

# Self-consistent population synthesis of active galactic nuclei from observational constraints in the X-rays

Dimitra Gerolymatou<sup>1,\*</sup>, Stéphane Paltani<sup>1</sup>, Claudio Ricci<sup>2,3</sup>, and Manon Regamey<sup>1</sup>

<sup>1</sup> Department of Astronomy, University of Geneva, 1290 Versoix, Switzerland

<sup>2</sup> Instituto de Estudios Astrofísicos, Universidad Diego Portales, Chile

<sup>3</sup> Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, China

Received 24 January 2025 / Accepted 16 June 2025

## ABSTRACT

The cosmic X-ray background (CXB) is produced by the emission of unresolved active galactic nuclei (AGNs), thus providing key information about the properties of the primary and reprocessed X-ray emission components of the AGN population. Equally important are studies of individual sources that provide additional constraints on the properties of AGNs, such as their luminosity and obscuration. Until now, these constraints have not been self-consistently addressed by intrinsically linking emission, absorption, and reflection. Here, we performed numerical simulations with the ray-tracing code REFLEX, which allows us to self-consistently model the X-ray emission of AGNs with flexible geometries for the circumnuclear medium. Using the REFLEX-simulated emission of an AGN population, we attempted to simultaneously reproduce the CXB and absorption properties measured in the X-rays, namely the observed fraction of  $N_{\text{H}}$  in bins of  $\log(N_{\text{H}})$  and the fraction of absorbed AGNs, including their redshift and luminosity evolution. We sampled an intrinsic X-ray-luminosity function and constructed gradually more complex, physically motivated geometrical models. We examined how well each model can match all observational constraints using a simulation-based inference (SBI) approach. We find that, while the simple unification model can reproduce the CXB, a luminosity-dependent dusty torus is needed to reproduce the absorption properties. When adding an accretion disc, the model best matches all constraints simultaneously. Our synthetic population is able to reproduce the dependence of the covering factor on luminosity, the AGN number counts from several surveys, and the observed correlation between reflection and obscuration. Finally, we derived an intrinsic Compton-thick fraction of  $21 \pm 7\%$ , which is consistent with local observations.

**Key words.** galaxies: active – X-rays: diffuse background – X-rays: galaxies

## 1. Introduction

Supermassive black holes (SMBHs) at the cores of galaxies are the engines of active galactic nuclei (AGNs) during an accreting phase. The SMBHs are closely linked to essential properties of their host galaxies, including mass, velocity dispersion, and luminosity (e.g. Magorrian et al. 1998; Gebhardt et al. 2000; Kormendy & Ho 2013). Such correlations hint at the potential for SMBHs to regulate star formation rates and play a significant role in galaxy evolution. To comprehensively probe the evolution of SMBHs, AGN demographics have been examined through population studies, allowing us to build their luminosity function (e.g. Ueda et al. 2014; Aird et al. 2015; Ananna et al. 2019) and other properties such as the black-hole mass function and Eddington-ratio distribution function.

X-ray radiation is a universal feature of AGNs that typically outshines the emission of the host galaxy (e.g. Brandt & Alexander 2015). Furthermore, X-ray observations are very efficient for detecting obscured AGNs, as X-ray photons can penetrate the high column density of the surrounding dusty material, thus allowing us to study the innermost regions of AGNs. The primary X-ray emission is attributed to the inverse Comptonisation of accretion-disc photons in a hot corona close to the SMBH (e.g. Haardt & Maraschi 1991), while reprocessed

components come from reflection on the circumnuclear material (e.g. Pounds et al. 1990; Ghisellini et al. 1994). The circumnuclear material is thought to be in the form of a dusty torus (see Hickox & Alexander 2018 for an overview) and acts as an absorber and reflector at the same time, while other dense structures such as the accretion disc can also act as reflectors. The integrated emission from the AGN population is imprinted on the cosmic X-ray background (CXB; e.g. Brandt & Yang 2022). A clear peak in the CXB spectrum is observed between 20–30 keV, which is attributed to the X-ray radiation reflected on the circumnuclear material. The CXB is often used as one of the constraints for AGN population-synthesis models (e.g. Gilli et al. 2007; Akylas et al. 2012; Ueda et al. 2014; Ananna et al. 2019); however, it can be reproduced by different AGN populations, for instance by assuming different AGN spectral features, the shape of the X-ray luminosity function (XLF), the intrinsic neutral hydrogen column-density ( $N_{\text{H}}$ ) distribution, and the fraction of Compton-thick (CT) AGNs (which we define as the fraction of AGNs with  $N_{\text{H}} \geq 10^{24} \text{ cm}^{-2}$ ), as demonstrated in Treister et al. (2009), Akylas et al. (2012), and Ananna et al. (2019). Some degeneracy between these parameters could be lifted by using AGN number counts obtained in X-ray surveys (e.g. Ueda et al. 2014; Aird et al. 2015; Ananna et al. 2019).

X-ray surveys of AGNs also offer insights into their absorption properties, which strongly affect the emission from AGNs and thus the modelling of the CXB. The fraction of absorbed AGNs (usually defined as the fraction of AGNs with  $N_{\text{H}} \geq$

\* Corresponding author: [dimitra.gerolymatou@unige.ch](mailto:dimitra.gerolymatou@unige.ch)

$10^{22} \text{ cm}^{-2}$ ) has been consistently found to be anti-correlated with luminosity (e.g. Lawrence 1991; Ueda et al. 2003; Hasinger 2008; Ueda et al. 2014; Aird et al. 2015; Buchner et al. 2015), suggesting that there are intrinsic physical differences between the populations of obscured and unobscured AGNs. The intrinsic fraction of CT AGNs is still debated, with previous studies finding values from 10 to 50% (e.g. Brightman & Nandra 2011; Akylas et al. 2012; Ueda et al. 2014; Ricci et al. 2015; Ananna et al. 2019). As the CXB peak at 20–30 keV can be reproduced by either a high fraction of CT AGNs or stronger reflected X-ray emission from all AGNs, the amount of reflection must be anti-correlated with the fraction of CT AGNs. Studies with stronger reflection components find a CT fraction as low as 10% (e.g. Treister et al. 2009; Ricci et al. 2011), while the most recent population synthesis of AGNs in Ananna et al. (2019), which is able to simultaneously reproduce the CXB and AGN number counts from multiple X-ray surveys, predicts an intrinsic CT fraction of  $56 \pm 7\%$  for  $z \lesssim 1$ .

An important caveat of all previous studies is that the absorption of AGNs is assumed and is not physically linked with any other AGN properties. Until now, X-ray observables concerning the absorption properties of AGNs, such as the observed column-density distribution and the dependence of the absorbed fraction on X-ray luminosity and redshift, have not been self-consistently addressed by physically linking obscuration and reflection. In this study, we create a fully self-consistent population-synthesis model of AGNs, where the X-ray obscuration and reflection are linked through the distribution of matter. In order to do so, we carried out numerical simulations using the ray-tracing code REFLEX (Paltani & Ricci 2017), which allows us to simulate the X-ray emission from AGNs with different geometries. We sampled an intrinsic X-ray-luminosity function to create a synthetic AGN population and used gradually more complex and physically motivated geometrical models. With each model, we attempted to reproduce the CXB and the observed absorption properties. We used a simulation-based inference approach (Tejero-Cantero et al. 2020; Greenberg et al. 2019) to compare the models to the observations and infer the model parameters that best match all constraints simultaneously. Finally, we examined the luminosity dependence of absorption, the AGN number counts at different energy bands, the reflection-obscuration correlation, and the impact of galactic obscuration on our population, and we derived the intrinsic CT fraction of the AGN population. Throughout this paper, all logarithmic values ( $\log$ ) refer to logarithms with base 10, and we adopted the following cosmological parameters:  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m = 0.3$ , and  $\Omega_\Lambda = 0.7$ .

## 2. X-ray observational constraints

To constrain the geometrical and physical properties of the AGN population, we had to gather several observational AGN properties. In Fig. 1, we show the X-ray observations used as constraints in this work, and we describe each of them in this section. The constraints are the CXB and several absorption properties from studies of X-ray surveys, namely the observed  $N_H$  distribution and the dependence of the observed fraction of absorbed AGNs on observed X-ray luminosity and on redshift. We chose observed properties because the intrinsic ones depend on the assumptions made in other papers. Instead, we applied the selection function of each survey on our synthetic population.

### 2.1. Cosmic X-ray background

The CXB is formed by the integrated emission of all AGN across cosmic time (see Brandt & Yang 2022 for an overview). As our first constraint, we use some of the most recent measurements of the CXB, comprehensively covering our target energy range, obtained from five instruments: Chandra, INTEGRAL (JEM-X, IBIS/ISGRI, SPI), RXTE/PCA, *Swift*/BAT and *Swift*/XRT (Cappelluti et al. 2017; Churazov et al. 2007; Revnivtsev et al. 2003; Ajello et al. 2008; Moretti et al. 2009; Fig. 1 top). All measurements find the CXB spectral index to be  $\sim 1.4$  below 10 keV. Any normalisation discrepancies between data sets may mostly be the result of inaccurate cross-calibration of spectra, unclear instrumental backgrounds, and cosmic variance (e.g. Moretti et al. 2009). We note the pioneering role of the HEAO-1 mission in the discovery and early characterisation of the CXB shape (Gruber et al. 1999). The statistical disagreement in normalisation between HEAO-1 and our chosen measurements is well documented and discussed in the literature (see, e.g. Churazov et al. 2007; Ajello et al. 2008; Moretti et al. 2009, and references therein). We considered the energy range 2–100 keV, as at  $E > 100 \text{ keV}$  there is a significant contribution from blazars (e.g. Draper & Ballantyne 2009), and at  $E < 2 \text{ keV}$  the AGN emission becomes quite complex because of absorption, soft-excess emission, and non-negligible contribution from galaxies (e.g. Cappelluti et al. 2012).

### 2.2. Absorption properties

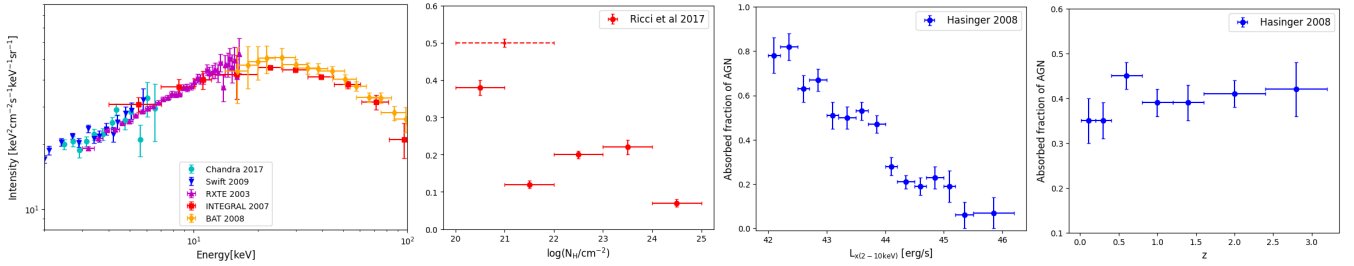
#### 2.2.1. Column-density distribution

The second constraint is the observed line-of-sight  $N_H$  distribution of local (median redshift  $z = 0.0367$ ) AGNs from Ricci et al. (2017a; hereafter R17). Their sample includes 731 non-blazar AGNs detected in the 14–195 keV energy band in the *Swift*/BAT 70-month all-sky survey. Being less biased against high- $N_H$  AGNs, it consists of  $\sim 50\%$  obscured AGN. The hard X-ray sensitivity of BAT allows us to investigate the distribution of absorption up to the CT regime, albeit with significant selection bias against CT AGN.

The  $N_H$  distribution presented in R17 shows some evidence of bimodality with the highest peak for  $\log(N_H/\text{cm}^{-2}) \leq 21$  and the secondary in the  $\log(N_H/\text{cm}^{-2}) = 23\text{--}24$  bin (see Fig. 1, bottom left). The R17 distribution is largely consistent with NuSTAR observations in the 8–24 keV band, which also show two peaks at the same  $N_H$  (e.g. Zappacosta et al. 2018). This bimodality suggests that the obscuration may be caused by different structures, possibly due to the plane of the galaxy hosting the AGN at lower  $N_H$  and due to the denser circumnuclear material at higher  $N_H$  (e.g. Paltani et al. 2008). Considering the purpose of this study, we do not model the host galaxy, resulting in a lack of mildly obscured ( $N_H < 10^{22} \text{ cm}^{-2}$ ) AGNs. Hence, we compared the number of simulated unobscured objects to all AGNs with  $N_H < 10^{22} \text{ cm}^{-2}$ , assuming galactic column densities of  $\log(N_H/\text{cm}^{-2}) < 22$  (e.g. Güver & Özel 2009).

#### 2.2.2. Fraction of absorbed AGNs

The last two constraints for this work are the observed fraction of absorbed AGNs as a function of the observed X-ray luminosity and as a function of redshift from Hasinger (2008; hereafter H08). In H08, 1290 AGNs detected in the 2–10 keV range from ten independent samples (summarised in Table 1 in their paper) are classified as type I (unabsorbed) or type II (absorbed) by



**Fig. 1.** X-ray observables used as constraints in this work. *From the left:* CXB measured by Chandra (cyan; Cappelluti et al. 2017), INTEGRAL (red; Churazov et al. 2007), RXTE (magenta; Revnivtsev et al. 2003), *Swift*/BAT (orange; Ajello et al. 2008), and *Swift*/XRT (blue; Moretti et al. 2009). The observed  $N_{\text{H}}$  distribution of 731 AGNs in the 70-month *Swift*/BAT catalogue, detected in the 14–195 keV range, from Ricci et al. (2017a). The point with the dashed error bars represents the sum of the first two bins used in this work (see Sect. 2.2 for more information). The absorbed fraction of 1290 AGNs from multiple surveys in the 2–10 keV range is shown as a function of observed luminosity and as a function of redshift from Hasinger (2008).

combining optical spectroscopy and X-ray-hardness ratios. Optical/UV broad lines are used to classify sources into broad-line AGNs and narrow-line AGNs. The distinction found between the two classes corresponds to  $N_{\text{H}} \approx 3 \times 10^{21} \text{ cm}^{-2}$ , above which only narrow lines are observable. Thus, for the calculation of the absorbed fraction, we considered the AGNs with  $\log(N_{\text{H}}/\text{cm}^{-2}) > 21.5$  [instead of the usual  $\log(N_{\text{H}}/\text{cm}^{-2}) > 22$ ] as absorbed in order to ensure a meaningful comparison with the H08 results.

In H08, the fraction of absorbed AGNs is found to decrease with increasing logarithmic luminosity with a slope of  $-0.226 \pm 0.014$  (fitted with a linear function) for  $\log(L_{\text{X}}/\text{erg s}^{-1}) \geq 42$ . The presence of the anti-correlation has been confirmed by multiple studies carried out in the last three decades (e.g. Lawrence 1991; Ueda et al. 2003, 2014; Aird et al. 2015; Buchner et al. 2015). At the same time, in H08, the fraction of absorbed AGNs is found to increase with redshift as  $(1+z)^{0.48}$  over the 0–3.2 redshift range at fixed luminosity. Such evolution was confirmed in later studies (Ueda et al. 2014; Aird et al. 2015; Liu et al. 2017; Iwasawa et al. 2020). Here, however, we considered the fraction integrated over the entire range of luminosity to avoid larger uncertainties caused by the low number of sources when the samples are divided into luminosity bins. In this case, the dependence of the fraction of absorbed AGNs on redshift is flat (Fig. 1 bottom right).

We opted for H08 instead of the more recent works cited above because it provides the absorbed fraction across broad luminosity and redshift ranges, in addition to using observed X-ray luminosities. This reduces the uncertainties that can arise from dividing into narrow bins with limited detected sources and avoids the additional assumptions made in other studies to infer intrinsic luminosities. Moreover, this is a consistent (as some studies do not provide the absorbed fraction as functions of both luminosity and redshift) and sufficient (as we do not expect that our simplistic models can explain all detailed features) set of constraints. Finally, in contrast to R17, the H08 sample contains both local and high redshift AGNs, meaning that our synthetic population has to be able to simultaneously match the absorption properties of AGNs detected at different redshifts as well as in different energy ranges (2–10 keV and 14–195 keV).

### 3. Modelling of the AGN population

We sought to create a self-consistent synthetic population of AGNs that can simultaneously reproduce the CXB and the observed absorption properties of AGNs. In our work, we looked for the geometrical model to best match the observables. In this section, we present the tools, models, and assumptions that we

used to create the synthetic AGN population and simulate their emission.

#### 3.1. *Reflex*

We used the ray-tracing code REFLEX<sup>1</sup> (Paltani & Ricci 2017) to self-consistently simulate the X-ray emission of AGNs with different geometries. REFLEX performs realistic propagation of X-ray (0.1–1 MeV) photons through the circumnuclear material of AGNs using Monte Carlo methods. Arbitrary geometries with spatially varying composition and density can be selected by the user. Moreover, the geometry and emission spectrum of the X-ray source can also be selected. The most important physical processes in the X-rays are implemented, namely Compton and Rayleigh scattering, photoelectric absorption, and fluorescence. In REFLEX 3.0 (Ricci & Paltani 2023), dust absorption and scattering is also included following the Milky Way model from Draine (2003). We used the latest version of REFLEX to generate spectra at specific inclinations and calculate fluxes in selected energy bands.

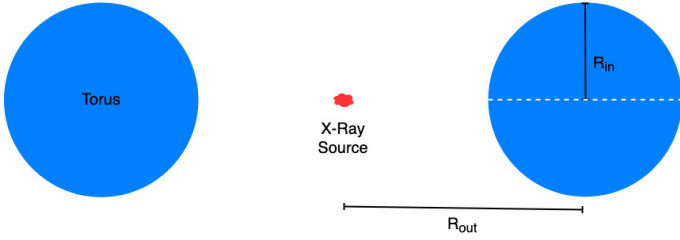
#### 3.2. X-ray luminosity function

To create a population of AGNs for our simulations, we needed an X-ray luminosity function, which provides the AGN number density in  $\text{Mpc}^{-3} \text{ dex}^{-1}$  as a function of redshift and logarithmic X-ray luminosity. We chose the intrinsic (i.e. absorption-corrected) rest-frame 2–10 keV luminosity function derived in Ueda et al. (2003; hereafter U03), which was frequently adopted in previous population syntheses models of the CXB (e.g. Gilli et al. 2007; Akylas et al. 2012; Esposito & Walter 2016). The U03 XLF is built from a sample containing 247 HEAO-1-, ASCA-, and Chandra-detected sources over the 2–10 keV range; thus, they assumed that there was a negligible percentage of CT AGNs in their sample.

We opted for U03 instead of more recent studies, where the XLFs are non-parametric (e.g., Buchner et al. 2015; Ananna et al. 2019), derived separately for each AGN type (e.g. Aird et al. 2015), or their sample contains CT AGNs (e.g. Ueda et al. 2014). As shown in Ueda et al. (2014) and Ananna et al. (2019), the  $\log(N_{\text{H}}/\text{cm}^{-2}) < 24$  XLFs mostly agree with U03, even though they assume different  $N_{\text{H}}$  distribution functions.

Our synthetic AGN population automatically creates a number of CT AGNs. Therefore, our LF is normalised so that the

<sup>1</sup> <https://www.astro.unige.ch/reflex/>



**Fig. 2.** Torus geometry in REFLEX. The inner radius,  $R_{\text{in}}$ , is defined as the radius of the cross-section, while the outer radius,  $R_{\text{out}}$ , is measured from the centre of the torus to the centre of the cross-section. The equatorial column density,  $N_{\text{H,eq}}$ , is measured across the diameter of the cross-section (white dashed line). The X-ray point source is placed at the centre of the torus.

number density of AGN with  $N_{\text{H}} < 10^{24} \text{ cm}^{-2}$  matches the U03 XLF. The fraction of CT AGNs is then determined self-consistently from the model parameters obtained from the posterior distributions. It is therefore an advantage to use an XLF derived from a sample containing a negligible amount of CT AGNs, especially considering that X-ray surveys are biased against detecting heavily CT AGNs even at  $>10 \text{ keV}$ .

The assumptions made in U03 (or later XLFs) affect our parent distribution and thus our results. Although it would be possible to include the fitting of the parameters of the XLF, for the purposes of this work we used the best-fit values derived in Table 3 of U03. In their work, the XLF,  $\phi(L_{\text{X}}^{\text{int}}, z)$ , is described by a luminosity-dependent density evolution (LDDE), starting with a smoothly connected double power law at  $z = 0$ :

$$\phi(L_{\text{X}}^{\text{int}}, z = 0) = \frac{dN(L_{\text{X}}^{\text{int}}, z = 0)}{d \log L_{\text{X}}^{\text{int}}} = K \left[ \left( \frac{L_{\text{X}}^{\text{int}}}{L_*} \right)^{\gamma_1} + \left( \frac{L_{\text{X}}^{\text{int}}}{L_*} \right)^{\gamma_2} \right]^{-1}, \quad (1)$$

where  $K$  is the normalisation in units of  $10^{-6} \text{ Mpc}^{-3}$ ,  $L_*$  the characteristic break luminosity in  $\text{erg s}^{-1}$ ,  $\gamma_1$  and  $\gamma_2$  the faint-end and bright-end slopes, respectively. The XLF evolves as a function of redshift and luminosity:

$$\phi(L_{\text{X}}^{\text{int}}, z) = \phi(L_{\text{X}}^{\text{int}}, z = 0) e(z, L_{\text{X}}^{\text{int}}). \quad (2)$$

The evolution term,  $e(z, L_{\text{X}}^{\text{int}})$ , is described by a power-law function of  $(1+z)$  with a cut-off redshift that is a function of luminosity, as defined in Eq. 16 and 17 in U03.

### 3.3. Models

The physical models of the AGN population that we considered and their geometries constructed in REFLEX are explained below and summarised in Table 1. All of our models comprise a cold, dusty homogeneous torus, the geometry of which is shown in Fig. 2. The torus is fully described in REFLEX by its inner,  $R_{\text{in}}$ , and outer,  $R_{\text{out}}$ , radii, as well as its equatorial column density,  $N_{\text{H,eq}}$ , and its composition. We set the molecular hydrogen fraction to 0.5 (e.g. Wada et al. 2009), the dust fraction (defined as the fraction of Fe atoms locked into dust grains) to 1, and the matter composition to Lodders (2003); we assumed solar metallicity. As shown in McKaig et al. (2022), varying standard abundances do not result in significant differences in the simulated spectra at the level of accuracy we want to achieve here.

For the two models that include only the torus, we placed the X-ray point source at the centre of the torus as a good approximation. In all cases, we chose a power-law emission spectrum for the X-ray source with typical values for the photon index,  $\Gamma = 1.9$ , and high-energy exponential cut-off,  $E_{\text{C}} = 200 \text{ keV}$  (e.g. Ricci et al. 2017a, 2018).

#### 3.3.1. Simple torus

We started from the simple unification of AGNs, which suggests that the observed differences between type-I and type-II AGNs are caused only by orientation effects depending on the observer's line of sight (Antonucci 1993). Our model is composed of an obscuring torus whose  $R_{\text{in}}/R_{\text{out}}$  ratio is a free parameter independent of any other AGN properties and remains the same throughout the population. Hence, in this population the only difference between different AGNs is the torus  $N_{\text{H,eq}}$ . Because the geometry of this model is only dependent on the ratio  $R_{\text{in}}/R_{\text{out}}$ , which sets the covering factor, the simulation can be scale-free. We fixed  $R_{\text{in}} = 0.5$  (in arbitrary units) and let  $R_{\text{out}}$  vary as  $R_{\text{out}} = R_{\text{in}} + R_1$ , with  $R_1$  being a free positive parameter.

#### 3.3.2. Luminosity-dependent torus

We chose our next physically motivated model to include a relationship between the location of the inner edge of the torus and the luminosity. Thus, we constructed a model based on the 'receding torus' (Lawrence 1991), which is often invoked to explain the decrease of the fraction of absorbed AGNs as a function of luminosity. In fact, it applies to a dusty torus, where the dust is destroyed when it reaches a temperature of  $\sim 1500 \text{ K}$  (see Nenkova et al. 2008) due to irradiation by the central engine. In the receding-torus model, the inner edge of the obscuring torus,  $r_{\text{dust}} = R_{\text{out}} - R_{\text{in}}$ , is set by the dust-sublimation radius, which increases with luminosity as  $r_{\text{dust}} \propto L^{0.5}$ . Therefore, if dust-free gas behaves as dusty gas, the fraction of absorbed AGNs is expected to scale as  $\propto L^{-0.5}$  at high luminosities.

Our luminosity-dependent (LD) model introduces a dependence of the covering factor of the scaleless tori on the intrinsic X-ray luminosity of the AGNs, with the assumption that gas behaves similarly to dust. We let the slope of the dependence vary because the 0.5 index of the receding torus model does not agree with observations of the absorbed fraction in the X-rays, which find a much flatter slope (e.g. H08, Burlon et al. 2011). We discuss the possible mechanisms for this in Sect. 5.1. For practical reasons, we again fixed  $R_{\text{in}} = 0.5$  (in arbitrary units) and let  $R_{\text{out}}$  scale with luminosity as

$$r_{\text{dust}} = R_{\text{out}} - R_{\text{in}} = R_0 \left( \frac{L_{\text{X}}^{\text{int}}}{L_0} \right)^{\alpha}, \quad (3)$$

where  $L_0 = 10^{44} \text{ erg s}^{-1}$  is arbitrarily fixed, and  $R_0$  and  $\alpha$  are free parameters.

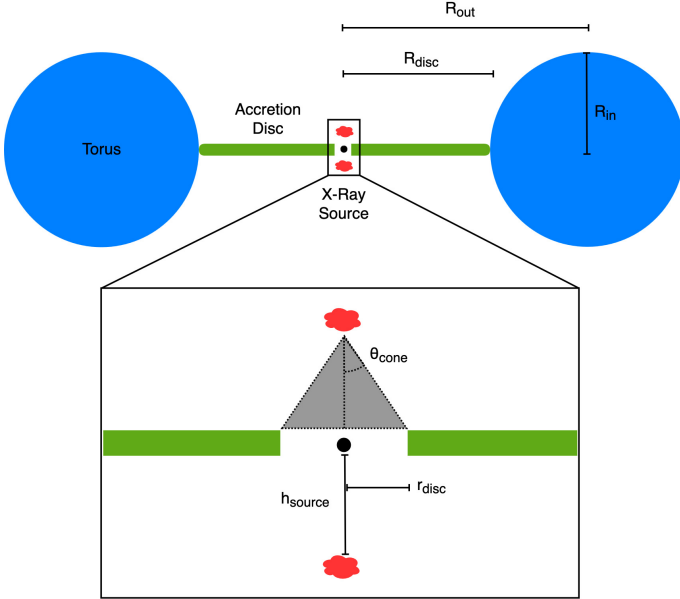
#### 3.3.3. Accretion disc

To make our model more realistic, we added a flat, optically thick, dust-free annulus to the previous model to simulate the presence of an accretion disc (AD). The outer radius of the annulus is set as  $R_{\text{disc}} \equiv r_{\text{dust}} = R_{\text{out}} - R_{\text{in}}$ , so that the torus is its physical continuation (Fig. 3). This model (hereafter LD+AD) cannot be scaleless because the height of the X-ray source is not related to other distances, such as the inner edge of the torus. Instead, we fixed  $R_{\text{in}} = 1 \text{ pc}$  (e.g. Ramos Almeida & Ricci 2017) and the inner radius of the annulus to  $r_{\text{disc}} = 6r_{\text{g}}$  (e.g. Wilkins & Fabian 2011), where  $r_{\text{g}} = 2GM_{\text{BH}}/c^2$  is the Schwarzschild gravitational radius and assuming a typical SMBH mass of  $M_{\text{BH}} = 10^7 M_{\odot}$  (e.g. Koss et al. 2017). We set H and He to be fully ionised, although since the disc is completely optically thick, this choice has a negligible effect on the simulated spectra.

**Table 1.** Overview of the models explored, their free parameters, and priors as discussed in Sect. 3.3.

Model <sup>1</sup>	$\mu$ (log/cm <sup>-2</sup> ) <sup>2</sup>	$\sigma$ (log/cm <sup>-2</sup> ) <sup>2</sup>	$R_0$ <sup>3,a</sup>	$\alpha$ <sup>3</sup>	$R_1$ <sup>4,a</sup>
Simple torus	21.5–25.5	0.01–2.5	–	–	0.01–1.5
LD torus	21.5–25.5	0.01–2.5	0.01–1.5	0.01–1.0	–
LD+AD	21.5–25.5	0.01–2.5	0.01–1.5	0.01–1.0	–

**Notes.** <sup>1</sup>Simple torus, luminosity-dependent (LD) torus, and LD torus with an accretion disc (AD). <sup>2</sup>Mean,  $\mu$ , and standard deviation,  $\sigma$ , of the tori  $N_{\text{H,eq}}$  distribution. <sup>3</sup>Normalisation,  $R_0$ , and slope,  $\alpha$ , in Eq. (3). <sup>4</sup>Scaling parameter,  $R_1$ , of the simple torus. <sup>a</sup>Arbitrary units for models without AD and measured in parsec for the AD model.



**Fig. 3.** Accretion-disc and double-X-ray source geometry in REFLEX (not to scale). The outer radius of the thin annulus representing the AD is set as  $R_{\text{disc}} = R_{\text{out}} - R_{\text{in}}$ , with  $R_{\text{in}} = 1\text{pc}$ , and the inner radius of the annulus as  $r_{\text{disc}} = 6r_g$ . We place a double X-ray point source above and below the centre of the annulus at  $h_{\text{source}} = 10r_g$ . The grey area represents the cone defined by  $h_{\text{source}}$  and  $r_{\text{disc}}$ , inside which we do not allow the sources to emit, as in reality emission through the centre of the annulus would be blocked by the SMBH. The half-opening angle of the cone,  $\theta_{\text{cone}} = \arctan(r_{\text{disc}} / h_{\text{source}})$ , is denoted by the dashed lines.

For this model, we placed a double X-ray point source located above and below the centre of the annulus at  $h_{\text{source}} = 10r_g$ , following the lamppost geometry (Martocchia & Matt 1996; Miniutti & Fabian 2004). For the purposes of this study, we adopted the lamppost geometry as a widely used and long-standing approximation of the X-ray source despite its known limitations (e.g. Tagliacozzo et al. 2023), since the geometry of the compact corona has negligible effects on the properties of the reflected radiation. To prevent the escape of emission through the centre of the annulus, which in reality would be blocked by the SMBH, we forced the emission of the sources to remain outside the cones defined by  $h_{\text{source}}$  and  $r_{\text{disc}}$  (as shown in Fig.3), and we corrected the intrinsic luminosity of each object in our population by the factor  $(1 - \cos \theta_{\text{cone}}) / 2$ , where  $\theta_{\text{cone}} = \arctan(r_{\text{disc}} / h_{\text{source}})$  is the half-opening angle of the cones.

### 3.4. Synthetic population

After defining the models, we created a large sample of AGNs for which we simulate the X-ray emission and determined

absorption properties. For each object in our synthetic population, we drew the redshift and intrinsic 2–10 keV X-ray luminosity of the X-ray source by sampling the XLF derived in U03, the observation angle,  $\theta_{\text{obs}}$ , from a cosine distribution between  $0^\circ$  and  $90^\circ$ , and the equatorial column density,  $N_{\text{H,eq}}$ , of the torus (Fig. 2) from a log-normal distribution,  $\log(N_{\text{H,eq}}/\text{cm}^{-2}) \sim \mathcal{N}(\mu, \sigma)$ , where  $\mu$  and  $\sigma$  are the mean and the standard deviation of the distribution, respectively, and free parameters in all our models.

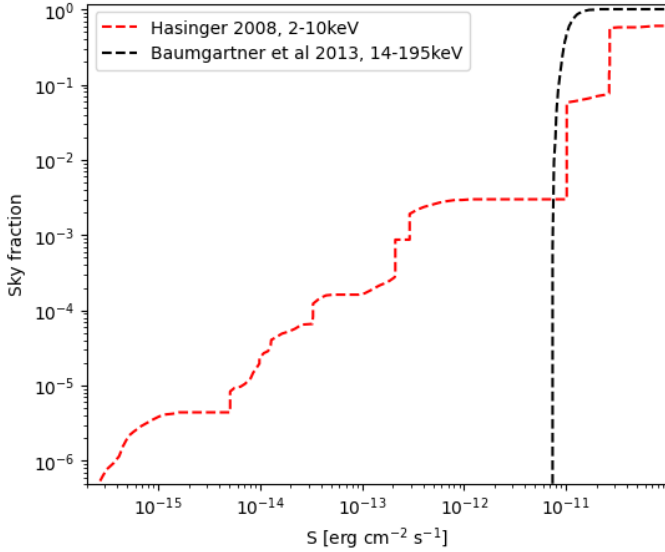
We calculated the line-of-sight column density of the homogeneous torus,  $N_{\text{H}}$ , at a certain observation angle,  $\theta_{\text{obs}}$ , as shown in Paltani & Ricci (2017):

$$N_{\text{H}} = \begin{cases} N_{\text{H,eq}} \left(1 - \left(\frac{R_{\text{out}}}{R_{\text{in}}}\right)^2 \cos^2 \theta_{\text{obs}}\right)^{1/2}, & \cos \theta_{\text{obs}} < \frac{R_{\text{in}}}{R_{\text{out}}} \\ 0 \text{ cm}^{-2}, & \cos \theta_{\text{obs}} > \frac{R_{\text{in}}}{R_{\text{out}}} \end{cases} \quad (4)$$

We performed numerical simulations using REFLEX to create a grid of models of AGN X-ray spectra in the 2–100 keV energy range, with  $\log(N_{\text{H,eq}}/\text{cm}^{-2})$  between 21 and 26 with a step of 0.1, and  $\log(R_{\text{in}}/R_{\text{out}})$  between  $-2$  and  $0$  with a step of 0.05. We then interpolated from the grid to obtain the spectrum of any AGN in our population. For practical reasons, the grid was computed at  $z = 0.0001$ , and the final spectra were obtained by redshifting the ones interpolated from the grid and rescaling the flux according to the luminosity distance.

To represent the diverse populations from which each observable is derived, we created subsamples of our synthetic population. For the reproduction of the CXB, we drew objects with  $L_{\text{X}}^{\text{int}} \geq 10^{41} \text{ erg s}^{-1}$  and  $z \leq 3$ , as non-blazar AGNs beyond this redshift are expected to contribute only  $\sim 1\%$  (Figs. 14 and 15 in Ueda et al. 2014). For the  $N_{\text{H}}$  distribution from R17, we sampled up to  $z = 0.3$ , since BAT is only sensitive to objects in the local Universe. Drawing objects at random from the XLF would result in large statistical fluctuations at low redshift, where the volume is small, and at high luminosity, where objects are rare. To avoid these stochastic effects and to remove the effect of cosmic variance in the simulations, we divided the redshift ranges considered for each observable into ten redshift bins. We then drew the same large number of objects,  $N_{\text{sample}}$ , in each redshift bin. As a different number of objects,  $N_{\text{XLF}}$ , is expected when integrating the  $\log(N_{\text{H}}/\text{cm}^{-2}) < 24$  XLF, each object is assigned a weight  $N_{\text{XLF}}/(N_{\text{sample}} - N_{\text{CT}})$ , where  $N_{\text{CT}}$  is the number of CT objects in our population within each redshift bin.

To correct our synthetic sample for the effect of absorption and compare our results to the observational constraints from R17 and H08, we used the sensitivity curves of the *Swift*/BAT 70-month catalogue (Baumgartner et al. 2013) and of the ten samples combined in H08, respectively. We expressed the sensitivity curves as the sky fraction as a function of 14–195 keV and 2–10 keV flux limit for each study, which we show in Fig. 4. Using REFLEX, we obtained the fluxes of each simulated source



**Fig. 4.** Sky fraction as a function of 14–195 keV and 2–10 keV flux limits representing the sensitivity curves of the *Swift*/BAT 70-month survey used in R17 (black line) and the multiple surveys used in H08 (red line), respectively.

in the respective energy bands and applied a weight determined by the probability of such an object being detected according to the sky fraction. This effectively applies the selection functions of these studies to our simulated population.

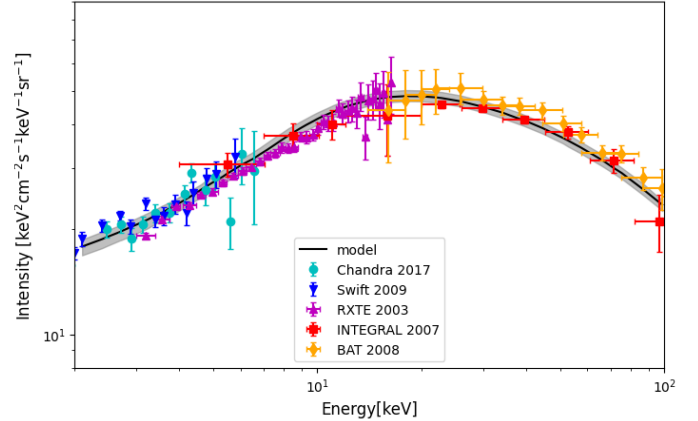
## 4. Results

Following the creation of a synthetic AGN population and the simulation of their X-ray spectra, we determined the parameter posterior distributions of our models based on three data sets: (a) the CXB; (b) the three observed absorption properties –namely the fraction of  $N_{\text{H}}$  in bins of  $\log(N_{\text{H}})$  and the fraction of absorbed AGNs as a function of observed 2–10 keV luminosity, and as a function of redshift; and (c) all four constraints combined.

### 4.1. Simulation-based inference

We determined model parameters using the simulation-based inference (SBI) approach, as implemented in the Python package *sbi* (Tejero-Cantero et al. 2020). SBI maps the relationship between simulated data and model parameters by constructing their joint density in the data and parameter space using a deep neural network. This approach is completely likelihood-free. We added noise to our practically noiseless simulated data to match the uncertainties in the observed data. For the CXB, we added Gaussian noise in each bin using the variance observed in the data. For the absorption properties, the noise was applied in each bin by drawing from a binomial distribution, with the success probability given by the model and the number of trials given by the number of ‘observed’ objects.

We used the sequential neural posterior estimation (SNPE) (Greenberg et al. 2019), which allows us to directly sample the posterior distribution of our parameters given the data. In the SNPE, the density is initially learned across the parameter priors (see Table 1) with a relatively small number of training samples. We then constructed the covariance matrix of the posteriors, which we used to provide a proposal prior in the form of a multivariate normal distribution for a second iteration of SNPE. This process accelerates the convergence of the deep neural-



**Fig. 5.** Median model (black line) for CXB in the 2–100 keV range using the simple torus applied only to the CXB data set. The grey region represents the 68% confidence interval of the model based on the posterior distribution of the parameters.

network training. Depending on the number of free parameters in each model (three or four), we trained the neural estimator using 2000 and 5000 simulations, respectively, to construct the proposal prior; then, we retrained with 5000 and 10 000 simulations, respectively, using the proposal prior. We find that these sizes for the training sets provide stable posterior estimates.

While model comparison can be performed in principle with any parameter inference algorithm that generates samples from the posterior (e.g. McEwen et al. 2021), we adopted a simple metric here to compare the performance of the different models: namely the  $\chi^2$  statistic of the median models drawn from the posteriors. Indeed, we do not expect that the simple models we explore are able to fully represent the data, so that a statistically rigorous model comparison is not warranted.

### 4.2. CXB

We first tested the ability of our models to reproduce the CXB data set. As shown in Fig. 5, the simple torus model is sufficient to match the CXB, proving that the CXB can be reproduced by even the simplest AGN unification models and, consequently, does not possess a strong constraining power. In Table 2, we summarise the median values of the parameter posteriors and the corresponding reduced  $\chi^2$  values. All posterior distributions are presented in Appendix A. In this AGN population, if observed edge-on, almost 92% of the AGNs would be CT, since the log-normal distribution of  $N_{\text{H,eq}}$  is centred around  $\log(N_{\text{H,eq}}/\text{cm}^{-2}) = 24.4$  with a scatter of only a factor of  $\sim 1.4$ .

The LD torus model can reproduce the CXB as well with model parameters close to those of the simple torus model and the posterior distribution of the slope in Eq. (3) showing a median value near 0.

### 4.3. Absorption properties

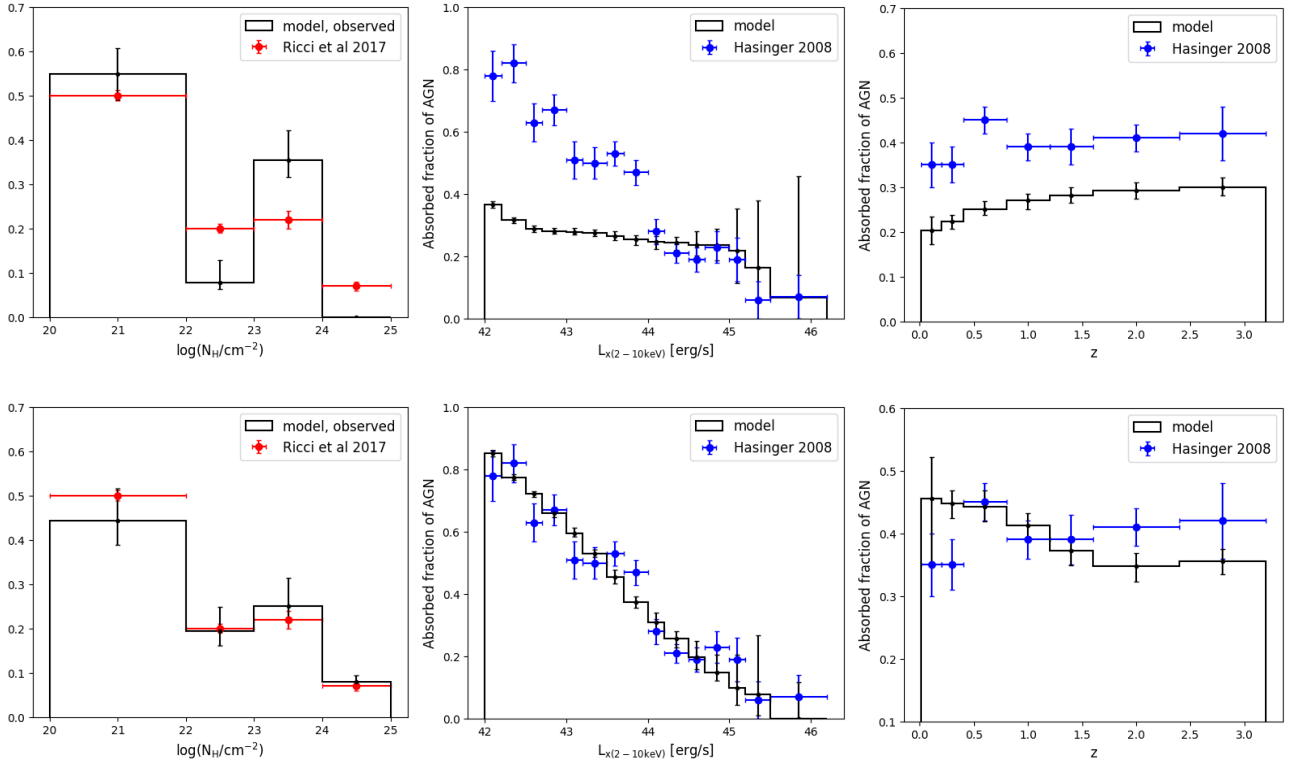
We next applied our models to the three absorption data sets. The limitations of the simple torus model become evident when applied to the absorption properties, even when the CXB constraint is ignored. In this case, the model fails to reproduce the observed anti-correlation between the absorbed fraction and luminosity (top row of Fig. 6).

In contrast, the LD torus model reproduces the three constraints very well (bottom row of Fig. 6). In this case, the posterior distribution of the slope in Eq. (3) yields a median value

**Table 2.** Median values of the parameter posterior distributions and reduced  $\chi^2$  values for each model applied to the three observational data sets.

Model	Constraints	$\mu$ (log/cm <sup>-2</sup> )	$\sigma$ (log/cm <sup>-2</sup> )	$R_0$ <sup>a</sup>	$\alpha$	$R_1$ <sup>a</sup>	$\chi^2$ /d.o.f.
Simple torus	CXB	24.43 <sup>+0.01</sup> <sub>-0.01</sub>	0.31 <sup>+0.01</sup> <sub>-0.02</sub>	–	–	0.24 <sup>+0.01</sup> <sub>-0.01</sub>	226/84
Simple torus	Absorption	23.27 <sup>+0.01</sup> <sub>-0.01</sub>	0.10 <sup>+0.01</sup> <sub>-0.03</sub>	–	–	0.98 <sup>+0.03</sup> <sub>-0.05</sub>	742/22
LD torus	Absorption	23.77 <sup>+0.08</sup> <sub>-0.07</sub>	1.06 <sup>+0.08</sup> <sub>-0.09</sub>	0.37 <sup>+0.02</sup> <sub>-0.01</sub>	0.41 <sup>+0.01</sup> <sub>-0.02</sub>	–	122/21
LD torus	Combined	24.24 <sup>+0.03</sup> <sub>-0.02</sub>	0.68 <sup>+0.02</sup> <sub>-0.01</sub>	0.21 <sup>+0.01</sup> <sub>-0.01</sub>	0.22 <sup>+0.01</sup> <sub>-0.01</sub>	–	1231/109
LD+AD	Combined	23.56 <sup>+0.05</sup> <sub>-0.04</sub>	0.98 <sup>+0.09</sup> <sub>-0.11</sub>	0.93 <sup>+0.04</sup> <sub>-0.03</sub>	0.24 <sup>+0.01</sup> <sub>-0.01</sub>	–	804/109
LD+AD+ $E_C$ <sup>b</sup>	Combined	23.56 <sup>+0.05</sup> <sub>-0.04</sub>	0.98 <sup>+0.09</sup> <sub>-0.11</sub>	0.93 <sup>+0.04</sup> <sub>-0.03</sub>	0.24 <sup>+0.01</sup> <sub>-0.01</sub>	–	540/109

**Notes.** <sup>a</sup>Arbitrary units for models without AD and measured in parsec for the AD model. <sup>b</sup>The same parameter set as before, but with a fixed high energy cut-off of  $E_C = 300$  keV (see Sect. 4.4 for more information).



**Fig. 6.** Median models (black lines) for the three observed absorption properties using the simple torus model (top row) and the LD torus (bottom row) applied only to the three absorption data sets. *Left:* Fraction of  $N_H$  in bins of  $\log(N_H)$ . *Middle:* Absorbed fraction of AGNs as a function of observed  $L_{X(2-10\text{keV})}$ . *Right:* Absorbed fraction of AGNs as a function of redshift. The error bars on the models represent their 68% confidence interval based on the posterior distribution of the parameters.

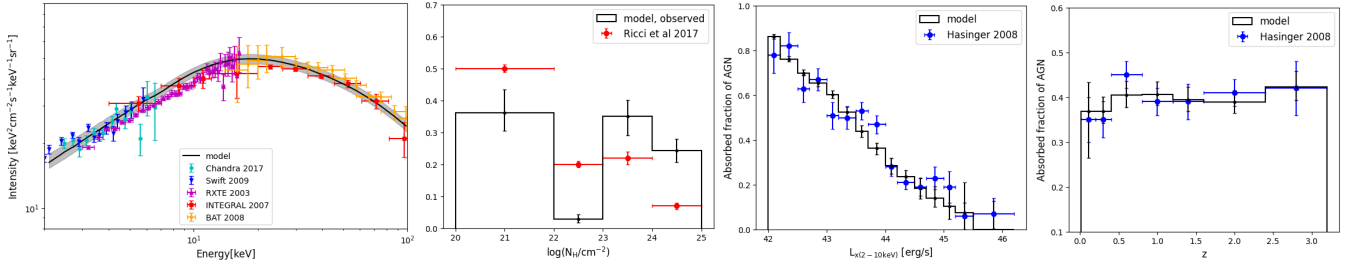
of  $\alpha = 0.41^{+0.01}_{-0.02}$ , which is close to, but still different from, the predictions of the receding torus model. Furthermore, matching the absorption properties requires the log-normal distribution of  $N_{H,\text{eq}}$  to be more than three times broader and centred 0.66 dex lower compared to the optimal match on the CXB. This, combined with the strong luminosity dependence of the covering factor, results in a distinctly different AGN population than the one that best reproduces the CXB. Specifically, for the CXB, approximately 55% of the synthetic population needs to be CT and only  $\sim 13\%$  Compton-thin, whereas for the absorption properties, the fractions shift to around 33% CT and 51% Compton-thin.

#### 4.4. Combined constraints

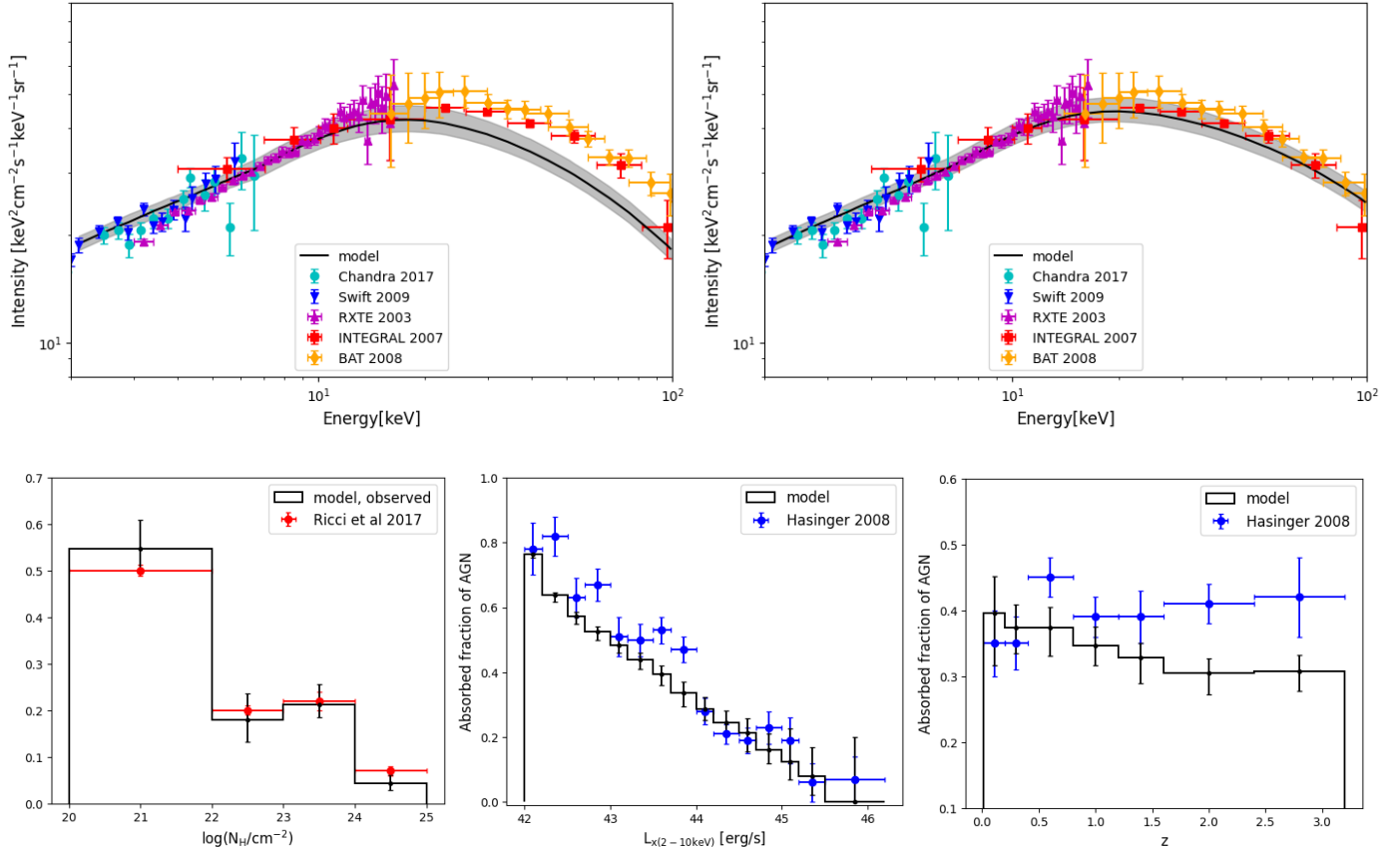
Finally, after demonstrating the distinct differences between the two AGN populations that best match the CXB and the absorption properties separately, we sought one population that satis-

fies all four constraints simultaneously. Given that the simple torus model clearly fails to reproduce the absorption properties, we did not apply it to the combined data sets, focusing instead on the LD torus model. For simplicity, we calculated the combined  $\chi^2$  without applying any particular weight to the data sets, despite the heteroscedasticity of the combined set. In Fig. 7, we show the CXB and the absorption properties reproduced by the LD torus. While most constraints are adequately matched, the model completely fails to explain the shape of the  $N_H$  fraction in bins of  $\log(N_H)$ , underestimating the number of less obscured objects and overestimating that of heavily obscured objects. The failure of the LD model to reproduce our combined constraints highlights the need for a more complex model to better capture the observed properties of the AGN population.

Building on the limitations of the LD torus model, we now turn to the LD+AD model to assess whether it can provide a better match to the combined observational constraints. While



**Fig. 7.** Median models (black lines) using the LD torus model applied to the combined constraints. *From the left:* CXB, the observed fraction of  $N_{\text{H}}$  in bins of  $\log(N_{\text{H}})$ , the absorbed fraction of AGNs as a function of observed  $L_{\text{X}(2-10\text{keV})}$ , and the absorbed fraction of AGNs as a function of redshift. The grey region and error bars represent the 68% confidence interval of the model based on the posterior distribution of the parameters.



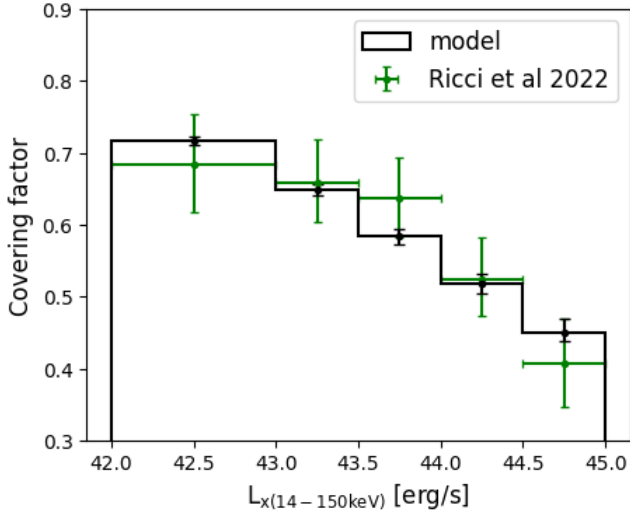
**Fig. 8.** Same as Fig. 7, but using the LD+AD model. *Top left:* CXB with fixed high-energy cut-off  $E_{\text{C}} = 200$  keV; *top right:* that with  $E_{\text{C}} = 300$  keV (see Sect. 4.4 for more information).

this model provides the best simultaneous match to all four constraints (shown in Fig. 8), it significantly underpredicts the CXB at high energies and predicts an  $\sim 25\%$  lower absorbed AGN fraction at high redshifts. The posterior distribution of the slope in Eq. (3) yields a median value of  $\alpha = 0.24^{+0.01}_{-0.01}$ , substantially deviating from the predictions of the receding torus model, but in close agreement with X-ray observations (e.g. Burlon et al. 2011; Sazonov et al. 2015). The log-normal distribution of  $N_{\text{H,eq}}$  of this population aligns very closely with that of the best match on the absorption properties –rather than the CXB– further demonstrating that the CXB does not carry enough constraining power. Even then, the posterior distributions of the normalisation,  $R_0$ , and the slope,  $\alpha$ , in Eq. (3) have medians that are more than twice as high and nearly twice as flat, respectively, compared to the best match to the absorption properties. The combined constraints result in the lowest fraction of CT AGNs ( $\sim 21\%$ ) and a Compton-thin fraction of  $\sim 48\%$ , leaving an unobscured fraction of  $\sim 31\%$ .

To address the underpredicted CXB at high energies, we increased the fixed high-energy cut-off of the X-ray source from  $E_{\text{C}} = 200$  keV to  $E_{\text{C}} = 300$  keV, as low-accretion AGNs are more common and tend to have higher  $E_{\text{C}}$  (R17; Ricci et al. 2018; Baloković et al. 2020). Since variations in  $E_{\text{C}}$  affect the high-energy region of the AGN spectra while having minimal impact on the lower energy range, we find that redetermining model parameters does not result in significant changes to the parameter posteriors. Using the same parameter posteriors, we find that the match to the CXB improves considerably (top right in Fig. 8), while the absorption properties remain consistent.

## 5. Discussion

Our results demonstrate that the CXB can be reproduced not only by more complex models, but even by the simple unification model of AGNs. However, our simple torus model lacks



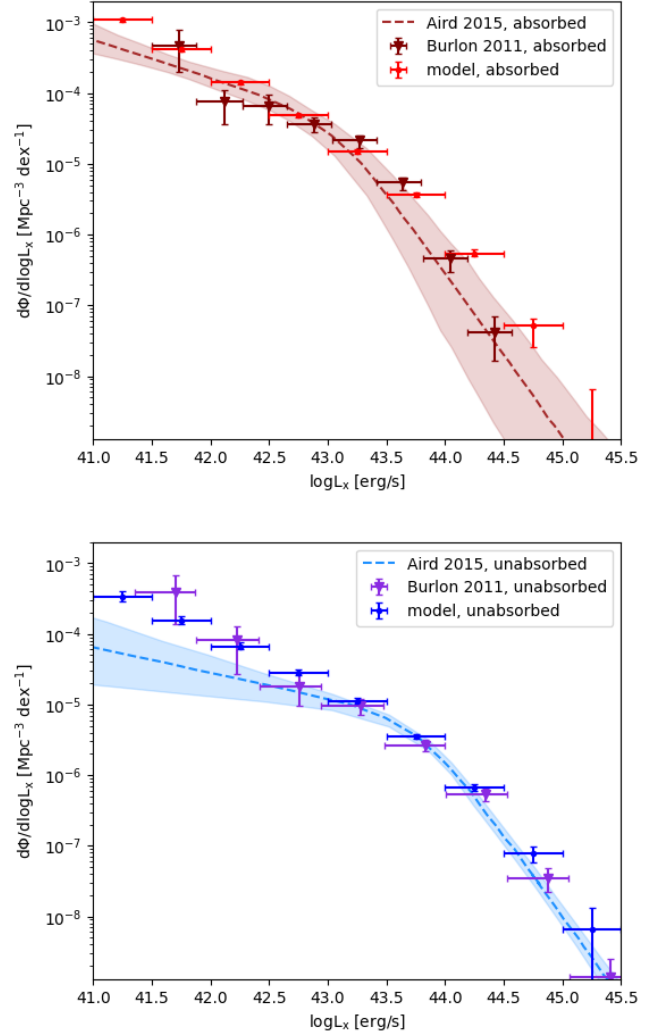
**Fig. 9.** Average torus covering factor of our local synthetic population as a function of intrinsic 14–150 keV luminosity using the LD+AD model (black). The error bars represent the 68% confidence interval of the model based on the posterior distribution of the parameters. For comparison, the covering factor of *Swift*/BAT-detected AGNs from Ricci et al. (2022; green).

some necessary ingredients to accurately represent the absorption properties. First, the covering factor of the absorber must have a clear luminosity dependence, which cannot be constrained by the CXB alone. Second, the inclusion of the AD component is essential to provide enough reflection given the constraint of the  $N_{\text{H}}$  distribution. Hence, the LD+AD model is the only realistic model capable of simultaneously matching both the CXB and the absorption data sets. Nevertheless, all models discussed in this work are simplistic and cannot be expected to explain the full phenomenology of AGNs in detail.

### 5.1. Luminosity dependence of absorbed fraction

As shown in Sect. 4.4, when applying the LD+AD model to all observational constraints simultaneously, we were able to reproduce the dependence of the absorbed fraction on the X-ray luminosity. From parameter determination, we can derive that the inner edge of typical dusty tori,  $r_{\text{dust}} = R_{\text{out}} - R_{\text{in}}$ , spans a range of 0.2–2.8 pc in the  $L_{\text{X}}^{\text{int}}$  range  $10^{41} - 10^{46} \text{ erg s}^{-1}$ , which is consistent with dust sublimation radius sizes found in infrared interferometry studies (e.g. Tristram et al. 2014; GRAVITY Collaboration 2020). However, we find a shallow median slope in Eq. (3),  $\alpha = 0.24_{-0.01}^{+0.01}$ , which is compatible with observational studies in the X-rays (e.g. Burlon et al. 2011; Sazonov et al. 2015), but in disagreement with the receding torus model. As a further investigation, Fig. 9 shows the average covering factor of the detected tori in our local AGN population as a function of intrinsic 14–150 keV luminosity to compare with the *Swift*/BAT results from Ricci et al. (2022). We find that the covering factor decreases by 27% across the  $L_{\text{X}(14-150\text{keV})}$  range of  $10^{42} - 10^{45} \text{ erg s}^{-1}$ , in very good agreement with Ricci et al. (2022) despite our model being constrained by the absorbed fraction from H08 that comprises a population up to  $z = 3.2$  detected in the 2–10 keV.

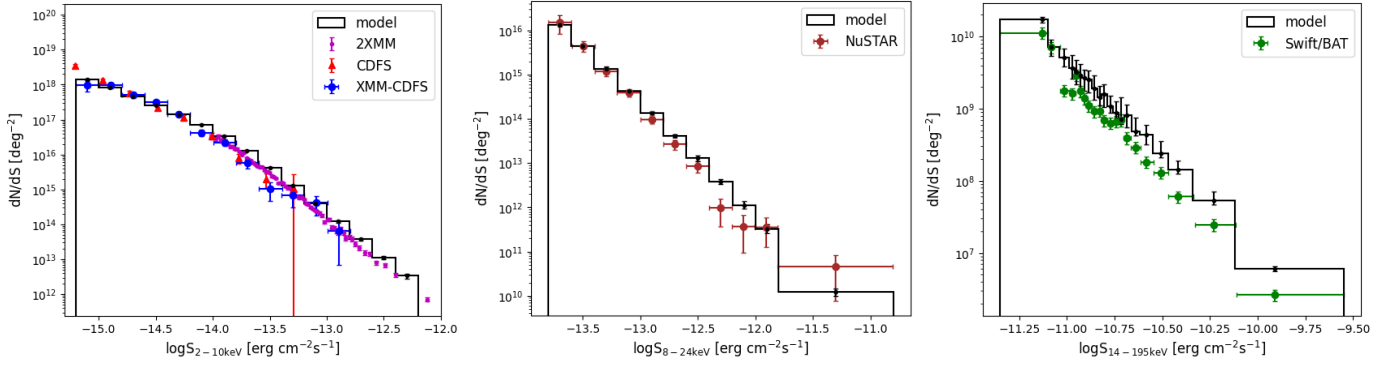
Other mechanisms have been proposed to explain the luminosity dependence of the absorbed fraction with a slope flatter than that predicted by the receding torus model. Additional X-ray obscuration could arise from the broad-line region (BLR), which is located within the dust sublimation radius. The pres-



**Fig. 10.** X-ray luminosity function of absorbed (red points) and unabsorbed (blue points) AGNs in our population up to  $z = 0.1$ , using the LD+AD model with our combined constraints. The error bars represent the 68% confidence interval of the model based on the posterior distribution of the parameters. *Top:* Comparison of absorbed XLF with Fig. 13 of Aird et al. (2015), where they extrapolated their model at  $z=0$  (brown dashed line) and present the local observations of Burlon et al. (2011) converted to 2–10 keV luminosities (brown triangles). *Bottom:* Same comparison of unabsorbed XLF with Aird et al. (2015) (blue dashed line) and Burlon et al. (2011) (purple triangles).

ence of absorption in the BLR could increase the fraction of absorbed AGNs at high luminosity, thus flattening the relationship between the absorbed fraction and luminosity. Another important mechanism is the clumpy dusty-torus model (e.g. Nenkova et al. 2008), from which Hönig & Beckert (2007) predicted that the absorbed fraction decreases as  $\propto L^{-0.25}$  due to the interplay between radiation pressure and the clump distribution. It must be pointed out that the recent studies of Ricci et al. (2017b, 2022) have revealed that when the AGNs are binned by Eddington ratio, the luminosity dependence of the absorbed fraction completely flattens. Instead, they propose a radiation-regulated unification model, where the absorbed fraction depends on the Eddington ratio rather than luminosity.

As a direct consequence of the relation between absorption and luminosity, our LD+AD model predicts that absorbed and unabsorbed AGNs have different XLFs. In Fig. 10, we present the two XLFs of our population up to  $z = 0.1$ .



**Fig. 11.** Differential number counts as a function of flux produced by the LD+AD model (black) in the *left*: 2–10 keV energy range; *centre*: 8–24 keV range; and *right*: 14–195 keV range. The error bars represent the 68% confidence interval of the model based on the posterior of the parameters. For comparison, we present the differential number counts of the 70-month *Swift*/BAT (Ananna et al. 2022, green), combined NuSTAR (Harrison et al. 2016, brown), 4Ms CDFS (Lehmer et al. 2012, red), XMM-CDFS (Ranalli et al. 2013, blue), and 2XMM (Mateos et al. 2008, magenta) surveys.

Our model predicts well the increase of absorbed AGNs below  $\log(L_X) = 43.5$ , as found by the model of Aird et al. (2015) extrapolated to  $z = 0$  (taken from their Fig. 13) and the local observations in Burlon et al. (2011; converted to 2–10 keV luminosities). Our model confirms that absorbed AGNs are intrinsically less luminous. Both our absorbed and unabsorbed XLFs agree very well with the results of Burlon et al. (2011), but our unabsorbed XLF deviates from the results of Aird et al. (2015) at low luminosities. That may be because the latter is mostly constrained by high-redshift sources.

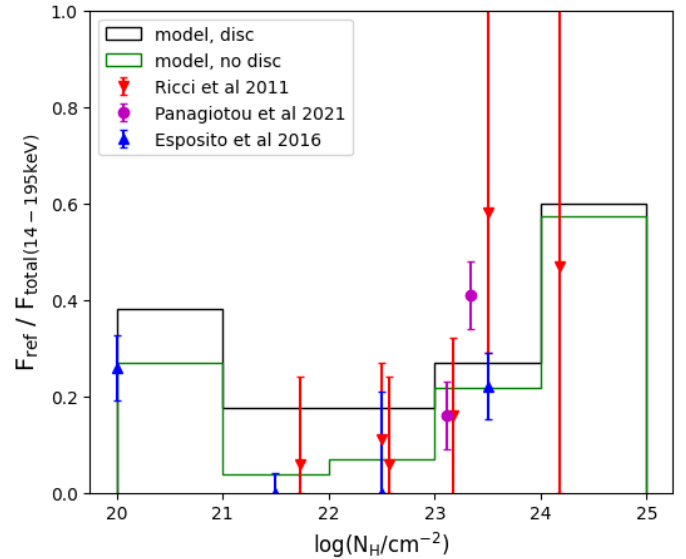
### 5.2. AGN number counts

In Fig. 11, we present AGN differential number counts as a function of flux in the 2–10 keV, 8–24 keV, and 14–195 keV energy bands as reproduced by the LD+AD model. The model predicts the 2–10 keV number counts from the 4Ms CDFS (Lehmer et al. 2012), XMM-CDFS (Ranalli et al. 2013), and 2XMM (Mateos et al. 2008) surveys well, particularly at fainter fluxes. Similarly, in the 8–24 keV band, the model can reproduce the number counts from the combined surveys of NuSTAR (Harrison et al. 2016).

By contrast, the model consistently overpredicts the 14–195 keV *Swift*/BAT differential number counts (Ananna et al. 2022) within  $1\sigma$ – $2\sigma$ . We note that this is in agreement with several other AGN population-synthesis models based on the CXB (e.g. Gilli et al. 2007; Ueda et al. 2014; Aird et al. 2015), which also produce higher number counts than *Swift*/BAT, as shown in Ananna et al. (2019) and Harrison et al. (2016). This tension is still under investigation (e.g. Akylas & Georgantopoulos 2019; Delaney et al. 2023).

### 5.3. Reflection versus obscuration

In Fig. 12, we compare the predicted reflection of our models to the observed trend between reflection and line-of-sight  $N_H$  reported in Panagiotou et al. (2021), Ricci et al. (2011), and Esposito & Walter (2016). We computed the fraction of reflected-to-total emission ( $F_{\text{ref}} / F_{\text{total}}$ ) in the 14–195 keV band using REFLEX for our LD and LD+AD models. Although these previous studies suggest an increase of reflection with obscuration, large uncertainties and unknown selection effects make this correlation debatable and not statistically conclusive. We therefore used the data for qualitative comparison only, by replacing



**Fig. 12.** Observed correlation between the fraction of reflected emission over total emission,  $F_{\text{ref}} / F_{\text{total}}$ , in the 14–195 keV energy band and the line-of-sight  $N_H$ ; these were produced with our combined constraints’ median parameters for the LD+AD model (black line) and the LD torus model (green line). The data are from Panagiotou et al. (2021), Esposito & Walter (2016), and Ricci et al. (2011) (magenta, blue, and red points, respectively) after replacing the reflection parameter of the *pexrav* model in XSPEC with the fraction  $F_{\text{ref}} / F_{\text{total}}$  in the same energy band.

the reflection parameter of the *pexrav* model in XSPEC (Arnaud 1996) with  $F_{\text{ref}} / F_{\text{total}}$  in the same energy band.

Both models broadly follow the trend of increasing reflection with obscuration, while they also predict the increase of reflection for unobscured sources [ $\log(N_H / \text{cm}^{-2}) < 21$ ], which was suggested in Esposito & Walter (2016) and Panagiotou & Walter (2019). The key difference between our models arises at  $\log(N_H / \text{cm}^{-2}) < 23$ , where the LD+AD model consistently provides optically thick material necessary to produce reflection. In comparison, the LD model lacks such material in this  $N_H$  range, resulting in nearly half the reflection predicted by the LD+AD model.

While Ricci et al. (2011) and Panagiotou et al. (2021) attributed enhanced reflection in heavily obscured sources to a clumpy dusty torus, we demonstrate that such a trend can

also emerge from our homogeneous torus models. Nonetheless, a clumpy torus could alter this relationship by introducing variability in the line of sight, as it allows the encountering of different  $N_{\text{H}}$  for the same orientation at different azimuthal angles. This may contribute to the observed scatter and could be explored in future work.

#### 5.4. Intrinsic fraction of Compton-thick AGNs

We derived the intrinsic CT fraction ( $\log(N_{\text{H}}/\text{cm}^{-2}) \geq 24$ ) with the LD and LD+AD models and find that it drops from  $50 \pm 2\%$  (68% confidence based on the posteriors) without the AD to  $21 \pm 7\%$  with the AD. As shown in Fig. 12, when adding the AD, the unobscured and Compton-thin AGNs produce more reflection without additional absorption, reducing the amount of CT AGNs needed to reproduce the CXB peak. As the LD+AD model is more realistic and best reproduces all constraints, the CT fraction of  $21 \pm 7\%$  is likely closer to the true value. This aligns very well with the findings of local X-ray surveys (e.g. Burlon et al. 2011; Ricci et al. 2015; Boorman et al. 2025) and with the predictions of previous AGN population syntheses at low redshift (e.g. Ueda et al. 2014; Aird et al. 2015). However, at  $z = 1$ , Ueda et al. (2014) and Ananna et al. (2019) predicted an intrinsic CT fraction of  $37 \pm 2\%$  and  $56 \pm 7\%$ , respectively. The discrepancy with our value likely arises at least in part because our model does not take into account redshift evolution, causing the CT fraction of the parent population to remain constant. This limitation may also explain why our model struggles to accurately match the absorbed AGN fraction at higher redshifts (however, see Sect. 5.5 for an alternative explanation). To incorporate evolution in our model, an additional parameter would be needed to evolve the size or density of the dusty torus, which is beyond the scope of this work.

In Fig. 13, we further compare the contribution of unobscured [ $\log(N_{\text{H}}/\text{cm}^{-2}) < 22$ ], Compton-thin [ $22 \leq \log(N_{\text{H}}/\text{cm}^{-2}) < 24$ ], and CT sources [ $\log(N_{\text{H}}/\text{cm}^{-2}) \geq 24$ ] to the CXB with the LD and LD+AD models. The total contribution of CT objects to the CXB in the 14–195 keV range is found to be  $\sim 28\%$  for the LD model, while the CT contribution drops to about 7% with the LD+AD model. The latter value is in close agreement with that found by Treister et al. (2009).

Finally, in Fig. 14, we present the intrinsic line-of-sight  $N_{\text{H}}$  distribution of our local population of AGNs using the LD+AD model. The distribution reveals the same bimodality seen in the observed one, with peaks in the  $\log(N_{\text{H}}/\text{cm}^{-2}) = 20$ –21 and 23–24 bins. Our model recovers the results from Ricci et al. (2015) well, although they assumed a fixed covering factor when correcting the observed  $N_{\text{H}}$  distribution into the intrinsic one. For reference, we also included the observed distribution of our model from Fig. 8, with the first two bins separated. The selection bias is most significant for completely unobscured and CT AGNs, as expected.

#### 5.5. Origin of mildly obscured AGNs

As discussed in Sect. 2.2, our models do not contain the ingredients to generate a population with a significant fraction of mildly obscured AGNs [ $\log(N_{\text{H}}/\text{cm}^{-2}) < 22$ ]. One possible origin of obscuration in these AGNs is the gas and dust in the plane of the host galaxy around our dusty torus, which would cause completely unobscured objects to be subject to additional absorption of  $\log(N_{\text{H}}/\text{cm}^{-2}) < 22$ . Gilli et al. (2022) constructed a clumpy interstellar medium (ISM) whose characteristic cloud surface density evolves with redshift and used it in combination with the

obscuring torus to reproduce the fraction of absorbed AGNs as a function of redshift from deep X-ray studies (e.g. Buchner et al. 2015; Iwasawa et al. 2020). They find that the median  $N_{\text{H}}$  of the ISM evolves as  $\propto (1+z)^{3.3}$ , assuming a median  $N_{\text{H}} = 10^{21} \text{ cm}^{-2}$  for local galaxies (Willingale et al. 2013).

Hence, if we incorporate this component into our model, the ISM  $N_{\text{H}}$  could increase by up to  $\sim 2.4$  times within our local ( $z < 0.3$ ) synthetic population, which is used to reproduce the observed fraction of  $N_{\text{H}}$  in bins of  $\log(N_{\text{H}})$ . This increase would be sufficient to shift a significant fraction of completely unobscured objects into the  $\log(N_{\text{H}}/\text{cm}^{-2}) = 21$ –22 bin. Furthermore, in the redshift range  $z = 1.2$ –3.2, the ISM  $N_{\text{H}}$  would increase from  $\sim 13$  times to more than 100 times. This could result in the increase of the absorbed fraction at higher redshift, which our current model underpredicts. Regarding our model parameters, incorporating the ISM would likely cause the posterior of the mean,  $\mu$ , of the  $N_{\text{H,eq}}$  distribution to shift to lower values. Our results on the CXB would likely not be significantly affected, since the biggest difference would be high redshift sources that are obscured (see right of Fig. 13).

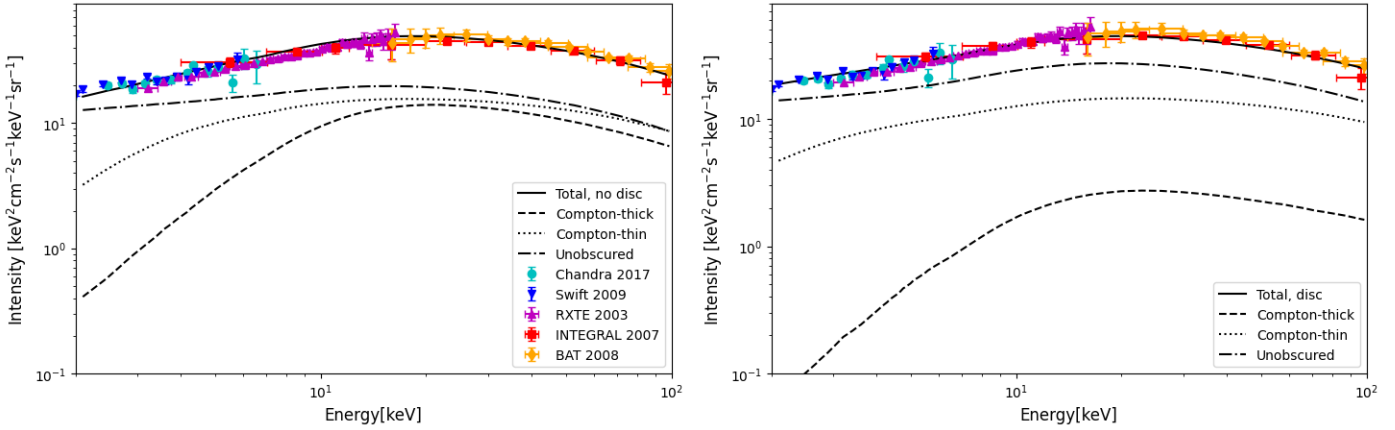
## 6. Conclusions

Until now, population-synthesis AGN models have not treated absorption and reflection self-consistently, but rather assumed values for the reflection at different obscurations and for the intrinsic fraction of CT AGNs. We created a fully self-consistent population-synthesis AGN model with numerical simulations of their X-ray emission using the ray-tracing code REFLEX. Our constraints include the CXB as well as absorption properties of AGNs detected in different X-ray surveys, namely the fraction of  $N_{\text{H}}$  in bins of  $\log(N_{\text{H}})$  and the fraction of absorbed AGNs as a function of observed  $L_{\text{X}}$  and as a function of redshift. We sampled an intrinsic XLF and parameterised the geometry and density of the circumnuclear material of the AGN population. We constructed geometrical models based on the simple unification of AGNs and the receding torus; lastly, we included an AD. We determined the parameter posterior distributions for each model using an SBI method.

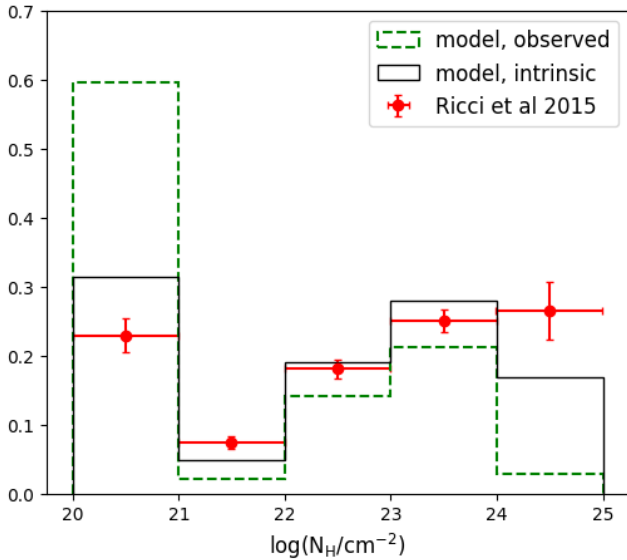
We find that the CXB can be well reproduced, even by the simple torus model (Fig. 5). However, matching the absorption properties requires the LD torus model (Fig. 6). The LD+AD model best satisfies all observational constraints simultaneously (Fig. 8). In this case, the posterior distribution of the slope in Eq. (3) has a median of  $\alpha = 0.24_{-0.01}^{+0.01}$ , in agreement with X-ray observations.

We derive an intrinsic CT fraction of  $21 \pm 7\%$  (68% confidence based on the posteriors), in agreement with observations of the local Universe, and find that the inner edge of the dusty torus spans a range of 0.2–2.8 pc in the  $L_{\text{X}}^{\text{int}}$  range  $10^{41}$ – $10^{46} \text{ erg s}^{-1}$ ; this is consistent with dust sublimation radius sizes from infrared interferometry studies. Our synthetic population can reproduce the luminosity dependence of the torus covering factor and XMM-Newton, Chandra, and NuSTAR number counts, as well as the observed correlation between reflection and obscuration (Figs. 9, 11, and 12). Additionally, it can recover the XLFs of absorbed and unabsorbed AGNs and the intrinsic  $N_{\text{H}}$  distribution of local AGNs. However, our AGN modelling lacks the obscuration originating at the host galaxy as well as evolution, resulting in a lower fraction of absorbed AGNs at high redshifts and a constant CT fraction.

This is the first work in which the CXB is reproduced with a fully self-consistent synthetic AGN-population model, without assuming the reflection features of the spectra, the intrinsic  $N_{\text{H}}$



**Fig. 13.** *Left:* Contribution of unobscured ( $\log(N_{\text{H}}/\text{cm}^{-2}) < 22$ ; dashed-dotted line), Compton-thin ( $22 \leq \log(N_{\text{H}}/\text{cm}^{-2}) < 24$ ; dotted line), and CT sources ( $\log(N_{\text{H}}/\text{cm}^{-2}) \geq 24$ ; dashed line) to the CXB (solid line) reproduced by the median parameters of the LD torus model. *Right:* As before, but for the LD+AD model.



**Fig. 14.** Intrinsic  $N_{\text{H}}$  distribution (black line) of our local synthetic population produced using the LD+AD model with the combined constraints. Our observed distribution (green line) from Fig. 8 and the intrinsic distribution from Ricci et al. (2015) (red points) are shown for comparison.

distribution, or the CT fraction, and without modifying the XLF. Modelling and linking the emission, absorption, and reflection through the matter distribution allowed us to break the degeneracy of fitting the CXB by simultaneously matching observational constraints concerning absorption, which has not been done before. Overall, our results emphasise that a better understanding of AGNs requires models that account for both the geometric structure and the physical mechanisms of the absorbers in order to explain their observed properties. In the future, this methodology can be extended to include more complex models that can account for effects such as the clumpiness of the obscuring material, evolution, or additional relationships between AGN parameters, introducing in particular the intrinsic properties of SMBHs, such as their mass and Eddington ratio.

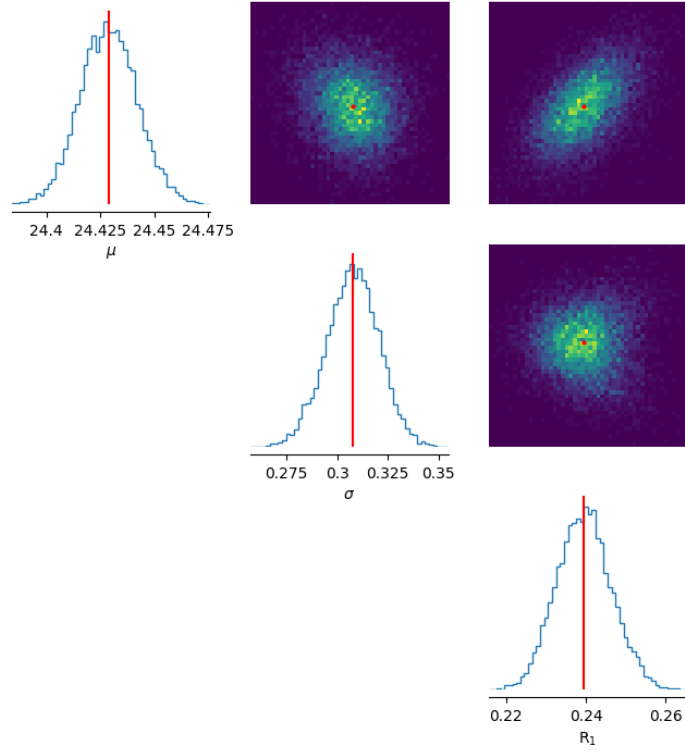
*Acknowledgements.* CR acknowledges support from Fondecyt Regular grant 1230345, ANID BASAL project FB210003 and the China-Chile joint research fund.

## References

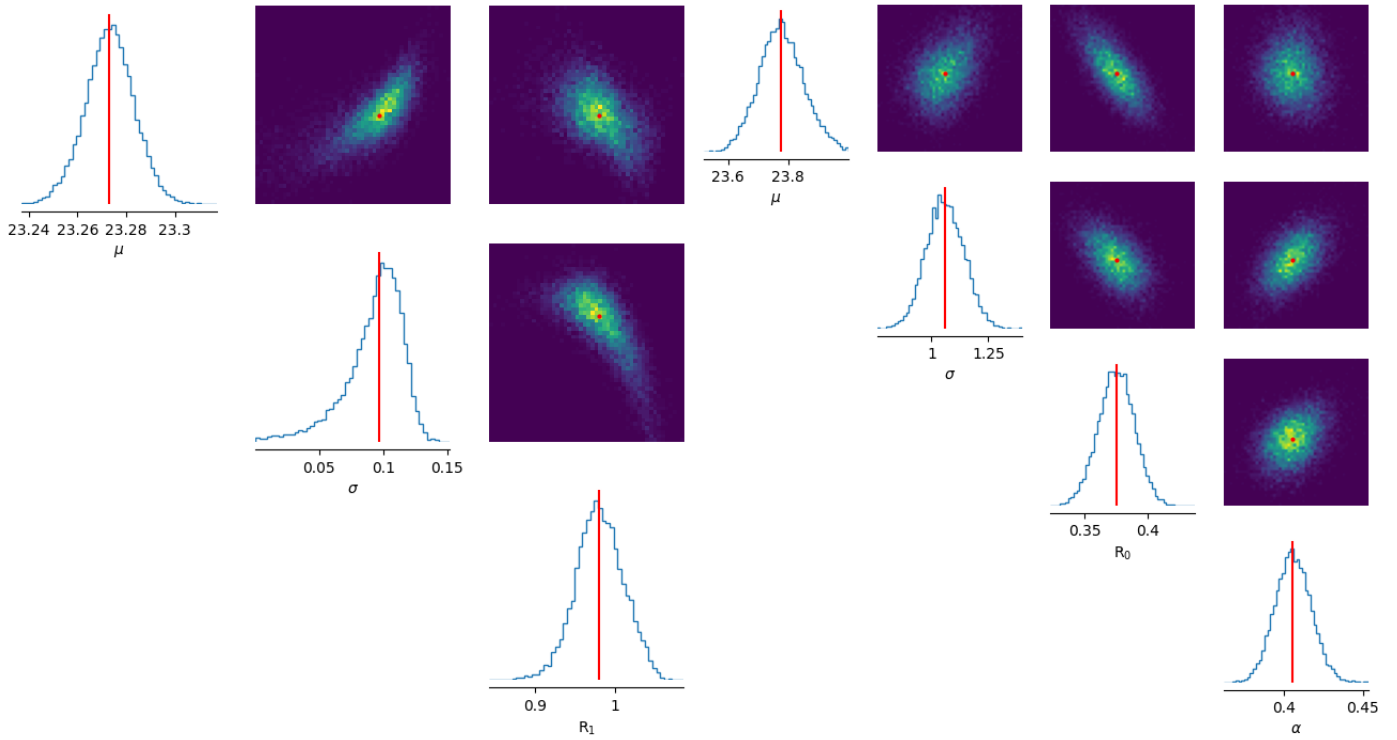
- Aird, J., Coil, A. L., Georgakakis, A., et al. 2015, *MNRAS*, 451, 1892  
 Ajello, M., Greiner, J., Sato, G., et al. 2008, *ApJ*, 689, 666  
 Akylas, A., & Georgantopoulos, I. 2019, *A&A*, 625, A131  
 Akylas, A., Georgakakis, A., Georgantopoulos, I., Brightman, M., & Nandra, K. 2012, *A&A*, 546, A98  
 Ananna, T. T., Treister, E., Urry, C. M., et al. 2019, *ApJ*, 871, 240  
 Ananna, T. T., Weigel, A. K., Trakhtenbrot, B., et al. 2022, *ApJS*, 261, 9  
 Antonucci, R. 1993, *ARA&A*, 31, 473  
 Arnaud, K. A. 1996, *ASP Conf. Ser.*, 101, 17  
 Baloković, M., Harrison, F. A., Madejski, G., et al. 2020, *ApJ*, 905, 41  
 Baumgartner, W. H., Tueller, J., Markwardt, C. B., et al. 2013, *ApJS*, 207, 19  
 Boorman, P. G., Gandhi, P., Buchner, J., et al. 2025, *ApJ*, 978, 118  
 Brandt, W. N., & Alexander, D. M. 2015, *A&ARv*, 23, 1  
 Brandt, W. N., & Yang, G. 2022, in *Handbook of X-ray and Gamma-ray Astrophysics*, eds C. Bambi, & A. Sanganello, 78  
 Brightman, M., & Nandra, K. 2011, *MNRAS*, 413, 1206  
 Buchner, J., Georgakakis, A., Nandra, K., et al. 2015, *ApJ*, 802, 89  
 Burlon, D., Ajello, M., Greiner, J., et al. 2011, *ApJ*, 728, 58  
 Cappelluti, N., Ranalli, P., Roncarelli, M., et al. 2012, *MNRAS*, 427, 651  
 Cappelluti, N., Li, Y., Ricarte, A., et al. 2017, *ApJ*, 837, 19  
 Churazov, E., Sunyaev, R., Revnivtsev, M., et al. 2007, *A&A*, 467, 529  
 Delaney, J. N., Aird, J., Evans, P. A., et al. 2023, *MNRAS*, 521, 1620  
 Draine, B. T. 2003, *ApJ*, 598, 1026  
 Draper, A. R., & Ballantyne, D. R. 2009, *ApJ*, 707, 778  
 Esposito, V., & Walter, R. 2016, *A&A*, 590, A49  
 Gebhardt, K., Bender, R., Bower, G., et al. 2000, *ApJ*, 539, L13  
 Ghisellini, G., Haardt, F., & Matt, G. 1994, *MNRAS*, 267, 743  
 Gilli, R., Comastri, A., & Hasinger, G. 2007, *A&A*, 463, 79  
 Gilli, R., Norman, C., Calura, F., et al. 2022, *A&A*, 666, A17  
 GRAVITY Collaboration (Dexter, J., et al.) 2020, *A&A*, 635, A92  
 Greenberg, D. S., Nonnenmacher, M., & Macke, J. H. 2019, ArXiv e-prints [arXiv:1905.07488]  
 Gruber, D. E., Matteson, J. L., Peterson, L. E., & Jung, G. V. 1999, *ApJ*, 520, 124  
 Güver, T., & Özel, F. 2009, *MNRAS*, 400, 2050  
 Haardt, F., & Maraschi, L. 1991, *ApJ*, 380, L51  
 Harrison, F. A., Aird, J., Civano, F., et al. 2016, *ApJ*, 831, 185  
 Hasinger, G. 2008, *A&A*, 490, 905  
 Hickox, R. C., & Alexander, D. M. 2018, *ARA&A*, 56, 625  
 Hönig, S. F., & Beckert, T. 2007, *MNRAS*, 380, 1172  
 Iwasawa, K., Comastri, A., Vignali, C., et al. 2020, *A&A*, 639, A51  
 Kormendy, J., & Ho, L. C. 2013, *ARA&A*, 51, 511  
 Koss, M., Trakhtenbrot, B., Ricci, C., et al. 2017, *ApJ*, 850, 74  
 Lawrence, A. 1991, *MNRAS*, 252, 586  
 Lehmer, B. D., Xue, Y. Q., Brandt, W. N., et al. 2012, *ApJ*, 752, 46  
 Liu, T., Tozzi, P., Wang, J.-X., et al. 2017, *ApJS*, 232, 8  
 Loders, K. 2003, *ApJ*, 591, 1220  
 Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, *AJ*, 115, 2285  
 Martocchia, A., & Matt, G. 1996, *MNRAS*, 282, L53  
 Mateos, S., Warwick, R. S., Carrera, F. J., et al. 2008, *A&A*, 492, 51  
 McEwen, J. D., Wallis, C. G. R., Price, M. A., & Spurio Mancini, A. 2021, ArXiv e-prints [arXiv:2111.12720]  
 McKaig, J., Ricci, C., Paltani, S., & Satyapal, S. 2022, *MNRAS*, 512, 2961

- Miniutti, G., & Fabian, A. C. 2004, *MNRAS*, **349**, 1435
- Moretti, A., Pagani, C., Cusumano, G., et al. 2009, *A&A*, **493**, 501
- Nenkova, M., Sirocky, M. M., Nikutta, R., Ivezić, Ž., & Elitzur, M. 2008, *ApJ*, **685**, 160
- Paltani, S., & Ricci, C. 2017, *A&A*, **607**, A31
- Paltani, S., Walter, R., McHardy, I. M., et al. 2008, *A&A*, **485**, 707
- Panagiotou, C., & Walter, R. 2019, *A&A*, **626**, A40
- Panagiotou, C., Walter, R., & Paltani, S. 2021, *A&A*, **653**, A162
- Pounds, K. A., Nandra, K., Stewart, G. C., George, I. M., & Fabian, A. C. 1990, *Nature*, **344**, 132
- Ramos Almeida, C., & Ricci, C. 2017, *Nat. Astron.*, **1**, 679
- Ranalli, P., Comastri, A., Vignali, C., et al. 2013, *A&A*, **555**, A42
- Revnivtsev, M., Gilfanov, M., Sunyaev, R., Jahoda, K., & Markwardt, C. 2003, *A&A*, **411**, 329
- Ricci, C., & Paltani, S. 2023, *ApJ*, **945**, 55
- Ricci, C., Walter, R., Courvoisier, T. J. L., & Paltani, S. 2011, *A&A*, **532**, A102
- Ricci, C., Ueda, Y., Koss, M. J., et al. 2015, *ApJ*, **815**, L13
- Ricci, C., Trakhtenbrot, B., Koss, M. J., et al. 2017a, *ApJS*, **233**, 17
- Ricci, C., Trakhtenbrot, B., Koss, M. J., et al. 2017b, *Nature*, **549**, 488
- Ricci, C., Ho, L. C., Fabian, A. C., et al. 2018, *MNRAS*, **480**, 1819
- Ricci, C., Ananna, T. T., Temple, M. J., et al. 2022, *ApJ*, **938**, 67
- Sazonov, S., Churazov, E., & Krivonos, R. 2015, *MNRAS*, **454**, 1202
- Tagliacozzo, D., Marinucci, A., Ursini, F., et al. 2023, *MNRAS*, **525**, 4735
- Tejero-Cantero, A., Boelts, J., Deistler, M., et al. 2020, *J. Open Source Soft.*, **5**, 2505
- Treister, E., Urry, C. M., & Virani, S. 2009, *ApJ*, **696**, 110
- Tristram, K. R. W., Burtscher, L., Jaffe, W., et al. 2014, *A&A*, **563**, A82
- Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, *ApJ*, **598**, 886
- Ueda, Y., Akiyama, M., Hasinger, G., Miyaji, T., & Watson, M. G. 2014, *ApJ*, **786**, 104
- Wada, K., Papadopoulos, P. P., & Spaans, M. 2009, *ApJ*, **702**, 63
- Wilkins, D. R., & Fabian, A. C. 2011, *MNRAS*, **414**, 1269
- Willingale, R., Starling, R. L. C., Beardmore, A. P., Tanvir, N. R., & O'Brien, P. T. 2013, *MNRAS*, **431**, 394
- Zappacosta, L., Comastri, A., Civano, F., et al. 2018, *ApJ*, **854**, 33

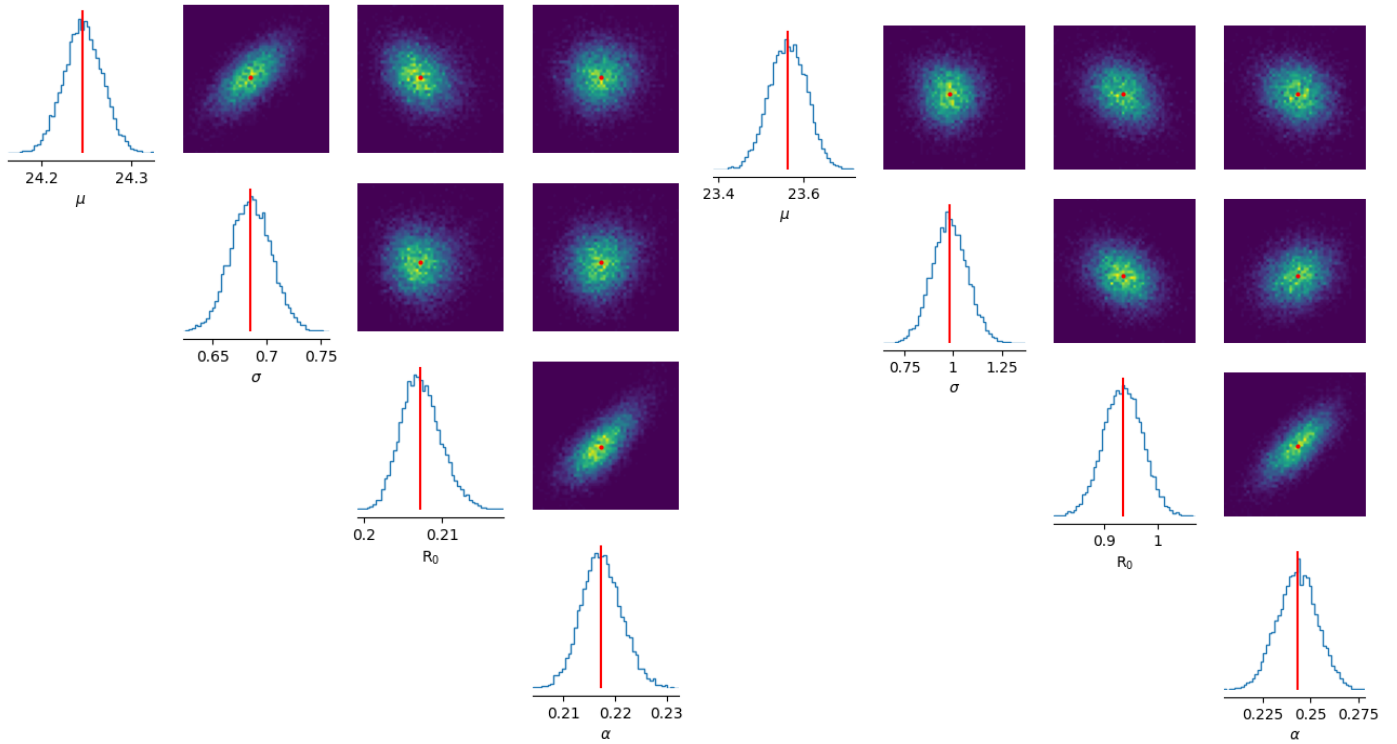
Appendix A: SBI posteriors



**Fig. A.1.** Parameter posteriors using the simple torus model when applied on the CXB data sets. The median of each distribution reported in Table 2 is denoted in red.



**Fig. A.2.** Parameter posteriors using *left*: the simple torus model and *right*: the LD torus model, when applied on the absorption properties. The median of each distribution reported in Table 2 is denoted in red.



**Fig. A.3.** Parameter posteriors using *left*: the LD torus model and *right*: the LD+AD model, when applied on all constraints. The median of each distribution reported in Table 2 is denoted in red.