first 5 rows of the catalogue available on MNRAS. Table A2 lists the descriptions of the available data columns. While the SDSS-V optical data and eRASS:1 X-ray data are separately in the public domain, the optical cross-match for this subsample is not yet published. As such, we do not present the X-ray information for sources with ‘eROSITA (eRASS:1)’ in the Instrument column of Table A1. See the Data Availability section for more information.

Table A1. The first 5 rows of the table containing optical and X-ray information for our sample. The full table is available online. See Table A2 for a description of the columns.

SDSS.ID	RACAT	DECCAT	zfinal	Instrument	BI	BI_err	BLVMAX	BLVMIN	BLVEXT	BLPOSMIN
70345276	7.898	0.572	2.236	Chandra	5556.876	1.726	11719.405	3027.526	8691.879	6202.113
70346150	8.987	1.280	1.732	Chandra	0.000	0.000	0.000	0.000	0.000	0.000
70837514	351.894	0.376	1.491	XMM	0.000	0.000	0.000	0.000	0.000	0.000
101611989	136.440	-1.538	1.615	eROSITA (eFEDS)	0.000	0.000	0.000	0.000	0.000	0.000
55777672	138.105	3.988	1.636	eROSITA (eFEDS)	229.533	2.440	5167.040	3027.526	2139.514	3027.526

AI	AL_err	ALVMAX	ALVMIN	ALVEXT	ALPOSMIN	CIV_EW	CIV_blue	ERO_ID	eFEDS_ERO_ID
5906.696	3.537	20597.718	956.732	19640.986	6202.113	71.107	1204.664	-	-
133.793	4.242	7168.049	6202.113	965.935	6892.081	48.667	740.697	-	-
0.000	0.000	0.000	0.000	0.000	0.000	70.718	194.348	-	-
0.000	0.000	0.000	0.000	0.000	0.000	63.500	221.314	-	3211
840.909	5.857	5167.040	197.408	4969.632	1854.099	57.143	1412.510	-	10482

SWIFT_ID	XMM.ID	CSC21P_ID	L0.2-2.3keV	L2keV	L2500	alpha_ox	Daosx	eddratio
-	-	2CXO_J003135.5+003421	45.151	44.728	46.372	-1.630	0.046	0.452
-	-	2CXO_J003556.8+011649	44.721	44.298	45.372	-1.411	0.125	0.322
8796114954158265	4XMM J232734.6+002233	2CXO_J232734.7+002234	44.863	44.812	46.007	-1.458	0.167	0.134
-	-	-	44.718	44.373	45.201	-1.317	0.195	0.110
-	-	-	44.587	44.242	46.064	-1.699	-0.066	0.231

Table A2. Description of the column names in Table A1.

Column Name	Description
SDSS_ID	Unique identifier for targets in SDSS.
RACAT, DECCAT	RA and DEC of the objects (degrees).
zfinal	The redshift of the objects, see Section 2 for how these values were obtained.
Instrument	The instrument from which the X-ray information is obtained.
BI, AI	The Balnicity Index and Absorption Index determined using equations (1) & 2 along with 'BI_err' (the error in BI), 'BI_VMAX' (maximum velocity at which C IV absorption occurs), 'BI_VMIN' (the minimum velocity), 'BI_VEXT' (width of the absorption trough), 'BI_POSMIN' (velocity of the deepest absorption) and equivalently for AI. All have units of km s^{-1} .
CIV_EW	The C IV emission line EW in angstroms.
CIV_blue	The C IV emission line blueshift in km s^{-1} .
ERO_ID	eRASS:1 source ID.
eFEDS_ERO_ID	eFEDS source ID.
CSC21P_ID	<i>Chandra</i> source ID.
XMM_ID	The IAU name associated with the <i>4XMM</i> source.
Swift_ID	<i>Swift</i> source ID (2SXPS_ID).
L0.2-2.3keV	Logarithm of rest frame 0.2–2.3 keV X-ray luminosity [$\log_{10}(\text{erg s}^{-1})$].
L2keV	Logarithm of monochromatic X-ray luminosity at 2 keV [$\log_{10}(\text{erg s}^{-1})$].
L2500	Logarithm of monochromatic optical luminosity at 2500 Å [$\log_{10}(\text{erg s}^{-1})$].
alpha_ox	The X-ray-to-optical spectral index, see equation (5).
Daos	The $\Delta\alpha_{\text{ox}}$ determined using equation (8).
eddratio	The Eddington ratio determined using black hole mass estimates from the Mg II emission line.

APPENDIX B: BI BAL AND AI > 0 REPRESENTATIVE SPECTRA

We present six representative spectra for the BI BALs and AI quasars identified in our X-ray-selected sample. Fig. B1 shows the rest frame spectra in black and the reconstructions in red for quasars with BI and AI-defined absorption systems using the definitions in equations (1) and (2). The C IV distance of the sources increases as we move vertically downwards and the troughs highlighted using horizontal blue lines indicate the absorption troughs classified and used to determine the BI or AI values (km s^{-1}) for these quasars.

The left panel of Fig. B1 shows the representative spectra at different C IV distances for the X-ray-selected BI BAL quasars. The horizontal blue lines in this panel highlight the 3000 km s^{-1} threshold set for the BI definition in equation (1) as many of the troughs extend to lower velocities which are not captured in the BI metric. As noted for the composites in Fig. 4, the absorption troughs blueward of the C IV line become wider and deeper with increasing C IV distance, as expected and previously observed for the optically selected BALs studied by R20.

The AI quasars in the right panel show the presence of NALs (as previously seen in composite spectra, see Fig. 7) as well as

narrow troughs that may be associated with mini-BALs (velocity widths of $\sim 1000\text{--}2000 \text{ km s}^{-1}$; Trump et al. 2006). For example, the middle trough of the fifth AI spectrum has a trough with a velocity width of 1173 km s^{-1} . Studies such as Knigge et al. (2008) and R20 have previously noted the sub-populations that arise in the AI-defined quasar population. In detail, the AI definition overestimates the number of BAL quasars in the sample by the inclusion of NALs. The NALs have absorption troughs that are narrow enough ($\geq 450 \text{ km s}^{-1}$) to be included in the AI definition while excluded in the BI; therefore resulting in classically defined non-BALs with NAL signatures being categorized as AI quasars. As discussed in Section 4, the NAL absorbers have very different internal kinematics relative to the BI-defined absorbers.

The contrast in the absorption troughs in the left and right panels of Fig. B1, suggests that the AI quasars are not representative of the 'classical' BAL quasars with strong and broad absorption lines, as also noted by Knigge et al. (2008). We reiterate from previous studies that while the BI definition may lead to mild underestimation of BAL quasars, it is better at identifying characteristic BAL troughs and limiting contamination from other absorbers with different properties, e.g. NALs.

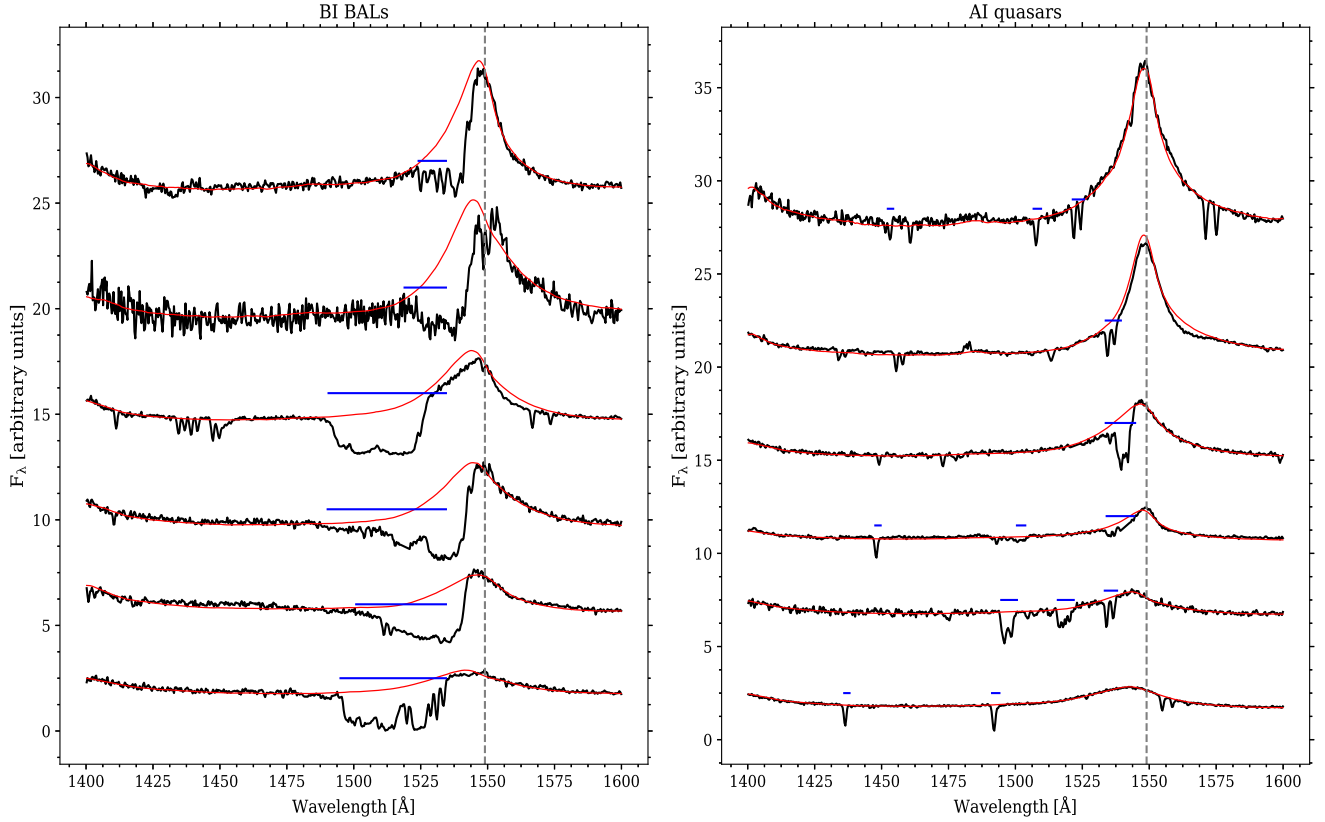


Figure B1. Six representative spectra for the X-ray selected quasars identified as BI BALs (*left panel*) and AI quasars (*right panel*) in our sample. The rest-frame spectra are in black, their ICA reconstructions in red and the blue horizontal lines show where the absorption troughs were identified [see equations (1) and (2)]. The dashed vertical line is at the C IV $\lambda 1549$ rest-frame wavelength, and the C IV distance of the sources increases downwards in the panels. The BI BAL spectra show the expected broader and deeper absorption troughs as you move downwards in the plot. The AI quasar spectra show the presence of NALs and mini-BALs.

APPENDIX C: N_H AND EDDINGTON RATIO WITH CHANGING C IV DISTANCE

Fig. C1 shows the observed trends for N_H (left panel) and Eddington ratio (right panel) with C IV distance. While noting the caveats presented in the main text, we observe that N_H does not change with the increasing C IV distance.

The Eddington ratios, calculated using black hole mass estimates via the Mg II emission line, are observed to increase with C IV distance. These trends in Fig. C1 are seen for all populations in our X-ray selected sample regardless of categorization as BAL or non-BAL, AI quasar or otherwise.

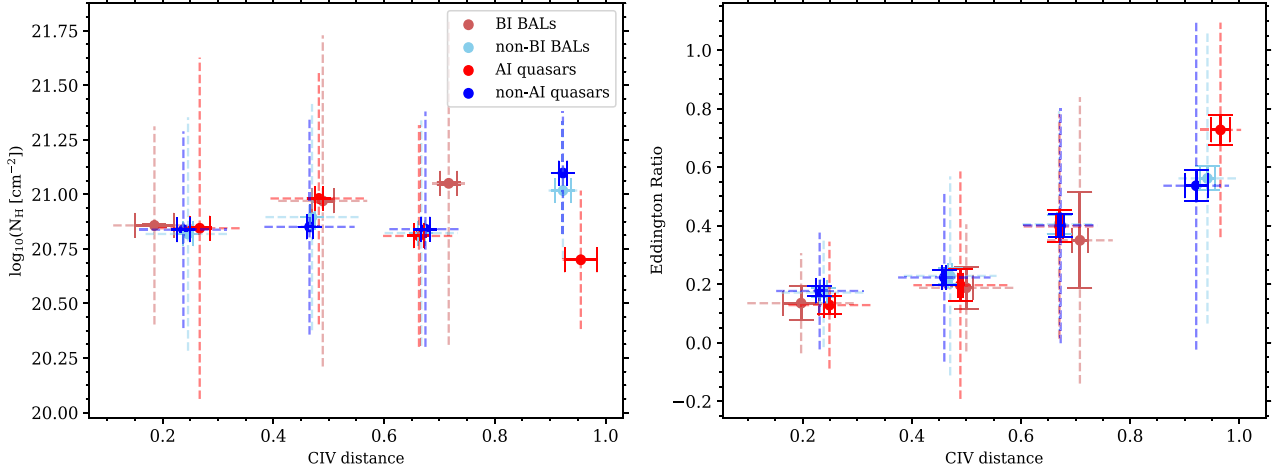


Figure C1. N_{H} versus C IV distance (*left panel*) and the Eddington ratio versus C IV distance (*right panel*) for all the subpopulations in our X-ray selected sample, with each point showing the median N_{H} value/Eddington ratio in different C IV distance bins. The dashed error bars indicate the standard deviation, and the capped error bars are the standard errors in the medians. As noted in Fig. 14, the BI BALs are not found in the last bin with the highest C IV distance. For the smaller eFEDS sample in our paper we find that there is no trend for N_{H} with C IV distance, suggesting that the intrinsic X-ray properties do not change with C IV distance. The Eddington ratio increases with C IV distance as also previously observed in optically selected studies. These trends are observed for all the subpopulations independent of their categorization.

This paper has been typeset from a $\text{\LaTeX}$ file prepared by the author.