Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **504,** 428–443 (2021) doi:10.1093/mnras/stab839
Advance Access publication 2021 March 24

BAT AGN Spectroscopic Survey XXVII: scattered X-Ray radiation in obscured active galactic nuclei

K. K. Gupta,[1★] C. Ricci,[1,2,3] A. Tortosa,[1] Y. Ueda,[4] T. Kawamuro,[1,5] M. Koss,[6] B. Trakhtenbrot,[7] K. Oh,[4,8] F. E. Bauer,[9,10,11] F. Ricci,[9] G. C. Privon,[12,13] L. Zappacosta,[14] D. Stern,[15] D. Kakkad,[16] E. Piconcelli,[14] S. Veilleux,[17] R. Mushotzky,[17] T. Caglar,[18] K. Ichikawa,[19] A. Elagali,[20,21] M. C. Powell,[22] C. M. Urry[23] and F. Harrison[24]

[1]*Núcleo de Astronomía de la Facultad de Ingeniería, Universidad Diego Portales, Av. Ejército Libertador 441, Santiago 22, Chile*
[2]*Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, People's Republic of China*
[3]*Department of Physics & Astronomy, George Mason University, MS 3F3, 4400 University Drive, Fairfax, VA 22030, USA*
[4]*Department of Astronomy, Kyoto University, Kitashirakawa-Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan*
[5]*National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan*
[6]*Eureka Scientific, 2452 Delmer Street Suite 100, Oakland, CA 94602-3017, USA*
[7]*School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978, Israel*
[8]*Korea Astronomy & Space Science institute, 776, Daedeokdae-ro, Yuseong-gu, Daejeon 34055, Republic of Korea*
[9]*Instituto de Astrofísica and Centro de Astroingeniería, Facultad de Física, Pontificia Universidad Católica de Chile, Casilla 306, Santiago 22, Chile*
[10]*Millennium Institute of Astrophysics, Nuncio Monseñor Sótero Sanz 100, Providencia, Santiago, Chile*
[11]*Space Science Institute, 4750 Walnut Street, Suite 205, Boulder, Colorado 80301, USA*
[12]*Department of Astronomy, University of Florida, 211 Bryant Space Science Center, Gainesville, FL 32611, USA*
[13]*National Radio Astronomy Observatory, 520 Edgemont Rd, Charlottesville, VA 22903, USA*
[14]*Osservatorio Astronomico di Roma, via di Frascati 33, I-00078 Monte Porzio Catone, Italy*
[15]*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, MS 169-224, Pasadena, CA 91109, USA*
[16]*European Southern Observatory, Alonso de Cordova 3107, Casilla 19, Santiago 19001, Chile*
[17]*Department of Astronomy, University of Maryland, College Park, MD 20742, USA*
[18]*Leiden Observatory, PO Box 9513, NL-2300 RA Leiden, the Netherlands*
[19]*Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, Sendai 980-8578, Japan*
[20]*International Centre for Radio Astronomy Research, The University of Western Australia, M468, 35 Stirling Highway, Crawley, WA 6009, Australia*
[21]*Australia Telescope National Facility, CSIRO Astronomy and Space Science, PO Box 76, Epping, NSW 1710, Australia*
[22]*Institute of Particle Astrophysics and Cosmology, Stanford University, 452 Lomita Mall, Stanford, CA 94305, USA*
[23]*Physics Department, Yale Center for Astronomy & Astrophysics, New Haven, CT 06520, USA*
[24]*Cahill Center for Astronomy and Astrophysics, California Institute of Technology, Pasadena, CA 91125, USA*

Accepted 2021 March 15. Received 2021 February 16; in original form 2020 November 19

ABSTRACT

Accreting supermassive black holes (SMBHs), also known as active galactic nuclei (AGN), are generally surrounded by large amounts of gas and dust. This surrounding material reprocesses the primary X-ray emission produced close to the SMBH and gives rise to several components in the broadband X-ray spectra of AGN, including a power-law possibly associated with Thomson-scattered radiation. In this work, we study the properties of this scattered component for a sample of 386 hard-X-ray-selected, nearby ($z \sim 0.03$) obscured AGN from the 70-month *Swift*/BAT catalogue. We investigate how the fraction of Thomson-scattered radiation correlates with different physical properties of AGN, such as line-of-sight column density, X-ray luminosity, black hole mass, and Eddington ratio. We find a significant negative correlation between the scattering fraction and the column density. Based on a large number of spectral simulations, we exclude the possibility that this anticorrelation is due to degeneracies between the parameters. The negative correlation also persists when considering different ranges of luminosity, black hole mass, and Eddington ratio. We discuss how this correlation might be either due to the angle dependence of the Thomson cross-section or to more obscured sources having a higher covering factor of the torus. We also find a positive correlation between the scattering fraction and the ratio of [O III] λ5007 to X-ray luminosity. This result is consistent with previous studies and suggests that the Thomson-scattered component is associated with the narrow-line region.

Key words: galaxies: active – X-rays: general – galaxies: Seyfert – quasars: general – quasars: supermassive black holes.

1 INTRODUCTION

Supermassive black holes (SMBHs), present in the centres of virtually all massive galaxies (e.g. Kormendy & Richstone 1995),

★ E-mail: kriti.gupta@mail.udp.cl

during their accreting phase are known as active galactic nuclei (AGN). The growth of SMBHs and that of their host galaxies is thought to be closely connected since several of their properties are strongly correlated (e.g. Magorrian et al. 1998; Marconi & Hunt 2003). These include the tight correlations found between the stellar velocity dispersion of the galaxy and the mass of the central black hole (e.g. Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002; Caglar et al. 2020). Hence, AGN are believed to play an important role in the evolution of their host galaxies (e.g. Kormendy & Ho 2013).

AGN are usually surrounded by a large amount of gas and dust, distributed in a structure referred to as the torus (e.g. Urry & Padovani 1995). This torus is thought to be responsible for reprocessing a fraction of the light emitted by the accretion flow (e.g. Antonucci 1993). The accretion disc, which is the main source of radiation in non-jetted AGN, emits optical and ultraviolet (UV) photons, which in some sources can be completely obscured by the surrounding torus. In fact, previous studies have shown that a majority of AGN are obscured (e.g. Maiolino et al. 1998; Risaliti, Maiolino & Salvati 1999; Ricci et al. 2015; Marchesi et al. 2018) and are thus difficult to account for based on optical or UV surveys. These obscured AGN are believed to contribute significantly to the cosmic X-ray background (CXB; e.g. Comastri et al. 1995; Ueda et al. 2003, 2014; Gilli, Comastri & Hasinger 2007; Ananna et al. 2020). Therefore, the study of obscured SMBHs is essential to improve our understanding of the entire AGN population and their evolution. This can be done by making use of a characteristic feature of AGN, which is their strong X-ray emission, which is produced by the Comptonization of optical and UV photons in a corona of hot electrons located close to the SMBH (e.g. Haardt & Maraschi 1991). This energetic radiation, and in particular that emitted in the hard-X-ray band ($E > 10$ keV), can be used to find obscured AGN, and to study their properties, because: (a) they are significantly less absorption-biased than lower-energy X-rays, hence can penetrate large column densities ($N_{\rm H}$), up to Compton-thick values ($N_{\rm H} > 10^{24}$ cm^{-2}; see fig. 1 of Ricci et al. 2015) and (b) they are substantially less contaminated by the light associated with the host galaxy than softer X-rays.

The broadband X-ray spectra of AGN have already been analysed in great detail in many previous studies (e.g. Dadina 2007; Winter et al. 2009; Koss et al. 2016; Ricci et al. 2017a). Typically, in unobscured AGN, the spectrum is dominated by a primary X-ray continuum associated with radiation produced by the corona (e.g. Dadina 2008; Beckmann et al. 2009). Apart from this, the spectrum also shows several distinctive features that can be attributed to the absorption and reprocessing of the primary X-ray radiation by the circumnuclear material. For example, for column densities above 10^{22} cm^{-2}, the primary X-ray continuum is significantly suppressed due to photoelectric absorption. Signatures of reprocessed X-ray radiation by the circumnuclear material are also seen in the X-ray spectrum in the form of a broad 'Compton hump' peaking around 30 keV (e.g. Lightman & White 1988; Ghisellini, Haardt & Matt 1994; Krolik, Madau & Zycki 1994) and a fluorescent Fe Kα line at 6.4 keV (e.g. Fabian et al. 2000; Yaqoob & Padmanabhan 2004; Shu, Yaqoob & Wang 2010; Fukazawa et al. 2011; Ricci et al. 2014). These features provide important information about the geometry, as well as the physical properties of the circumnuclear material (e.g. George & Fabian 1991; Matt, Perola & Piro 1991; Ponti et al. 2013).

When the X-ray continuum is obscured ($N_{\rm H} \gtrsim 10^{22}$ cm^{-2}), a component possibly related to Thomson scattering of the X-ray continuum by circumnuclear photoionized gas also becomes visible. This component is believed to be produced in the Compton-thin circumnuclear material and is observed as a weak, unabsorbed, power-law continuum emerging at energies lower than the photoelectric cutoff. The scattered X-ray radiation can help us to understand the effects of obscuration and can shed light on the structure of the inner regions of AGN. The parameter generally used to study the scattered radiation in AGN is the 'scattering fraction' ($f_{\rm scatt}$; e.g. Turner et al. 1997; Cappi et al. 2006; Ueda et al. 2007; Winter et al. 2009; Yamada et al. 2020). This parameter denotes the fraction of radiation scattered with respect to the primary X-ray emission of the AGN. Its intensity depends on both the covering angle of the surrounding torus and the amount of gas available for scattering (e.g. Levenson et al. 2002; Eguchi et al. 2009).

Using broadband X-ray observations of hard-X-ray-selected AGN, Ueda et al. (2007) discovered a new type of AGN with extremely low values of scattering fraction ($f_{\rm scatt} < 0.5\%$). They concluded that either the central SMBH in these sources is buried in a geometrically thick torus with a small opening angle, or the amount of gas available for scattering is exceptionally low, or both. Noguchi, Terashima & Awaki (2009) later presented an improved sample of these 'buried' AGN, and investigated their multiwavelength properties (Noguchi et al. 2010), finding that these objects also tend to have lower values of the ratio of [O III] λ5007 to X-ray luminosity ($L_{\rm [O\,III]}/L_{\rm X}$). They interpreted this as evidence of these AGN having a small opening angle of the torus that collimates the narrow-line region (NLR), a result further supported by Ueda et al. (2015) and Kawamuro et al. (2016). Studies have already shown that the soft X-ray ($E < 10$ keV) emission in obscured sources can extend up to hundreds of parsecs (e.g. Sako et al. 2000; Young, Wilson & Shopbell 2001; Greene et al. 2014) and has a morphology similar to that of the optical NLR, as traced by the [O III] λ5007 emission (e.g. Bianchi, Guainazzi & Chiaberge 2006; Bianchi et al. 2010; Dadina et al. 2010; Gómez-Guijarro et al. 2017; Fabbiano et al. 2018). The detailed analysis of high-resolution X-ray spectra of obscured AGN (e.g. NGC 1068 by Kinkhabwala et al. 2002 and Brinkman et al. 2002; Mrk 3 by Sako et al. 2000) showed that the $<$ 3–4 keV range is dominated by photoionization lines. This points towards both the soft X-ray and the optical emission lines being produced in the photoionized gas of the NLR. As a result, low values of $L_{\rm [O\,III]}/L_{\rm X}$ for buried AGN suggest that the possible locus of Thomson scattering of X-ray radiation is also in the NLR. However, Ichikawa et al. (2012), who presented a mid- to far-infrared study of a sample of local AGN, attributed the low scattering values of AGN to strong starburst activities of the host galaxy. On the other hand, Hönig et al. (2014) found that objects with low scattering fractions reside in high-inclination galaxies or merger systems. Therefore, they concluded that the low values of $f_{\rm scatt}$ are related to the obscuration of scattered emission due to host galactic gas and dust.

This work aims to study the properties of Thomson-scattered radiation in a sample of hard-X-ray-selected AGN and to constrain the possible region of origin of this radiation. The objects in our sample were detected by the Burst Alert Telescope (BAT: Barthelmy et al. 2005; Krimm et al. 2013) onboard the NASA mission, *Neil Gehrels Swift Observatory*. BAT operates in the 14–195 keV energy band and is therefore sensitive to heavily obscured AGN. Since 2005, BAT has been continuously surveying the entire sky and has proved very useful to study AGN in the local Universe ($z < 0.1$). The BAT AGN Spectroscopic Survey (BASS[1]) has provided high-quality multiwavelength data for the BAT AGN, including black hole mass measurements (Koss et al. 2017), X-ray spectroscopy and modelling (Ricci et al. 2017a), near-infrared (NIR; 1–2.4 μm) spectra and

[1] www.bass-survey.com

modelling (Lamperti et al. 2017), NIR AO imaging (Koss et al. 2018), extensive continuum modelling of the far-infrared emission (Ichikawa et al. 2019b), radio emission (Smith et al. 2020), and molecular gas measurements (Koss et al. 2021). From this data, numerous correlations have been found such as between X-ray emission and high-ionization optical lines (e.g. Berney et al. 2015) and outflows (e.g. Rojas et al. 2020) with the Eddington ratio being a key parameter in these trends (e.g. Oh et al. 2017; Ricci et al. 2017b, 2018).

We use the 70-month *Swift*/BAT catalogue (Baumgartner et al. 2013), with revised counterpart classifications from Ricci et al. (2017a), to understand the properties of scattered X-ray radiation in nearby obscured AGN and how it relates with the physical properties of AGN. To do so, we study possible correlations between scattering fraction and other physical properties of accreting SMBHs, such as their column density ($N_{\rm H}$), X-ray luminosity ($L_{\rm X}$), black hole mass ($M_{\rm BH}$), and Eddington ratio ($\lambda_{\rm Edd}$). We also investigate the dependence of $f_{\rm scatt}$ on the ratio of [O III] λ5007 luminosity to X-ray luminosity ($L_{\rm [O\,III]}/L_{\rm X}$), as well as the ratio of [O II] λ3727 luminosity to X-ray luminosity ($L_{\rm [O\,II]}/L_{\rm X}$).

The structure of the paper is as follows. In Section 2, we describe the sample used in this work, while in Section 3 we present the correlation we found between scattering fraction and column density. In Section 3.1, we simulate and fit spectra for a dummy population of obscured AGN, to verify that the correlation found in this section is intrinsic to our sample, and not due to parameter degeneracy. The correlations with $L_{\rm [O\,III]}/L_{\rm X}$ and $L_{\rm [O\,II]}/L_{\rm X}$ are discussed in Section 4, while potential correlations with black hole mass, X-ray luminosity, and Eddington ratio are investigated in Section 5. Finally, we discuss and summarize our findings in Sections 6 and 7. Throughout the paper, we assume a cosmological model with $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$, $\Omega_{\rm M} = 0.3$, and $\Omega_\Lambda = 0.7$.

2 SAMPLE AND DATA

We use here the 70-month *Swift*/BAT catalogue, which includes 838 hard-X-ray-selected, nearby (median redshift = 0.037) AGN. Of these, 386 (∼53%) sources are classified as obscured AGN based on strong intrinsic absorption required in the initial absorbed power-law fits to their X-ray spectra (Ricci et al. 2017a; see Section 2.1). The line-of-sight column densities ($N_{\rm H}$) of these sources lie in the range $10^{21.5}\,{\rm cm^{-2}}$ to $10^{25}\,{\rm cm^{-2}}$. Our analysis focuses on these obscured sources, which show features of Thomson-scattered radiation in their X-ray spectra. Of these, nine sources are flagged as blazars by Ricci et al.(2017a; see also Paliya et al. 2019). However, they are included in the analysis as their X-ray spectra showed signatures of reprocessed X-ray emission (Ricci et al. 2017a). The various parameters used in our work have been obtained from the optical (Section 2.2) and X-ray spectroscopic analysis (Section 2.1) of these sources, carried out as part of the first data release (DR1) of BASS (e.g. Koss et al. 2017; Ricci et al. 2017a).

2.1 X-ray data

The *Swift*/BAT AGN have been extensively followed up by X-ray observatories covering the 0.3–10 keV energy range. By combining observations from *XMM–Newton* (Jansen et al. 2001), *Swift*/XRT (Burrows et al. 2005), *ASCA* (Tanaka, Inoue & Holt 1994), *Chandra* (Weisskopf et al. 2000), and *Suzaku* (Mitsuda et al. 2007) in the soft X-ray band ($E < 10$ keV), together with the *Swift*/BAT data, Ricci et al. (2017a; hereafter R17) carried out a detailed broadband X-ray (0.3–150 keV) spectral analysis of all 838 AGN from the 70-month *Swift*/BAT catalogue, using XSPEC (Arnaud 1996; see Section 3.1). First, they fitted the X-ray spectra of all sources using a simple power-law model, including Galactic absorption along the line of sight (TBABS$_{\rm Gal}$.[ZPOW] in XSPEC). The residuals from these fits were then visually inspected to check for signatures of absorption from neutral matter to classify the sources as obscured or unobscured. After this preliminary exercise, 24 different spectral models were used for the X-ray spectral fitting.

The sample of 386 obscured, non-blazar AGN were fitted using nine models (B1 to B9; see R17 for more details) of increasing complexity. The different components included in these models are: (1) an X-ray continuum with a high energy cutoff (CUTOFFPL) representing the primary X-ray emission from the corona, (2) absorption of the X-ray radiation by neutral material via photoelectric absorption (ZPHABS) and Compton scattering (CABS), (3) reflection of the primary X-ray continuum by optically thick, neutral circumnuclear material (PEXRAV; Magdziarz & Zdziarski 1995), and (4) Thomson scattering of part of the primary X-ray radiation by Compton-thin circumnuclear material.

For 272/386 (70%) obscured sources, model B1 was sufficient and hence adopted by R17 for the spectral fitting. This model consisted of an absorbed primary X-ray continuum, along with an unobscured reflection component and a Thomson-scattered component. For the latter, a cutoff power-law was used, with values of the free parameters (photon index [Γ], cutoff energy [$E_{\rm C}$], normalization [K]) set to be the same as those of the primary X-ray continuum. A multiplicative constant corresponding to the scattering fraction ($f_{\rm scatt}$) was added to this component, as a free parameter, to renormalize the flux. For the remaining objects, model B1 was modified by adding multiple collisionally ionized plasma components or partially covering absorbers, as and when such a modified model improved the fit significantly.

The best-fit values of all free parameters in different models, employed for the spectral fitting of the AGN sources, are reported by R17. They also provide absorption-corrected fluxes of the primary continuum emission in the 2–10 keV, 14–150 keV, and 14–195 keV energy bands (F_{2-10}, F_{14-150}, and F_{14-195}, respectively) for all obscured sources. The main spectral parameters used in our analysis, such as $f_{\rm scatt}$, $N_{\rm H}$, and $F_{\rm X}$, have all been obtained from their fitting analysis. Best-fit values of column densities, along with upper and lower bounds (of 90% confidence level), are available for all 386 obscured sources. In the case of the scattering fraction, best-fit values along with upper and lower errors (at 90% confidence level) are available for 250/386 (∼65%) objects, while for the rest (136/386 ∼ 35%), upper limits are provided. We also calculate the intrinsic (i.e. absorption-corrected and k-corrected) luminosities for the obscured sources in the 2–10 keV, 14–150 keV, and 14–195 keV energy bands (L_{2-10}, L_{14-150}, and L_{14-195}, respectively). Spectroscopic redshifts and corresponding distance values are available for 382/386 (∼99%) sources, and therefore, we can calculate the intrinsic luminosity for these 382 sources.

2.2 Optical data

Koss et al. (2017; BASS DR1) and Koss et al. (in preparation; BASS DR2) analysed the optical spectra of the 838 hard-X-ray-selected AGN in the 70-month *Swift*/BAT catalogue. They performed multiple spectral measurements to understand the general properties of *Swift*/BAT AGN, such as broad and narrow emission-line diagnostics to classify these AGN. They also fitted the host galaxies using stellar templates to calculate their stellar velocity dispersions. They presented the fluxes and strengths of optical emission lines, as well as black hole mass estimates and accretion rates, for both obscured

and unobscured AGN in their sample. Out of the 386 sources in our sample, reliable black hole mass estimates that we adopt in this work are available for 273 (∼71%) sources. For 261/273 (∼96%) sources, the black hole masses have been derived from the correlation between the black hole mass and the stellar velocity dispersion of the host galaxy given by Kormendy & Ho (2013). For the remaining (12/273 ∼ 4%) sources, we use black hole mass values from the literature. We calculate the Eddington ratio ($\lambda_{\rm Edd}$) for these 273 sources by incorporating their black hole masses and the 2–10 keV bolometric correction ($L_{\rm bol} = 20 \times L_{2-10}$) following Vasudevan et al. (2009):

$$\lambda_{\rm Edd} = \left[\frac{20 \times L_{2-10}\,({\rm erg\,s^{-1}})}{1.5 \times 10^{38} \times M_{\rm BH}\,({\rm M_\odot})}\right]. \quad (1)$$

To include the effects of using more advanced bolometric corrections in our work, we repeat part of the analysis described in Section 5 with Eddington ratio-dependent bolometric corrections (Vasudevan & Fabian 2009) and get similar results.

Finally, in our analysis, we use the emission line measurements presented by Koss et al. (2017) and Oh et al. (in preparation) using the spectral line fitting codes described by Sarzi et al. (2006) and improved by Oh et al. (2011, 2015). We adopt a Gaussian amplitude to noise ratio (A/N) of 3 as the threshold for reliable line measurements and hence, do not include line strengths with A/N <3 in our analysis. We use the observed fluxes of the [O II] λ3727 line to calculate the observed [O II] λ3727 luminosity ($L_{\rm [O\,II]}$) for 295/386 (∼76%) sources. We also use equivalent widths and observed fluxes of the [O III] λ5007 line for 359/386 (∼93%) sources and fluxes of the narrow H α and H β lines for 318/386 (∼82%) sources and calculate the corresponding observed [O III] λ5007 luminosity ($L_{\rm [O\,III]}$) as well as the extinction-corrected [O III] λ5007 luminosity ($L^{\rm corr}_{\rm [O\,III]}$) using the Balmer decrement, as described in Ueda et al. (2015).

3 THE CORRELATION BETWEEN SCATTERING FRACTION AND COLUMN DENSITY

We study here the relation between the fraction of Thomson-scattered X-ray radiation and the line-of-sight column density of AGN. In Fig. 1, we plot the scattering fraction as a function of column density (in grey) for the 386 obscured sources in our sample. The figure clearly shows a negative correlation between these parameters, which we quantify by fitting a linear regression line (in red) to the data points. We initially excluded the upper limits while calculating the regression line here since they can affect our results considerably if not handled properly. However, we do include them in later parts of this analysis to properly quantify their contribution to the correlation. We use the `linregress`[2] correlation function of the `scipy.stats` module in PYTHON to fit the data with a linear regression. This function minimizes the sum of squares of the difference between the actual data points and that predicted by the regression model. Throughout this paper, we use this technique to fit linear regression lines to our data. Slope and intercept of the linear regression line, Pearson's correlation coefficient, and the probability of the data arising from an underlying uncorrelated distribution (null-hypothesis) are reported in Table 1. To better visualize the decrease of the scattering fraction as the column density increases, we divide the column density into six bins (of width $\log[N_{\rm H}/{\rm cm^{-2}}] = 0.5$, except for the last one) and plot the median of the scattering fraction, along with the standard error (in black), for each bin. The median and error in $f_{\rm scatt}$ in each $\log N_{\rm H}$ bin are calculated using survival analysis (SA; e.g. Feigelson & Nelson 1985; Shimizu et al. 2017) to include the contributions from the upper limits. The median $f_{\rm scatt}$ is calculated using the 50th percentile, and the uncertainties for each bin are calculated with respect to the median value in that bin using the 16th and 84th percentile. For $\log N_{\rm H}$, we simply plot the midpoint of each bin with error bars = 0.5 (= 1 for the last bin).

The SA approach efficiently includes the upper limits in $f_{\rm scatt}$ while calculating its median value in specified bins of column density. However, it does not take into account the errors in $f_{\rm scatt}$ and $\log N_{\rm H}$. In order to include these errors in our analysis, we make use of Monte Carlo (MC) simulations. To apply this technique, we handled the best-fit values with errors and upper limits separately, as discussed below:

(i) Best-fit values with errors: We created an asymmetric Gaussian distribution (normalized) to simulate the values of parameters $f_{\rm scatt}$ and $\log N_{\rm H}$. We used an asymmetric distribution because the upper and lower errors in both the parameters can be unequal. To create this distribution, we first assumed two different symmetric Gaussian distributions with the mean of both Gaussians equal to the best-fit value of the parameter, whereas the standard deviation (1σ) was chosen to be $\frac{\rm Upper\ Error}{1.64}$ and $\frac{\rm Lower\ Error}{1.64}$ for the first and second Gaussian, respectively. The errors were divided by 1.64 because they are 2σ errors. The final Gaussian was obtained by combining these two Gaussian distributions, such that the probability distribution of values greater than the mean is governed by the first Gaussian, while the probability distribution of values lower than the mean is decided by the second Gaussian. Hence, corresponding to each source with a best-fit $f_{\rm scatt}$ and $\log N_{\rm H}$ value, we created two asymmetric Gaussians depicting the probability distribution functions of $f_{\rm scatt}$ and $\log N_{\rm H}$, by incorporating their best-fit values as well as their upper and lower errors.

(ii) Best-fit upper limits: Throughout our analysis, upper limits are only present for scattering fraction. In this particular case, we employed a normalized uniform distribution to describe the probability distribution of $f_{\rm scatt}$. The distribution, used for each source with an upper limit, ranges from 0 to the corresponding upper limit value of $f_{\rm scatt}$, with each value in the distribution assigned equal probability.

After finalizing the probability distribution functions for both parameters ($f_{\rm scatt}$ and $\log N_{\rm H}$), we performed 10,000 runs of MC simulations. In each run, a random value of the scattering fraction and column density was taken for each source from the probability distributions described above. In the case of the column density, values greater than $\log(N_{\rm H}/{\rm cm^{-2}}) = 26$ were not permitted in the simulations, and for the scattering fraction, negative values were not considered. Thus, we end up with 10,000 simulated values of the scattering fraction and column density for each of the 386 sources. Next, we bin the 3,860,000 simulated column density values into six separate bins (similar to the $\log N_{\rm H}$ bins created for SA) and calculate the median column density for each bin. Values of $\log(N_{\rm H}/{\rm cm^{-2}}) > 25$ are included in the last bin. Similarly, the corresponding simulated values of $f_{\rm scatt}$ are also binned, and the median value of $f_{\rm scatt}$ in each bin is calculated. The standard error in each bin is calculated using the 16th and 84th percentile, as was done during the SA method. MC results are shown in blue in Fig. 1. As evident from the figure, median and error values from SA and MC simulations are completely consistent within their error bars. Although, median $f_{\rm scatt}$ values calculated using SA are slightly lower than those calculated from MC simulations. However, unlike SA,

[2] https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.linregress.html

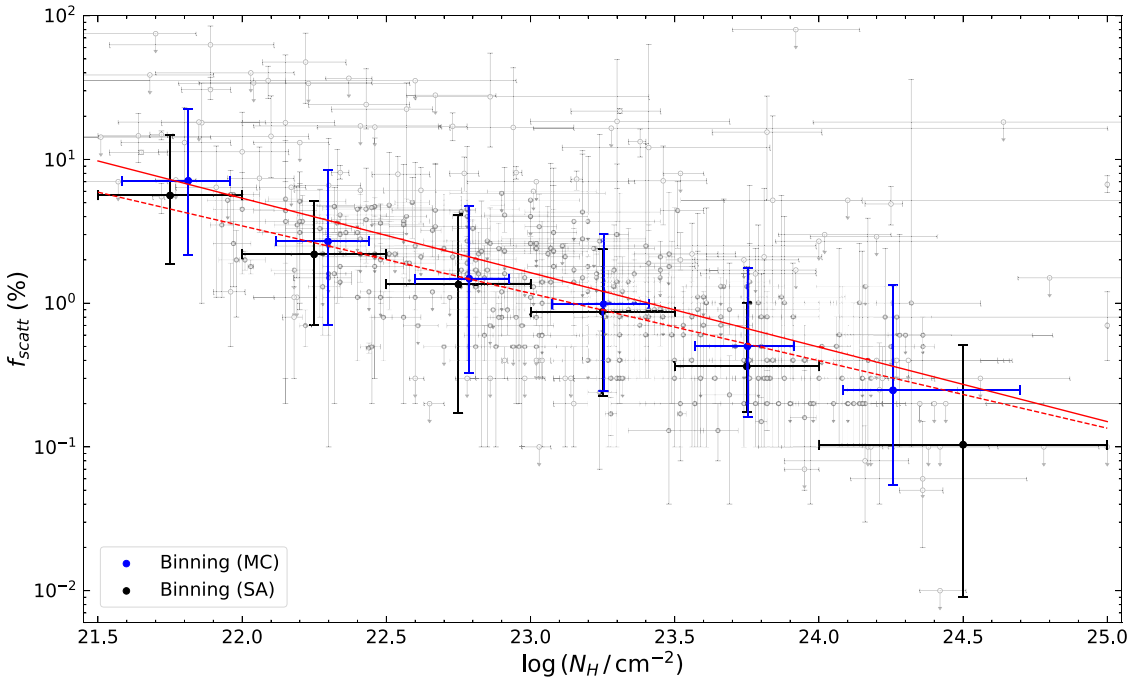


Figure 1. Scattering fraction versus column density: The plot shows the scattering fraction and the column density values (from R17) for our sample of 386 obscured AGN in the background as grey open circles (upper limits as downward arrows and best-fit values with error bars showing errors at 90% confidence level). Black circles with error bars correspond to the survival analysis (SA) results for each log $N_{\rm H}$ bin ($f_{\rm scatt}$ median calculated using the 50th percentile and the 1σ uncertainty calculated using the 16th and 84th percentile). Blue circles with error bars (1σ) depict the results obtained from 10,000 Monte Carlo (MC) simulations. The solid red line is the linear regression obtained from the data points excluding upper limits, while the dashed red line is the linear regression calculated using MC simulations. Two sources in our sample have $\log(N_{\rm H}/{\rm cm}^{-2}) > 25$, but for display purpose only they are plotted at $\log(N_{\rm H}/{\rm cm}^{-2}) = 25$. The plot clearly shows a negative correlation between $f_{\rm scatt}$ and log $N_{\rm H}$. Correlation parameters for the two methods are given in Table 1.

Table 1. Correlation results for Fig. 1 (Section 3), Fig. 4 (Section 4), and Fig. 5(c) (Section 5).

Parameter (in log) (versus $f_{\rm scatt}$)	Slope[a]		Intercept[a]		R-Value[b]		P-Value[c]	
	Data[d]	MC[e]	Data[d]	MC[e]	Data[d]	MC[e]	Data[d]	MC[e]
$N_{\rm H}$	-0.52 ± 0.04	-0.47 ± 0.03	12.1 ± 1.0	10.9 ± 0.7	-0.63	-0.50	1.9×10^{-28}	1.5×10^{-25}
$L_{[\rm O\,III]}/L_{2-10}$	0.26 ± 0.05	0.33 ± 0.02	0.76 ± 0.11	0.79 ± 0.06	0.35	0.36	2.5×10^{-8}	2.8×10^{-12}
$L_{[\rm O\,III]}/L_{14-195}$	0.22 ± 0.05	0.27 ± 0.02	0.75 ± 0.15	0.76 ± 0.07	0.27	0.28	2.0×10^{-5}	9.0×10^{-8}
$L^{\rm corr}_{[\rm O\,III]}/L_{2-10}$	0.27 ± 0.05	0.36 ± 0.03	0.65 ± 0.10	0.70 ± 0.05	0.33	0.34	4.9×10^{-7}	3.2×10^{-10}
$L^{\rm corr}_{[\rm O\,III]}/L_{14-195}$	0.20 ± 0.06	0.27 ± 0.03	0.62 ± 0.14	0.66 ± 0.07	0.23	0.24	6.0×10^{-4}	1.4×10^{-5}
$L_{[\rm O\,II]}/L_{2-10}$	0.31 ± 0.05	0.40 ± 0.03	1.01 ± 0.16	1.12 ± 0.08	0.39	0.38	1.4×10^{-8}	1.8×10^{-11}
$L_{[\rm O\,II]}/L_{14-195}$	0.26 ± 0.06	0.33 ± 0.03	0.97 ± 0.20	1.08 ± 0.10	0.29	0.29	2.7×10^{-5}	4.4×10^{-7}
$\lambda_{\rm Edd}$	-0.14 ± 0.05	-0.16 ± 0.02	-0.22 ± 0.09	-0.40 ± 0.05	-0.23	-0.20	2.1×10^{-3}	1.1×10^{-3}

[a]Of the linear regression line.
[b]The Pearson's correlation coefficient.
[c]The probability of the data set appearing if the null hypothesis is correct.
[d]Excluding upper limits.
[e]Monte Carlo simulations.

MC simulations include the already calculated uncertainties in both parameters (reported by R17) and thus give better estimates of the final median and error in each column density bin.

We also used the simulated values of scattering fraction and column density to better quantify the correlation found between them. For each run of the MC simulations, we fitted a linear regression line to the simulated data and obtained the regression parameters such as slope, intercept, correlation coefficient, and the probability of the null hypothesis. The final slope, intercept, and Pearson's correlation coefficient (reported in Table 1) are calculated from the median of these 10,000 values, and the corresponding regression line is shown as a dashed red line in Fig. 1. Compared to the linear regression line fitted initially to the data after excluding the upper limits, the regression line found through MC simulations is more accurate since these simulations are a better representation of the range of values that the scattering fraction and the column density could take for each source. Hence, the corresponding regression parameters better define the observed negative correlation between $f_{\rm scatt}$ and log $N_{\rm H}$. To confirm the reliability of this correlation, we examined the probability distribution of the linear regression obtained from MC simulations. As shown in Table 1, for the correlation between scattering fraction and column density, we get a probability $= 1.5 \times 10^{-25}$ for the null hypothesis. Such a low p-value further confirms the significance of the inverse correlation obtained. We also fitted various higher degree polynomial regressions to the data points shown in Fig. 1 but did not find any significant improvement in

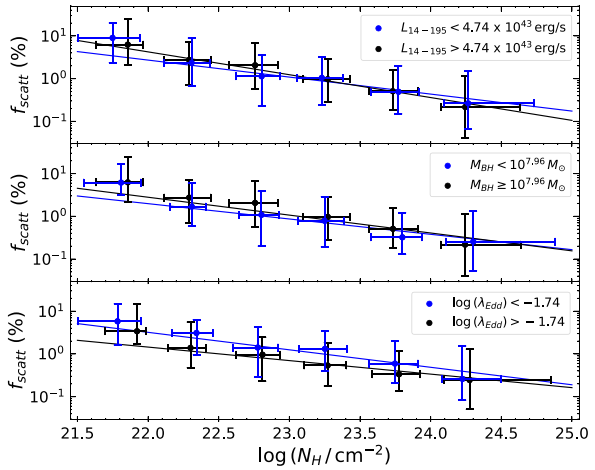


Figure 2. Scattering fraction versus column density for objects corresponding to two bins in: (a) Top panel: intrinsic 14–195 keV luminosity (L_{14-195}), (b) Middle panel: black hole mass ($M_{\rm BH}$), and (c) Bottom panel: Eddington ratio ($\lambda_{\rm Edd}$). The linear regression line as well as the median $f_{\rm scatt}$ and log $N_{\rm H}$ values with 1σ uncertainty for each log $N_{\rm H}$ bin are calculated using MC simulations. The plots show a distinct negative correlation between scattering fraction and column density. The correlation parameters are given in Table 2.

the chi-square values as compared to the one for the linear regression fit.

We also check if the correlation we found between scattering fraction and column density holds for objects in different ranges of L_{14-195}, $M_{\rm BH}$, and $\lambda_{\rm Edd}$ in our sample. By specifying certain constraints on these physical properties of AGN, we aim to investigate how these external factors modify the correlation we found. To do so, we create two bins each of these three parameters, as follows: (a) $L_{14-195} < 4.74 \times 10^{43}$erg s^{-1} and $> 4.74 \times 10^{43}$erg s^{-1}, (b) $M_{\rm BH} < 10^{7.96}$ M$_\odot$ and $\geq 10^{7.96}$ M$_\odot$ and (c) log($\lambda_{\rm Edd}$) < -1.74 and > -1.74. These ranges were determined from the median values of these parameters. In Fig. 2, we show the linear regression line obtained using MC simulations for each of these intervals. We also show the median and uncertainty in the values of scattering fraction and column density for each bin calculated using MC simulations. The correlation parameters for these fits and the number of sources in each parameter window are reported in Table 2. Very low values of probability (of the null hypothesis) establish the significance of the negative correlation between $f_{\rm scatt}$ and log $N_{\rm H}$ even when different conditions related to the physical properties of the accreting system are considered. Therefore, we can conclude that X-ray luminosity, black hole mass, and Eddington rate do not affect the negative correlation we found.

The X-ray spectra of 75/386 ($\sim$19%) sources in our sample were also fitted using torus models (see R17 for details). We have verified that including $f_{\rm scatt}$ and log $N_{\rm H}$ values obtained from those fittings in our original data does not affect the anticorrelation found between these parameters. One should note that scattering fraction values $\geq 5\% - 10\%$ can arise due to a partial covering absorber instead of Thomson scattering. To exclude their contributions, we apply an uppercut on the value of scattering fraction at 5% and 10%. In Appendix A, we show that applying such criteria does not modify the correlation significantly (Fig. A1[a]). The correlation we found may be biased by the typically low count rates of sources with high column densities. To make sure that these sources do not regulate the observed trend, we impose a lower limit on the number of counts per source at 200 counts. Fig. A1(b) in Appendix A shows that, even after removing the sources with low counts, we recover the same correlation between $f_{\rm scatt}$ and log $N_{\rm H}$, using MC simulations.

3.1 Testing parameter degeneracies

An important step in our analysis is to check if the correlation we found is due to degeneracy between the two spectral parameters of interest, $f_{\rm scatt}$ and log $N_{\rm H}$. Hence, in this section, we perform simulations to check for degeneracy between the scattering fraction and the column density. We simulated 100 dummy populations of obscured AGN, with the same size as our sample, which corresponds to more than 38,000 simulations, using a predefined model to describe the X-ray spectrum of an AGN. The model was decided based on its ability to provide a statistically significant fit to the X-ray spectra of the majority of sources in our sample. We used the same model to fit the simulated spectra and obtain simulated values of the parameters column density and scattering fraction. These two parameters were then plotted against each other to check if they show any inherent correlation and degeneracy.

The X-ray spectra are simulated using the `FAKEIT` command in XSPEC v12.10.1f (Arnaud 1996). As input, this command requires a background file, a response matrix file (RMF), and an auxiliary response file (ARF). For our analysis, we use the *XMM–Newton*/EPIC-pn background, RMF, and ARF files. As a test, we repeat a set of simulations with *Swift*/XRT background and response files and get similar results. We also redo the simulations with multiple sets of background and response files to accommodate any changes due to source positioning in the field of view and get similar results as to when we use a single set of background and response files. `FAKEIT` also needs an exposure time, which we randomly select from the distribution of the actual exposure times of the 386 sources in our sample. Finally, the `FAKEIT` command requires us to specify a model that will be used to generate fake spectra. For our simulations, we use the model B1 from R17 (described in detail in Section 2.1), which was used to fit the X-ray spectra of 272/386 ($\sim$ 71%) sources of our sample (see R17 for details). In XSPEC terminology, the model is defined by the following expression:

`ZPHABS` × `CABS` × `CUTOFFPL` + `PEXRAV` + *constant* × `CUTOFFPL`

The various parameters in the model, their priors, and the conditions applied to them are explained in detail here:

(i) Scattering fraction ($f_{\rm scatt}$): This parameter is represented by the *constant* term in the model. Its input value is selected randomly from a uniform distribution ranging from 0.1% to 5%.

(ii) Column density ($N_{\rm H}$): This parameter appears twice in the model (in components `ZPHABS` and `CABS`) and its value in these two components is tied. The default range acceptable for this parameter in the model needs to be modified to $0.1 < N_{\rm H}/10^{22}\,{\rm cm}^{-2} < 1000$ to describe the obscured AGN in our sample. The input value of this parameter is randomly selected from a uniform distribution of values between 10^{22} cm^{-2} and $10^{24.5}$ cm^{-2}.

(iii) Photon index (Γ): This parameter appears three times in the model (in the two power-laws and the `PEXRAV` component), and its value in all three components is tied. The input value of Γ is selected randomly from a Gaussian distribution centred around 1.8, with a standard deviation of 0.2 (Ricci et al. 2017a).

(iv) High energy cutoff ($E_{\rm C}$): This parameter also appears in the same three components as Γ, and its input value is selected randomly from a Gaussian distribution centred around 210 keV, with

Table 2. Correlation results from MC simulations for Fig. 2 (Section 3).

Parameter	Number of sources	Bin	Slope[a]	Intercept[a]	R-Value[b]	P-Value[c]
L_{14-195}	191/382	$< 4.74 \times 10^{43}$ erg s^{-1}	-0.40 ± 0.04	9.2 ± 0.9	-0.43 ± 0.04	8.5×10^{-10}
	191/382	$> 4.74 \times 10^{43}$ erg s^{-1}	-0.54 ± 0.04	12.4 ± 0.9	-0.57 ± 0.04	1.2×10^{-17}
M_{BH}	136/273	$< 10^{7.96}$ M$_\odot$	-0.36 ± 0.05	8.2 ± 1.1	-0.42 ± 0.05	4.6×10^{-7}
	137/273	$\geq 10^{7.96}$ M$_\odot$	-0.42 ± 0.05	9.7 ± 1.2	-0.43 ± 0.05	1.2×10^{-7}
λ_{Edd}	136/273	$<10^{-1.74}$	-0.41 ± 0.05	9.5 ± 1.2	-0.42 ± 0.05	2.8×10^{-7}
	137/273	$>10^{-1.74}$	-0.32 ± 0.05	7.1 ± 1.1	-0.38 ± 0.05	5.7×10^{-6}

[a]Of the linear regression line.
[b]The Pearson's correlation coefficient.
[c]The probability of the data set appearing if the null hypothesis is correct.

a standard deviation of 40 keV (Ricci et al. 2018). The value of this parameter for the three components is tied and not allowed to vary.

(v) Normalization (K): This parameter appears in the same three components as Γ and E_C, and its value in the three components is tied. It is calculated from flux values in the 2–10 keV band. Once the normalization values corresponding to all sources in our sample are calculated, the input value of this parameter in the model is selected randomly from the distribution of those values.

(vi) Reflection parameter (R): The default range acceptable for this parameter is set to $-1 < R < 0$, and its value is not allowed to vary. The input value is selected randomly from a Gaussian distribution centred around -0.4, with a standard deviation of 0.1 (Ricci et al. 2017a). The reflection parameter and luminosity could be anticorrelated (e.g. Zappacosta et al. 2018) with a preference for luminous AGN to have low to null reflection parameter (e.g. Vignali et al. 1999; Reeves & Turner 2000; Page et al. 2005). Furthermore, Ricci et al. (2011) found different levels of reflection as a function of AGN type. However, we do not consider any such correlation between R and K or R and N_H in these simulations.

(vii) All the remaining parameters, such as redshift, iron abundance, and inclination angle, are set to their default values and not allowed to vary during the fitting process.

After feeding the model and the input values of all parameters to the software, we simulated 100 spectra each for all sources (total 38,600 spectra). All spectra were rebinned to have at least 20 counts per bin to use chi-squared statistics. We only generated the spectra in the 0.3–10 keV interval, matching the energy range covered by *XMM–Newton*/EPIC-pn. All spectra were then fitted using the model defined above to obtain the best-fit values of all free parameters. For each set of simulated data, we then checked for the presence of a correlation between f_{scatt} and log N_H with a linear fit, as it was done in Section 3 for our sample. To do so, we create a distribution of the slopes of the linear regression lines for all 100 f_{scatt} versus log N_H plots, as shown in Fig. 3(a). The figure clearly shows a large difference between the slopes obtained from the data/MC simulations and the XSPEC simulations. Hence, we can safely infer from this distribution that the negative correlation we found between the scattering fraction and the column density for our sample of obscured AGN is not due to parameter degeneracy. We also show the probability distribution of the null hypothesis in Fig. 3(b), which reinforces the validity of the inverse correlation we found. To confirm that these simulations do not miss sources with low f_{scatt} and low log N_H values, that are otherwise not present in the original data (as visible in Fig. 1), we looked at the ratio of simulated f_{scatt} to input f_{scatt} as a function of input log N_H. The plot did not show any trend, hence, proving the ability of these simulations to recover such sources. An example of a simulated spectrum, along with the model used and the residuals after the fit, is shown in Fig. B1(a) in Appendix B. We also show an example of a simulated correlation between scattering fraction and column density values obtained from these simulations in Fig. B1(b) in Appendix B.

One should note that the scattered X-ray radiation could be absorbed by the host galaxy. To take this into account, we repeat a set of XSPEC simulations with a modified model including an absorbed scattered component. The column density of the absorbing material is assumed to be randomly distributed between 10^{19} cm^{-2} and 10^{21} cm^{-2}. The simulated spectral parameters, column density and scattering fraction, hence obtained, do not appear to be correlated, thus confirming the lack of degeneracy between them. Since, throughout our simulations, we use the same model to create and fit the spectra, the validity of our approach to look for degeneracies may be debated. Therefore, we repeat a set of XSPEC simulations using two models, one to create the spectra and another to fit them. We employ the `RXTorus`[3] model developed by Paltani & Ricci (2017) to simulate the fake spectra and the model B1 from R17 to fit them. Since we want to simulate the X-ray spectra of obscured sources, we assume an input value of 80° for the viewing angle (where 0° is face-on), and the ratio of the inner-to-outer radius of the torus is assumed to be 0.5. The input value of the equatorial column density is randomly selected from a uniform distribution of values between 10^{22} cm^{-2} to $10^{24.5}$ cm^{-2} and is connected to the line-of-sight column density via the inclination angle and the torus covering fraction. All the other parameters and their priors are the same as model B1 (described earlier). The resultant simulated parameters (log N_H and f_{scatt}) are then checked for any correlation using a linear regression fit, and we obtain a slope $= 0.1 \pm 0.04$. Hence, we can conclusively confirm the absence of parameter degeneracy and the validity of the anticorrelation between the scattering fraction and column density.

4 THE CORRELATION BETWEEN SCATTERING FRACTION AND $L_{[O\,III]}/L_X$, $L^{corr}_{[O\,III]}/L_X$, AND $L_{[O\,II]}/L_X$

Having confirmed the negative correlation between the scattering fraction and the column density for our sample of obscured AGN, we now investigate if the scattering fraction relates with the ratio of [O III] λ5007 to X-ray luminosity ($L_{[O\,III]}/L_X$). In the case of buried AGN with extremely low scattering fractions (< 0.5%), several studies have shown that these sources exhibit relatively lower values of $L_{[O\,III]}/L_X$ (e.g. Noguchi et al. 2010; Ueda et al. 2015). As a further test, we also check for possible correlations between f_{scatt} and the

[3]https://www.astro.unige.ch/reflex/xspec-models

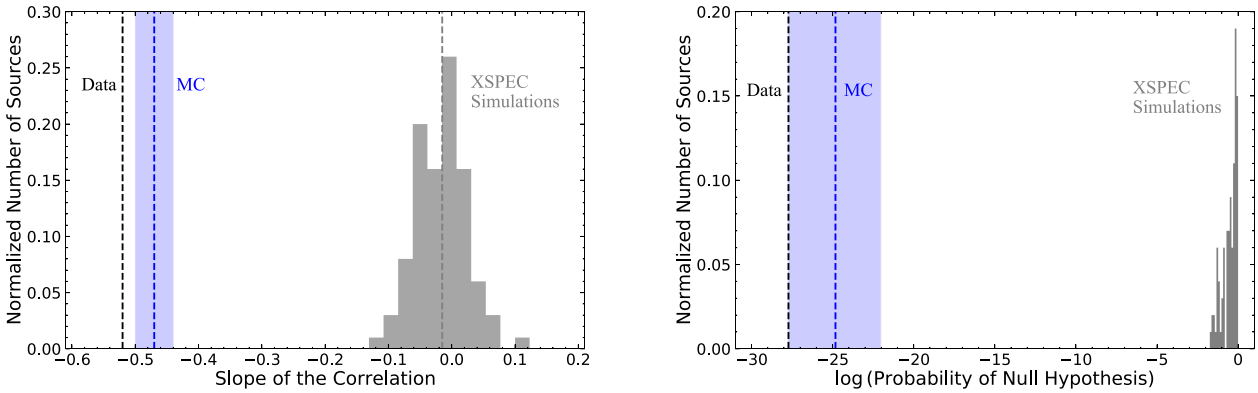


Figure 3. (a) Distribution in values of the slope of the correlation between scattering fraction and column density. (b) Distribution in the probability of the null hypothesis for the correlation between scattering fraction and column density. The distribution obtained from the XSPEC simulations is shown in grey, with the median slope marked by a dashed grey line (in the left-hand panel). The value obtained from the data (excluding upper limits) is shown as a dashed black line. The median value obtained from MC simulations (including upper limits) is shown by a dashed blue line with the shaded blue region showing the possible range of values. The difference in the values obtained from the XSPEC simulations and the data/MC simulations, for the slope of the correlation and the probability of the null hypothesis, shows that any degeneracy between the parameters ($f_{\rm scatt}$ and $\log N_{\rm H}$) does not produce an artificial correlation.

ratio of [O II] λ3727 to X-ray luminosity ($L_{[\rm O\,II]}/L_{\rm X}$), another tracer of the NLR. For our analysis, we adopt the observed [O II] λ3727 luminosity ($L_{[\rm O\,II]}$) and the observed as well as extinction-corrected [O III] λ5007 luminosity ($L_{[\rm O\,III]}$ and $L^{\rm corr}_{[\rm O\,III]}$, respectively). For $L_{\rm X}$, we use the intrinsic X-ray luminosity in the 2–10 keV and the 14–195 keV energy bands. Although, L_{2-10} and L_{14-195} are related quantities, L_{2-10} can be strongly affected by high column densities. Therefore, we include the luminosity in the harder energy band (14–195 keV) because it is less biased by high column densities in our sample.

Fig. 4(a) shows $f_{\rm scatt}$ as a function of $L_{[\rm O\,III]}/L_{2-10}$ for 359 sources (in grey) for which we have the $L_{[\rm O\,III]}$ and L_{2-10} measurements. Similar to the procedure followed in Section 3, we initially fit a linear regression line (in red) to all data points, excluding the upper limits. This fitting suggests that there might be a positive correlation between $f_{\rm scatt}$ and $L_{[\rm O\,III]}/L_{2-10}$ that needs to be analysed in more detail. To do so, we divide the data into six bins of $L_{[\rm O\,III]}/L_{2-10}$, such that each bin contains at least 20 sources. The median and standard error of $f_{\rm scatt}$ in each bin (shown in black), after including the upper limits, is calculated using the SA method. For the x-axis, we just plot the midpoint of each bin with error bars = 0.25 or 0.5, depending on the size of the bin. Furthermore, to get better estimates of the median and standard deviation in $f_{\rm scatt}$ in each bin, we use MC simulations. As described in detail in Section 3, these simulations include the already calculated errors in the parameters to generate their probability distribution functions. However, it is important to note that we do not have error estimates for $L_{[\rm O\,III]}/L_{2-10}$. Therefore, in this case, we only created these probability distributions for $f_{\rm scatt}$. Whereas, for $L_{[\rm O\,III]}/L_{2-10}$, we just represent the midpoints and the size of each bin, as done for the SA method. Following the procedure in Section 3, we obtained 10,000 values of scattering fraction for each source. For each run, depending on the $L_{[\rm O\,III]}/L_{2-10}$ bins, we classify the $f_{\rm scatt}$ values and calculate their median in each bin. The resulting median of each bin is the average of all 10,000 medians in that bin. The error in $f_{\rm scatt}$ in each bin is the standard deviation in the median values with respect to the final median (shown in blue). Therefore, these errors calculated from MC simulations are significantly smaller (by a factor of $\sqrt{N}$, where N is the sample size in each bin) compared to those calculated from SA. Apart from this, the results from the two methods are consistent with each other. We also show the linear regression line calculated from the MC simulations (as a dashed line) in Fig. 4(a). All the correlation parameters are reported in Table 1.

In a similar manner as described above, the correlation between scattering fraction and $L_{[\rm O\,III]}/L_{14-195}$ is also inspected and is shown in Fig. 4(b). It is clear from both plots (Figs 4[a] and [b]) that scattering fraction correlates with $L_{[\rm O\,III]}/L_{\rm X}$ (where X = 2–10 and 14–195). The significance of these two correlations is determined from their low p-values ($\approx 10^{-12}$ for $L_{[\rm O\,III]}/L_{2-10}$ and $\approx 10^{-8}$ for $L_{[\rm O\,III]}/L_{14-195}$). We also explore if this correlation persists when we correct the [O III] λ5007 luminosities for reddening from the Balmer decrement (e.g. Ueda et al. 2015). The relation between the scattering fraction and $L^{\rm corr}_{[\rm O\,III]}/L_{\rm X}$ for 318 sources is shown in Figs 4(c) and (d). The figures show a positive correlation similar to the one found between the scattering fraction and the ratio of the observed [O III] λ5007 to X-ray luminosity. The null hypothesis probabilities for the correlation between $f_{\rm scatt}$ and $L^{\rm corr}_{[\rm O\,III]}/L_{\rm X}$ are quite low ($\approx 10^{-10}$ for X = 2–10 and $\approx 10^{-5}$ for X = 14–195; see Table 1). Thereby, highlighting the significance of the positive correlation between these two parameters. Finally, we examine the dependence of scattering fraction on the ratio of the observed [O II] λ3727 to X-ray luminosity for 295 sources in our sample for which we could simultaneously obtain values of the observed [O II] λ3727 luminosity and the intrinsic X-ray luminosity in the 2–10 and 14–195 keV bands. (Figs 4[e] and [f]). As reported in Table 1, the correlations observed in this case have slightly higher slopes compared to those with $L_{[\rm O\,III]}/L_{\rm X}$. However, for both sets of correlations we get similar p-values ($< 10^{-3}$), which confirms their significance.

For the three sets of correlations evaluated in this section, it is worth noting that not always the narrow emission lines, such as [O II] λ3727, [O III] λ5007, H α and H β can be ascribed solely to the AGN. In some AGN hosts, there might be a non-negligible contribution from star formation that cannot be ignored (e.g. Maddox 2018). To solve this issue, we checked the emission-line classification of all sources in our sample from Koss et al. (2017). We used the [O III] λ5007/H β versus [N II] λ6583/H α and [S II] λ6717/H α classifications of Veilleux & Osterbrock (1987), revised by Kewley et al. (2006), to determine which sources lie in the star-forming regions of these line diagnostics diagrams. After removing those sources ($\sim$ 15%), we recovered all

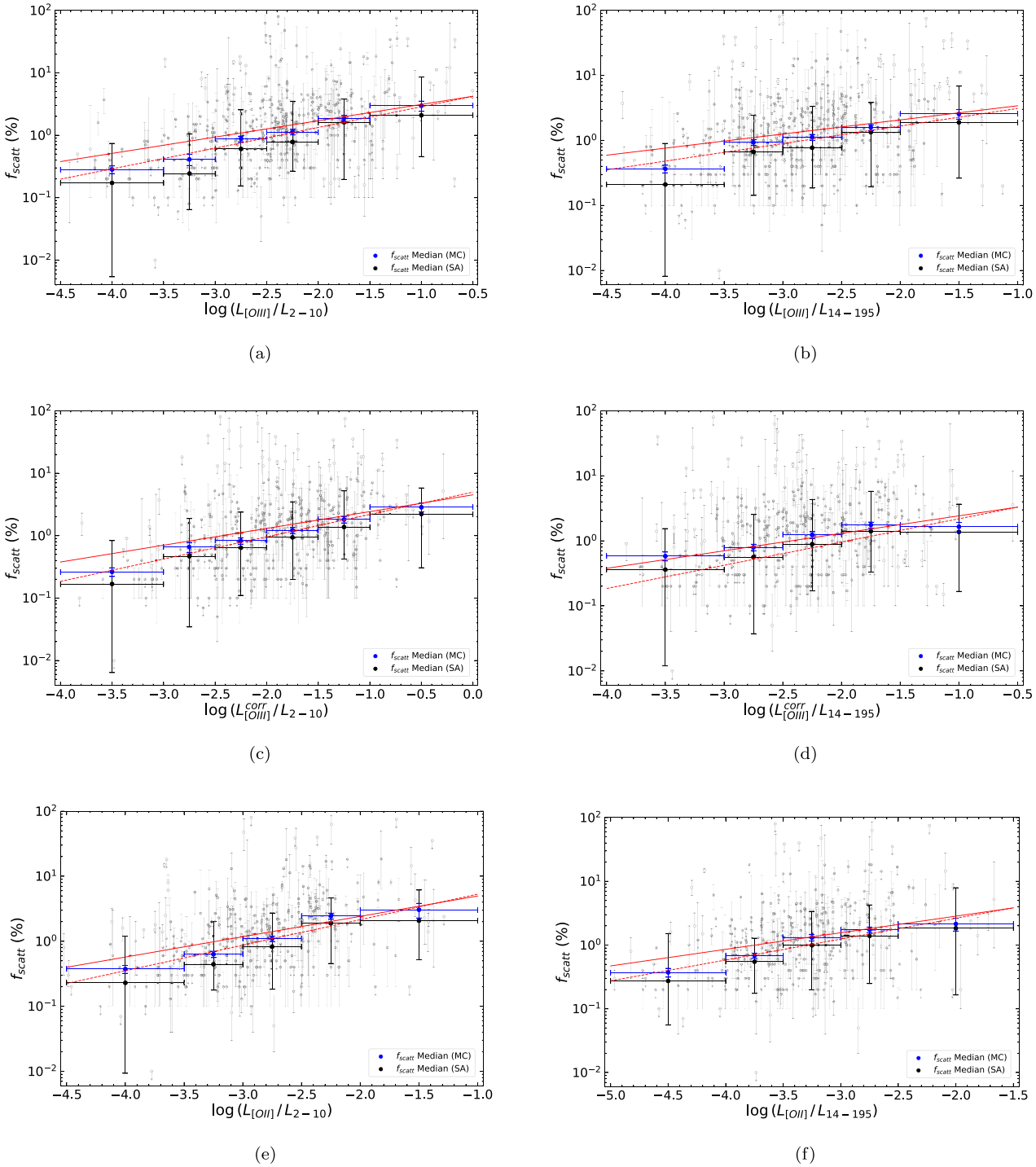


Figure 4. (a) Scattering fraction versus the ratio of observed [O III] λ5007 to intrinsic 2–10 keV X-ray luminosity ($L_{[\mathrm{O\,III}]}/L_{2-10}$). (b) Scattering fraction versus the ratio of observed [O III] λ5007 to intrinsic 14–195 keV X-ray luminosity ($L_{[\mathrm{O\,III}]}/L_{14-195}$). (c) Scattering fraction versus the ratio of extinction-corrected [O III] λ5007 to intrinsic 2–10 keV X-ray luminosity ($L^{\mathrm{corr}}_{[\mathrm{O\,III}]}/L_{2-10}$). (d) Scattering fraction versus the ratio of extinction-corrected [O III] λ5007 to intrinsic 14–195 keV X-ray luminosity ($L^{\mathrm{corr}}_{[\mathrm{O\,III}]}/L_{14-195}$). (e) Scattering fraction versus the ratio of observed [O II] λ3727 to intrinsic 2–10 keV X-ray luminosity ($L_{[\mathrm{O\,II}]}/L_{12-10}$). (f) Scattering fraction versus the ratio of observed [O II] λ3727 to intrinsic 14–195 keV X-ray luminosity ($L_{[\mathrm{O\,III}]}/L_{14-195}$). Our data are shown as grey open circles in the background (upper limits as downward arrows and best-fit values with error bars showing errors at 90% confidence level). The black and blue circles with error bars correspond to the median and 1σ uncertainty in f_{scatt} calculated using SA and MC simulations, respectively. The solid red linear regression line is obtained from the data by excluding upper limits, while the dashed red regression line is obtained from MC simulations. The plots show a positive correlation between f_{scatt} and $L_{[\mathrm{O\,III}]}/L_{\mathrm{X}}$, $L^{\mathrm{corr}}_{[\mathrm{O\,III}]}/L_{\mathrm{X}}$ and $L_{[\mathrm{O\,II}]}/L_{\mathrm{X}}$. Correlation parameters are presented in Table 1.

the positive correlations between $f_{\rm scatt}$ and $L_{[\rm O\,III]}/L_{\rm X}$, $L^{\rm corr}_{[\rm O\,III]}/L_{\rm X}$, and $L_{[\rm O\,II]}/L_{\rm X}$.

5 THE RELATION BETWEEN SCATTERING FRACTION AND $M_{\rm BH}$, $L_{\rm X}$, AND $\lambda_{\rm Edd}$

In this section, we investigate how the scattering fraction evolves with different physical properties of AGN, such as X-ray luminosity, black hole mass, and Eddington ratio (Equation 1). We have already seen that the correlation between scattering fraction and column density is not affected by any of these parameters. However, we want to check if the fraction of Thomson-scattered radiation is influenced by these properties. In Fig. 5(a), we show the scattering fraction as a function of the intrinsic 14–195 keV luminosity (L_{14-195}), for 382 sources (in grey). Following the procedure described in Section 4, we have divided the x-axis into six logarithmic bins of L_{14-195} and plotted the median and error in $f_{\rm scatt}$, calculated using SA (in black) and MC simulations (in blue), for each bin. The plot shows no correlation (p-value = 0.11) between $f_{\rm scatt}$ and L_{14-195}. Fig. 5(b) shows scattering fraction as a function of black hole mass for 273 sources, for which we used black hole mass estimates from the BASS DR2 (Section 2.2). The plot clearly shows a lack of correlation between $f_{\rm scatt}$ and $M_{\rm BH}$ (p-value = 0.24).

Finally, we plot scattering fraction as a function of Eddington ratio for 273 sources in our sample (Fig. 5[c]). The Eddington ratio for these sources is calculated from Equation 1, using the 2–10 keV bolometric correction from Vasudevan et al. (2009) and the black hole masses. We find a weak negative correlation between $f_{\rm scatt}$ and $\lambda_{\rm Edd}$ (p-value ≈ 0.001). Similar to the methods employed previously, we fit two linear regression lines to the data, one excluding the upper limits in $f_{\rm scatt}$ and the other including upper limits and errors in $f_{\rm scatt}$ using MC simulations. We have reported the regression parameters calculated from both methods in Table 1. Vasudevan & Fabian (2009) showed that bolometric correction factors could be a function of the Eddington ratio. However, the correlation we found between $f_{\rm scatt}$ and $\lambda_{\rm Edd}$ does not change if we take into account such dependence.

6 DISCUSSION

In the previous sections, we found various correlations between the scattering fraction and some physical properties of AGN, such as the line-of-sight column density, the ratio of the observed and extinction-corrected [O III] λ5007 luminosity to X-ray luminosity, the ratio of the observed [O II] λ3727 luminosity to X-ray luminosity, and Eddington ratio. In the following, we explore various interpretations of these correlations. In Section 6.1, we discuss two possible reasons behind the negative correlation between $f_{\rm scatt}$ and $\log N_{\rm H}$. In Section 6.2, we discuss the dependence of $f_{\rm scatt}$ on $L_{[\rm O\,III]}/L_{\rm X}$ and other similar relations. We also explain how these correlations could shed light on the possible locus of Thomson scattering. And finally, in Section 6.3, we discuss about the weak negative correlation between $f_{\rm scatt}$ and $\lambda_{\rm Edd}$.

6.1 The effect of inclination angle and torus covering factor on scattering fraction

For our sample of 386 hard-X-ray-selected, nearby, obscured AGN from the 70-month *Swift*/BAT catalogue, we find a negative correlation between the scattering fraction and the column density (Section 3 and Fig. 1). This trend is consistently observed even when we consider splitting the sample into different ranges of some of the fundamental parameters of accreting SMBHs, such

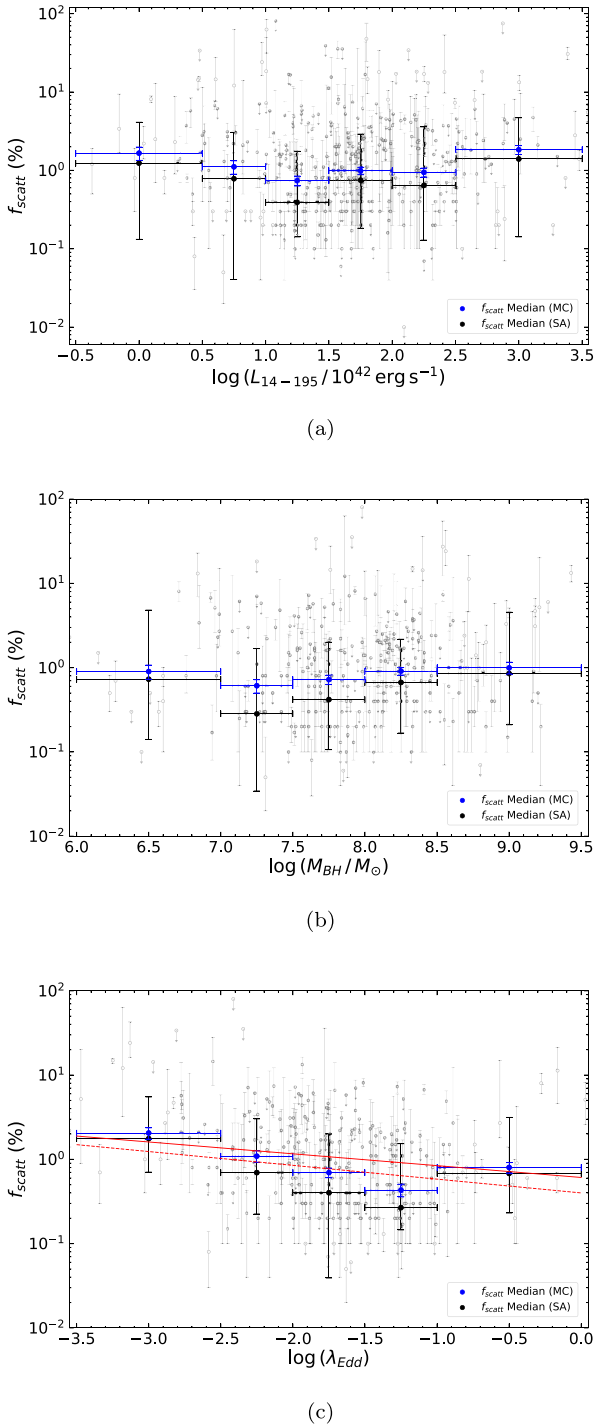


Figure 5. (a) Scattering fraction versus intrinsic 14–195 keV luminosity. (b) Scattering fraction versus black hole mass. (c) Scattering fraction versus Eddington ratio. Our data are shown as grey open circles in the background (upper limits as downward arrows and best-fit values with error bars showing error at 90% confidence level). The black and blue circles with error bars correspond to the median and 1σ uncertainty in $f_{\rm scatt}$ calculated using SA and MC simulations, respectively. The solid red line (only in the bottom panel) depicts the linear regression fit to the data, excluding upper limits. The dashed red line (only in the bottom panel) is the regression fit obtained from MC simulations. The plots show no correlation between $f_{\rm scatt}$ and L_{14-195} (r-value = 0.081) and $M_{\rm BH}$ (r-value = 0.071). A weak negative correlation is detected between $f_{\rm scatt}$ and $\lambda_{\rm Edd}$. Correlation parameters are reported in Table 1.

as their X-ray luminosity, black hole mass, and Eddington ratio (Fig. 2). We exclude the possibility that this correlation is due to the degeneracy between different spectral parameters by simulating dummy populations of obscured AGN (Section 3.1). We fitted all 38,600 simulated spectra and explored possible correlations between $f_{\rm scatt}$ and $N_{\rm H}$. The slope and probability distributions (Fig. 3) confirm the absence of any correlation between the simulated values, thereby implying that the correlation we found is intrinsic. Very high values of scattering fraction ($\geq 5\% - 10\%$) can be either due to partially covering absorbers or due to significant contributions from radio jets (see Ricci et al. 2017a for a discussion). Therefore, we also check and confirm the existence of this correlation when an uppercut at $f_{\rm scatt} = 5\%$ and 10% is applied (Fig. A1[a]). We also verify that this trend is not due to sources with low counts: by considering a lower limit of 200 counts per source, we recover the same correlation we find for the complete sample (Fig. A1[b]).

We can therefore confirm the existence of a significant negative correlation (p-value $\approx 10^{-25}$) between the scattering fraction and the column density. We can also infer from our analysis that the dependence of the scattering fraction on the column density is not, in any way, affected by the X-ray luminosity of the accreting system, the mass of the central SMBH, or the Eddington ratio (Table 2). This implies that the main parameter driving the correlation is indeed the column density. To explain this inverse correlation between the scattering fraction and column density, we discuss here two possible explanations: (a) the inclination angle dependence of the Thomson cross-section and (b) the covering factor of the surrounding torus of AGN.

6.1.1 *Inclination angle*

To comprehend the reason behind the anticorrelation between scattering fraction and column density, we first consider the inclination angle dependence of the Thomson cross-section. The cross-section of Thomson scattering has a cosine square dependence on the inclination angle (e.g. Rybicki & Lightman 1986). We also assume an average AGN model where the torus geometries and properties are similar for all sources, and the only parameter changing between the different sources is the inclination angle. A schematic of the assumed structure of the torus is shown in Fig. 6, where the column density increases with the inclination angle. We can also predict the covering factor ($f_{\rm cov}$) of different layers of the torus, based on the fraction of sources with line-of-sight column density in a specific range, following statistical arguments given by Ricci et al. (2015) and Ramos Almeida & Ricci (2017). We start with a covering factor of 70% for the torus, corresponding to a column density of $10^{22.5\pm0.5}$ cm^{-2} and calculate the inclination angle ($\theta_{\rm inc}$) to be approximately 46° (where $\theta_{\rm inc} = \cos^{-1} f_{\rm cov}$). As we consider higher values for the column density of the obscurer, at $10^{23.5\pm0.5}$ cm^{-2} and $10^{24.5\pm0.5}$ cm^{-2}, the covering factor of the torus relative to that column density reduces to 52% and 27%. Therefore, the viewing angles corresponding to those column densities increase to 59° and 74°. Since the fraction of Thomson-scattered radiation (denoted by $f_{\rm scatt}$) will be proportional to the cross-section of Thomson scattering ($[\frac{{\rm d}\sigma}{{\rm d}\Omega}]_{\rm unpol} = \frac{r_{\rm o}^2}{2} \times [1 + \cos^2\theta_{\rm inc}]$, where $r_{\rm o} = e^2/mc^2$), we used $\cos^2\theta_{\rm inc}$ as a proxy for the scattering fraction and plotted it against the column density of the torus. We obtained an inverse correlation between these parameters with a slope $=-0.42 \pm 0.07$. This slope is consistent with what we found from our data using the MC simulations (-0.47 ± 0.03). Hence, we can conclude that the observed trend could be explained by the inclination angle dependence of the Thomson cross-section. Some studies have shown that the equivalent width of the [O III] λ5007 line could be an indicator of inclination (e.g. Risaliti, Salvati & Marconi 2011; Bisogni, Marconi & Risaliti 2017; Vietri et al. 2018). As a further test, we checked for a possible relation between scattering fraction and the equivalent width of the [O III] λ5007 line and did not find them to be correlated with each other.

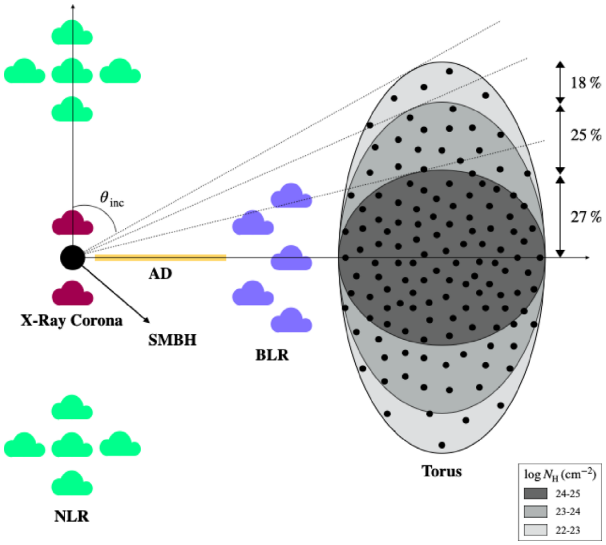


Figure 6. Schematic representation of an AGN showing a particular structure and geometry of the torus (also see fig. 4[b] of Ramos Almeida & Ricci 2017). The inclination angle is measured from the normal to the accretion disc and relative to the observer. The intrinsic column density of the obscuring torus increases with inclination. The torus covering factor is determined from the fraction of sources within a specific column density range (details are given in Ricci et al. 2015 and Ramos Almeida & Ricci 2017).

6.1.2 *Torus covering factor*

Another possible explanation for the anticorrelation between $f_{\rm scatt}$ and $\log N_{\rm H}$ could be that sources with high column densities tend to have, on average, a higher covering factor of the torus (e.g. Ricci et al. 2011 Elitzur 2012; Mateos et al. 2016; Tanimoto et al. 2018; Panagiotou & Walter 2019). This would reduce the fraction of Thomson-scattered radiation and result in lower values of scattering fraction. This interpretation also puts into perspective the conclusions made for the population of buried AGN identified by Ueda et al. (2007). They suggested that one of the plausible explanations for such low values of scattering fraction ($< 0.5\%$) would be the presence of an obscuring torus with a very high covering factor. However, we also need to keep in mind the inherent degeneracy between the torus covering factor and the amount of material available for Thomson scattering. Due to this degeneracy, a low scattering fraction can also arise due to the deficiency of circumnuclear material needed for the scattering to occur (e.g. Ueda et al. 2007).

6.2 X-ray scattering in the narrow-line region

In Section 4, we detect a positive correlation between the scattering fraction and the ratio of the observed [O III] λ5007 luminosity to the intrinsic X-ray luminosity (Figs 4[a] and [b]). This correlation also persists when the [O III] λ5007 luminosity is corrected for extinction (Figs 4[c] and [d]). Therefore, suggesting that the soft scattered X-ray radiation observed in obscured AGN may originate in the NLR.

It is well established that emission lines present in the soft X-ray spectrum of obscured AGN are produced in a gas photoionized by the central active nucleus, rather than a collisionally ionized gas (e.g. Kinkhabwala et al. 2002; Schurch et al. 2004; Guainazzi & Bianchi 2007; Bianchi et al. 2010; Nucita et al. 2010; Braito et al. 2017). We also know that the NLR is composed of material produced by photoionization, based on which, Bianchi et al. (2006) conducted a high-resolution spectral analysis of a sample of Seyfert 2 galaxies and found a striking similarity in the extension and the overall morphology of the soft X-ray and the [O III] λ5007 emission. They showed, using photoionization models, that it is possible to have a gas photoionized by the central AGN, extending over hundreds of parsec and producing both the [O III] λ5007 and the soft X-ray emission. The correlation we found further supports the claim that the soft X-rays and the [O III] λ5007 emission are produced by the same gas, i.e. the NLR. This correlation is also in agreement with previous results in the literature that found low values of $L_{[O\,III]}/L_X$ for sources with very low scattering fraction (e.g. Noguchi et al. 2010; Ueda et al. 2015).

Compared to $L_{[O\,III]}/L_X$ (slope $= 0.33 \pm 0.02$), we recover a slightly steeper correlation (slope $= 0.40 \pm 0.03$) when we consider $L_{[O\,II]}/L_X$ (Figs 4[e] and [f]). An even steeper dependence (slope $=$ 0.98) is quoted by Kawamuro et al. (2016) for the ratio of [O IV] λ24.89 μm luminosity to 10–50 keV X-ray luminosity. Considering the fact that the [O IV] λ24.89 μm emission is less extincted by dust, compared to the [O III] λ5007 emission, it is expected that the scattering fraction correlates more with its ratio to the X-ray luminosity. Consequently, the various correlations of f_{scatt} with ratios of optical line luminosities to X-ray luminosity further strengthen our conclusion that Thomson scattering occurs in the NLR. We also investigated if the correlations we found are affected by the different aperture sizes of the sources in our sample (Appendix C). To do so, we checked if the $L_{[O\,III]}/L_X$ correlates with the redshift or the physical width of the slits used for [O III] λ5007 measurements (reported by Koss et al. 2017 and Koss et al., in preparation). As shown in Fig. C1 in Appendix C, we do not find any significant correlation between $L_{[O\,III]}/L_X$ and redshift or $L_{[O\,III]}/L_X$ and slit sizes. Therefore, we can exclude any aperture effects on the trends we observe. As the column density is much better constrained compared to the scattering fraction, we also checked if the ratio of the [O III] λ5007 to X-ray luminosity is correlated with the column density. However, we only obtain a very weak negative correlation between $\log N_H$ and $L_{[O\,III]}/L_X$, with a slope $= -0.20 \pm 0.01$ and p-value $= 1.1 \times 10^{-4}$ for the 2–10 keV band, and a slope $= -0.12 \pm 0.01$ and p-value $= 2.0 \times 10^{-2}$ for the 14–195 keV range. This weak correlation disappears when we consider two bins in f_{scatt} around median f_{scatt} $= 1.3$. Hence, we can conclude that the ratio of the [O III] λ5007 to X-ray luminosity is indeed correlated with the scattering fraction and not with the column density.

The scatter visible in all plots of Fig. 4 could be attributed to different variability time-scales of the X-rays and the optical emission lines (e.g. Mushotzky, Done & Pounds 1993; Schawinski et al. 2015; Oh et al. 2017; Ichikawa et al. 2019a). This correlation between scattering fraction and [O III] λ5007 to X-ray luminosity ratio could be useful in removing the large scatter in the L_X versus $L_{[O\,III]}$ relation (e.g. Ueda et al. 2015). Fig. 4(a) shows that the ratio of the optical to X-ray bolometric corrections could change by a factor of ∼1000 since a change in scattering fraction from 0.1% to 5% causes a variation of almost three orders of magnitude in $L_{[O\,III]}/L_X$. As we show in Section 4 and 5, it is $L_{[O\,III]}$ which changes with f_{scatt} and not L_X, therefore the ratio between the [O III] λ5007 and the X-ray bolometric corrections for different objects could vary depending on their scattering fractions. However, exploring these effects in detail is beyond the scope of this paper. Finally, given the fact that low f_{scatt} AGN, which have higher column densities, have typically low ionized optical line luminosity with respect to the X-ray luminosity, we would expect optical surveys, which rely on narrow emission lines to identify obscured AGN, could miss a significant fraction of the population of heavily obscured AGN.

6.3 The role of Eddington ratio

For our sample, we find a weak negative correlation between the scattering fraction and the Eddington ratio (Section 5 and Fig. 5[c]). A possible interpretation of this correlation can be acquired from the inverse correlation between f_{scatt} and $\log N_H$. We discuss in Section 6.1 that the low scattering fractions could be, at least in part, due to a higher covering factor of the torus. This would imply that sources with high accretion rates having lower scattering fractions tend to be surrounded by a thicker torus. A similar argument was proposed by Noguchi et al. (2010), who also found the Eddington ratio to be anticorrelated with the scattering fraction. Another potential explanation of this inverse correlation is explored by Ricci et al. (2017b), who presented a negative correlation between the covering factor of the Compton-thin material and λ_{Edd}. As a result, we would expect the scattering fraction to also decrease with increasing accretion rates since Thomson scattering occurs in the Compton-thin circumnuclear material. This explanation is supported by the idea that the amount of material surrounding an AGN is regulated by the radiation pressure (e.g. Fabian, Celotti & Erlund 2006; Fabian, Vasudevan & Gandhi 2008; Fabian et al. 2009). Hence, it is possible that more rapidly accreting AGN tend to have a lower fraction of Thomson-scattered radiation, either due to removal of Compton-thin circumnuclear material by radiation pressure or due to a geometrically thick torus.

7 SUMMARY AND CONCLUSION

In this work, we used the 70-month *Swift*/BAT catalogue to study the properties of Thomson-scattered X-ray radiation in a sample of 386 hard-X-ray-selected, obscured AGN in the local Universe. We used X-ray spectral parameters reported by Ricci et al. (2017a), black hole masses estimated by Koss et al. (2017) and Koss et al. (in preparation), and narrow-line fluxes calculated by Oh et al. (in preparation) to investigate possible correlations between the fraction of Thomson-scattered radiation and other physical properties of SMBHs, such as the line-of-sight column density, X-ray luminosity, black hole mass, Eddington ratio, [O III] λ5007 to X-ray luminosity ratio, and [O II] λ3727 to X-ray luminosity ratio. Here, we summarize our main findings:

(i) We found a significant negative correlation between the scattering fraction and the column density (see Section 3 and Fig. 1), with a slope $= -0.47 \pm 0.03$ (Table 1). We excluded the possibility that this correlation is due to sources with lower counts (Fig. A1[b]) or due to partially covering absorbers ($f_{scat} \geq 5\% - 10\%$; Fig. A1[a]). We also verified that the correlation persists when considering different bins of X-ray luminosity, black hole mass, and accretion rate (Fig. 2 and Table 2).

(ii) To confirm that the correlation we found is intrinsic to our sample and not due to degeneracy between the parameters, we simulated and fitted more than 38,000 obscured AGN spectra (see Section 3.1 and Fig. B1[a]). We then checked for possible correlations between the simulated parameters. The probability and slope distribution plots

(Fig. 3) demonstrate that parameter degeneracy is not responsible for the observed trend (Fig. B1[b]).

(iii) To understand the physical mechanism responsible for the anticorrelation between scattering fraction and column density, we discussed two possible explanations. First, we considered the inclination angle dependence of the Thomson cross-section and assumed a certain average geometry of the circumnuclear material for all AGN (see Fig. 6 and Section 6). Based on these assumptions, we were able to explain the trend we observe. A second explanation for this trend is that, in general, sources with high line-of-sight column densities tend to have higher covering factors of the surrounding torus. As a result, the amount of Thomson-scattered radiation could decrease, in turn reducing the scattering fraction for these sources.

(iv) We found a positive correlation between the scattering fraction and the ratio of [O III] λ5007 to X-ray luminosity (see Section 4 and Fig. 4), with a slope $= 0.33 \pm 0.02$ (0.27 ± 0.02) for the intrinsic 2–10 (14–195) keV X-ray luminosity. A similar, but slightly steeper, correlation was observed for the ratio of [O II] λ3727 to X-ray luminosity. Both the trends suggest that Thomson-scattered radiation could originate in the same region responsible for the optical [O III] λ5007 and [O II] λ3727 emission, i.e. the NLR. This conclusion is also supported by several previous studies (e.g. Bianchi et al. 2006; Ueda et al. 2015; Kawamuro et al. 2016).

(v) We did not find any dependence of the fraction of Thomson-scattered radiation on the X-ray luminosity or black hole mass of the AGN. However, we detected a weak negative correlation between scattering fraction and Eddington ratio (see Section 5 and Fig. 5) with a slope $= -0.16 \pm 0.02$. This relation could imply that rapidly growing SMBHs are either surrounded by a thicker torus (Noguchi et al. 2010) or have radiatively driven away the Compton-thin material needed for scattering (Ricci et al. 2017b).

ACKNOWLEDGEMENTS

We would like to thank the anonymous referee for the prompt and constructive report that strengthened this paper. We thank Matthew Temple for his valuable comments. This work made use of data from the NASA/IPAC Infrared Science Archive and NASA/IPAC Extragalactic Database (NED), which are operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This research has made use of data and/or software provided by the High Energy Astrophysics Science Archive Research Center (HEASARC), which is a service of the Astrophysics Science Division at NASA/GSFC and the High Energy Astrophysics Division of the Smithsonian Astrophysical Observatory. KKG was supported by a 2018 grant from the ESO – Government of Chile Joint Committee. We acknowledge support from FONDECYT Iniciacion grant 11190831 (CR), ANID grants CATA-Basal AFB-170002 (FEB), FONDECYT Regular 1190818 and 1200495 (FEB), and Millennium Science Initiative ICN12_009 (FEB). MK acknowledges support from NASA through ADAP award 80NSSC19K0749. KO acknowledges support from the National Research Foundation of Korea (NRF-2020R1C1C1005462). EP acknowledges financial support under ASI/INAF contract 2017-14-H.0.

DATA AVAILABILITY

The data underlying this article are available in Ricci et al. (2017a), Koss et al. (2017), Koss et al. (in preparation), and Oh et al. (in preparation). Additional data will be shared on a reasonable request to the corresponding author.

REFERENCES

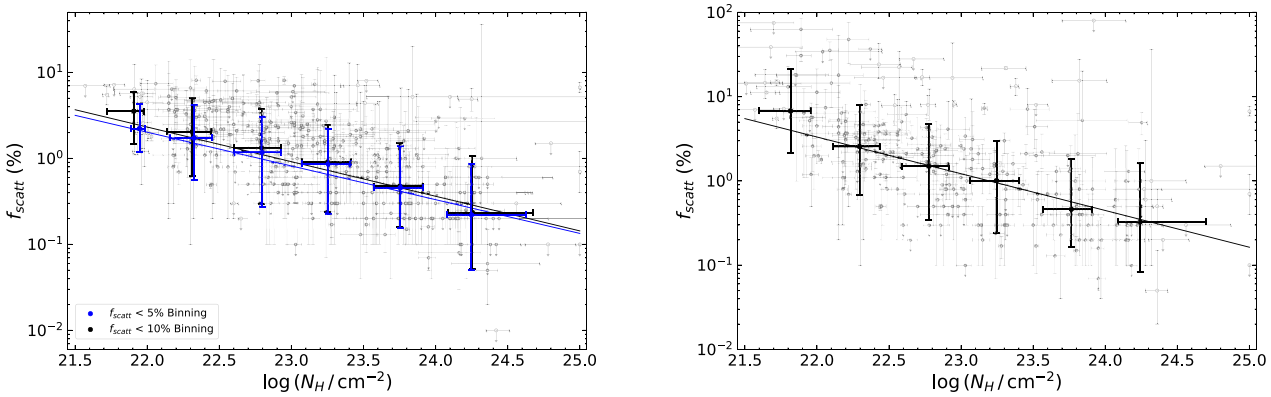
Ananna T. T., Treister E., Urry C. M., Ricci C., Hickox R. C., Padmanabhan N., Marchesi S., Kirkpatrick A., 2020, ApJ, 889, 17
Antonucci R., 1993, ARA&A, 31, 473
Arnaud K. A., 1996, in Jacoby G. H., Barnes J., eds, ASP Conf. Ser. Vol. 101, Astronomical Data Analysis Software and Systems V. Astron. Soc. Pac., San Francisco, p. 17
Barthelmy S. D. et al., 2005, Space Sci. Rev., 120, 143
Baumgartner W. H., Tueller J., Markwardt C. B., Skinner G. K., Barthelmy S., Mushotzky R. F., Evans P. A., Gehrels N., 2013, ApJS, 207, 19
Beckmann V. et al., 2009, A&A, 505, 417
Berney S. et al., 2015, MNRAS, 454, 3622
Bianchi S., Guainazzi M., Chiaberge M., 2006, A&A, 448, 499
Bianchi S., Chiaberge M., Evans D. A., Guainazzi M., Baldi R. D., Matt G., Piconcelli E., 2010, MNRAS, 405, 553
Bisogni S., Marconi A., Risaliti G., 2017, MNRAS, 464, 385
Braito V., Reeves J. N., Bianchi S., Nardini E., Piconcelli E., 2017, A&A, 600, A135
Brinkman A. C., Kaastra J. S., van der Meer R. L. J., Kinkhabwala A., Behar E., Kahn S. M., Paerels F. B. S., Sako M., 2002, A&A, 396, 761
Burrows D. N. et al., 2005, Space Sci. Rev., 120, 165
Caglar T. et al., 2020, A&A, 634, A114
Cappi M. et al., 2006, A&A, 446, 459
Comastri A., Setti G., Zamorani G., Hasinger G., 1995, A&A, 296, 1
Dadina M., 2007, A&A, 461, 1209
Dadina M., 2008, A&A, 485, 417
Dadina M., Guainazzi M., Cappi M., Bianchi S., Vignali C., Malaguti G., Comastri A., 2010, A&A, 516, A9
Eguchi S., Ueda Y., Terashima Y., Mushotzky R., Tueller J., 2009, ApJ, 696, 1657
Elitzur M., 2012, ApJ, 747, L33
Fabbiano G., Paggi A., Karovska M., Elvis M., Maksym W. P., Wang J., 2018, ApJ, 865, 83
Fabian A. C., Iwasawa K., Reynolds C. S., Young A. J., 2000, PASP, 112, 1145
Fabian A. C., Celotti A., Erlund M. C., 2006, MNRAS, 373, L16
Fabian A. C., Vasudevan R. V., Gandhi P., 2008, MNRAS, 385, L43
Fabian A. C., Vasudevan R. V., Mushotzky R. F., Winter L. M., Reynolds C. S., 2009, MNRAS, 394, L89
Feigelson E. D., Nelson P. I., 1985, ApJ, 293, 192
Ferrarese L., Merritt D., 2000, ApJ, 539, L9
Fukazawa Y. et al., 2011, ApJ, 727, 19
Gebhardt K. et al., 2000, ApJ, 539, L13
George I. M., Fabian A. C., 1991, MNRAS, 249, 352
Ghisellini G., Haardt F., Matt G., 1994, MNRAS, 267, 743
Gilli R., Comastri A., Hasinger G., 2007, A&A, 463, 79
Gómez-Guijarro C., González-Martín O., Ramos Almeida C., Rodríguez-Espinosa J. M., Gallego J., 2017, MNRAS, 469, 2720
Greene J. E., Pooley D., Zakamska N. L., Comerford J. M., Sun A.-L., 2014, ApJ, 788, 54
Guainazzi M., Bianchi S., 2007, MNRAS, 374, 1290
Haardt F., Maraschi L., 1991, ApJ, 380, L51
Hönig S. F., Gandhi P., Asmus D., Mushotzky R. F., Antonucci R., Ueda Y., Ichikawa K., 2014, MNRAS, 438, 647
Ichikawa K., Ueda Y., Terashima Y., Oyabu S., Gandhi P., Matsuta K., Nakagawa T., 2012, ApJ, 754, 45
Ichikawa K. et al., 2019a, ApJ, 883, L13
Ichikawa K. et al., 2019b, ApJ, 870, 31
Jansen F. et al., 2001, A&A, 365, L1
Kawamuro T., Ueda Y., Tazaki F., Ricci C., Terashima Y., 2016, ApJS, 225, 14
Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, MNRAS, 372, 961

Kinkhabwala A. et al., 2002, ApJ, 575, 732
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511
Kormendy J., Richstone D., 1995, ARA&A, 33, 581
Koss M. J. et al., 2016, ApJ, 825, 85
Koss M. et al., 2017, ApJ, 850, 74
Koss M. J. et al., 2018, Nature, 563, 214
Koss M. J. et al., 2021, ApJS, 252, 29
Krimm H. A. et al., 2013, ApJS, 209, 14
Krolik J. H., Madau P., Zycki P. T., 1994, ApJ, 420, L57
Lamperti I. et al., 2017, MNRAS, 467, 540
Levenson N. A., Krolik J. H., Życki P. T., Heckman T. M., Weaver K. A., Awaki H., Terashima Y., 2002, ApJ, 573, L81
Lightman A. P., White T. R., 1988, ApJ, 335, 57
Maddox N., 2018, MNRAS, 480, 5203
Magdziarz P., Zdziarski A. A., 1995, MNRAS, 273, 837
Magorrian J. et al., 1998, AJ, 115, 2285
Maiolino R., Salvati M., Bassani L., Dadina M., della Ceca R., Matt G., Risaliti G., Zamorani G., 1998, A&A, 338, 781
Marchesi S., Ajello M., Marcotulli L., Comastri A., Lanzuisi G., Vignali C., 2018, ApJ, 854, 49
Marconi A., Hunt L. K., 2003, ApJ, 589, L21
Mateos S. et al., 2016, ApJ, 819, 166
Matt G., Perola G. C., Piro L., 1991, A&A, 247, 25
Mitsuda K. et al., 2007, PASJ, 59, S1
Mushotzky R. F., Done C., Pounds K. A., 1993, ARA&A, 31, 717
Noguchi K., Terashima Y., Awaki H., 2009, ApJ, 705, 454
Noguchi K., Terashima Y., Ishino Y., Hashimoto Y., Koss M., Ueda Y., Awaki H., 2010, ApJ, 711, 144
Nucita A. A., Guainazzi M., Longinotti A. L., Santos-Lleo M., Maruccia Y., Bianchi S., 2010, A&A, 515, A47
Oh K., Sarzi M., Schawinski K., Yi S. K., 2011, ApJS, 195, 13
Oh K., Yi S. K., Schawinski K., Koss M., Trakhtenbrot B., Soto K., 2015, ApJS, 219, 1
Oh K. et al., 2017, MNRAS, 464, 1466
Page K. L., Reeves J. N., O'Brien P. T., Turner M. J. L., 2005, MNRAS, 364, 195
Paliya V. S. et al., 2019, ApJ, 881, 154
Paltani S., Ricci C., 2017, A&A, 607, A31
Panagiotou C., Walter R., 2019, A&A, 626, A40
Ponti G. et al., 2013, A&A, 549, A72
Ramos Almeida C., Ricci C., 2017, Nat. Astron., 1, 679
Reeves J. N., Turner M. J. L., 2000, MNRAS, 316, 234
Ricci C., Walter R., Courvoisier T. J. L., Paltani S., 2011, A&A, 532, A102
Ricci C., Ueda Y., Paltani S., Ichikawa K., Gand hi P., Awaki H., 2014, MNRAS, 441, 3622
Ricci C., Ueda Y., Koss M. J., Trakhtenbrot B., Bauer F. E., Gandhi P., 2015, ApJ, 815, L13
Ricci C. et al., 2017a, ApJS, 233, 17 (R17)
Ricci C. et al., 2017b, Nature, 549, 488
Ricci C. et al., 2018, MNRAS, 480, 1819
Risaliti G., Maiolino R., Salvati M., 1999, ApJ, 522, 157
Risaliti G., Salvati M., Marconi A., 2011, MNRAS, 411, 2223
Rojas A. F. et al., 2020, MNRAS, 491, 5867
Rybicki G. B., Lightman A. P., 1986, Radiative Processes in Astrophysics. Wiley, New York
Sako M., Kahn S. M., Paerels F., Liedahl D. A., 2000, ApJ, 543, L115
Sarzi M. et al., 2006, MNRAS, 366, 1151
Schawinski K., Koss M., Berney S., Sartori L. F., 2015, MNRAS, 451, 2517
Schurch N. J., Warwick R. S., Griffiths R. E., Kahn S. M., 2004, MNRAS, 350, 1
Shimizu T. T., Mushotzky R. F., Meléndez M., Koss M. J., Barger A. J., Cowie L. L., 2017, MNRAS, 466, 3161
Shu X. W., Yaqoob T., Wang J. X., 2010, ApJS, 187, 581
Smith K. L. et al., 2020, MNRAS, 492, 4216
Tanaka Y., Inoue H., Holt S. S., 1994, PASJ, 46, L37
Tanimoto A., Ueda Y., Kawamuro T., Ricci C., Awaki H., Terashima Y., 2018, ApJ, 853, 146
Tremaine S. et al., 2002, ApJ, 574, 740
Turner T. J., George I. M., Nandra K., Mushotzky R. F., 1997, ApJS, 113, 23
Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, ApJ, 598, 886
Ueda Y. et al., 2007, ApJ, 664, L79
Ueda Y., Akiyama M., Hasinger G., Miyaji T., Watson M. G., 2014, ApJ, 786, 104
Ueda Y. et al., 2015, ApJ, 815, 1
Urry C. M., Padovani P., 1995, PASP, 107, 803
Vasudevan R. V., Fabian A. C., 2009, MNRAS, 392, 1124
Vasudevan R. V., Mushotzky R. F., Winter L. M., Fabian A. C., 2009, MNRAS, 399, 1553
Veilleux S., Osterbrock D. E., 1987, ApJS, 63, 295
Vietri G. et al., 2018, A&A, 617, A81
Vignali C., Comastri A., Cappi M., Palumbo G. G. C., Matsuoka M., Kubo H., 1999, ApJ, 516, 582
Weisskopf M. C., Tananbaum H. D., Van Speybroeck L. P., O'Dell S. L., 2000, in Truemper J. E., Aschenbach B., eds, SPIE Conf. Ser. Vol. 4012, X-Ray Optics, Instruments, and Missions III. SPIE, Bellingham, p. 2
Winter L. M., Mushotzky R. F., Reynolds C. S., Tueller J., 2009, ApJ, 690, 1322
Yamada S., Ueda Y., Tanimoto A., Oda S., Imanishi M., Toba Y., Ricci C., 2020, ApJ, 897, 107
Yaqoob T., Padmanabhan U., 2004, ApJ, 604, 63
Young A. J., Wilson A. S., Shopbell P. L., 2001, ApJ, 556, 6
Zappacosta L. et al., 2018, ApJ, 854, 33

APPENDIX A: THE $f_{\rm scatt} - N_{\rm H}$ CORRELATION WITH DIFFERENT CUTS

High values of scattering fraction ($\geq 5\% - 10\%$) in obscured AGN can be attributed to partially covering absorbers or to jet emission (Ricci et al. 2017a). To assure that these sources do not affect our analysis, we apply an uppercut on the values of scattering fraction at 5% and 10% and hence drop 67 and 38 sources, respectively. We plot the remaining sources as a function of column density in Fig. A1(a). Using MC simulations (see Section 3), we calculate the correlation parameters and obtain a linear regression line with a slope $= -0.40 \pm 0.03$ for both cases, and p-value $= 6.5 \times 10^{-17}$ (2.5×10^{-19}), for the 5% (10%) cut on $f_{\rm scatt}$. As evident from the plot, we find a trend similar to the one we obtain for the entire sample.

To verify that the anticorrelation we find between scattering fraction and column density is not due to sources with low counts, we apply a cut at 200 counts per source on our sample. It implies removing 136 sources from our original sample of 386 sources. We use MC simulations to quantify the dependence of $f_{\rm scatt}$ on $\log N_{\rm H}$ for the modified sample. Fig. A1(b) demonstrates that these sources do not affect the original correlation, as we recover a similar correlation with a linear regression slope $= 0.44 \pm 0.03$ and p-value $= 7.0 \times 10^{-16}$.

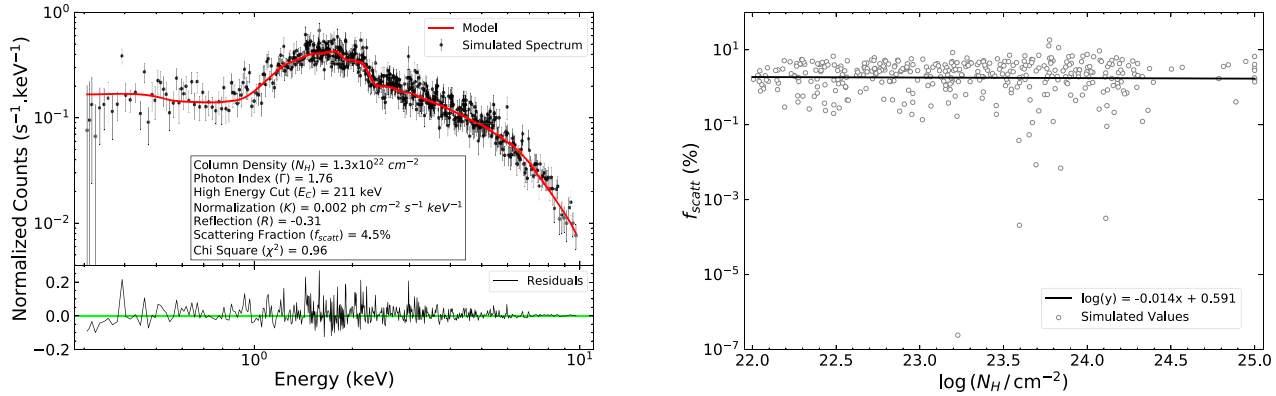


(a) (b)

Figure A1. (a) Scattering fraction versus column density with an uppercut on $f_{\rm scatt}$ at 5% (blue) and 10% (black). The grey open circles in the background show the data points (with error bars showing errors at 90% confidence level) for objects with $f_{\rm scatt} < 10\%$. The black and blue circles are the median and uncertainty (1σ) in $f_{\rm scatt}$ and log $N_{\rm H}$ calculated using MC simulations for each log $N_{\rm H}$ bin and both the linear regression lines are also obtained using the same method. (b) Scattering fraction versus column density for objects with > 200 counts. The data points for these objects are shown as grey open circles in the background (error bars show values in the 90% confidence level). The black circles correspond to the median and uncertainty (1σ) in $f_{\rm scatt}$ and log $N_{\rm H}$ calculated using MC simulations for each log $N_{\rm H}$ bin, while the black line is the linear regression line obtained using the same method.

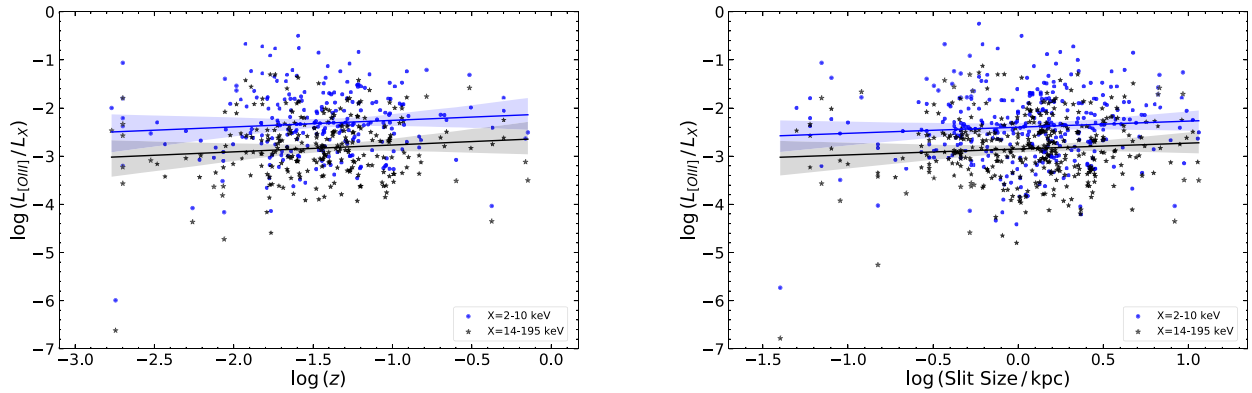
APPENDIX B: X-RAY SPECTRAL SIMULATIONS

We ran multiple spectral simulations to exclude the possibility that the inverse correlation we find between the fraction of Thomson-scattered radiation and the line-of-sight column density is due to the degeneracy between these parameters (discussed in Section 3.1). In Fig. B1(a), we show an example of a simulated spectrum, along with the model used for simulating and fitting it and the residuals after the fitting. We also show in Fig. B1(b) a correlation between the set of simulated parameters.



(a) (b)

Figure B1. (a) An example of a simulated spectrum created using XSPEC. The model used to simulate and fit the spectrum is shown in red. The text box shows the best-fit values obtained for the various free parameters, while the bottom panel shows the fit residuals. (b) An example of a correlation between the simulated values of scattering fraction and column density obtained from XSPEC simulations. The black line shows the linear regression fit to the simulated data set.



(a) (b)

Figure C1. (a) The ratio of observed [O III] λ5007 luminosity to intrinsic X-ray luminosity versus redshift. (b) The ratio of observed [O III] λ5007 luminosity to intrinsic X-ray luminosity versus physical slit size (kpc) used for [O III] λ5007 measurements. The blue and black lines show the linear regression fit to our data for intrinsic X-ray luminosity in the 2–10 keV and 14–195 keV energy bands, respectively. The shaded regions show the uncertainty in the regression lines. The lack of correlation in both cases (r-value = 0.087 and 0.077, respectively) suggests that the [O III] λ5007 measurements are not significantly affected by the aperture size of the sources.

APPENDIX C: APERTURE EFFECTS ON [O III] λ5007 MEASUREMENTS

We have adopted observed flux measurements of the [O III] λ5007 emission line in our analysis. To eliminate possible aperture effects on these measurements, we investigated potential correlations between the ratio of observed [O III] λ5007 to intrinsic X-ray luminosity and the redshift (Fig. C1[a]). We obtain a linear regression with slope = 0.14 ± 0.10 for both energy bands and p-value = 0.18 (0.13), for 2–10 keV (14–195 keV) X-ray luminosity. We also checked, in Fig. C1(b), if $L_{[O\,III]}/L_X$ correlates with the physical size of the slits used for the [O III] λ5007 measurements (reported in Koss et al. 2017 and Koss et al., in preparation). We get a linear regression line with slope = 0.13 ± 0.0.09 (0.12 ± 0.09) for 2–10 keV (14–195 keV) X-ray luminosity and p-value = 0.16 for both cases. Based on these results, we can conclude that no significant correlation exists between the parameters, and we can exclude aperture effects from our analysis.

This paper has been typeset from a TEX/LATEX file prepared by the author.