Publication Year	2022
Acceptance in OA	2025-01-10T14:55:04Z
Title	A New Emulated Monte Carlo Radiative Transfer Disk-Wind Model: X-Ray Accretion Disk-wind Emulator - XRADE
Authors	Matzeu, G. A., Lieu, M., Costa, M. T., Reeves, J. N., Braito, V., Dadina, M., NARDINI, Emanuele, Boorman, P. G., Parker, M. L., Sim, S. A., Barret, D., Kammoun, E., Middei, R., Giustini, M., Brusa, M., Pérez Cabrera, J., MARCHESI, STEFANO
Publisher's version (DOI)	10.1093/mnras/stac2155
Handle	http://hdl.handle.net/20.500.12386/35618
Journal	MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY
Volume	515

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **515,** 6172–6190 (2022) https://doi.org/10.1093/mnras/stac2155
Advance Access publication 2022 August 1

A new emulated Monte Carlo radiative transfer disc-wind model: X-Ray Accretion Disc-wind Emulator – XRADE

G. A. Matzeu [1,2,3] M. Lieu [4]★ M. T. Costa,[5] J. N. Reeves,[6,7]★ V. Braito [6,7] M. Dadina,[2] E. Nardini [8] P. G. Boorman,[9,10] M. L. Parker [11] S. A. Sim,[12] D. Barret,[13] E. Kammoun [8,13] R. Middei [14] M. Giustini,[15] M. Brusa,[1,2] J. Pérez Cabrera[16] and S. Marchesi[2,17]

[1]*Department of Physics and Astronomy (DIFA), University of Bologna, Via Gobetti 93/2, I-40129 Bologna, Italy*
[2]*INAF – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Gobetti 93/3, I-40129 Bologna, Italy*
[3]*European Space Agency (ESA), European Space Astronomy Centre (ESAC), E-28691 Villanueva de la Cañada, Madrid, Spain*
[4]*School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, UK*
[5]*Astrophysics Group, School of Physical and Geographical Sciences, Keele University, Keele, Staffordshire ST5 5BG, UK*
[6]*Department of Physics, Institute for Astrophysics and Computational Sciences, The Catholic University of America, Washington, DC 20064, USA*
[7]*INAF – Osservatorio Astronomico di Brera, Via Bianchi 46, I-23807 Merate, LC, Italy*
[8]*INAF – Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, I-50125 Firenze, Italy*
[9]*Astronomical Institute of the Czech Academy of Sciences, Boční II 1401/1A, CZ-14100 Praha 4, Czech Republic*
[10]*School of Physics and Astronomy, University of Southampton, Highfield, Southampton SO17 1BJ, UK*
[11]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[12]*School of Mathematics and Physics, Queen's University Belfast, Belfast BT7 1NN, UK*
[13]*IRAP, Université de Toulouse, CNRS, UPS, CNES 9, Avenue du Colonel Roche, BP 44346, F-31028 Toulouse Cedex 4, France*
[14]*Space Science Data Center – ASI, Via del Politecnico, snc, I-00133 Roma, Italy*
[15]*Centro de Astrobiología (CSIC-INTA), Camino Bajo del Castillo s/n, Villanueva de la Cañada, E-28692 Madrid, Spain*
[16]*Aurora Technology B.V., European Space Astronomy Centre (ESAC), Camino Bajo del Castillo s/n, Villanueva de la Cañada, E-28692 Madrid, Spain*
[17]*Department of Physics and Astronomy, Clemson University, Kinard Lab of Physics, Clemson, SC 29634, USA*

Accepted 2022 July 27. Received 2022 July 27; in original form 2022 April 15

ABSTRACT

We present a new X-Ray Accretion Disc-wind Emulator (XRADE) based on the 2.5D Monte Carlo radiative transfer code that provides a physically motivated, self-consistent treatment of both absorption and emission from a disc wind by computing the local ionization state and velocity field within the flow. XRADE is then implemented through a process that combines X-ray tracing with supervised machine learning. We develop a novel emulation method consisting in training, validating, and testing the simulated disc-wind spectra into a purposely built artificial neural network. The trained emulator can generate a single synthetic spectrum for a particular parameter set in a fraction of a second, in contrast to the few hours required by a standard Monte Carlo radiative transfer pipeline. The emulator does not suffer from interpolation issues with multidimensional spaces that are typically faced by traditional X-ray fitting packages such as XSPEC. XRADE will be suitable to a wide number of sources across the black hole mass, ionizing luminosity, and accretion rate scales. As an example, we demonstrate the applicability of XRADE to the physical interpretation of the X-ray spectra of the bright quasar PDS 456, which hosts the best-established accretion disc wind observed to date. We anticipate that our emulation method will be an indispensable tool for the development of high-resolution theoretical models, with the necessary flexibility to be optimized for the next generation microcalorimeters onboard future missions, like *X-Ray Imaging and Spectroscopy Mission* (*XRISM*)/Resolve and *Athena*/X-ray Integral Field Unit (X-IFU). This tool can also be implemented across a wide variety of X-ray spectral models and beyond.

Key words: radiative transfer – methods: numerical – techniques: spectroscopic – galaxies: active – galaxies: individual: PDS 456.

1 INTRODUCTION

Accretion disc winds are generally observed through blueshifted absorption features at rest-frame energies >7 keV, imprinted in the X-ray spectra of active galactic nuclei (AGNs; Chartas et al. 2002; Pounds et al. 2003; Reeves, O'Brien & Ward 2003). Their degree of blueshift from the lab energies of Fe XXV Heα (He-like) and/or FeXXVI Lyα (H-like) implies mildly relativistic outflow velocities, typically falling in the range $\sim$0.1–0.4c (e.g. Reeves et al. 2009; Gofford et al. 2014; Nardini et al. 2015; Matzeu et al. 2016, 2019; Middei et al. 2020; Parker et al. 2020). Their frequent detection, in approximately 35–40 per cent of local AGNs (Tombesi et al. 2010;

★ E-mail: maggie.lieu@nottingham.ac.uk (ML); james.n.reeves456@gmail.com (JNR)

© 2022 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

Gofford et al. 2013; Igo et al. 2020), suggests that the wind geometry is characterized by a large covering factor (Ω). This was confirmed by the direct measurement of $\Omega \gtrsim 2\pi$ in the luminous quasar PDS 456 (Nardini et al. 2015, hereafter N15). With such a high covering factor, coupled with high column densities ($N_{\rm H} \gtrsim 10^{23}$ cm^{-2}; Tombesi et al. 2011; Gofford et al. 2013) and high velocities, a large amount of kinetic power can be transported, possibly exceeding the 0.5–5 per cent of the bolometric luminosity required for significant AGN feedback (King 2003; King & Pounds 2003; Di Matteo, Springel & Hernquist 2005; Hopkins & Elvis 2010).

Measuring the intrinsic physical properties of these winds can provide important insights into the mechanism through which they are driven (launched, accelerated). There are currently three known physical mechanism responsible for driving accretion disc winds: gas pressure, radiation pressure, and magnetic fields. While gas pressure (thermal driving) is unable to explain the large velocities observed in accretion disc winds in AGN, the two other mechanisms are in principle able to do so. Three possible scenarios might therefore be able to explain the observations of AGN accretion disc winds: (i) *radiatively driven* winds (e.g. Proga, Stone & Kallman 2000; Kallman & Bautista 2001; Proga & Kallman 2004; Giustini & Proga 2019); (ii) *magnetically driven* (hereafter MHD) winds (e.g. Emmering, Blandford & Shlosman 1992; Ohsuga et al. 2009; Fukumura et al. 2010, 2015; Kazanas et al. 2012); and/or (iii) to some extent a likely combination of the two (e.g. de Kool & Begelman 1995; Everett 2005; Matzeu et al. 2016).

In the radiatively driven scenario, the AGN radiation pressure launches a wind from the accretion disc from tens to thousands of gravitational radii from the supermassive black hole (SMBH; the gravitational radius $r_{\rm g} = GM_{\rm BH}/c^2$, with G the gravitational constant, c the speed of light, and $M_{\rm BH}$ the black hole mass). The detection of strongly blueshifted broad absorption lines (BALs), associated with the ultraviolet (UV) transitions (e.g. Weymann et al. 1991; Matthews et al. 2016; Rankine et al. 2020), demonstrates that substantial momentum can be transferred from a powerful radiation field to the gas, thus accelerating mass outflows. These type of radiatively driven outflows are described as *line-driven* winds, as their strength depends on the opacity of the absorption lines, which acts as a *force multiplier* to the radiation pressure and can make the bound–bound absorption cross-section considerably larger than the Thomson cross-section for electron scattering (i.e. $\sigma_{\rm line} \gg \sigma_{\rm T}$; e.g. Castor, Abbott & Klein 1975; Stevens & Kallman 1990; Dannen et al. 2019). The strength of line-driven disc winds depends on the ionization state of the gas $\xi = L/nR^2$, where n is the gas density, L is the ionizing luminosity, and R is the distance between the gas and the source of the ionizing luminosity. As demonstrated by Dannen et al. (2019), for a typical AGN spectral energy distribution (SED) the effects of the force multiplier drop at $\log \xi > 3$, where all the relevant opacity is lost. Line-driven winds are therefore likely more relevant for sub-Eddington sources,[1] where the ionizing luminosity is not as large as completely ionize the illuminated gas.

In AGN close to Eddington or super-Eddington, the ionization state of the gas is so high that the dominant interaction between the outflowing gas and the radiation field is likely Thomson (and Compton) scattering (King & Pounds 2003; King 2010). In this case, a direct correlation between the momentum rate of the outflow and the momentum rate of the radiation field, i.e. $\dot{p}_{\rm out} \sim \dot{p}_{\rm rad} (= L/c)$, would be expected if the optical depth to electron scattering is $\tau \sim 1$. This indeed appears to be the case in many observations of fast, highly ionized winds (Tombesi et al. 2013; Gofford et al. 2015; Nardini, Lusso & Bisogni 2019), but it does require the AGN to radiate at a considerable fraction of its Eddington luminosity, $L_{\rm Edd}$ (King & Pounds 2003).

Most theoretical outflow studies are mainly concentrated on radiatively driven winds in both AGNs (Sim et al. 2008, 2010; Hagino et al. 2015, 2016, 2017; Matthews et al. 2016, 2020; Nomura & Ohsuga 2017; Luminari et al. 2018; Nomura, Ohsuga & Done 2020; Quera-Bofarull et al. 2020; Mizumoto et al. 2021), X-ray binaries (XRBs; Higginbottom et al. 2019, 2020; Tomaru et al. 2020a,b), and cataclysmic variables (e.g. Matthews et al. 2015). Nevertheless, MHD wind models have been successfully applied to both AGNs (Fukumura et al. 2010, 2015, 2018) and XRBs (Fukumura et al. 2017, 2021; Ratheesh et al. 2021). These findings suggest that both driving/launching mechanisms apply across the black hole mass and luminosity scales.

The development of physical models for accretion disc winds and a self-consistent test of their predictions are among the primary goals in modern X-ray astronomy. Predictions can be tested by using grids of spectral simulations generated for different values of the physical parameters of interest, such as the ionization state and column density of the gas. Up until now, astronomers had to compromise between the sampling resolution and the extent of the parameter space covered in the model, due to the extremely demanding computational times involved. Although grids generated with coarser sampling generally allow one to explore a broader parameter space, they are more susceptible to interpolation issues (Arnaud 1996) that may affect the degree of accuracy of the measurements. The next generation of instruments, onboard *X-Ray Imaging and Spectroscopy Mission* (*XRISM*) and *Athena* (planned to be launched in 2023 and early 2030, respectively), will provide a significant increase in spectral resolution, with $\Delta E \sim 5$ eV for *XRISM*/Resolve (Tashiro et al. 2020) and $\Delta E \sim 2.5$ eV for *Athena*/X-ray Integral Field Unit (X-IFU; Barret et al. 2018). Such advances in technology will inevitably require the development of higher resolution grids to match the improved spectral information.

Machine learning techniques, which allow us to learn the mapping from an input space to an output space, can play a fundamental role in speeding up this process. In supervised machine learning, a sample of both the input and the output is known, and the objective is to learn a mapping that is able to best reproduce the output for a given *loss function*. In our case, the loss function is a measure of how close the machine learning emulated X-ray spectra are to *ground truth*, the simulated spectral values. This is the training of the model, and the data sample is known as the training data. Machine learning benefits from large data samples and although such process can be computationally expensive to train, the trained data are capable of efficiently computing the mapping.

Consequently, supervised learning methods can be useful to reduce the computational cost of large and complex models, provided that a representative training set can be obtained (Kasim et al. 2022). Trained machine learning models can be used as surrogate models to approximate computationally expensive models such as the weather and climate (Watson-Parris 2021). These frameworks are known as *emulators*, and they can be developed as *artificial neural networks* (ANNs hereafter). The ANN architecture is loosely based on the human brain and consists of interconnected neurons organized into layers. ANNs are also quickly becoming popular in astronomy to approxi-

[1]The Eddington luminosity is defined as $L_{\rm Edd} = 4\pi GMm_{\rm p}c/\sigma_{\rm T}$, with $m_{\rm p}$ the proton mass and $\sigma_{\rm T}$ the Thomson cross-section, and it is the luminosity for which the radiation pressure and the gravitational pull are equal, for a given mass M.

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

mate simulations and interpolate between them (see e.g. Chardin et al. 2019; He et al. 2019). Kerzendorf et al. (2021) created an emulator to replace expensive radiative transfer codes for modelling supernova spectral time series, while Alsing et al. (2020)'s stellar population synthesis (SPS) model emulator accurately generates galaxy spectra and photometry from SPS parameters.

The aim of this paper is to present the description, development, and application of a new extended Monte Carlo radiative transfer ($\mathcal{MCRT}$ hereafter) accretion disc-wind code initially developed by Sim et al. (2008, 2010, hereafter S08, S10): X-Ray Accretion Disc-wind Emulator (XRADE). The novel approach in the development of XRADE is twofold. (i) First, we compute a new set of synthetic X-ray spectra in order to explore the physical conditions of accretion disc winds in a larger AGN population. (ii) Secondly, as the synthetic spectra are fed into a purposely built ANN, the data will undergo a process of training, validation, and testing with the aim of (a) accelerating the process of synthetic spectra simulations, and (b) solving the multidimensional interpolation problems[2] that arise when multiplicative tables are adopted in spectral fitting packages such as XSPEC (Arnaud 1996). Our ANN allows the user to generate customized XRADE tables at their requirement. On this basis, we generated two new large $\mathcal{MCRT}$ tables, namely `slow64` and `fast32`, which cover a larger parameter space (CPU time: ∼7–8 months with 480 50 GB cores), than the one generated in S08, S10, and Reeves et al. (2014). In the future, the spectral resolution of the wind grids will also be increased, in order to match the next-generation calorimeter data. Hence to reduce the computational demands for our future tables, machine learning is a very important tool.

This paper is organized as follows. In Section 2, we give an overview of the $\mathcal{MCRT}$ methods used to simulate disc-wind synthetic spectra from the code originally developed in S08 and S10, and we describe the physical assumptions adopted in the disc-wind `slow64` and `fast32` models. We also discuss the input parameters and we present a brief description of the main input parameters. In Section 3, we describe in detail the methods adopted for the development of XRADE. In Section 4, we apply XRADE to the quasar PDS 456, which hosts one of the most powerful and persistent accretion disc winds discovered to date. We specifically test XRADE on the *XMM–Newton* and *NuSTAR* 2013 observation of PDS 456, as the X-ray spectrum is characterized by the most prominent and best studied P Cygni-like profile yet observed. In Section 5, we draw our conclusions and discuss further work.

2 RADIATIVE TRANSFER CODE OVERVIEW

The development of XRADE is based on training the synthetic wind spectra simulated with the $\mathcal{MCRT}$ code by S08 and S10 into an ANN. Note that a more detailed description of the input model set-up can be found in S08 and S10. In this section, we present, for completeness, an overview of the physical basis and approach adopted in generating the input $\mathcal{MCRT}$ wind spectra for XRADE.

Initially, S08 carried out multidimensional (2.5D) Monte Carlo radiative transfer simulations in a biconical wind structure (see Fig. 1). The simulated spectra were calculated over grid points with coordinates x, y, z, under the assumption that the system is axisymmetric about the polar (z) axis in the azimuthal direction. S10 extended the atomic data base with the inclusion of the L- and M-shell transitions. As a result, the simulated synthetic spectra

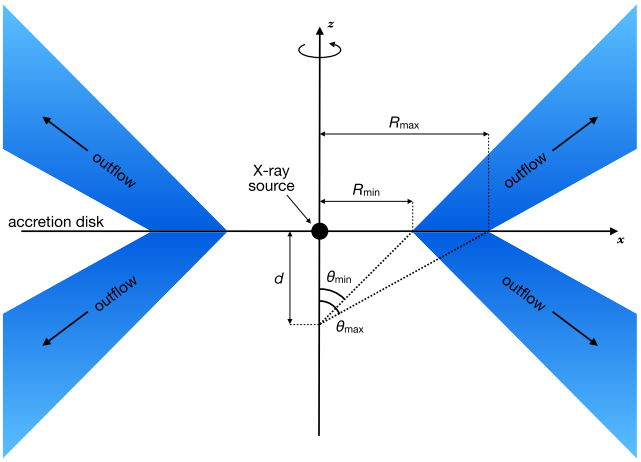

Figure 1. Schematic representation of the biconical structure adopted for the disc-wind model. The 2D geometry is defined with three parameters: R_{min}, R_{max}, and d. R_{min} and R_{max} correspond to the radii at which the inner and outer edges of the disc wind intercept the disc plane, respectively. The blue-shaded area represents the physical extent of the outflow, while d defines the focus point of the wind below the disc plane, which controls the degree of collimation and opening angle of the wind (where $\tan\theta_{min} = \frac{R_{min}}{d}$). This 2D structure is then rotated around the vertical z-axis and mirrored with respect to the accretion disc plane to produce an axisymmetric 2.5D wind geometry.

were more accurate over a larger range of photon energies (i.e. 0.2–10 keV). Additionally, the Monte Carlo ray-tracing method described in Lucy (2002, 2003) was implemented in the code. This allowed the treatment of ionization and radiative heating of the gas by means of self-consistent calculations of the heating/cooling of electrons based on the *photon packets* (the computational structure used in the simulations) that propagate throughout the wind. A temperature gradient is then calculated in order to provide a more physical representation of the ionization structure of the wind. This process is then reiterated multiple times to accurately define the heating and cooling rates for the wind until they reach equilibrium. In `slow64` and `fast32`, we set 160 000 photon packets each, which are then collected and grouped into 10 000 energy bins (and binned up by a factor of 10× giving a total of 1000 energy bins), based on their viewing (observer) angle (θ) from the z-axis.

The $\mathcal{MCRT}$ code creates tables of simulated wind spectra that take into account the effects of the radiation transmitted through the wind, which include the scattering and reflected emission from the flow. An interesting outcome of this model is that the accretion disc wind itself can give rise to Fe K emissions, with line widths up to $\sigma_{width} \sim 1$ keV (Parker et al. 2022). Such profiles are obtained from the combination of (i) velocity shear in the flow, (ii) its rotation around the polar axis, and (iii) the Compton scattering of the Fe Kα line photons in the wind. Thus the disc-wind model can provide a physically motivated, self-consistent treatment of both the absorption and emission produced in the wind by computing the ionization state and velocity field within the flow. In other words, the disc-wind model calculates the ionization at each point in the wind over a wide range of states, thus describing a more realistic (non-uniform) ionization structure and velocity field throughout the outflow. As for iron, the code covers charge states from Fe X–XXVII, and the output spectra include not just the absorption and emission from Fe K, but also from L-shell iron and the K-shell lines of lighter elements in the soft X-ray band.

[2] https://heasarc.gsfc.nasa.gov/xstar/docs/html/node95.html

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

Routines that take into account special relativistic aberration of angles and Doppler shifts between the comoving and observer frames are included in our $\mathcal{MCRT}$ code, so that the model fully accounts for special relativistic effects. Such an implementation provides realistic and accurate estimates of the mass outflow rate and overall energetics, as the local radiative pressure might require non-negligible special relativistic corrections (e.g. Luminari et al. 2020, 2021). The number of energy packets used in each simulation is chosen so that the Monte Carlo noise in the estimators is <3 per cent. This level of precision is sufficient given the quality of the observational data available at present (see Section 4).

2.1 Geometry

The assumed biconical structure of the inner disc-wind geometry is shown in Fig. 1. The x-axis corresponds to the plane of the accretion disc, and the z-axis to the polar (rotational) direction. The black hole is located at the origin and the X-ray source is located within $6\,r_{\rm g}$ from it (see Section A2). $R_{\rm min}$ and $R_{\rm max}$ are, respectively, the distances from the origin to the inner and outer edge of the wind at the interception with the equatorial (xy) plane. The radii $R_{\rm min}$ and $R_{\rm max}$ (expressed in gravitational units) then enclose the disc-wind launch region, and set the collimation and the overall opening angle (equatorial or polar) together with the parameter d, which represents the distance of the focal point of the wind along the z-axis below the origin. The overall wind inclination angle θ is measured with respect to the z-axis, with the polar opening angle defined as $\theta_{\rm min} = \arctan R_{\rm min}/d$. Here we set $d/R_{\rm min} = 1$, so the wind opening angle is 45° from the pole. The observer's polar angle is included in the code through $\mu = \cos\theta$, where any line of sight with $\mu < 0.7$ intercepts the wind. The terminal velocity[3] attained by the wind is $v_\infty = f_{\rm v} v_{\rm esc}$, where $v_{\rm esc} = (2/R_{\rm min})^{1/2} c$, and the factor $f_{\rm v}$ parameter allows the user to vary the terminal velocity for a given launch radius (see below). The lines that extend from d and intercept the xy-plane in $R_{\rm min}$ and $R_{\rm max}$ produce the first quarter of the biconical wind, which is made axisymmetric under rotation in the azimuthal direction and reflected with respect to the disc plane (see Fig. 1). The difference between the outermost and innermost launch radii ($\Delta R = R_{\rm max} - R_{\rm min}$) of the flow off the disc plane defines the overall thickness of the wind streamline.

2.2 Velocity

Having set up the geometric framework in which the wind is simulated, we now describe the properties and key parameters of the synthetic spectra that will be subsequently fed into the `ANN`. Note that the emulation process itself will be described in more detail in Section 3. For the purpose of this work we generated two $\mathcal{MCRT}$ disc-wind tables named `fast32` and `slow64` (see Table 1 for the summary of their parameter space). The former is tuned to the fastest disc-wind cases like PDS 456, where typically $v_{\rm out}/c = 0.25$–0.35 (e.g. Matzeu et al. 2017), with $R_{\rm min} = 32\,r_{\rm g}$; thus, for $f_{\rm v} = 1$, $v_\infty = -0.25c$. The latter is instead tuned to slower winds, e.g. MCG-03-58-007 (e.g. Braito et al. 2022) or PG 1211+143 (e.g. Pounds et al. 2016), with $R_{\rm min} = 64\,r_{\rm g}$; for $f_{\rm v} = 1$, $v_\infty = -0.177c$ (see Section A2). Our input choice of $R_{\rm min}$ is related to the range of outflow velocities typically observed in AGNs, in the range $v_{\rm w} \sim 0.05$–$0.4c$ (e.g. Tombesi et al. 2010; Gofford et al. 2013; Reeves et al. 2018a; Igo et al. 2020; Chartas et al. 2021). The terminal velocity parameter $f_{\rm v}$ can be considered a fine-tuning factor of the outflow velocity, which allows the user to adjust v_∞ to match their observations. So v_∞ is regulated by changing the $f_{\rm v}$ parameter, for a given launch radius. Note that for these $\mathcal{MCRT}$ simulations a black hole mass of $M_{\rm BH} = 10^9\,{\rm M}_\odot$ is assumed. However as most of the units are normalized, e.g. radii to the gravitational radius, mass outflow rate, and X-ray luminosity to the Eddington value (see below), the output table parameters are black hole mass invariant.

Note that these versions of these $\mathcal{MCRT}$ tables are newly generated in this work and they will be made publicly available. Hence, the new range of parameters is tabulated in Table 1. The spectral properties of these grids, in particular in relation to the inclination and launch radius, are discussed further in Appendix A. Both the `slow64` and `fast32` tables were generated with $f_{\rm v}$ ranging between 0.25 and 2 in steps of $\Delta f_{\rm v} = 0.25$. As a result, the following ranges of v_∞ are covered:

$$v_\infty/c = \begin{cases} -0.500 \lesssim v_\infty/c \lesssim -0.0625 & R_{\rm min}/\,r_{\rm g} = 32, \\ -0.354 \lesssim v_\infty/c \lesssim -0.0442 & R_{\rm min}/\,r_{\rm g} = 64. \end{cases} \quad (1)$$

For simplicity, for both the `fast32` and `slow64` tables, the geometric thickness of the outflow is set to be $R_{\rm max}/R_{\rm min} = 1.5$, but in principle this could be variable. The outer boundary of the simulations is set as $\log(R_{\rm out}/\,r_{\rm g}) = 4.53$ (i.e. $\sim 34\,000\,r_{\rm g}$), whereas the X-ray source is set to originate from a central region of $6\,r_{\rm g}$ in radius. Both the `slow64` and `fast32` tables are generated with 5 grid points for the photon index (Γ; see Section 2.5), 8 for the terminal velocity parameter ($f_{\rm v}$), 12 for the normalized mass outflow rate ($\dot{\mathcal{M}}_{\rm w}$; see Section 2.3), 9 for the ionizing luminosity ($\mathcal{L}_{\rm X}$; see Section 2.4), and 20 angular bins (μ). The combination of these parameters produces $5 \times 8 \times 12 \times 9 \times 20 = 86\,400$ synthetic spectra in each table, for a total of 172 800. Each spectrum is simulated over 1000[4] spectral points, uniform in log-space, and subsequently used in the emulation process described below.

2.3 Mass outflow rate

The mass within the flow is determined by the normalized mass outflow rate parameter, which is expressed in Eddington units as $\dot{\mathcal{M}}_{\rm w} = \dot{M}_{\rm out}/\dot{M}_{\rm Edd}$ (a radiative efficiency for a Schwarzschild black hole of $\eta = 0.06$ is assumed; Shapiro & Teukolsky 1983). Hence, $\dot{\mathcal{M}}_{\rm w}$ is not directly dependent upon the black hole mass of the source. An increase in $\dot{\mathcal{M}}_{\rm w}$ affects the mass density in each cell by increasing the opacity of the medium thereby yielding a higher column density through the wind and deeper absorption lines (see Section A4 in Appendix A). Additionally, as scattering of photons increases with opacity, the relative strength of the component scattered out of the flow would also increase proportionally with $\dot{\mathcal{M}}_{\rm w}$. In both tables, the $\dot{\mathcal{M}}_{\rm w}$ parameter covers the $0.020 < \dot{\mathcal{M}}_{\rm w} < 0.676$ range in 12 equally spaced logarithmic steps (see Table 1). This range covers the bulk of the typical measurements carried out in the literature, i.e. $-2 \lesssim \log(\dot{\mathcal{M}}_{\rm w}) \lesssim 0$ (see fig. 2 in Tombesi et al. 2012 and fig. 1 in Gofford et al. 2015). Note that for future grids it is our intention to extend the $\dot{\mathcal{M}}_{\rm w}$ parameter space to super-Eddington values, $\dot{\mathcal{M}}_{\rm w} \gtrsim 1$.

[3] See Section A3 in Appendix A for a calculation of the velocity field through the streamline, up to the maximum terminal velocity, v_∞.

[4] Note that 1000 energy bins are adopted when simulating CCDs resolution spectra, i.e. $\Delta E = 60\,{\rm eV}$ at 6 keV, over the 0.1–511 keV range. For future microcalorimeter resolution we will increase the binning by at least one order of magnitude.

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

Table 1. `fast32` and `slow64` input parameters. Note that the 172 800 output spectra will be available for developing XRADE. For the quantities flagged with † or ‡, more details are provided in Sections A3 and A4 (Appendix A), respectively. Note that in the bottom five rows the ANN input parameters are also the measurable output when `fast32` and `slow64` tables are loaded into X-ray fitting packages such as XSPEC.

Input parameter	Values	
	`fast32`	`slow64`
Range of source photon energies in simulation	0.1–511 keV	
Photon packets	160 000	
Size of X-ray emission region (r_{er})	$6\,r_g$	
Inner radius of the accretion disc (r_d)	$6\,r_g$	
Inner launch radius (R_{min})	$32\,r_g$	$64\,r_g$
Outer launch radius (R_{max})	$1.5R_{min}$	
Distance to wind focus (d)	R_{min}	
Velocity scale length (R_v)†	R_{min}	
Velocity exponent (β)†	1.0	
Launch velocity (v_0)†	0.0	
Mass-loss exponent (κ)‡	−1.0	
Outer radius of simulation grid	$33876\,r_g$	
3D Cartesian RT grid cells	$180 \times 180 \times 180$	
2D wind grid zones	100×100	
Input parameter	Values	
Source power-law photon index (Γ)	{1.6, 1.8, 2.0, 2.2, 2.4}	
Terminal velocity parameter (f_v)	{0.25, 0.50, 0.75, 1.0, 1.25, 1.50, 1.75, 2.0}	
Source luminosity $\left(\mathcal{L}_X = \frac{L_{2-10\,keV}}{L_{Edd}}\right)$	$\{0.0252, 0.0475, 0.0796, 0.1415, 0.2516, 0.4475, 0.7958, 1.4151, 2.5165\} \times 10^{-2}$	
Wind mass-loss rate $\left(\dot{\mathcal{M}}_w = \frac{\dot{M}_{out}}{\dot{M}_{Edd}}\right)$	{0.0196, 0.0270, 0.0373, 0.0515, 0.0710, 0.0980, 0.1352, 0.1866, 0.2575, 0.3552, 0.4901, 0.6762}	
Angular bins ($\mu = \cos\theta$)	{0.025, 0.075, 0.125, 0.175, 0.225, 0.275, ···0.725, 0.775, 0.825, 0.875, 0.925, 0.975}	

2.4 Ionizing X-ray luminosity

The ionizing luminosity parameter is defined as the fraction of X-ray luminosity, calculated in the 2–10 keV band, with respect to the Eddington luminosity, i.e. $\mathcal{L}_X = L_{2-10\,keV}/L_{Edd}$. As per $\dot{\mathcal{M}}_w$, with this normalization the $\mathcal{L}_X$ parameter keeps the same meaning across the black hole mass scale. $\mathcal{L}_X$ measures the overall degree of ionization of the material within the flow, where lower values of $\mathcal{L}_X$, typically <1 per cent of L_{Edd}, lead to the wind being less ionized and more opaque to X-rays. In contrast, an increase in $\mathcal{L}_X$ will lead to winds that are more ionized and transparent to X-rays, to the extent that the spectrum becomes completely featureless. In the disc-wind code, the ionization of the plasma is self-consistently computed at each point in the wind, whilst both shielding and scattering of photons are also accounted for in the calculations. As a result, the overall ionization is stratified along the wind, whereby the innermost surface of the wind is almost fully ionized (mainly Fe XXVI), as expected, being fully exposed to the X-ray source. The denser base of the wind is, not surprisingly, less ionized (with charge states down to Fe X–XVI). The decrease in ionization occurs both along the flow and across the base of the wind. More details regarding the input spectrum and its effect upon the wind ionization will be discussed in Section 2.5.

Compared to other models (e.g. Hagino et al. 2015, see Section 4.2), the disc-wind code has access to more extensive atomic data, which cover a wide range in ionization; ions from Fe X–XXVI and from lighter elements such as C–Si are included. Thus, for any given observation of an AGN, the $\mathcal{L}_X$ parameter can be calculated by comparing the intrinsic 2–10 keV luminosity to the (known) Eddington luminosity, and it is not a degenerate parameter in the modelling. The synthetic spectra for XRADE were simulated over a range of $2.5 \times 10^{-4} < \mathcal{L}_X < 2.5 \times 10^{-2}$ (or 0.025 per cent to 2.5 per cent of L_{Edd}) over nine equally spaced logarithmic increments (see Table 1). It is worth briefly discussing how such a range compares to the observed distributions of Eddington ratios ($\lambda_{Edd} = L_{bol}/L_{Edd}$) and bolometric corrections ($k_{bol} = L_{bol}/L_{2-10\,keV}$) as, by definition, $\mathcal{L}_X = \lambda_{Edd}/k_{bol}$. These two quantities are known to correlate with each other, and their ratio typically falls in the range $\approx 10^{-3}$–10^{-2} for the majority of type 1 AGNs (e.g. Vasudevan & Fabian 2009; Lusso et al. 2012). We conservatively adopt for $\mathcal{L}_X$ a more extended range, especially at the low end, based on the evidence that the strongest winds are usually observed in sources that are relatively weak in the X-rays compared to the UV (hence a larger k_{bol}), which is interpreted as a requirement for effective line driving (e.g. Castor et al. 1975; Giustini & Proga 2019).

2.5 The input spectrum

The choice of the initial input spectrum is a crucial step for setting the $\mathcal{MCRT}$ simulations, required for the development of XRADE, as the intrinsic spectrum can profoundly affect the observable disc-wind parameters. Steep (i.e. $\Gamma > 2$) spectral slopes of the X-ray continuum are, in fact, critically responsible for producing strong absorption profiles. On the other hand, harder spectra (i.e. $\Gamma < 2$) likely overionize the obscuring medium, leading to a considerable attenuation or disappearance of the absorption profiles (e.g. Pinto et al. 2018).

Various surveys on Seyfert galaxies and quasars (e.g. Porquet et al. 2004; Piconcelli et al. 2005; Bianchi et al. 2009; Scott & Stewart 2014; Marchesi et al. 2016; Williams, Gliozzi & Rudzinsky 2018; Chartas et al. 2021) established the diverse nature of the primary continuum slope in AGNs. The vast majority of objects studied in the above samples are type 1 sources, hence they provide a reliable measure of their intrinsic spectral shape due to the general lack of obscuration. These studies show that ∼80 per cent of AGNs are characterized by an intrinsic slope distribution ranging between $\Gamma = 1.6$ and 2.4 and peaking at $\Gamma \sim 2$.

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

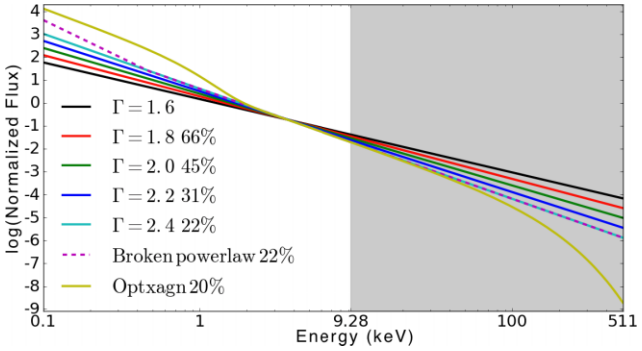

Figure 2. Seven input SEDs, calculated in the 0.1–511 keV range, corresponding to five power laws with slope in the range $\Gamma = 1.6$–2.4 and two more physically motivated input spectra, such as a double broken power law and the disc-corona Comptonization model `optxagnf` (see text). The spectra have been normalized to unity in the 2–10 keV band for comparison purposes. The shaded area indicates the energy band above the ionization threshold of Fe XXVI, at $E = 9.28$ keV.

A power-law SED is assumed to be a reasonable first-order approximation of the intrinsic X-ray continuum of AGNs, but in reality we know it to be much more complex. In Fig. 2, we show seven possible input spectra that correspond to five power laws with $\Gamma =$ 1.6–2.4 along with two more complex SED models, such as a broken power law and `optxagnf`[5] (Done et al. 2012), where the integrated 2–10 keV fluxes of the input spectra are normalized to unity. In this plot, the fraction of luminosity radiated above $E_{\rm lab} = 9.28$ keV (i.e. the ionization threshold of Fe XXVI, shaded area) compared to the hardest ($\Gamma = 1.6$) power law is calculated for each of the input continua. The percentages of the integrated photon flux in the 9.28–511 keV band, corresponding to each of the seven input spectra, are also noted in Fig. 2. As the input spectrum becomes steeper, the number of photons above $E_{\rm lab} = 9.28$ keV decreases, leading to a lower mean charge of iron within the flow. On the other hand, harder spectra would induce a higher ionization of the gas, possibly overionizing iron for its K-shell to be significantly populated.

In Fig. 3, we show the output spectra corresponding to the different $\Gamma =$ 1.6–2.4 in Fig. 2, which illustrate how a change in ionization affects the spectra. Note that the `optxagnf` and broken power-law continuum, which both adopted a $\Gamma = 2.4$ photon index at hard X-rays, produced a very similar Fe K absorption line depth as per the corresponding simple power-law case. In other words, the cases with a more complex continuum (`optxagnf`, broken power law) produced consistent results compared to the equivalent power-law case ($\Gamma = 2.4$). Subsequently, to generate our `slow64` and `fast32` tables, we choose a power-law SED with a photon-index range of $\Gamma =$ 1.6–2.4 with five linear steps of $\Delta\Gamma = 0.2$ between 0.1 and 511 keV. The above results in Fig. 3 suggest that the strongest lines from disc winds should occur in steep spectrum X-ray sources. For the case of the simulations in Fig. 3, the equivalent width of the Fe XXVI line increases fourfold from $\Gamma = 1.6$ (EW $\sim$ 110 eV) to $\Gamma = 2.4$ (EW $\sim$ 420 eV). This could be the case observationally, where strong (EW $\gtrsim$ 100 eV) blueshifted Fe K absorption lines are apparent in AGNs with steep ($\Gamma > 2$) photon indices or when they

[5] `optxagnf` is a self-consistent Comptonized disc emission model in XSPEC, and it was adopted in generating XSTAR (Bautista & Kallman 2001; Kallman et al. 2004) tables for PDS 456 (see section 4.2 in Matzeu et al. 2016, for more details). Note that in this exercise we adopted a $\Gamma = 2.4$ and a hot coronal temperature of $kT_{\rm e} = 100$ keV.

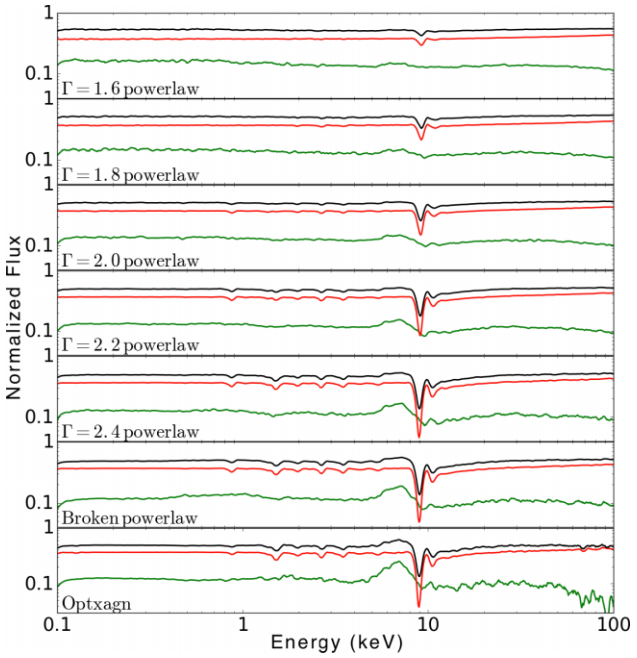

Figure 3. The total (black), direct (red), and scattered (green) simulated wind spectra normalized to their corresponding input SED from Fig. 2. The blueshifted absorption profile at $E \sim 9$ keV increases in strength and shifts to lower energies as the input SED becomes steeper. On the other hand, the hardening of the spectra would lead to an increase of ionizing photons that would eventually overionize the wind. As a consequence the overionized material would lose its opacity, which translates into a shallow absorption feature as shown in the top panel. For this example, the disc-wind simulations were carried out by assuming a 2–10 keV luminosity of 1.24 per cent of $L_{\rm Edd}$, an outflow rate of $\sim$40 per cent of $\dot{M}_{\rm Edd}$, $f_{\rm v} = 1.25$ ($v_{\rm w} = 0.3125c$ for $R_{\rm min} = 32\,r_{\rm g}$) and an inclination of $\mu = 0.675$.

are intrinsically X-ray weak (low $\mathcal{L}_{\rm X}$), e.g. PDS 456 (Reeves et al. 2021), PG 1211+143 (Pounds et al. 2003), and IRAS 13224−3809 (Parker et al. 2017; Pinto et al. 2018).

3 ARTIFICIAL NEURAL EMULATOR

ANNs are machine learning algorithms consisting of a set of neurons organized into layers. Each neuron is a distinct mathematical operation. They take an input $\boldsymbol{x}$and apply an affine transformation followed by a threshold function a, known as the activation function, to ensure the mathematical operation is non-linear. This then allows several neurons to be applied sequentially, thus forming a network. If the network is fully connected, then the output of each neuron in a given layer becomes the input to every neuron in the next layer:

$$x_m^n = a(\boldsymbol{x}^{n-1} \cdot \boldsymbol{W}_m + \boldsymbol{b}_m), \tag{2}$$

where $\boldsymbol{x}^{n-1}$ is the input to the nth layer and x_m^n is the mth neuron in the nth layer. $\boldsymbol{W}$ and $\boldsymbol{b}$ are the trainable weight and bias (i.e. analogue role to a constant value in a linear function) parameters that are updated during the training phase of the model.

The universal approximation theorem (Hornik, Stinchcombe & White 1989) states that ANNs with just a single layer can approximate any continuous function with a finite number of neurons. Here we train a simple feed-forward neural network (FFNN; Bebis & Georgiopoulos 1994) to map physical parameters to simulated disc-wind spectra (y), using both the `fast32` and `slow64` disc winds.

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

The inputs to the first layer are the parameters describing the AGN spectra $\boldsymbol{x}^0 = \{\Gamma,\ \dot{\mathcal{M}}_{\rm w},\ f_{\rm v},\ \mathcal{L}_{\rm X},\ \mu,\ R_{\rm in}\}$ and the outputs of the final layer $\boldsymbol{x}^N$ are the predicted spectral values ($\hat{y}$). The trainable parameters ($\mathcal{NP}$) of the network are updated to optimize the loss function by comparing the predicted spectral values with the true spectral values. We explored the use of various loss functions and we found that the mean square error loss function,

$$\mathrm{L2} = \sum_i (y_i - \hat{y}_i)^2, \tag{3}$$

was most suited to this problem, as it is simple to compute and sensitive to outliers: an important characteristic to ensure absorption and emission lines are conserved. Furthermore, we experimented with the use of various activation functions:[6] linear, tanh, Exponential Linear Unit (ELU), and sigmoid (see e.g. Nwankpa et al. 2018). Additionally, we tested the activation functions outlined in Alsing et al. (2020), which was developed specifically to reproduce well both to smooth and sharp features – again, an important feature for spectra. Nevertheless, we found that these activation functions underperformed compared to the Rectified Linear Unit (ReLU) activation function on our data set,

$$a(x) = \max(0, x). \tag{4}$$

This activation ensures that the outputs are positive, which is a key requirement for spectra. Under this same constraint, it is not possible to fit spectra in log units, where values can be $\leq$0. In the case of log spectra, the network would have to be redesigned with some other activation function such as tanh, and/or a linear final activation function. Our emulator network consists only of fully connected layers, the best of which used three dense layers, each with 1000 neurons (Fig. 4). In ANN a dense (or hidden) layer is located between inputs and outputs of the algorithm and performs non-linear transformations (i.e. fitting complex data) of the inputs and directs them into the outputs. They are referred to as dense (or hidden) because the 'true' values of their neurons are unknown. In total, this results in $\mathcal{NP} = 2009\,000$ trainable parameters (see Appendix B for the derivation).

The network was trained over 1500 epochs, where each epoch comprises the entire data cycle, however for improved efficiency (i.e. not to feed the data at the same time), the parameters of the network are updated in batches. For the training, we use an incremental batch size updating from 1 to 100, to 1000 at each 500 epoch interval. The batch size is the size of the training data subsample that is used to optimize the weights at a time. Larger batch sizes require more memory to load that can result in slower training, but smaller batch sizes give more stochastic loss that can will also take longer for the network to reach global minima. We use an increasing batch size that is equivalent to decreasing the learning rate (Smith et al. 2017). The learning rate is another hyperparameter that determines the size of the changes made to weights at each step. This aids the network in reaching the minimum loss, because as you approach the minima you need to make smaller changes to the weights or you will overshoot. Similarly, slowly increasing the batch size provides more confidence in the direction of descent to the minima as opposed to the stochastic descent provided by a small batch size. Additionally, early stopping (Yao, Rosasco & Caponnetto 2007) is implemented to prevent overfitting. This ends the training process once the model is no longer improving. We use the adaptive optimizer Adam (Kingma & Ba 2014) to update the weights with learning rate of 10^{-3}.

In addition to training data that are used to optimize the network, additional data are required to validate the model, to ensure that it will generalize to new data. These data are seen during the training of the network to determine when to stop training. The performance of the trained network was then evaluated on additional test data that are *not seen* during the training of the network. In total we have 172 800 $\mathcal{MCRT}$ synthetic spectra available for the ANN and we choose a train–validation–test split of 0.8–0.1–0.1. This equates to 138 240 spectra for training, 17 280 for validation and 17 280 for testing. The training set was checked to ensure a good representation of all parameters was included. The final L1 (absolute error),

$$\mathrm{L1} = \sum_i |y_i - \hat{y}_i|, \tag{5}$$

and L2 (mean square error) loss on the validation data was 0.0071 and 0.0002, respectively. The L1 and L2 statistics for the test data set are 0.0071 and 0.0001, respectively. Fig. 5 shows some examples of the spectra predicted by the emulator from the test data set.

3.1 Mitigating interpolation issues with emulation

Until now, the data used to train and test the network are $\mathcal{MCRT}$ simulations from two grids of parameters. But we need to know if the emulator is capable of reproducing parameter values between all the grid points. To do this, the trained network is further tested against 2000 new $\mathcal{MCRT}$ simulations, where 100 of each of the parameters are drawn from uniform distributions: $\Gamma \sim \mathcal{U}(1.6, 2.4)$, $\dot{\mathcal{M}}_{\rm w} \sim \mathcal{U}(0.0196, 0.6762)$, $f_{\rm v} \sim \mathcal{U}(0.25, 2)$, $\mathcal{L}_{\rm X} \sim \mathcal{U}(2.52 \times 10^{-4}, 2.52 \times 10^{-2})$, and the launch radius from a binomial distribution $R_{\rm in} \sim \mathcal{B}(1, 0.5)$, corresponding to the `fast32` and `slow64` winds. For each spectrum, we have corresponding μ values of 0.025 to 0.975 in steps of 0.05. Fig. 6 shows the fractional offset of the predicted from the ground truth spectra,

$$\text{fractional error} = \left| \frac{\hat{y} - y}{y} \right|. \tag{6}$$

The fractional error is in most cases smaller than the noise on the simulated spectra, and we find no bias with respect to any particular parameter. Typical errors are of per cent level across the entire energy range (Fig. 7), although a non-negligible error is seen in the 7–8 keV band, corresponding to Fe XXV–XXVI transitions in both emission and absorption.

To investigate the influence of the fractional error in the 7–8 keV band on the parameters (Fig. 7), we take the fractional error on the flux values at from the test simulations at 8 keV. We order the error values and take the parameter set corresponding to the 70 per cent, 75 per cent, 80 per cent, 85 per cent, and 90 per cent error value as shown in Fig. 8. From these five parameters sets we emulate spectra using XRADE and create CCD observations, by using the *XMM–Newton* European Photon Imaging Camera (EPIC)-pn response and background files corresponding to the PDS 456 ObsCD observation in 2013 (see Section 4), between the 0.3 and 10 keV energy range (i.e. the *XMM–Newton* bandpass), using XSPEC. The observation is then fit using the $\mathcal{MCRT}$ tables. The parameters are generally well recovered despite the differences between the $\mathcal{MCRT}$ tables and XRADE. We find that the recovery of $\boldsymbol{\mu}$ is the only parameter that is affected by the uncertainty on the Fe K bandpass. This parameter is the one that affects the shape of the spectral features the most (see Section A1 in Appendix A).

[6]The activation function is a mathematical function that is added to an ANN in order to ensure non-linearity. In this way the ANN can learn the complex patterns of the training data 'fed' into it.

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

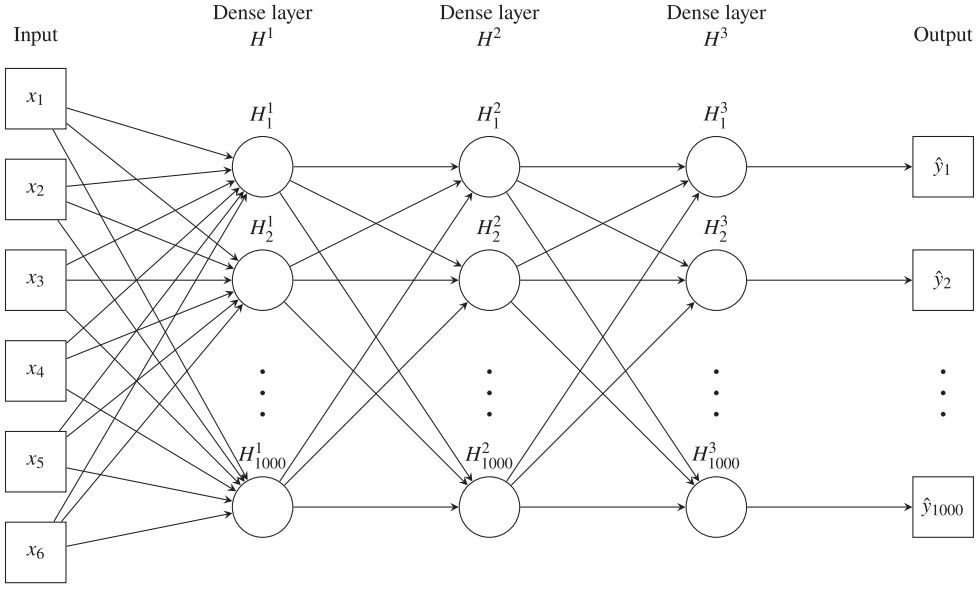

Figure 4. The FFNN architecture of the emulator constructed for XRADE. Each circle represents a neuron. Each connection is marked as a right arrow.

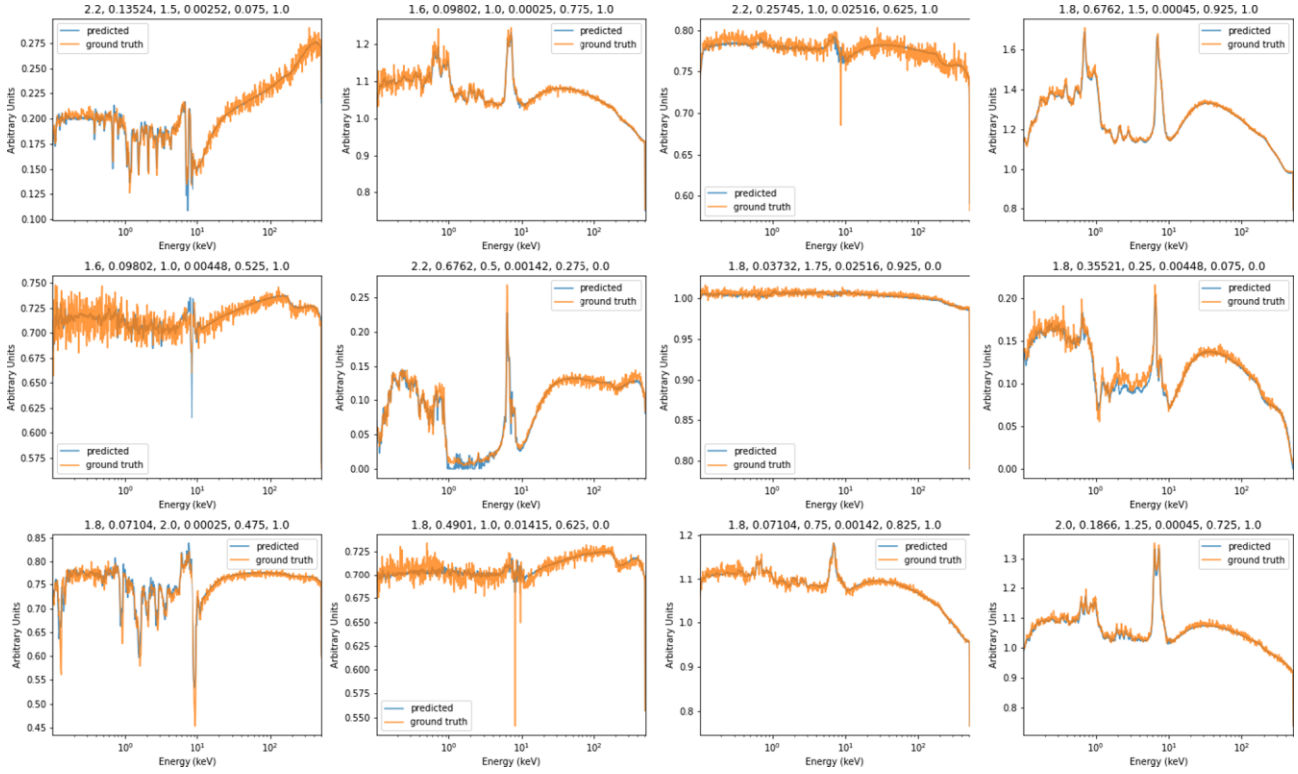

Figure 5. Randomly chosen examples of 12 ground truth spectra (orange) out of the 17 280 available from the test data (10 per cent of total) with the corresponding emulated predicted spectra (blue). Note the ANN had not seen the spectra from the test set (10 per cent of the total) during training. The input parameters corresponding to $\{\Gamma, \dot{\mathcal{M}}_{\rm w}, f_{\rm v}, \mathcal{L}_{\rm X}, \mu, R_{\rm min} = 64\,r_{\rm g}\,(0.0)$ or $32\,r_{\rm g}\,(1.0)\}$ are listed above each plot.

The fractional error seems to increase, not as severely, at energies corresponding to other wind features e.g. O VIII, Ne X, and Si XIV. The network could benefit from training data with more spectral points in these regions, and ideally a more finely sampled grid of parameters (see Section 5). In particular, in the near future, we also aim to regenerate steps for the `fast32` and `slow64` grids in linear (rather than log) space for the $\mathcal{L}_{\rm X}$ and $\dot{\mathcal{M}}_{\rm w}$ parameters. This will likely increase the accuracy of mapping these parameters through the emulator and this could be especially important in training the emulator at the higher $\dot{\mathcal{M}}_{\rm w}$ range, which is currently more sparsely sampled in logarithmic space. As a consequence, this may also reduce the fractional error seen over the Fe K bandpass in Fig. 7.

This test demonstrates that not only are we able to use the emulator for parameter values within the training range, but also on parameter

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

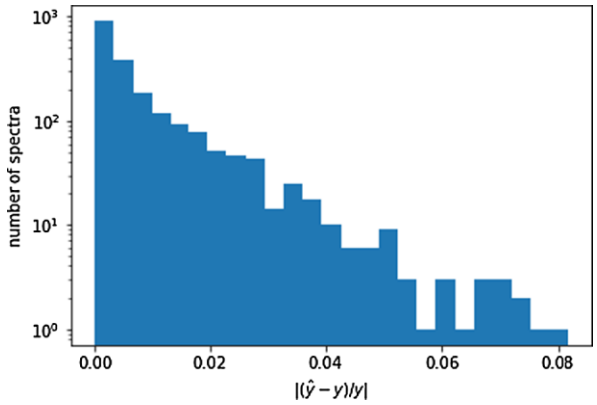

Figure 6. Histogram of absolute fractional error of the emulated spectra on the 2000 additional simulations.

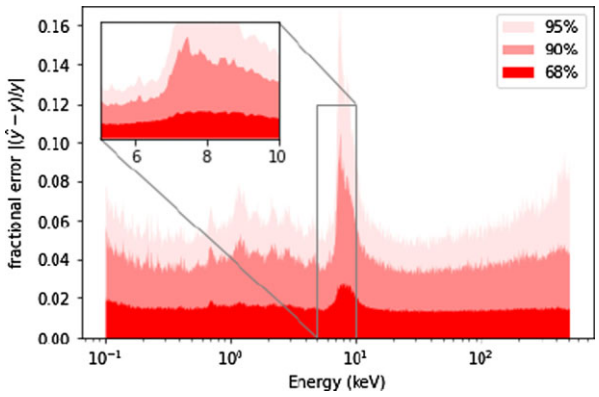

Figure 7. Fractional error on flux at different energies for the 2000 additional test simulations from 0.1 to 511 keV. We show 68 per cent, 90 per cent, and 95 per cent of the sample.

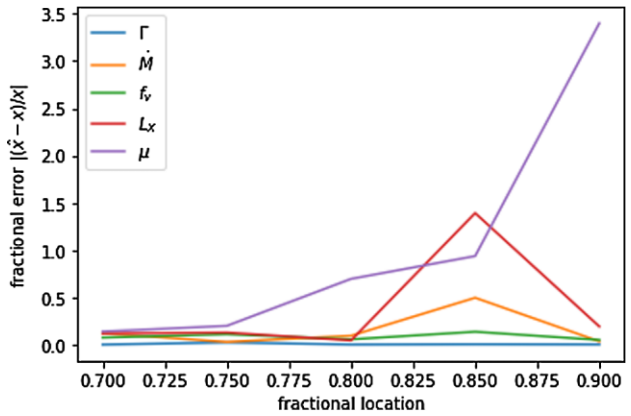

Figure 8. The fractional error on individual parameters fit using the $\mathcal{MCRT}$ tables based on five different spectra generated from XRADE with increasing fractional error at the 8 keV band (x-axis).

values that lie between points on the simulated grid. The trained emulator can predict spectra for a particular parameter set in ∼0.04 s in comparison to ∼2–3 h when using the $\mathcal{MCRT}$ pipeline, which allows us to emulate finer grids of models more efficiently. In this light, we test whether the predictions of our emulation process are able to reproduce a true (or ground) spectrum. We then compare them with spectra arising from standard interpolation between grid values, which are normally occurring in X-ray fitting packages such as XSPEC. For this test, a true spectrum can be selected from any of the 17 280 available test spectra. We chose two test cases in Fig. 9, one (upper panels) where there is one free wind parameter (μ) and one where two parameters (μ and f_v) are varied (lower panels).

For the 1D test we considered the case where the true spectrum has the following parameters: $\boldsymbol{x}_{\rm true} = [\Gamma = 2.0,\ \dot{\mathcal{M}}_{\rm w} = 0.257,\ f_{\rm v} = 1.0,\ \mathcal{L}_{\rm X} = 2.52 \times 10^{-4},\ \mu = 0.625,\ R_{\rm min} = 64\,r_{\rm g}]$. Here the 1D (Fig. 9, panel A, upper-left) interpolation (blue) between $\mu = 0.575$ and 0.675 (in order to reproduce a real value of $\mu = 0.625$). Here the interpolation does a reasonable job in reproducing the true spectrum with the following conditions: $\boldsymbol{x}_{\rm true} = [\Gamma = 2.0,\ \dot{\mathcal{M}}_{\rm w} = 0.257,\ f_{\rm v} = 1.0,\ \mathcal{L}_{\rm X} = 2.52 \times 10^{-4},\ \mu = 0.625,\ R_{\rm min} = 64\,r_{\rm g}]$ (orange), but underestimates the profile depth. The emulated true spectrum plotted in panel (B) (blue) is better at reproducing the depth of the absorption trough at ∼8 keV, but slightly worse at estimating the higher order transition at ∼8.3 keV. Overall, both methods reasonably predict the true spectrum between 5 and 12 keV.

In panels (C) and (D) we compare a 2D interpolation (i.e. two parameters of interest) between $\mu = (0.575, 0.675)$ and $f_{\rm v} = (0.75, 1.25)$ to respectively reproduce $\mu = 0.625$ and $f_{\rm v} = 1.0$. In this scenario it is much harder for the interpolated spectrum to reproduce $\boldsymbol{x}_{\rm true}$ as these parameters produce both a shift in energy and depth simultaneously. Clearly the interpolated spectrum fails to reproduce the true spectrum. On the other hand the emulated spectrum is a closer match to $\boldsymbol{x}_{\rm true}$ (panel D). It is worth noting that any interpolation issues in $\dot{\mathcal{M}}_{\rm w}$ nor $\mathcal{L}_{\rm X}$ are not as dramatic, given that they are mainly affecting the profile depth and do not tend to produce a shift in energy between the points.

4 OBSERVATIONAL DATA: FITTING THE POWERFUL DISC WIND IN PDS 456 WITH XRADE AND FAST32

The generated XRADE spectra are tabulated into FITS files and can be used as multiplicative grids within XSPEC. In this section, we want to compare the overall performance and reliability consistency check of our $\mathcal{MCRT}$ and XRADE tables with real CCD data. we also want to check and compare them in the ability predicting values between the grid points.

As a test case we consider the 'prototypical' (and most studied) disc wind hosted in the luminous quasar PDS 456. A large monitoring campaign, covering 6 months, was carried out between 2013 and 2014 and consisted of five joint *XMM–Newton* and *NuSTAR* observations (ObsA–ObsE) of ∼100 ks each. During these observations, a prominent and persistent P Cygni profile was revealed (N15). Such a feature is characterized by the combination of a broad emission and absorption profile, where the former is produced by scattered photons off the wind averaged from all angles and the latter from transmitted photons through the material. ObsC and ObsD were separated by only ∼3 d, so their spectra were virtually identical. As per N15, we subsequently combined them into a single ObsCD observation resulting into a total net exposure time of 195 ks, showing a P Cygni feature of unprecedented quality. The *XMM–Newton* and *NuSTAR* data considered here are the EPIC-pn (Strüder et al. 2001) and FPMA+FPMB (Harrison et al. 2013), respectively, and they are reduced following the procedure presented in N15.

From what was discussed in Section A2 of Appendix A, the initial setting of $R_{\rm min}$ has a direct impact on the range of outflow velocities that can be measured (see equation 1). In this paper, we chose `fast32` for our comparison with XRADE. Note that `fast32` was initially generated based on the range of velocities observed in

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

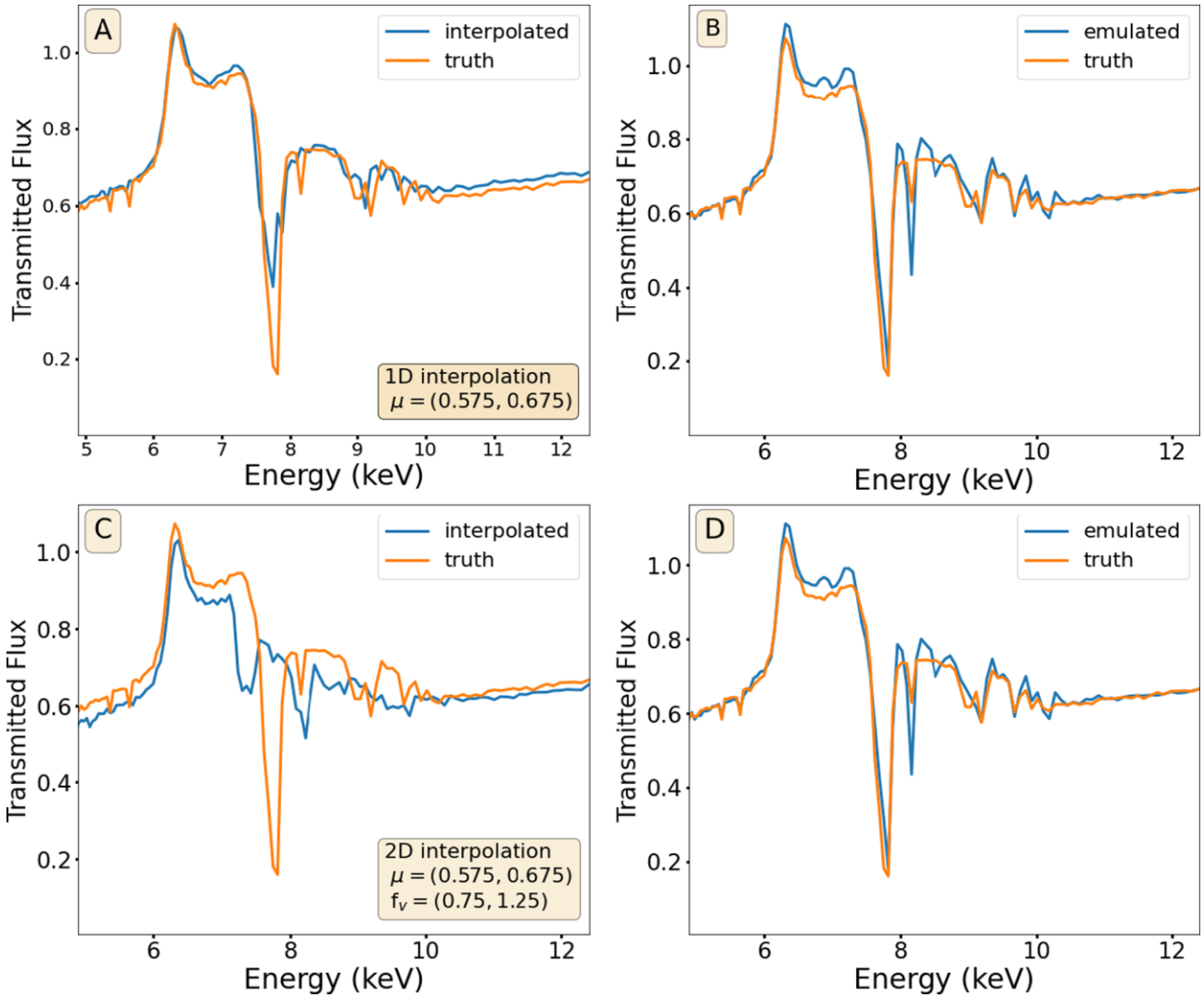

Figure 9. A comparison between interpolated spectra, normally produced in XSPEC, and those predicted by our trained ANN. Top row: in panel (A) we show the true spectrum of $\boldsymbol{x}_{\rm true} = [\Gamma = 2.0,\ \dot{\mathcal{M}}_{\rm w} = 0.257,\ f_{\rm v} = 1.0,\ \mathcal{L}_{\rm X} = 2.52 \times 10^{-4},\ \mu = 0.625,\ R_{\rm min} = 64\, r_{\rm g}]$ (orange) and the spectral prediction using interpolation between $\mu = (0.675$ and $0.575)$. Same in panel (B), but the true spectrum is predicted by the emulator (blue). Panel (C): same as panel (A) but with a 2D parameter interpolation involving μ and $f_{\rm v}$ (see text box). The resulting interpolated spectrum is not able to accurately recover the amplitude and shape of the spectral lines. Panel (D): the emulator is able to produce an accurate mapping of the ground truth spectrum.

PDS 456 (e.g. $v_{\rm out}/c = 0.25$–0.35; Matzeu et al. 2017; Reeves et al. 2018a) since its first detection with *XMM–Newton* in 2001 (Reeves et al. 2003). On the other hand, by following the same prescription in Section A2, `slow64` was successfully applied in modelling the powerful disc wind observed in the Seyfert 2 galaxy MCG-03-58-007 (Braito et al. 2022).

In Fig. 10, we show the unfolded *XMM–Newton* and *NuSTAR* spectra of PDS 456 (ObsCD) between 2 and 40 keV against a simple $\Gamma = 2$ power law. Once the continuum (cyan) is accounted for, there are strong residuals in the Fe K region that correspond to the P Cygni feature. From a visual inspection the centroid energies are located at $E_{\rm rest,em} \sim 7$ keV and $E_{\rm rest,abs} \sim 9$ keV for the emission and absorption component, respectively. The model in XSPEC is expressed as

`Tbabs` × `pcfabs` × (`powerlaw` × `highecut`) × XRADE (or `fast32`), (7)

where `Tbabs` is the Galactic absorption of $N_{\rm H}^{\rm Gal} = 2.9 \times 10^{21}$ cm^{-2} (Reeves et al. 2021). To model the soft X-ray spectral curvature we adopt a layer of neutral partial covering (`pcfabs` in XSPEC) with $N_{\rm H} = 7.9^{+2.1}_{-2.5} \times 10^{22}$ cm^{-2}, and covering fraction of $\mathcal{C}_{\rm frac} = 0.37^{+0.05}_{-0.02}$. A high-energy rollover (`highecut`) fixed at $E_{\rm cut} = 100$ keV was also adopted and a cross-normalization factor between the *XMM–Newton* and *NuSTAR* detectors was measure at $\mathcal{C}_{\rm cal} = 1.10 \pm 0.02$. For this test, we generated a customized XRADE grid with the values tabulated in Table 2.

Fitting the P Cygni profile with XRADE yielded a mass outflow rate of $\dot{\mathcal{M}}_{\rm w} = 0.318^{+0.014}_{-0.046}$, i.e. about 30 per cent of $\dot{M}_{\rm Edd}$. In PDS 456, with $M_{\rm BH} \sim 10^9\,{\rm M}_\odot$ and $L_{\rm Edd} \sim 1.3 \times 10^{47}$ erg s^{-1}, then $\dot{M}_{\rm Edd} \sim 40\ {\rm M}_\odot\,{\rm yr}^{-1}$ for $\eta = 0.06$, $\dot{M}_{\rm out} \sim 10\ {\rm M}_\odot\,{\rm yr}^{-1}$. The X-ray ionizing luminosity is $\mathcal{L}_{\rm X} = 0.272^{+0.090}_{-0.061} \times 10^{-2}$ or ∼0.3 per cent of $L_{\rm Edd}$, i.e. $L_{2-10\,{\rm keV}} \sim 4.0 \times 10^{44}$ erg s^{-1}. By comparison, the directly observed intrinsic 2–10 keV luminosity is of the order of $\sim 5 \times 10^{44}$ erg s^{-1} and hence consistent with the XRADE predicted value, see Table 3. A line-of-sight orientation angle of $\theta \sim 50^\circ$ (i.e. $\mu = 0.63^{+0.01}_{-0.02}$) with

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

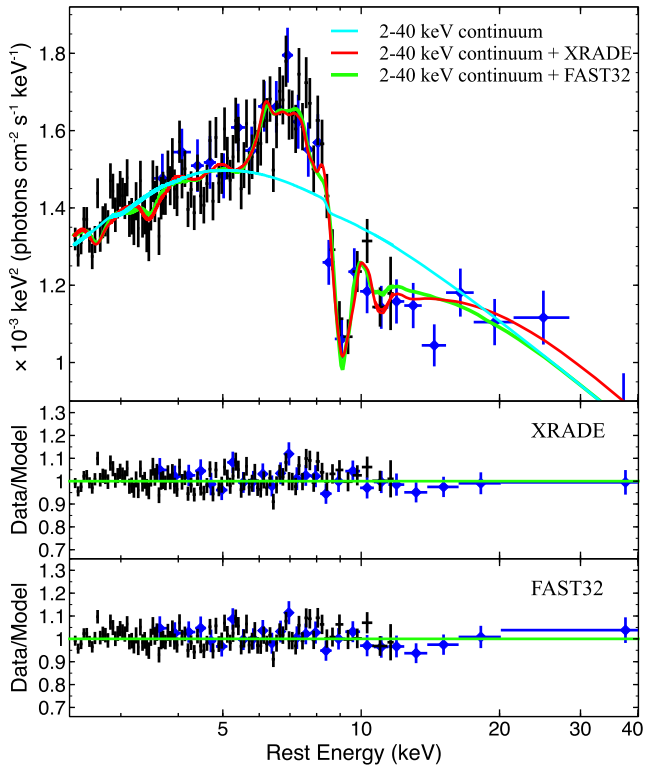

Figure 10. Top: unfolded *XMM–Newton* (black) and *NuSTAR* (blue) spectra (ObsCD) of PDS 456 (against a $\Gamma = 2$ power law) between 2 and 40 keV with the continuum-only model (cyan), XRADE (red), and `fast32` (green) superimposed. Bottom: the corresponding data/model ratio plots. Both models do an excellent job in fitting the P Cygni feature, of which they self-consistently fit the broad emission (scattered/reflected component) and absorption (direct component).

Table 2. Customized XRADE model values and ranges used for PDS 456 ObsCD.

Parameter	Value range	Δ value	Steps
Γ	1.6–2.4	0.1	9
$\dot{\mathcal{M}}_{\rm w}$	0.05–0.65	0.05	13
$f_{\rm v}$	0.25–2.0	0.097	19
$\mathcal{L}_{\rm X}$	$(0.05–1.5) \times 10^{-2}$	9.7×10^{-4}	16
μ	0.2–0.9	0.05	15
$R_{\rm min}$	$32\,r_{\rm g}$	–	–

Note. Number of emulated spectra: 533 520.

Table 3. XRADE and `fast32` model results for PDS 456 ObsCD. The uncertainties are calculated at a 90 per cent confidence level that correspond to a $\Delta\chi^2/\Delta\nu = 2.71/1$. The power-law normalization is in unit of photons $\rm keV^{-1}\,cm^{-2}\,s^{-1}$ at 1 keV.

Parameter	XRADE	`fast32`
Γ	$2.19^{+0.02}_{-0.05}$	$2.21^{+0.05}_{-0.04}$
$\dot{\mathcal{M}}_{\rm w}$	$0.318^{+0.014}_{-0.046}$	$0.243^{+0.056}_{-0.131}$
$f_{\rm v}$ (v_∞)	$1.33^{+0.03}_{-0.04}$ $\left(-0.33^{+0.01}_{-0.01}c\right)$	$1.31^{+0.08}_{-0.04}$ $\left(-0.33^{+0.02}_{-0.01}c\right)$
$\mathcal{L}_{\rm X}$	$0.273^{+0.090}_{-0.061} \times 10^{-2}$	$0.208^{+0.058}_{-0.130} \times 10^{-2}$
$\mu = \cos\theta$	$0.629^{+0.010}_{-0.017}$	$0.652^{+0.008}_{-0.022}$
$\log(N_{\rm H}/{\rm cm}^{-2})$	$22.8^{+0.4}_{-0.4}$	$22.9^{+0.4}_{-0.4}$
$C_{\rm frac}$	$0.39^{+0.04}_{-0.04}$	$0.37^{+0.03}_{-0.02}$
norm/10^{-3}	$3.2^{+0.05}_{-0.3}$	$2.7^{+0.3}_{-0.2}$
$C_{\rm cal}$	$1.10^{+0.02}_{-0.02}$	$1.10^{+0.02}_{-0.02}$
$\Delta\chi^2/\Delta\nu$	655.3/682	659.7/682

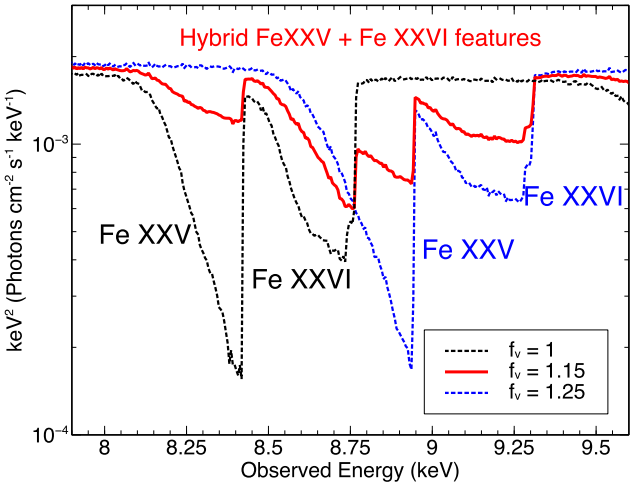

Figure 11. Three simulated microcalorimeter resolution (i.e. 2 eV at 6.4 keV) spectra, using a high-resolution `f32hires` grid, corresponding to $f_{\rm v} = 1$ (black) and $f_{\rm v} = 1.25$ (blue). The remaining disc-wind parameters are fixed to $\Gamma = 2$, $\dot{\mathcal{M}}_{\rm w} = 0.3$, $\mathcal{L}_{\rm X} = 0.5$, and $\mu = 0.625$. The interpolated spectrum between $f_{\rm v} = 1$ and $f_{\rm v} = 1.25$ corresponding to $f_{\rm v} = 1.15$ is shown in red. Note the asymmetry of the profiles visible at high resolution is a direct result of the acceleration of the gas along the streamline.

respect to the polar axis is required, suggesting that the sightline fully intercepts the innermost and fastest wind streamline, hence explaining the prominence (and high degree of blueshift) of the P Cygni feature. The terminal velocity parameter was measured at $f_{\rm v} = 1.33^{+0.03}_{-0.04}$ and, as XRADE was generated by assuming a launch radius of $R_{\rm min} = 32\,r_{\rm g}$, this translates into a terminal wind velocity of $v_\infty = -0.33^{+0.01}_{-0.01}c$.

Note that the input photon index of the XRADE model is tied to the `powerlaw` continuum at $\Gamma = 2.19^{+0.05}_{-0.02}$. The addition of XRADE resulted in a large improvement on the fit statistics by $\Delta\chi^2/\Delta\nu = -231.4/4$ (> 99.99 per cent), for an overall best-fitting $\chi^2/\nu = 655.3/682$. We subsequently replaced XRADE with our $\mathcal{MCRT}$ generated `fast32` in equation (7). We find that both fits are excellent and almost identical with $\chi^2/\nu = 659.7/682$ (see Fig. 10, bottom right). During the fitting procedure in XSPEC, the 'delta' value parameter has been set to be 0.001 (via the XSET command) so that a like-for-like comparison could have been achieved between XRADE and `fast32`. Moreover, the same best-fitting values were returned when restoring the original fixed delta values of the model (i.e. via the command XSET delta 0.0)

The values are largely consistent with XRADE, as shown in Table 3. This initial consistency test demonstrates that both physical models provide an excellent fit to the P Cygni-like profile in PDS 456 and that XRADE is able to reproduce the results obtained by the $\mathcal{MCRT}$ grid. Note that errors measured in both grids are indeed similar due to CCD spectral resolution of the data that illustrates that, at the resolution of the data, XSPEC interpolation upon the $\mathcal{MCRT}$ table models achieves an equally adequate parametrization of the data as per the emulated XRADE tables. However the limitations of the former and over-reliance of interpolation is more likely to have a significant impact for calorimeter resolution spectra, which we further discuss below.

In Fig. 11, we show three simulated *Athena*/X-IFU resolution (i.e. 2 eV at 6.4 keV) spectra using the high-resolution disc-wind grid (`f32hires`; see Parker et al. 2022 for details). `f32hires` was a $\mathcal{MCRT}$ generated table to match the microcalorimeter resolution data of *XRISM*/Resolve and *Athena*/X-IFU with a total of 10 000

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

energy bins (i.e. with an energy resolution of $\Delta E = 1.8$ eV) between 0.1 and 20 keV. Because of its high CPU cost, `f32hires` is in a preliminary stage and is limited to 2400 grid points, however it will be expanded in the near future.

Here we keep all the parameters fixed (see caption) whilst the changing the velocity factor parameter to $f_v = 1$ (black), $f_v = 1.25$ (blue) and the interpolated value of $f_v = 1.15$ (red) between the former two grid points. As expected, both the highly ionized (i.e. Fe XXV Heα and FeXXVI Lyα) absorption features are prominent in both $f_v = 1$ and $f_v = 1.25$ spectrum, although more blueshifted in the latter. The intermediate (interpolated) point seems to generate a spectrum that is characterized by some hybrid set of absorption feature caused by interpolation. The intermediate (interpolated) spectrum at $f_v = 1.15$ is characterized by a hybrid set of absorption features caused by interpolation in energy space, between the $f_v = 1$ and $f_v = 1.25$ grid points. In fact such an issue is already striking, unlike in the CCD resolution framework (see Fig. 9), in the simplest 1D interpolation discussed in Section 3.1. A more detailed set of experiments will be performed and reported on a following companion paper.

At this stage, the key contrast between these two tables is the vast difference in the CPU time required to generate these grids. In fact, to produce the 86 400 synthetic spectra in `fast32` required an overall CPU time of ∼4 months on 600 cores at 50 GB RAM (per core), against an impressive time-scale of ∼4 s for generating 533 520 emulated spectra for the XRADE table. Note that our emulator has the flexibility to generate parameter ranges with unprecedented resolutions within minutes.

4.1 Global parameter exploration

We sought to test the emulated parameter space created with XRADE via global parameter exploration. We use the same *XMM–Newton*/EPIC-pn and *NuSTAR* PDS 456 data sets as described in Section 4 and an identical model set-up. For the purposes of comparing the different parameter spaces, we use the `fast32` and XRADE table model as in equation (7) (see Table 2). We employ the Bayesian X-ray Analysis (BXA v2.10; Buchner et al. 2014) software platform that connects the nested sampling algorithm `MultiNest` (Feroz, Hobson & Bridges 2009) with the XSPEC fitting environment. In brief, nested sampling (see Buchner 2021 for a recent review) stores a set of parameter vectors drawn from the prior distribution. The lowest likelihood parameter vector is iteratively replaced with a new one of higher likelihood, until some termination condition is met. In this way, the algorithm scans the global prior-defined parameter space and is thus a useful tool for visually exploring and comparing the multidimensional parameter spaces associated with `fast32` and XRADE.

We assign uniform priors to all parameters apart from the partial covering absorber column density and intrinsic power-law normalization that were assigned log-uniform priors, and the multiplicative cross-calibration constant that was assigned a custom log-Gaussian prior with mean zero (i.e. a linear cross-calibration of unity) and 0.1 standard deviation. This choice of prior is useful for the cross-calibration to avoid negative values, whilst also peaking close to unity (e.g. Madsen et al. 2017). The same 10 free parameters were used in both models.

The result of the fits is shown in Fig. 12 with grey and blue contours for `fast32` and emulated XRADE tables, respectively. Shaded regions represent the 2σ level, though note that the percentage of points encompassed by the 2D contours is not the same as in the 1D histograms.[7] In general, the parameter space attained with XRADE appears to match the `fast32` parameter space well with good agreement within 2σ. The majority of individual posterior shapes also show good agreement, indicating that the emulation process is able to reliably map different regions of parameter space to spectral space.

There are some parameters that have different posterior shapes, e.g. μ. Disagreements between posterior shapes could indicate that particular regions of the emulated spectral/parameter space require more training data as input. Alternatively, even though both models were fit with XSPEC, the emulated parameter grid of XRADE was finer than `fast32`, hence with the corresponding interpolation between adjacent grid points performed over smaller parameter steps with XSPEC. We note that the higher resolution XRADE table does not necessarily mean that the confidence intervals should be smaller, since the aim of the emulator is to reproduce the multidimensional parameter space associated with the original `fast32` model as accurately as possible. The ultimate limitation to the confidence intervals is thus the data quality, since the emulated XRADE model was trained on `fast32` originally.

If the input training data was sufficient for the ANN to learn the complex mapping process involved, posterior differences could hint to alternative parameter estimation with emulation versus interpolation. However, since $\dot{\mathcal{M}}_w$, f_v, and μ have a very strong (and/or non-linear) relation to the observed spectral shape of the model, such parameters are most likely to suffer from interpolation issues, suggesting such parameters may require finer parameter resolution training grids in particular. None the less, testing future emulated XRADE tables on real data with BXA may be an efficient method to iteratively explore and check the emulated parameter space in detail.

Fig. 13 presents an alternative comparison between the spectral fits performed with BXA in Section 4.1. A total of 500 posterior parameter vectors from the `fast32` (left) and XRADE (right) model fits were loaded and overplotted with the unfolded spectral data. The models found for each data set (distinguished by the cross-calibration) are plotted with the same colour in each panel and shaded regions represent the overall 500 realizations. Clearly the spectral shapes are very similar apart from a small difference at ∼8 keV, in agreement with Fig. 7.

4.2 Other models

A model similar to XRADE is defined and used in Hagino et al. (2015) (MONACO – MONte Carlo simulation for Astrophysics and COsmology), which is then subsequently applied in Hagino et al. (2016, 2017) to fit the disc winds in PDS 456, 1H 0707−495, and APM 08279+5255. Here, the same biconical structure is used (see fig. 3 in Hagino et al. 2015). MONACO separates the wind structure into shells and then performs a series of XSTAR runs to ascertain the ionization balance and the luminosity leaving and entering each layer. The radiative transfer is then performed using the He-like and H-like iron and nickel transitions along with Compton scattering. This has the benefit of being less computationally expensive than our disc-wind code, as the higher the number of lines that are tracked, the more computationally intensive the simulation. Therefore, the limited number of transitions allows a quicker exploration of the parameter space. The argument for only tracking the highly ionized species is that high-velocity winds are typically highly ionized.

[7] See https://corner.readthedocs.io/en/latest/pages/sigmas.html

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

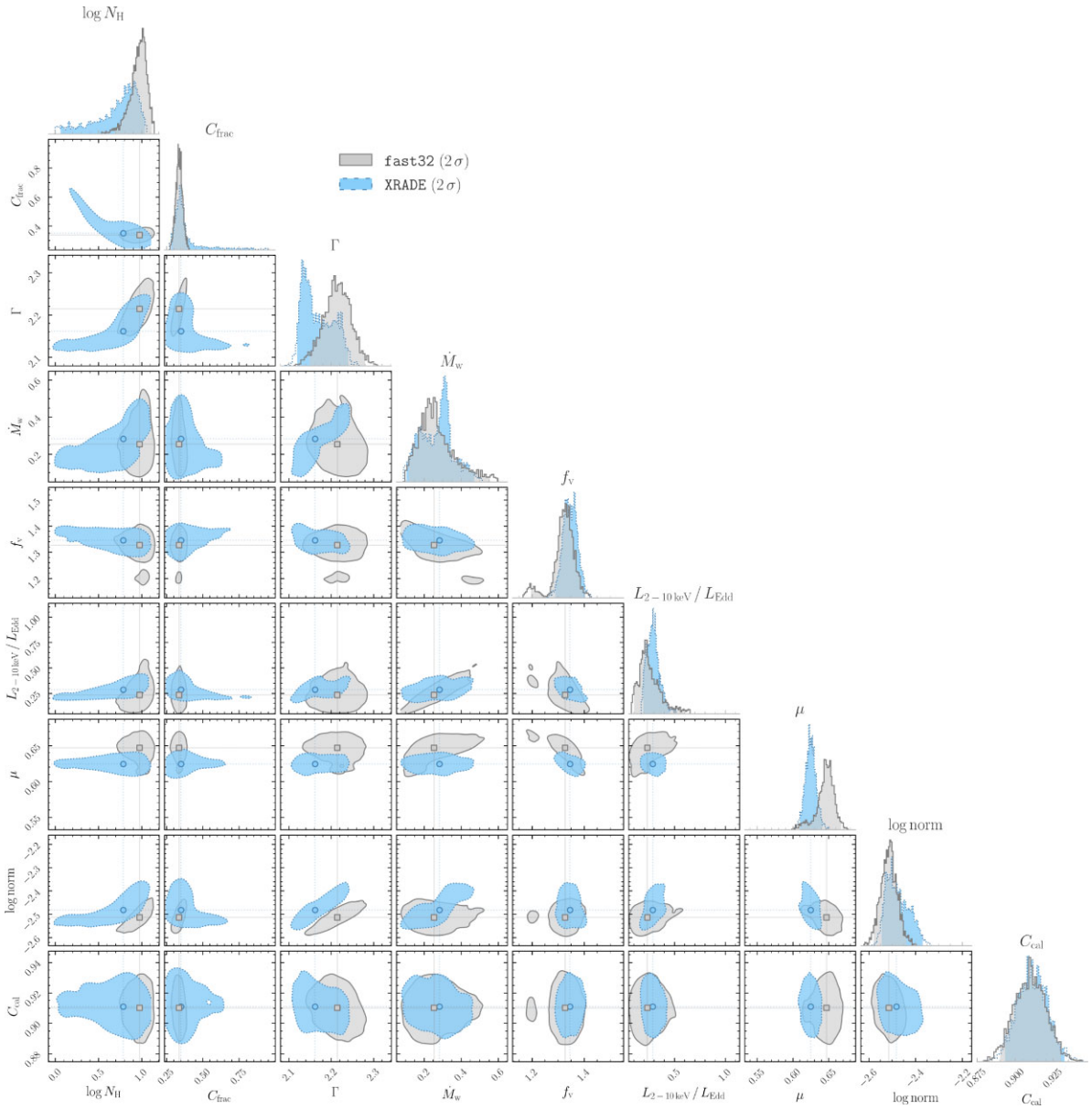

Figure 12. Corner plot showing the results obtained from fitting the *XMM–Newton* and *NuSTAR* spectra with `fast32` (grey) and XRADE (blue) using BXA. Shaded regions show the 2σ confidence level.

However, lower ionization species can survive in thicker winds and should be considered in a more general case. These lower ionization species may be observed at lower energies, such as the lower ionization lines observed in the *XMM–Newton* Reflection Grating Spectrometer (RGS) data of many AGNs. In PG 1211+143 (Pounds et al. 2016; Reeves, Lobban & Pounds 2018b) and PDS 456 (Reeves et al. 2016, 2020) these soft features appear to be physically associated with the highly ionized outflow. These features may be studied in more detail in the future by lowering the ionization in runs. This can be done by either lowering the source luminosity or increasing density through clumps within the streamlines.

It is thus important to stress that the faster winds will not just produce more highly blueshifted lines, but also produce intrinsically broader line profiles, both in emission and absorption. While in principle such profiles may be accounted for in other non-wind scenarios (e.g. by absorption through a corotating disc atmosphere; Gallo & Fabian 2011; Gallo et al. 2013; Fabian et al. 2020), in Section 4 we demonstrate that the broad P Cygni-like profile in PDS 456 can be self-consistently modelled by our solar abundance, `fast32` and XRADE table of models.

5 CONCLUSIONS AND FUTURE WORK

In this paper, we presented an improved version of the state-of-the-art disc-wind model obtained from a Monte Carlo multidimensional (2.5D) radiative transfer code initially developed by Sim et al. (2008, 2010). For this purpose, we generated two large $\mathcal{MCRT}$ tables, `slow64` and `fast32`, of 172 800 synthetic spectra, covering a much wider parameter space (see Table 1) than previously presented (S08; S10; Reeves et al. 2014; Reeves & Braito 2019). These will allow us to explore the physical conditions that characterize the accretion disc winds across a wide range of sources, as our measurements are black hole mass invariant. As mentioned above, `slow64` has been already

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

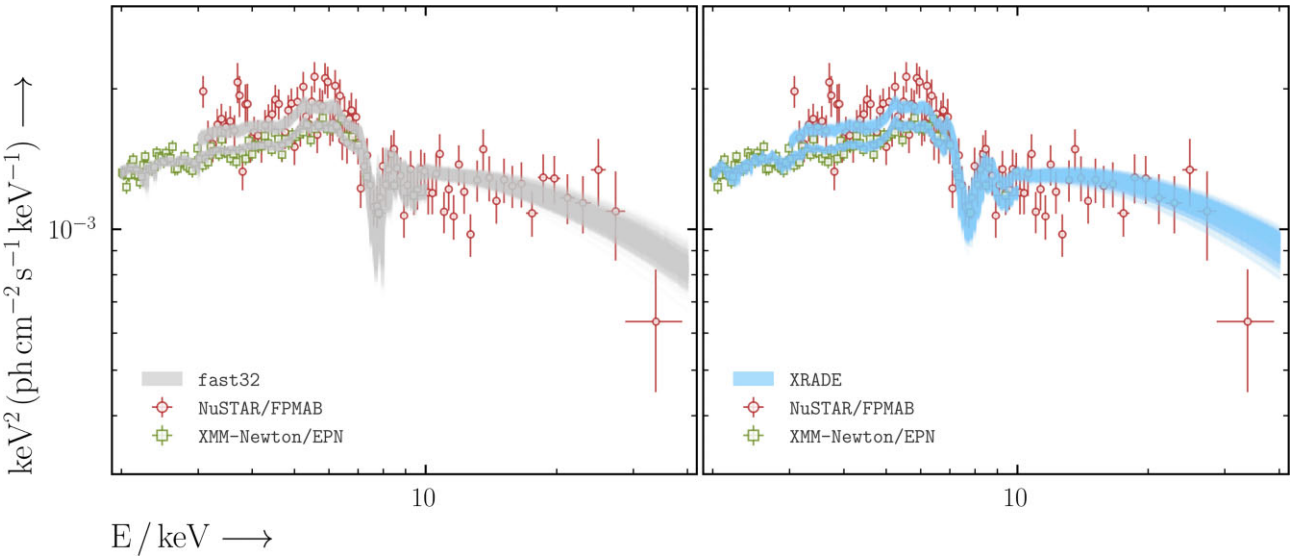

Figure 13. Unfolded model realizations from the BXA fits with the original `fast32` (grey) and XRADE (blue) multiplicative table models. The data adopted here are the same as in Fig. 10.

applied to MCG-03-58-007 (Braito et al. 2022), and the `fast32` will be applied to all the PDS 456 data from 2001 to 2019 (Reeves et al., in preparation).

We also presented the development and implementation of a novel emulator based on a purposely built ANN: X-Ray Accretion Disc-wind Emulator (XRADE). The method developed here works as follows. From the available $\mathcal{MCRT}$ generated spectra, we fed 80 per cent (or 138 240 spectra) into the ANN. A further 10 per cent (17 280) are used for validation and the remaining 10 per cent are exclusively kept for testing the emulated spectra. Our emulator is not only able to reproduce the `slow64` and `fast32` synthetic spectra, which required a total of ∼8 months (600 cores) to be generated, but also to emulate 533 520 spectra (see Table 2) within a 4 min time-scale, i.e. ∼5 orders of magnitude faster, with an average mean square error of just 1.4 per cent.

After the training and validation process, our built ANN can emulate synthetic $\mathcal{MCRT}$ spectra well within 10 per cent accuracy. As far as using XRADE in XSPEC, we are able to successfully produce finer tables than `slow64` and `fast32` as long as they are within the parameter boundaries set in the $\mathcal{MCRT}$ tables. Any user can easily build a fully customized XRADE multiplicative table that will be suitable for spectral analysis in XSPEC. A future test is, however, to explore whether a coarser and wider parameter grid can be used in order to localize regions of the parameter space to an acceptable level of precision, via e.g. the BXA process and error searches. Once the parameter space is mapped, then finer grids can be adopted.

We note that a finer XRADE table would still be susceptible to XSPEC interpolation issues. Our foreseeable goal is to exploit the ANN impressive emulation rate to be *directly* implemented in the fitting procedure. We aim at eventually bypassing interpolation based fitting programs such as XSPEC, as well as grid development, and use XRADE in the likelihood calculations for parameter inference in a Bayesian model. One solution is to integrate XRADE into the publicly available Bayesian software e.g. 3ML.[8] The advantage of such an approach is that we will be able to obtain more accurate parameter estimates and their full posterior distributions, all the while taking into account any principled prior information about the source.

The great advantages of XRADE are the following: (a) it avoids the need to rerun the initial time-consuming ray-tracing simulations, speeding up, in turn, the process of generating new spectra or even grids; (b) our ANN allows the user to generate fully customised XRADE tables at the user's specific requirements; (c) it produces very large XRADE tables, e.g. with much finer steps, over a much shorter computational time-scale, i.e. seconds–minutes; (d) it greatly mitigates interpolation issues within XSPEC between coarse grid points, while maintaining numerical accuracy to the 1 per cent level (see Fig. 10); and (e) the emulation process can be applied to a large variety of models (see text below) and can be easily implemented directly into Bayesian inference pipelines.

We presented a test case by applying XRADE and `fast32` to PDS 456, which hosts one of the most powerful, persistent accretion disc winds. We specifically tested XRADE on the combined *XMM–Newton* and *NuSTAR* 2013 September 17–21 observations of PDS 456, as the X-ray spectrum is characterized by the best-quality P Cygni feature observed to date, and compared the results with those from `fast32`. We found that both XRADE and `fast32` return an excellent fit to the data, providing measurements of $\dot{\mathcal{M}}_{\rm w}$, $\mathcal{L}_{\rm X}$, $f_{\rm v}$, μ, and Γ with <10 per cent discrepancy. We demonstrated that XRADE provides an excellent fit to the P Cygni profile in PDS 456.

The best-fitting values measured with both `fast32` and XRADE are loosely consistent with N15; in particular the $\dot{\mathcal{M}}_{\rm w}$ is a factor of ∼3 smaller than in N15. This difference can be simply attributed to an assumed launching radius being a factor of ∼3 larger, i.e. $R_{\rm min} = 100\,r_{\rm g} = 1.5 \times 10^{16}$ cm than here. It is important to note that since $R_{\rm min}$ is not yet a free parameter but fixed a priori, the 'true' mass outflow rate maintains a certain degree of uncertainty. For this reason, in future work it is our priority to make $R_{\rm min}$ a measurable parameter in XRADE, as well as to further explore the wind thickness ($R_{\rm max}/R_{\rm min}$) or even a variable d parameter (i.e. changes the wind opening angle). Note that another source of discrepancy for $\dot{\mathcal{M}}_{\rm w}$ can be also attributed, on a lower extent, to the assumed accretion efficiency value of $\eta = 0.06$ here with respect to that in N15 (i.e. $\eta = 0.1$).

[8] https://threeml.readthedocs.io/en/stable/xspec_users.html

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

The extended energy band from 0.1 keV up to 511 keV was adopted as in S10 in order to allow a comparison with observational measurements from future instruments with a significant effective area at relatively high photon energies, >100 keV. However, as such a milestone has not been achieved yet, a possibility for the near future would be to restrict the energy range of the calorimeter-resolution grids, so to optimize computational time and parameter space sampling over the region where this is most relevant (especially for covering the Fe K region).

At present, the major difference between XRADE and disc-wind tables (`slow64` and `fast32`) generated through a 'standard' X-ray tracing method is the enormous difference of CPU time involved in the process. To emulate one single spectrum we require a CPU time of 4.9×10^{-5} s, against 10–50 min (∼60 eV resolution) or 2–3 h (2 eV resolution). We also used BXA to perform a global exploration of the parameter spaces associated with the original `fast32` and finer XRADE tables whilst fitting PDS 456 (Section 4.1). We find good agreement between the overall best-fitting parameter contours, as well as individual posterior distribution shapes (see Fig. 12), indicating that the ANN is able to learn the complex mapping between parameter space and spectral space. Global parameter exploration algorithms thus represent a powerful tool to iteratively test the accuracy of emulation-based table models in the future.

Although XRADE is already a powerful alternative model to the computationally expensive $\mathcal{MCRT}$ simulations, there is still much room for improvement. Most notably, the increase in fractional error seen in the Fe K band will be improved by introducing finer sampling in the training process. Currently our training set is based on simulated spectra generated from a grid of parameters, however ideally we would train from spectra that have parameter values that are randomly sampled across the chosen parameter range. Using a random parameters allows the network to better map the domain and parameter space in comparison to the grid of parameters. Analogous to this is the extensive research that has shown that random search is superior over grid search methods for hyperparameter turning of machine learning algorithms (see e.g. Bergstra & Bengio 2012). Any future work must allow for a sampling of μ and, most importantly, $R_{\rm min}$ values, so that a more accurate energetics and eventually the launching/driving mechanism involved in the disc wind can be can be achieved. The real power of the emulation method is that the implementation of our ANN will undoubtedly be an indispensable tool in anticipation of future X-ray detectors, such as the microcalorimeters onboard *XRISM* and *Athena*. Our emulation method will not be only restricted to the development of XRADE, but it will be implemented in other wind models, such as magnetohydrodynamic (e.g. Fukumura et al. 2010) and WINd Emission (WINE) models (Luminari et al. 2020). This tool can be also applied to non-wind models and beyond X-ray astronomy studies.

ACKNOWLEDGEMENTS

We would like to thank the anonymous referee for helpful comments and feedback that helped us to improve the clarity of this paper. The authors would also like to thank Justin Alsing for helpful discussions. GAM acknowledges the financial support from Attività di Studio per la comunità scientifica di Astrofisica delle Alte Energie e Fisica Astroparticellare: Accordo Attuativo ASI-INAF no. 2017-14-H.0. GAM also acknowledges the Sciops technical IT Unit – SITU at the European Space Astronomy Centre for letting him use their cluster. ML acknowledges a Machine Learning in Science research fellowship from the University of Nottingham. JNR and VB acknowledge NASA-ADAP grant 80NSSC22K0474. PGB acknowledges financial support from the Czech Science Foundation project no. 22-22643S. EK acknowledges financial support from the Centre National d'Etudes Spatiales (CNES). MB is supported by the European Innovative Training Network (ITN) 'BiD4BEST' funded by the Marie Sklodowska-Curie Actions in Horizon 2020 (GA 860744). MG is supported by the 'Programa de Atracción de Talento' of the Comunidad de Madrid, grant number 2018-T1/TIC-11733.

DATA AVAILABILITY

`fast32` and `slow64` will be publicly available on https://gabrielematzeu.com/disk-wind/. XRADE models will be initially available on request to the authors and the XRADE generator will be publicly available in the foreseeable future. All *XMM–Newton* and *NuSTAR* data used in this work are publicly available from the corresponding archives.

REFERENCES

Alsing J. et al., 2020, ApJS, 249, 5
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
Barret D. et al., 2018, in den Herder J.-W. A., Nikzad S., Nakazawa K., eds, Proc. SPIE Vol. 10699, Space Telescopes and Instrumentation 2018: Ultraviolet to Gamma Ray. SPIE, Bellingham, p. 106991G
Bautista M. A., Kallman T. R., 2001, ApJS, 134, 139
Bebis G., Georgiopoulos M., 1994, IEEE Potentials, 13, 27
Behar E., 2009, ApJ, 703, 1346
Bergstra J., Bengio Y., 2012, J. Machine Learning Res., 13, 281
Bianchi S., Guainazzi M., Matt G., Fonseca Bonilla N., Ponti G., 2009, A&A, 495, 421
Braito V. et al., 2022, ApJ, 926, 219
Buchner J., 2021, preprint (arXiv:2101.09675)
Buchner J. et al., 2014, A&A, 564, A125
Castor J. I., Abbott D. C., Klein R. I., 1975, ApJ, 195, 157
Chardin J., Uhlrich G., Aubert D., Deparis N., Gillet N., Ocvirk P., Lewis J., 2019, MNRAS, 490, 1055
Chartas G., Brandt W. N., Gallagher S. C., Garmire G. P., 2002, ApJ, 579, 169
Chartas G. et al., 2021, ApJ, 920, 24
Dannen R. C., Proga D., Kallman T. R., Waters T., 2019, ApJ, 882, 99
de Kool M., Begelman M. C., 1995, ApJ, 455, 448
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Done C., Davis S. W., Jin C., Blaes O., Ward M., 2012, MNRAS, 420, 1848
Emmering R. T., Blandford R. D., Shlosman I., 1992, ApJ, 385, 460
Everett J. E., 2005, ApJ, 631, 689
Fabian A. C. et al., 2020, MNRAS, 493, 2518
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Fukumura K., Kazanas D., Contopoulos I., Behar E., 2010, ApJ, 715, 636
Fukumura K., Tombesi F., Kazanas D., Shrader C., Behar E., Contopoulos I., 2015, ApJ, 805, 17
Fukumura K., Kazanas D., Shrader C., Behar E., Tombesi F., Contopoulos I., 2017, Nat. Astron., 1, 0062
Fukumura K., Kazanas D., Shrader C., Behar E., Tombesi F., Contopoulos I., 2018, ApJ, 853, 40
Fukumura K., Kazanas D., Shrader C., Tombesi F., Kalapotharakos C., Behar E., 2021, ApJ, 912, 86
Gallo L. C., Fabian A. C., 2011, MNRAS, 418, L59
Gallo L. C. et al., 2013, MNRAS, 428, 1191
Giustini M., Proga D., 2019, A&A, 630, A94
Gofford J., Reeves J. N., Tombesi F., Braito V., Turner T. J., Miller L., Cappi M., 2013, MNRAS, 430, 60

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

Gofford J. et al., 2014, ApJ, 784, 77
Gofford J., Reeves J. N., McLaughlin D. E., Braito V., Turner T. J., Tombesi F., Cappi M., 2015, MNRAS, 451, 4169
Hagino K., Odaka H., Done C., Gandhi P., Watanabe S., Sako M., Takahashi T., 2015, MNRAS, 446, 663
Hagino K., Odaka H., Done C., Tomaru R., Watanabe S., Takahashi T., 2016, MNRAS, 461, 3954
Hagino K., Done C., Odaka H., Watanabe S., Takahashi T., 2017, MNRAS, 468, 1442
Harrison F. A. et al., 2013, ApJ, 770, 103
He S., Li Y., Feng Y., Ho S., Ravanbakhsh S., Chen W., Póczos B., 2019, Proc. Natl. Acad. Sci., 116, 13825
Higginbottom N., Knigge C., Long K. S., Matthews J. H., Parkinson E. J., 2019, MNRAS, 484, 4635
Higginbottom N., Knigge C., Sim S. A., Long K. S., Matthews J. H., Hewitt H. A., Parkinson E. J., Mangham S. W., 2020, MNRAS, 492, 5271
Hopkins P. F., Elvis M., 2010, MNRAS, 401, 7
Hornik K., Stinchcombe M., White H., 1989, Neural Networks, 2, 359
Igo Z. et al., 2020, MNRAS, 493, 1088
Kallman T., Bautista M., 2001, ApJS, 133, 221
Kallman T. R., Palmeri P., Bautista M. A., Mendoza C., Krolik J. H., 2004, ApJS, 155, 675
Kasim M. et al., 2022, MLS&T, 3, 015013
Kazanas D., Fukumura K., Behar E., Contopoulos I., Shrader C., 2012, Astron. Rev., 7, 92
Kerzendorf W. E., Vogl C., Buchner J., Contardo G., Williamson M., van der Smagt P., 2021, ApJ, 910, L23
King A., 2003, ApJ, 596, L27
King A. R., 2010, MNRAS, 402, 1516
King A. R., Pounds K. A., 2003, MNRAS, 345, 657
Kingma D. P., Ba J., 2014, preprint (arXiv:1412.6980)
Knigge C., Woods J. A., Drew J. E., 1995, MNRAS, 273, 225
Lucy L. B., 2002, A&A, 384, 725
Lucy L. B., 2003, A&A, 403, 261
Luminari A., Piconcelli E., Tombesi F., Zappacosta L., Fiore F., Piro L., Vagnetti F., 2018, A&A, 619, A149
Luminari A., Tombesi F., Piconcelli E., Nicastro F., Fukumura K., Kazanas D., Fiore F., Zappacosta L., 2020, A&A, 633, A55
Luminari A., Nicastro F., Elvis M., Piconcelli E., Tombesi F., Zappacosta L., Fiore F., 2021, A&A, 646, A111
Lusso E. et al., 2012, MNRAS, 425, 623
Madsen K. K., Beardmore A. P., Forster K., Guainazzi M., Marshall H. L., Miller E. D., Page K. L., Stuhlinger M., 2017, AJ, 153, 2
Marchesi S. et al., 2016, ApJ, 830, 100
Matthews J. H., Knigge C., Long K. S., Sim S. A., Higginbottom N., 2015, MNRAS, 450, 3331
Matthews J. H., Knigge C., Long K. S., Sim S. A., Higginbottom N., Mangham S. W., 2016, MNRAS, 458, 293
Matthews J. H., Knigge C., Higginbottom N., Long K. S., Sim S. A., Mangham S. W., Parkinson E. J., Hewitt H. A., 2020, MNRAS, 492, 5540
Matzeu G. A., Reeves J. N., Nardini E., Braito V., Costa M. T., Tombesi F., Gofford J., 2016, MNRAS, 458, 1311
Matzeu G. A., Reeves J. N., Braito V., Nardini E., McLaughlin D. E., Lobban A. P., Tombesi F., Costa M. T., 2017, MNRAS, 472, L15
Matzeu G. A. et al., 2019, MNRAS, 483, 2836
Middei R. et al., 2020, A&A, 635, A18
Mizumoto M., Nomura M., Done C., Ohsuga K., Odaka H., 2021, MNRAS, 503, 1442
Nardini E. et al., 2015, Science, 347, 860 (N15)
Nardini E., Lusso E., Bisogni S., 2019, MNRAS, 482, L134
Nomura M., Ohsuga K., 2017, MNRAS, 465, 2873
Nomura M., Ohsuga K., Done C., 2020, MNRAS, 494, 3616
Nwankpa C., Ijomah W., Gachagan A., Marshall S., 2018, preprint (arXiv:1811.03378)
Ohsuga K., Mineshige S., Mori M., Kato Y., 2009, PASJ, 61, L7
Parker M. L. et al., 2017, Nature, 543, 83
Parker M. L. et al., 2020, MNRAS, 498, L140
Parker M. L., Matzeu G. A., Matthews J. H., Middleton M. J., Dauser T., Jiang J., Joyce A. M., 2022, MNRAS, 513, 551
Piconcelli E., Jimenez-Bailón E., Guainazzi M., Schartel N., Rodríguez-Pascual P. M., Santos-Lleó M., 2005, A&A, 432, 15
Pinto C. et al., 2018, MNRAS, 476, 1021
Porquet D., Reeves J. N., O'Brien P., Brinkmann W., 2004, A&A, 422, 85
Pounds K. A., Reeves J. N., King A. R., Page K. L., O'Brien P. T., Turner M. J. L., 2003, MNRAS, 345, 705
Pounds K. A., Lobban A., Reeves J. N., Vaughan S., Costa M., 2016, MNRAS, 459, 4389
Proga D., Kallman T. R., 2004, ApJ, 616, 688
Proga D., Stone J. M., Kallman T. R., 2000, ApJ, 543, 686
Quera-Bofarull A., Done C., Lacey C., McDowell J. C., Risaliti G., Elvis M., 2020, MNRAS, 495, 402
Rankine A. L., Hewett P. C., Banerji M., Richards G. T., 2020, MNRAS, 492, 4553
Ratheesh A., Tombesi F., Fukumura K., Soffitta P., Costa E., Kazanas D., 2021, A&A, 646, A154
Reeves J. N., Braito V., 2019, ApJ, 884, 80
Reeves J. N., O'Brien P. T., Ward M. J., 2003, ApJ, 593, L65
Reeves J. N. et al., 2009, ApJ, 701, 493
Reeves J. N. et al., 2014, ApJ, 780, 45
Reeves J. N., Braito V., Nardini E., Behar E., O'Brien P. T., Tombesi F., Turner T. J., Costa M. T., 2016, ApJ, 824, 20
Reeves J. N., Braito V., Nardini E., Lobban A. P., Matzeu G. A., Costa M. T., 2018a, ApJ, 854, L8
Reeves J. N., Lobban A., Pounds K. A., 2018b, ApJ, 854, 28
Reeves J. N., Braito V., Chartas G., Hamann F., Laha S., Nardini E., 2020, ApJ, 895, 37
Reeves J. N., Braito V., Porquet D., Lobban A. P., Matzeu G. A., Nardini E., 2021, MNRAS, 500, 1974
Scott A. E., Stewart G. C., 2014, MNRAS, 438, 2253
Shapiro S. L., Teukolsky S. A., 1983, Black Holes, White Dwarfs, and Neutron Stars: The Physics of Compact Objects. Wiley, New York
Sim S. A., Long K. S., Miller L., Turner T. J., 2008, MNRAS, 388, 611 (S08)
Sim S. A., Miller L., Long K. S., Turner T. J., Reeves J. N., 2010, MNRAS, 404, 1369 (S10)
Smith S. L., Kindermans P.-J., Ying C., Le Q. V., 2017, preprint (arXiv:1711.00489)
Stevens I. R., Kallman T. R., 1990, ApJ, 365, 321
Strüder L. et al., 2001, A&A, 365, L18
Tashiro M. et al., 2020, in den Herder J.-W. A., Nikzad S., Nakazawa K., eds, Proc. SPIE Vol. 11444, Space Telescopes and Instrumentation 2020: Ultraviolet to Gamma Ray. SPIE, Bellingham, p. 1144422
Tatum M. M., Turner T. J., Sim S. A., Miller L., Reeves J. N., Patrick A. R., Long K. S., 2012, ApJ, 752, 94
Tomaru R., Done C., Ohsuga K., Odaka H., Takahashi T., 2020a, MNRAS, 494, 3413
Tomaru R., Done C., Ohsuga K., Odaka H., Takahashi T., 2020b, MNRAS, 497, 4970
Tombesi F., Cappi M., Reeves J. N., Palumbo G. G. C., Yaqoob T., Braito V., Dadina M., 2010, A&A, 521, A57
Tombesi F., Cappi M., Reeves J. N., Palumbo G. G. C., Braito V., Dadina M., 2011, ApJ, 742, 44
Tombesi F., Cappi M., Reeves J. N., Braito V., 2012, MNRAS, 422, L1
Tombesi F., Cappi M., Reeves J. N., Nemmen R. S., Braito V., Gaspari M., Reynolds C. S., 2013, MNRAS, 430, 1102
Vasudevan R. V., Fabian A. C., 2009, MNRAS, 392, 1124
Watson-Parris D., 2021, Philos. Trans. R. Soc. A, 379, 20200098
Weymann R. J., Morris S. L., Foltz C. B., Hewett P. C., 1991, ApJ, 373, 23
Williams J. K., Gliozzi M., Rudzinsky R. V., 2018, MNRAS, 480, 96
Yao Y., Rosasco L., Caponnetto A., 2007, Constructive Approximation, 26, 289

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

APPENDIX A: SPECTRAL PROPERTIES OF THE DISC WIND

A1 Influence of geometry on the disc-wind features

In the $\mathcal{MCRT}$ code, the photon packets are collected into 20 inclination bins and then processed into 1000 energy bins. The observer's line-of-sight inclination (θ) is measured with respect to the polar z-axis. Each angular bin is defined by $\mu = \cos\theta$ and determines the degree of line-of-sight interception through the wind. In both tables, the angular bins cover the range $0.025 < \mu < 0.975$ in 20 incremental linear steps of $\Delta\mu = 0.05$. As the geometric framework assumes a flow with an opening angle of 45°, the observer's line of sight does not directly intercept the wind when $\theta < 45°$, or $\mu \gtrsim 0.7$. In such a scenario, the corresponding spectra will be dominated by a reflection component via photons scattered off the wind (see Tatum et al. 2012 for examples of fitting the wind spectra to the Fe K emission profiles of bare Seyferts). Conversely, at high inclinations ($\mu \lesssim 0.7$), the line-of-sight intercepts the wind and, consequently, blueshifted absorption features, as well as scattered emission, will be imprinted on the spectra.

Depending on the range of the angular bin, the inclinations can be denoted as: *low* (polar; $\theta = 0°$–$45°$), *intermediate* (wind fully intercepted; $\theta = 45°$–$66°$), and *high* (edge-on or equatorial; $\theta = 66°$–$90°$). The different sightlines, from each angular bin, intercept material with increasing column densities (or optical depth). Fig. A1 illustrates how the column density of the obscuring medium (for a given $\dot{\mathcal{M}}_{\rm w} = 0.4$, i.e. 30 per cent of $\dot{M}_{\rm Edd}$ in `fast32`), rapidly reaches the optically thick regime (i.e. $N_{\rm H} = 1/\sigma_{\rm T} = 1.5 \times 10^{24}$ cm^{-2}) with increasing θ. The boundary at the opening angle $\theta = 45°$ would unavoidably create some discontinuity regions in the simulations.

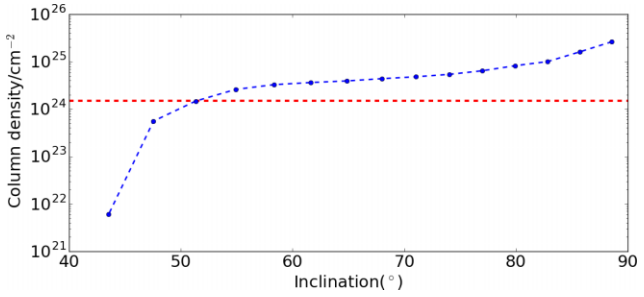

Figure A1. Dependence of the line-of-sight column density on the viewing angle bins for a given $\dot{\mathcal{M}}_{\rm w} = 0.4$ in `fast32`. Along a line of sight with $\theta < 45°$ the wind does not intercept the line of sight to the X-ray source so the column density is approaching zero, although the observed spectrum is modified by photons scattered into the line of sight. The column density increases for $\theta \gtrsim 45°$ as the line of sight becomes more edge-on. The dashed red line represents one Compton depth, corresponding to $N_{\rm H} = 1/\sigma_{\rm T} = 1.5 \times 10^{24}$ cm^{-2}, where the flux is suppressed to 38 per cent of the unattenuated flux.

*fig*A1

In Fig. A2, we show the output spectra from the different inclination ranges indicated above, where the total, direct (or transmitted), and scattered/reflected spectral components are denoted in black, red, and green, respectively. Specifically, at low inclination, e.g. $\theta \sim 30°$ (left), the line of sight does not intercept the wind, so the transmitted spectrum (red) is unaffected by the medium. The total spectrum (black) is dominated by the primary continuum and is supplemented by the scattered/reflected component (green) from the wind material. Distinct features such as the 'Compton hump' (peaking at 20–30 keV) and Fe Kα emission at ∼6.4 keV (blurred by the Doppler shifts within the flow) are prominent. At intermediate values, $\theta \sim 50°$ (middle), the line-of-sight intercepts the outflowing Compton-thick material with $N_{\rm H} \sim 2 \times 10^{24}$ cm^{-2}. The scattered component is similar to before, however the direct continuum is now suppressed by the obscuring medium with absorption features imprinted on the spectra. At high inclinations, $\theta \sim 75°$ (right), the scattered emission dominates over the transmitted component as the line of sight is intercepting material with $N_{\rm H} \sim 5 \times 10^{24}$ cm^{-2} (see Fig. A1), corresponding to a Thomson (or Compton) depth of $\tau \sim 3$.

The difference in shape and centroid energy of the absorption profiles, as seen in Fig. A2 (centre and right), reveals how the Fe K strength and degree of blueshift are strongly dependent on both the wind opening angle and the line-of-sight orientation. In fact, at intermediate inclination, the centroid energy is measured at ∼8.2 keV (centre), whilst at high inclination, the line is centred at ∼7.2 keV (right). In other words, as the viewing angles progressively become polar, the shift in velocity of the profile increases. Such variation arises from the line-of-sight projection of the velocity vector (see fig. 6, S08).

A2 Influence of the launch radius on wind features

The overall shape of the Fe K absorption profile in the simulated spectra changes upon the choice of the launch radius $R_{\rm min}$ and thickness of the flow ΔR. In Fig. A3, we show the broad-band simulated spectra for the `slow64` and `fast32` grids with fine-tuning velocity factor parameter fixed at $f_{\rm v} = 1$. Our findings are as follows.

(i) The effect of $R_{\rm min}$ on the terminal velocity, and thus the degree of blueshift of the profiles, is larger at smaller radii.

(ii) The width of the absorption lines depends on the range of the velocities intercepted. Broader profiles are thus naturally reproduced for faster (inner) winds – due to both the larger range in terminal velocity and the wider shear of velocities intercepted along the flow up to v_∞ (see Section A3).

(iii) An inner, faster wind has a greater opacity as its ΔR (in $r_{\rm g}$) is smaller, thus the density is higher for any given $\dot{\mathcal{M}}_{\rm w}$.

A3 Calculation of wind velocity

A prescription that stipulates the rotational velocity at every point in the wind, following the parametrization of Knigge, Woods & Drew (1995), was included in the S10 code. It is assumed that the specific angular momentum is conserved by the outflowing 'packets' of matter about the polar z-axis. At the base of the wind streamline, the angular momentum of the packets is assumed to be Keplerian for the radius at which the streamlines cross the xy-plane. Thus the rotational velocity is solely defined by choosing the wind geometry and black hole mass of the source.

The outflow radial velocity, which points away from the focus point of the wind d (see Fig. 1), is

$$v_l = v_0 + (v_\infty - v_0)\left(1 - \frac{R_v}{R_v + l}\right)^{\beta}, \quad \text{(A1)}$$

where l is the distance along the wind streamline, and R_v is the velocity scale length (set to be equal to $R_{\rm max}$), which defines how far the packet of matter has travelled before reaching halfway of the terminal speed in the streamline. The β exponent governs the acceleration rate and is usually set to 1 due to difficulties in

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

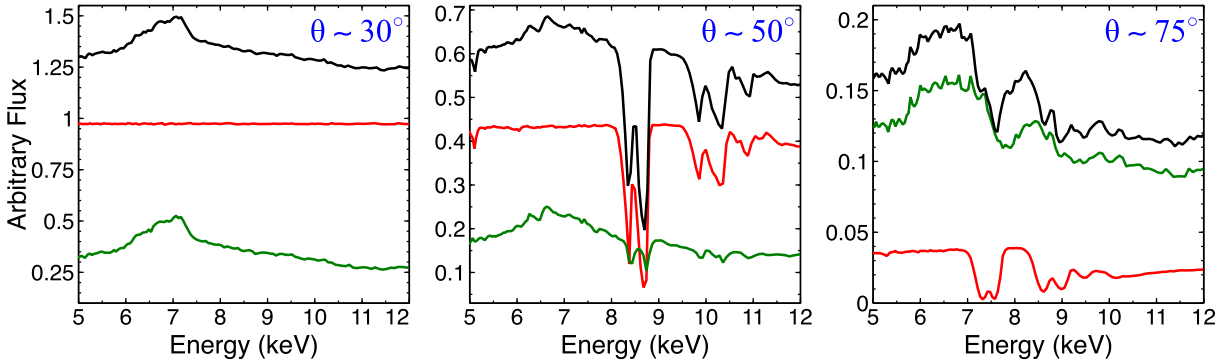

Figure A2. Simulated `fast32` spectra between 5 and 12 keV with given $\dot{\mathcal{M}}_{\rm w} = 0.5$, $\mathcal{L}_{\rm X}$ (per cent) = 0.2, and $f_{\rm v} = 1.0$ corresponding to low- (left; $\mu = 0.875$, or $\theta \sim 30°$), intermediate- (centre; $\mu = 0.625$, or $\theta \sim 50°$), and high inclinations (right; $\mu = 0.275$, or $\theta \sim 75°$). The total spectra (black) and their respective direct (red) and scattered (green) components are shown. Low inclinations: as the line of sight does not intercept the wind, the resulting spectrum is dominated by scattered photons from the inner edge of the flow, leading to the broad Fe K emission feature peaking at $E \sim 7$ keV. Intermediate inclinations: as the line of sight fully intercepts the wind, deep absorptions, e.g. from Fe XXV–XXVI, and broad emissions are imprinted on the spectra. High inclinations: at nearly equatorial line-of-sight orientation, the total spectrum is dominated by the scattering component whilst the primary continuum is heavily suppressed by the Compton-thick material at the base of the flow. Such orientation would lead to a broad, shallow absorption feature in the spectra.

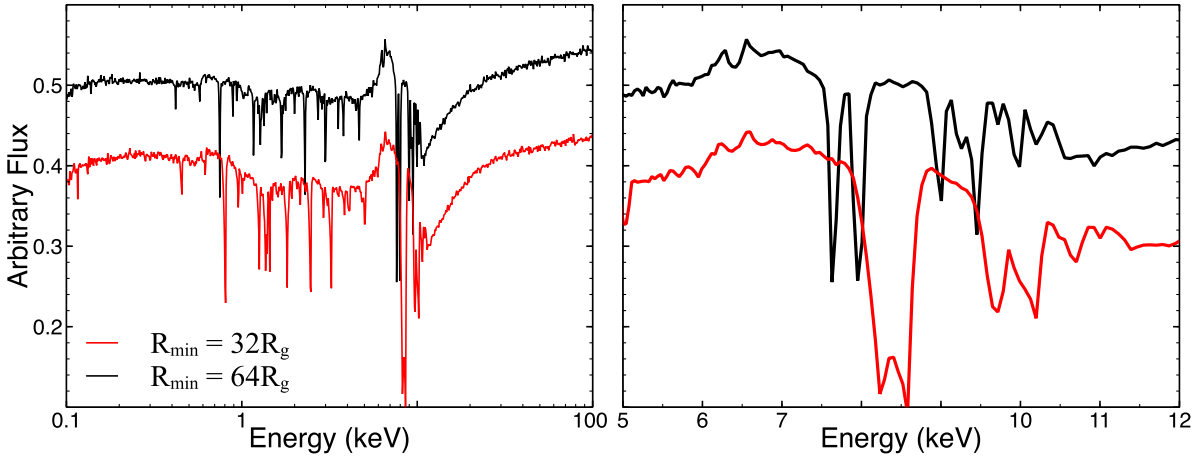

Figure A3. Left: total broad-band simulated spectra shown in the 0.1–100 keV band, with a given $\mu = 0.625$, corresponding to launch radii of $R_{\rm min} = 64\,r_{\rm g}$ (black) and $R_{\rm min} = 32\,r_{\rm g}$ (red). Right: zoom on the Fe K region. Regulated by the assumptions on $R_{\rm min}$, a velocity shift is observed in the synthetic spectra, whereby the black one ($R_{\rm min} = 64\,r_{\rm g}$) is slower than the red one ($R_{\rm min} = 32\,r_{\rm g}$). This is a natural consequence of the velocity along the streamline, which scales directly with the escape velocity at the base of the flow. As the terminal velocity $v_\infty = f_{\rm v}\sqrt{2/R_{\rm min}(r_{\rm g})}\,c$ is affected by both $R_{\rm min}$ and $f_{\rm v}$ (or an *effective* $R_{\rm min} \propto f_{\rm v}^2$), in our simulations we kept $f_{\rm v} = 1$ for clarity. Additionally, with a given mass outflow rate ($\dot{\mathcal{M}}_{\rm w} = 0.5$) and ionizing luminosity ($\mathcal{L}_{\rm X} = 2 \times 10^{-3}$, i.e. $L_{2-10\,\rm keV} = 0.2$ per cent of $L_{\rm Edd}$), the fastest wind spectrum (red) is overall more attenuated and with stronger line depths than its counterpart. The innermost winds have a larger column density as a consequence of the observer's line of sight crossing a larger portion of the wind than the outer flows (see Section A4).

constraining it with the X-ray data currently available. The initial velocity v_0 is set to 0, given that $v_0 \ll v_\infty$. Variations in the R_v and β terms in equation (A1) can result in a change of the width of the simulated wind feature profiles. Reducing β would increase the red-wing of the absorption feature, as the material would take longer to accelerate to v_∞. On the other hand, by reducing R_v, the packet of matter would travel a shorter distance along the streamline before v_∞ is reached, hence reducing the red-wing. Additionally, the presence and strength of a red-wing characterizes the probability of observing slower packets of matter along a given streamline.

A4 Mass density

The wind is assumed to be smooth and in a steady state, and it can be characterized by a mass-loss rate $\dot{M}_{\rm out}$ that corresponds to the total mass present within the flow. The local mass-loss rate per unit area as a function of R is defined as $\mathrm{d}\dot{m}/\mathrm{d}A \propto R^\kappa$. In these simulations the mass-loss exponent is set to $\kappa = -1$ (default value) that falls within the range expected in a continuous large-scale radial outflow, i.e. $-1.3 < \kappa < -1$ (Behar 2009). The integral of $\mathrm{d}\dot{m}/\mathrm{d}A$ has to equal the total mass-loss rate, such that $\dot{M}_{\rm out} = 4\pi \int_{R_{\rm min}}^{R_{\rm max}} (\mathrm{d}\dot{m}/\mathrm{d}A) R\,\mathrm{d}(R)$. Thus,

$$\frac{\mathrm{d}\dot{m}}{\mathrm{d}A} = \frac{\dot{M}_{\rm out}(\kappa+2)}{4\pi\left[R_{\rm max}^{\kappa+2} - R_{\rm min}^{\kappa+2}\right]} R^\kappa. \tag{A2}$$

A decrease of the mass-loss parameter has the effect of making the mass within the flow more centrally concentrated. The mass density for a given cell is $\rho = \mathrm{d}m/\mathrm{d}V$, the unit volume is $\mathrm{d}V = v_l\,\mathrm{d}t\,\mathrm{d}A$, where $v_l\,\mathrm{d}t$ is the distance travelled by a packet of matter at velocity v_l along the streamline. By combining these terms, the mass per unit volume is

$$\rho = \frac{1}{v_l}\frac{\mathrm{d}\dot{m}}{\mathrm{d}A}. \tag{A3}$$

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024

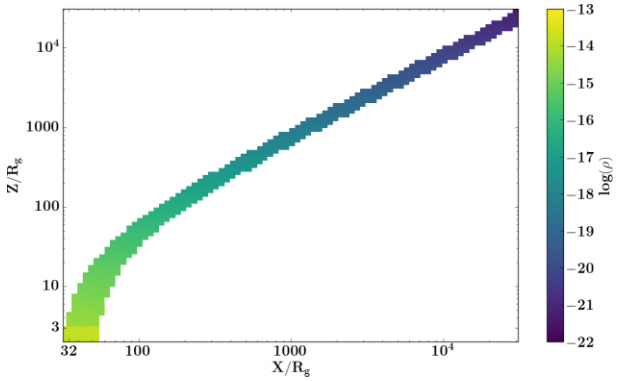

Figure A4. Colour map of an example model run showing the distribution of the mass number density through the wind cells that falls off with radius as a function of $R^{(\kappa=-1)}$. The corresponding density values in log scale are shown in the colour bar in units of g cm^{-3}. The x-axis represents the disc plane, while the z-axis is along the rotational axis. Both axes are in units of gravitational radii.

The above expression suggests that the mass density falls off faster than what is expected in equation (A2) at $\kappa = -1$, as the wind accelerates up to v_∞. Such occurrence can be seen in Fig. A4 where the mass number density at $R = 10^3\, r_g$ and $10^4\, r_g$ is $\log(\rho/\mathrm{g\ cm^{-3}}) = -17.6$ and $\log(\rho/\mathrm{g\ cm^{-3}}) = -20.1$, respectively, i.e. $\Delta \log(\rho/\mathrm{g\ cm^{-3}}) = -2.5$. Thus the mass density falls off quicker than what would be expected from the mass loading equation (A2) at $\kappa = -1$ due to the effect of the increasing velocity vector of the flow along the streamlines.

APPENDIX B: NETWORK PARAMETERS

In a FFNN framework the number of trainable parameters (N_p) is derived from the number of connection in between each layer plus the number of biases in each layer. Each dense layer contains one bias per neuron, so we have a total of 3000 biases. A general expression can be written as

$$N_p = \sum_{k=1}^{n} N_{H^k} N_{H^{(k-1)}} + N_{H^k}, \tag{B1}$$

where n is the number of dense layers, N_{H^0} is the number of inputs to the neural network. In our case $n = 3$ and $N_{H^0} = 6$. Each of the three dense layers H^1, H^2, and H^3 have 1000 neurons each so we have $N_p = (6 \times 1000 + 1000) + (1000 \times 1000 + 1000) + (1000 \times 1000 + 1000) = 2009\,000$.

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

Downloaded from https://academic.oup.com/mnras/article/515/4/6172/6653103 by guest on 04 December 2024