Determination of AGN torus morphology using X-ray spectra: A reliability study

Tathagata Saha[1], Alex G. Markowitz[1·2] and Johannes Buchner[3]

1. Nicolaus Copernicus Astronomical Center, Polish Academy of Sciences, Bartycka 18, 00–716 Warsaw, Poland
2. University of California, San Diego, Center for Astrophysics and Space Sciences, 9500 Gilman Drive, La Jolla, CA 92093-0424
3. Max Plank Institute for Extraterrestrial Physics, Giessenbachstrasse, D-85741 Garching, Germany

Several models of the X-ray reprocessing of the AGN torus have been recently released, (UXCLUMPY, CTORUS, MYTORUS etc.). These models assume a range of torus geometries and morphologies. The degeneracies in these models can limit reliable constraints of several parameters of interest, such as the intrinsic photon index and parameters of torus morphology. To investigate these effects, we simulate synthetic data under *XMM-Newton* and *NuSTAR* response files based on six different models. We use Bayesian methods to analyze the simulated datasets with the same set of models. Several geometrical parameters remain unconstrained for exposure times and fluxes typical of nearby Compton-thick AGN. In addition, a distinction of model or morphology using Bayesian methods is possible only if we have a high intrinsic value of flux for a typical exposure time. Our project aims to provide guidance for the X-ray community both in terms of the accuracy in applying the correct torus model (with implications for conclusions on the torus geometry and morphology) and the robustness of estimation of model parameters (with implications for limitations on precision of those parameters).

1 Introduction

Non-Blazar Active Galactic Nuclei (AGN) are classified into two types in optical waveband: Type-1s, which exhibit both broad and narrow emission lines, and Type-2s, which exhibit only narrow emission lines. This is explained by the existence of a circumnuclear obscurer popularly called the AGN torus. When the obscurer is at the line of sight of the observer it attenuates a significant fraction of continuum emission from the central engine along with broad (Balmar) emission lines from the broad-line region (BLR) (Antonucci, 1993) resulting in the Type-2 or obscured AGN phenomenon. Initially, the obscuring torus was depicted as a simple doughnut (Urry & Padovani, 1995), where the Type-2 phenomenon was thought to be the sole result of the orientation of the observer's line of sight concerning the doughnut axis. However, recent studies from the infrared (IR) and X-rays wavelength range suggest a more complex, and rather clumpy structure of the torus matter distribution. The evidence of clumpy tori was derived from spectral properties in the IR (e.g. Ramos Almeida et al., 2011) and variable line of sight absorption $N_{\rm H,los}$ in X-ray eclipse events (e.g. Markowitz et al., 2014). However, open questions remain regarding AGN tori and its connection to other AGN physics e.g. the torus and its connection to the accretion inflow-outflow and the cosmic X-ray background (XRB, Comastri et al., 1995), where a significant contributor to the XRB is thought to be obscured AGNs. Correct understanding of the properties e.g. gas density, morphology, etc. will help the community understand its connection to inflow-outflow and XRB.

Our work focuses on Compton thick (C-thick) X-ray obscured AGNs. In a C-thick AGN, a part of the direct power-law continuum from the corona is photoelectrically absorbed. The rest are Compton scattered (or reflected) by the torus. Over the last decade, several models simulating the X-ray spectra of the AGN torus have been developed. These models use complex radiative transfer codes to simulate torus radiation. Each of these models assumes a different morphology of the torus gas distribution. However, the question remains whether it is possible to distinguish these simplistic morphologies using currently available X-ray observatories and hence understand the feasibility in real-life studies.

To address the above question, we simulate data under the theoretical spectrum from different torus models using the instrument functions of *XMM–Newton* EPIC-pn (Jansen et al., 2001) and *NuSTAR* (Harrison et al., 2013) X-ray observatories. We then analyze the simulated data using Bayesian fitting methods. We perform two different kinds of analysis: (1) intramodel fits (IM-fits hereafter): where the data simulation and fitting model are the same, and (2) cross-model (CM-fits hereafter) fits, where the simulation and fitting model are different. In IM-fits we check the accuracy with which the model parameters are recovered. In CM-fits we check whether specific parameters are recovered and how the difference in assumed morphology can potentially affect our conclusions. We also test whether different models can be distinguished via Bayesian methods.

2 Methodology

2.1 The Compton-thick AGN model

The main components in the X-ray spectrum of an obscured AGN are :

1. The absorbed component or the zeroth-order continuum [$I_{\rm torus,abs}(E)$] is practically an absorbed power law. It exhibits a rollover for $E < 6$ keV and a strong Fe K edge, with shape dependent on the value of $N_{\rm H,los}$.
2. The scattered or reflected continuum [$I_{\rm torus,reflect}(E)$] shows a Compton reflection hump (CRH) in the hard X-rays and along with strong emission lines and their accompanying Compton shoulders e.g. the 6.4 keV iron line.
3. A small fraction (0.1 to 10%) of the intrinsic X-ray emission gets scattered in the diffuse gas or the inter-clump medium of a clumpy torus. To the first order, the component is modeled as a simple power-law (Buchner et al., 2019). This component [$I_{\rm SCPL}$] is the scattered power-law or the warm mirror.
4. Soft X-ray components might appear due to potential contamination from host galaxy structures such as star-forming regions and point sources such as ULXs or XRBs etc.

The generic model is shown in Fig. 1. We simulated data in two regimes characterized by the value of the input column density:

- A. Medium Compton Thick (MCT) regime: $N_{\rm H,los} < 200$ where the zeroth order continuum is quite strong. The ratio of the theoretical flux of the zeroth-order continuum ($F_{\rm T}$) to that of the scattered continuum ($F_{\rm R}$) in the 3-100 keV band is $\left[\frac{F_{\rm T}}{F_{\rm R}}\right]_{3-100\ \rm keV} = 1.8 - 5.0$.

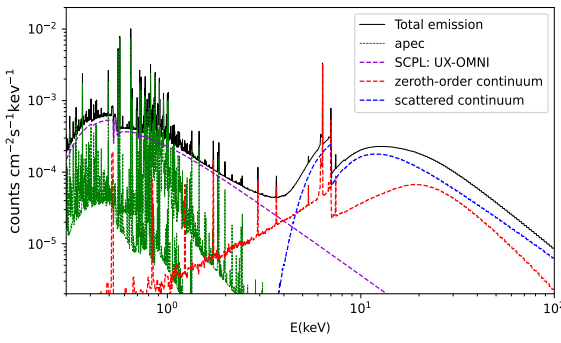


Fig. 1: Generic model of a spectrum of a X-ray obscured AGN. The total spectrum is shown in the solid black line and the contributing components are shown in the coloured lines.

- B. Heavy Compton Thick (HCT) regime: $N_{\rm H,los} > 200$ where the scattered component is dominant over the zeroth-order continuum. In this case, the ratio $\left[\frac{F_{\rm T}}{F_{\rm R}}\right]_{3-100\ \rm keV}$ is ~ 0.1.

2.2 Synthetic data simulation

We simulate data using the XSPEC tool FAKEIT and the instrument functions of *XMM–Newton* EPIC-pn and *NuSTAR* focal plane module instruments (FPM) A and B for a C-thick AGN model. An exemplary model is shown in Fig. 1. We assume an exposure of 100 ks and 50 ks for *XMM–Newton* and *NuSTAR* respectively. The intrinsic (absorbed) flux in 2-10 keV band ($F_{2-10\ \rm keV}$) is taken to be 0.5 mCrab for the major part of our analysis.

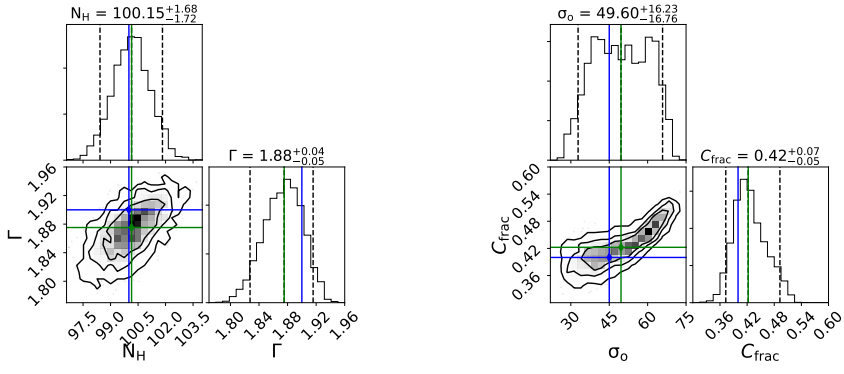
2.3 Fitting of synthetic data

We use the package Bayesian X-ray analysis (BXA, Buchner et al., 2014) which implements the nested-sampling package MULTINEST (Feroz et al., 2009) to analyse the simulated data. MULTINEST has two major advantages compared to traditional Goodman Weare Markov Chain Monte Carlo (GW-MCMC). Firstly, unlike GW-MCMC MULTINEST converges to the global maxima of the likelihood and thus does not require multiple burn-ins. It calculates Bayesian evidence Z (probability of the data: D, given a hypothesis: H) to determine the best among all competing models given a dataset (e.g. Buchner et al., 2014). In our work, we use BXA to study the posterior distributions of parameters and compare evidence values to distinguish models.

3 Results from the intramodel fits

For most datasets, we recovered the input parameters with 90% confidence, i.e., the input parameters lie in the 90% confidence region of the posteriors. We list the following properties :

- Parameters common across all models viz. photon index (Γ), line of sight absorption ($N_{\rm H,los}$) (see Fig. 2a). See (Saha et al., 2022) for more details.



(a) (b)

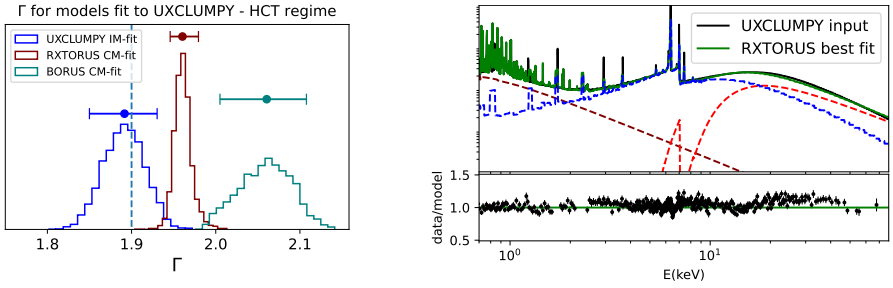
Fig. 2: Corner plots for (a) $N_{\rm H,los}$ of the torus and Γ (b) σ_o of the torus and $C_{\rm frac}$ of the inner ring (defined in Buchner et al., 2019) for the model UXCLUMPY. See Saha et al. (2022) for details.

- Parameters of morphology which determine the geometrical property of the torus, e.g., the standard deviation ($\sigma_{\rm o}$) of the spread of the torus clouds about the equator of the model UXCLUMPY were recovered. However, we notice from the contour plots that several parameters are correlated with each other (see Fig. 2).
- The constraints on parameters indicate dependence on regimes. A stronger presence of the zeroth-order continuum in the MCT regime allows tighter constraints on $N_{\rm H,los}$. In the HCT regime for a given 2.0-10.0 keV flux, the scattered continuum is the dominant component in the spectrum. The constraints on the morphological parameters are tighter in the scattered continuum-dominated HCT regime see Sec. 2.1 and Saha et al. (2022) for detailed discussion.

4 Results from the cross-model fits

Here we discuss the various aspects of the cross-model fits:

- Photon index (Γ): The photon index is determined incorrectly in the CM-fits for most models, i.e., they are systematically discrepant from the input value (see Fig. 3a). The discrepancy is due to the spectral shape difference for various models.
- LOS column density ($N_{\rm H,los}$): $N_{\rm H,los}$ was determined with almost correctly in the MCT regime because of a stronger presence zeroth-order continuum. For example, fitting data simulated under UXCLUMPY (clumpy model) in the MCT regime with MYTORUS (solid doughnut) returns $N_{\rm H,los,out} \simeq 92.0 \pm 2$ for input $N_{\rm H,los,in} = 100$ (see Saha et al. 2022 for details). However, in the HCT regime the absence of the zeroth-order continuum does not allow correct determination of $N_{\rm H,los}$ value, with huge discrepancies with respect to the input.



(a) (b)

Fig. 3: (a) The posteriors distributions of Γ when data simulated under UXCLUMPY is fit by other models. (b) Plot of RXTORUS best-fit spectrum and its components, for data simulated under UXCLUMPY. In the input model the transmitted component is heavily attenuated. In the fitting model the transmitted component increases to adjust the shape of the CRH region (see Saha et al. 2022 for details).

- The spectral fitting can be misleading as the contribution of the different components in replicating the Compton hump shape of the input spectrum can be discrepant from the input. For example, we find that, in the HCT regime, an incorrect fitting model arbitrarily increases the normalization of the zeroth-order continuum to adjust the shape of the CRH. However, the zeroth-order continuum is practically absent due to heavy attenuation in the HCT regime. (see Fig. 3b).

All these observations raise questions on the reliability of the models when analysis of data from observations is concerned.

5 The dependence of evidence on flux levels

We simulate ten spectra (S_n, n runs from 1 to 10) under the model CTORUS. The n-th spectrum (S_n)has a flux value of $f_n = 0.8/n$ mCrab. We perform both IM- and CM-fits using the UXCLUMPY and BORUS for all the above flux values. We define BF $= Z_{\rm fit}/Z_{\rm in}$, where Z is the Bayesian evidence of a fit. $Z_{\rm fit}$ is Bayesian evidence of the CM-fits using UXCLUMPY or BORUS. $Z_{\rm in}$ is the Bayesian evidence of the input. We study the variation of the Bayes-factor (BF) for intrinsic flux values. We adopt a cutoff value of BF $= 10^{-2}$ as the maximum value above which model distinction is not reliable. Figure 4 shows that the distinguishability of the models decreases with lower intrinsic flux and narrower energy band coverage as BF approaches the cutoff value and random fluctuations in the BF values increases (see Fig. 4). It also depends on the energy band coverage, i.e., logically, a broad energy band coverage implies better distinguishability (Saha et al., 2022).

6 Summary

From the IM-fits, we find that only a correct torus model allows us to determine the model parameters correctly, given a dataset with the typical exposure times on the instruments. The CM-fits indicate that the fit results are dependent on the model

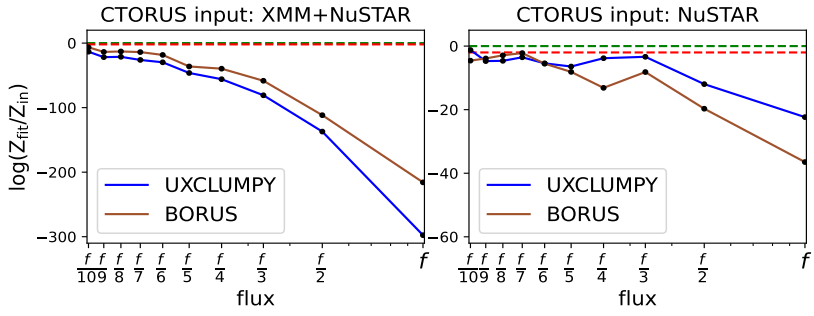


Fig. 4: Variation of Bayes Factor with source flux. *Left*: *XMM–Newton* + *NuSTAR*, *Right*: *NuSTAR* only. The blue and magenta dashed lines denote the flux level at which the relation $\log(Z_{\rm fit}/Z_{\rm in}) \leq -2$ holds for the wrong model. In this example, $Z_{\rm in} = Z_{\rm CTOR}$ and $Z_{\rm fit}$ corresponds to UXCLUMPY and BORUS. The black and green dashed lines marks the logBF=2 and logBF=0 level respectively. The grey line demonstrates the effect of random fluctuations, which brings the Bayes factor value close to the cutoff value of -2. Here, f = 0.8 mCrab.

we used for fitting. It is possible to get wrong values of parameters when the fitting model is different from the input model (see Fig. 3a). We show that Bayes factor-based model distinction is reliable only when the intrinsic flux of a given object is high enough. The presence of random fluctuations (see Fig. 4) in the low flux (<0.1 mCrab) regime or the absence of certain energy band coverage indicates the importance of synthetic data analysis given a dataset and models. These analyses will help understand the limitations. In general, conservative thresholds on Bayes-Factor values (far from the random fluctuation-dominated zone) should be applied to reduce the risks of crossing over to the Bayes factor > 1 (wrong model selection zone) region. More detailed discussion can be found in Saha et al. (2022).

Acknowledgements. This work is financially supported by Polish National Science Center (NCN) grants No. 2016/23/B/ST9/03123 and 2018/31/G/ST9/03224.

References

Antonucci, R., *ARA&A* **31**, 473 (1993)

Buchner, J., et al., *A&A* **564**, A125 (2014)

Buchner, J., et al., *A&A* **629**, A16 (2019)

Comastri, A., Setti, G., Zamorani, G., Hasinger, G., *A&A* **296**, 1 (1995)

Feroz, F., Hobson, M. P., Bridges, M., *MNRAS* **398**, 4, 1601 (2009)

Harrison, F. A., et al., *ApJ* **770**, 2, 103 (2013)

Jansen, F., et al., *A&A* **365**, L1 (2001)

Markowitz, A. G., Krumpe, M., Nikutta, R., *MNRAS* **439**, 2, 1403 (2014)

Ramos Almeida, C., et al., *ApJ* **731**, 2, 92 (2011)

Saha, T., Markowitz, A. G., Buchner, J., *MNRAS* **509**, 4, 5485 (2022)

Urry, C. M., Padovani, P., *PASP* **107**, 803 (1995)