

# MAXI J1820+070 X-ray spectral-timing reveals the nature of the accretion flow in black hole binaries

Tenyo Kawamura<sup>1</sup>,<sup>2</sup> Chris Done,<sup>2,3</sup> Magnus Axelsson<sup>4,5</sup>† and Tadayuki Takahashi<sup>1,2</sup>

<sup>1</sup>Department of Physics, University of Tokyo, Bunkyo, Tokyo 113-0033, Japan

<sup>2</sup>Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Kashiwa, Chiba 277-8583, Japan

<sup>3</sup>Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

<sup>4</sup>Oskar Klein Center for Cosmoparticle Physics, Department of Physics, Stockholm University, SE-10691 Stockholm, Sweden

<sup>5</sup>Department of Astronomy, Stockholm University, SE-10691 Stockholm, Sweden

Accepted 2022 December 19. Received 2022 November 18; in original form 2022 September 2

## ABSTRACT

Black hole X-ray binaries display significant stochastic variability on short time-scales (0.01–100 s), with a complex pattern of lags in correlated variability seen in different energy bands. This behaviour is generally interpreted in a model where slow fluctuations stirred up at large radii propagate down through the accretion flow, modulating faster fluctuations generated at smaller radii. Coupling this scenario with radially stratified emission opens the way to measure the propagation time-scale from data, allowing direct tests of the accretion flow structure. We previously developed a model based on this picture and showed that it could fit the *Neutron star Interior Composition Explorer* (NICER; 0.5–10 keV) data from the brightest recent black hole transient, MAXI J1820+070. However, here we show it fails when extrapolated to higher energy variability data from the *Insight-Hard X-ray Modulation Telescope* (HXMT). We extend our model so that the spectrum emitted at each radius changes shape in response to fluctuations (pivoting) rather than just changing normalization. This gives the strong suppression of fractional variability as a function of energy seen in the data. The derived propagation time-scale is slower than predicted by a magnetically arrested disc (MAD), despite this system showing a strong jet. Our new model jointly fits the spectrum and variability up to 50 keV, though still cannot match all the data above this. Nonetheless, the good fit from 3 to 40 keV means the quasi-periodic oscillation (QPO) can most easily be explained as an extrinsic modulation of the flow, such as produced in the Lense–Thirring precession, rather than arising in an additional spectral-timing component such as the jet.

**Key words:** accretion, accretion discs – black hole physics – X-rays: binaries – X-rays: individual: MAXI J1820+070.

## 1 INTRODUCTION

The nature and geometry of the X-ray emission region in black hole binaries are still controversial, especially in the low/hard state, where most of the power is emitted in a spectrum quite unlike a standard disc (Shakura & Sunyaev 1973). Spectral fitting alone is degenerate, with proposed geometries being a compact source on the spin axis (lamppost), extended emission along the jet direction (jet corona), extended coronal emission on top of an underlying accretion disc (sandwich), and extended coronal emission that replaces the accretion disc (truncated disc/hot inner flow; see e.g. Poutanen, Veledina & Zdziarski 2018). The truncated disc/hot inner flow model has the advantage that it gives a framework to explain the evolution of the spectrum and its fast variability properties together (Done, Gierliński & Kubota 2007), although there are persistent questions over the extent of disc truncation from modelling the reflected emission and its associated iron line (e.g. Buisson et al. 2019; but see Zdziarski et al. 2021a). Another way to track the extent of the disc is the quasi-thermal emission arising from the same X-ray irradiation of

the disc, which gives rise to the iron line and reflected emission (De Marco et al. 2015; Wang et al. 2022). Photons that are not reflected are reprocessed in the disc, producing a thermal reverberation signal. This gives a soft lag, where variations of soft photons follow those of hard photons with a light travel time delay. Reverberation size scales do indeed point to a truncated disc, with a truncation radius that decreases as the source spectrum softens (De Marco et al. 2021). Perhaps the most compelling evidence for a truncated disc is the new polarization results for the low/hard state of Cyg X-1. These rule out the X-ray emission region being aligned with the jet and, instead, require it to be aligned with the accretion flow (Krawczynski et al. 2022). Truncated disc/hot inner flow models are thus strongly favoured, motivating our work in exploring how we can derive the physical properties of the hot flow.

The fast variability (0.01–100 s) gives independent constraints on the accretion flow. It shows many complex properties that change as a function of energy and variability time-scale (see e.g. the review by Uttley et al. 2014). The most promising framework in which to explain these is with propagating fluctuations (Lyubarskii 1997; Kotov, Churazov & Gilfanov 2001). The idea is that variability is generated in the accretion flow (e.g. by the turbulent dynamo magnetorotational instability – MRI; Balbus & Hawley 1991), with a characteristic time-scale that is shorter at smaller radii. Fluctuations

\* E-mail: [tenyo.kawamura@ipmu.jp](mailto:tenyo.kawamura@ipmu.jp)

† Deceased.

generated at any radius in the hot flow propagate down so that slower fluctuations stirred up at larger radii propagate down to modulate the faster fluctuations produced at smaller radii. This produces correlated but lagged multi-time-scale variability in the entire hot flow (Lyubarskii 1997; Kotov et al. 2001). These lags can be seen directly in the data if the flow produces different spectra at different radii, with the observed ‘hard lags’ (fluctuations at 10–20 keV lagging behind the same fluctuation at 2–3 keV; Miyamoto et al. 1988; Nowak, Wilms & Dove 1999) requiring that smaller radii have harder spectra, which seems physically intuitive. Numerical models combining the propagating fluctuations process with a spectrally inhomogeneous hot flow have shown general agreement with the variability properties observed by the *Rossi X-ray Timing Explorer* (*RXTE*) in the 3–30 keV bandpass (Arévalo & Uttley 2006; Ingram & Done 2011, 2012).

However, accurately reproducing all the observed timing properties with the propagating fluctuations model turned out to be more difficult than expected in these first quantitative models, which considered only the variability in the hot flow. One approach is simply to bypass this complexity and use phenomenological models of the intrinsic variability and its lags. This is especially useful if the goal is simply to measure reverberation lags as in the *RELTRANS*code (Ingram et al. 2019). However, our goal is instead to make a physical model of the hot flow, self-consistently producing its spectrum and variability. Such models can then be used to measure physical properties, e.g. the propagation speed, which constrains the nature and geometry of the hot flow. Hence we extend the hot flow propagating fluctuation models to get a better match to the data. There are many potential ways to do this, but the goal is to identify the physical processes that make the most impact on the observed data.

One difficulty with models of propagating fluctuations through the hot flow is that this typically gives rise to a single-peaked power spectrum (Arévalo & Uttley 2006; Ingram & Done 2011, 2012), whereas the observed power spectra are often double peaked (Belloni, Psaltis & van der Klis 2002; Pottschmidt et al. 2003; Axelsson, Borgonovo & Larsson 2005; Grinberg et al. 2014). It is possible to change the time-scales and amplitude of variability with radius in the hot flow to match the data, but it seems fine-tuned (Mahmoud & Done 2018a,b). Another issue is that the observed power spectra often span a very broad range in frequencies, which is difficult to quantitatively match by the fairly small range of radii spanned by the hot flow without going to extreme parameters (Ingram & Done 2011, 2012; Mahmoud & Done 2018a,b).

A key to matching both the power spectral shape and width was the recognition that the disc generates considerable variability in the low/hard state, in addition to that expected from the hot flow (Wilkinson & Uttley 2009; Uttley et al. 2011). While the disc does not contribute to the *RXTE* bandpass ( $>3$  keV) in the low/hard state, its variability will propagate down into the hot flow, so it will strongly affect the variability properties. Rapisarda et al. (2016) proposed that the inner edge of the disc had a much longer variability time-scale than the outer edge of the hot flow due to its smaller scale height, and showed that this naturally produces double-peaked power spectra (but see Veledina 2016 for another potential mechanism). The slowly variable disc also widens the range of time-scales on which the X-ray flux varies even when the truncation radius is only a few tens of gravitational radii (Rapisarda et al. 2016).

Modelling of the broad-band X-ray variability demonstrates how the timing properties give additional information about the nature of the accretion flow. Combining these with spectra (spectral-timing studies) gives an even more powerful tool, as it uses all the information from the energy spectrum and its fluctuations (power spectra)

together with causal connections (lags/leads; e.g. Axelsson & Done 2018; Mahmoud, Done & De Marco 2019; De Marco et al. 2021; Wang et al. 2021). In our previous work (Kawamura et al. 2022, hereafter *K22*), we developed a spectral-timing model based on propagating fluctuations from a turbulent disc through a spectrally inhomogeneous (approximated by two Comptonization regions) flow that generates variability at each radius. We also incorporated reverberation of the variable Comptonization components illuminating the disc to perform a self-consistent spectral-timing analysis. We applied the model to the recently discovered black hole transient MAXI J1820+070 (Kawamuro et al. 2018; Tucker et al. 2018), which has been widely studied (e.g. Kara et al. 2019; Shidatsu et al. 2019; Bright et al. 2020; Homan et al. 2020; Axelsson & Veledina 2021; Ma et al. 2021; Tetarenko et al. 2021; Wang et al. 2021; You et al. 2021; Prabhakar et al. 2022) thanks to its exceptional brightness, low galactic absorption (Uttley et al. 2018), and intensive monitoring by multiple telescopes. *K22* fit the time-averaged energy spectrum for the *Neutron star Interior Composition Explorer* (*NICER*; 0.5–10 keV) + the *Nuclear Spectroscopic Telescope Array* (*NuSTAR*; 3–73 keV) and used this to develop a model for the variability below 10 keV seen in *NICER*. However, *NuSTAR* has less capability for fast timing, so *K22* could not investigate the variability at higher energies, which means that we could not fully probe the innermost parts of the hot flow. Better constraints on propagation require extending the bandpass for fast timing to higher energies.

Here we use contemporaneous data from the *Insight-Hard X-ray Modulation Telescope* (*HXMT*; Section 2) to test our model at higher energies. We predict the high-energy power spectra and phase lags and show how these fail to describe several key features of the data (Section 3). We give the model maximal freedom by fitting only the timing data rather than using the full spectral-timing data, and consider several ways to extend our propagating fluctuation model to better match the data (Section 4). In particular, the phase lags give a clear indication that the propagation time through the flow is slower than the time-scale on which fluctuations are generated (Rapisarda, Ingram & van der Klis 2017), but the full energy dependence of the variability is quite difficult to fit. The key to matching the power spectra is to allow the Comptonization spectra to pivot, so that they change in shape and normalization in response to the fluctuations (Mastroserio, Ingram & van der Klis 2018, 2019; Mastroserio et al. 2021). This is physically expected from Comptonization models (Veledina 2016, 2018) and is observed (Malzac et al. 2003; Gandhi et al. 2008; Yamada et al. 2013; Bhargava et al. 2022). We are able to get a good match to the timing properties (power spectra and phase-lag spectra) from 2.6 up to 48 keV by including spectral pivoting, as well as separation of generator and propagation time-scale. We implement this as a full spectral-timing model and find we can fit all the data in the 2.6–48 keV bandpass, though the model for both spectra and timing diverge from the data above this energy (Section 5). We discuss the physical properties of the accretion flow, comparing them with theoretical hot flow models (Section 6), and then conclude that the current data quality is still better than the best physical models of the flow, which motivates further development (Section 7). All of the technical details of the model formalism are given in the appendices so that the main text stresses the physical aspects of the model.

## 2 OBSERVATION AND DATA REDUCTION

We investigate the bright low/hard state of MAXI J1820+070 observed by *Insight-HXMT*: 2018 March 22 10:46:53 to 2018 March 24 02:49:49 (Obs. ID: P0114661003). The same data are studied in

Wang et al. (2020), Ma et al. (2021), and Yang et al. (2022). The observation time is slightly later than that we studied in K22 (Obs. ID: 1200120106; 2018 March 21), but there are simultaneous *NICER* data (Obs. ID: 1200120108; 2018 March 23) corresponding to these *Insight-HXMT* data. We checked that the energy spectrum, power spectra, and phase-lag spectra in these simultaneous *NICER* data are almost identical to K22.

We used the *Insight-HXMT* Data Analysis Software package (HXMTDAS) v2.04 to calibrate and screen the data using the same criteria as in Yang et al. (2022). We checked that the spectral and variability properties did not change substantially over the observation and then merged all the data to achieve high signal-to-noise ratio.

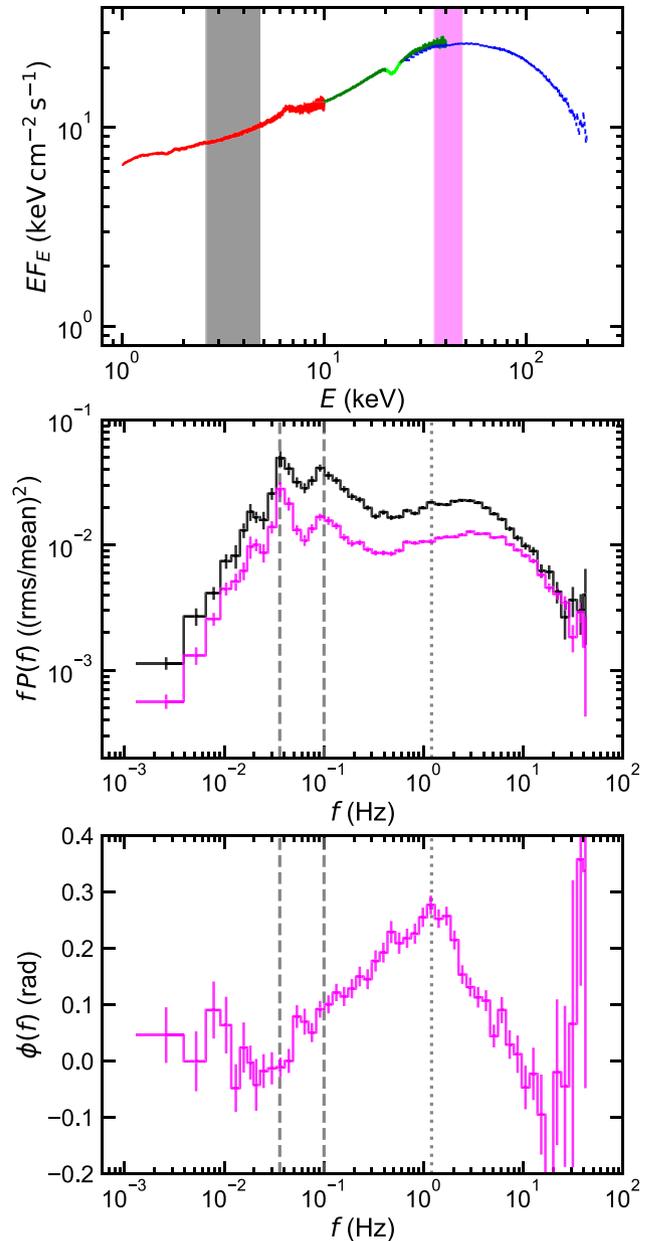
The resulting energy spectrum is shown in Fig. 1 (top). The different colours represent different telescopes (red: LE; green: ME; blue: HE; Zhang et al. 2020). The energy spectrum from ME has a dip of around 22 keV (light green), which is associated with silver fluorescent lines generated within the detector (Li et al. 2020). Following You et al. (2021), we added 1.5 per cent systematic errors to all spectral data.

To study the fast variability, we split the background-subtracted light curves into segments of 256 s with 1/128 s time bins ( $2^{15}$  points), where we avoided any data gaps. We only used the data where all telescopes were active to calculate light curves, using the same time selection for every energy band. We calculate the white-noise-subtracted power spectra and the cross-spectra from each 256 s segment and average them over different segments and logarithmically spaced Fourier frequencies (Uttley et al. 2014; Ingram 2019). All power spectra are normalized such that their integral over frequency corresponds to the fractional variance (Miyamoto et al. 1991; Vaughan et al. 2003). Phase-lag spectra are calculated from the cross-spectra, using the relation between the phase-lag spectrum  $\phi(f)$  and cross-spectrum  $C(f)$ ,  $\phi(f) = \tan^{-1}(\text{Im}[C(f)]/\Re[C(f)])$ , where  $\Re[\dots]$  and  $\text{Im}[\dots]$  denote the real and imaginary parts, respectively. Phase lag relates to time lag via the relation  $2\pi f t(f) = \phi(f)$ .

Fig. 1 (middle) shows the power spectra of the 2.6–4.8 keV (black) and 35–48 keV (magenta) light curves. These energy bands are marked in the energy spectrum with shaded regions. A quasi-periodic oscillation (QPO) and its harmonic exist around 0.036 and 0.1 Hz (shown with dashed lines), in addition to the broad-band variability.

Fig. 1 (bottom) shows the phase-lag spectrum between these two energy bands. The convention throughout this paper is that positive lags mean that the harder energy band lags behind the softer one. The phase lag peaks at  $\sim 1.2$  Hz (shown with a dotted line), which is not at the same frequency as the high-energy peak in the power spectra. This is unexpected as simple propagating fluctuations models have the same peak frequency both in the power spectrum and cross-spectrum (Ingram & van der Klis 2013; Rapisarda et al. 2016). The QPO fundamental appears to affect the phase lag between these two energy bands, creating a dip in the phase-lag spectrum around the corresponding frequency (Ma et al. 2021). The effect of the second harmonic on the phase lag is not so clear between these energy bands, but we note it does have an impact on different choices of energy bands (Ma et al. 2021).

For all of the data fits performed in this paper, we use XSPEC 12.12.1 (Arnaud 1996). We formatted variability data and created a diagonal dummy response such that XSPEC can import power spectra and phase-lag spectra as a function of Fourier frequency. We developed



**Figure 1.** Spectral-timing properties of MAXI J1820+070 observed by *Insight-HXMT*. Top: time-averaged energy spectrum. Red, green, and blue markers represent LE, ME, and HE telescopes, respectively. The dip around 22 keV (light green) is associated with fluorescent lines of silver generated within the ME detector. The coloured regions show the low (black: 2.6–4.8 keV) and high (magenta: 35–48 keV) used to extract light curves. Middle: power spectra calculated for low- and high-energy bands. Both these have the characteristic double peak shape, but with the QPO and its harmonic (marked with dashed lines) superimposed. The high-energy power spectrum is very similar in shape to that at low energy, but with lower normalization (compare to the model in Fig. 2f). Bottom: phase-lag spectrum between the light curves in the low- and high-energy bands. The lags are defined as positive if variations in higher energy bands lag behind those in lower energy bands (hard lags). The frequency at which the phase lag is maximum is marked with a dotted line. This is substantially lower than the characteristic frequency of the second peak in the power spectrum (compare to the model in Fig. 2h).

our model as an XSPEC model. Being able to perform timing fits with the common tool in spectral fits is beneficial in performing spectral-timing fits. For example, we will perform a joint fit to energy spectrum, six power spectra, and five phase-lag spectra in Section 5. We ignore variability below  $\sim 10^{-2}$  Hz because it behaves differently from other Fourier frequencies. Yang et al. (2022) interpreted this low-frequency variability as the QPO subharmonic.

### 3 PROPAGATION AND REVERBERATION IN OUR PREVIOUS MODEL

#### 3.1 Summary of our previous work

We start with a summary of our previous physical model. Fundamentally we assume that variability is generated by fluctuations in density in the flow, which propagate inwards as accretion rate fluctuations. Thus the variability generated at each radius propagates down with the accretion flow so that slower fluctuations generated at large radii imprint on faster fluctuations generated at small radii. These fluctuations in mass accretion rate change the luminosity emitted in the spectrum at that radius.

We then need additional assumptions to turn this into a quantitative model, as we have to assume the form of radial stratification for both the spectrum and variability and how fluctuations are generated and propagated. We expand on each of these below.

We assume a basic geometry that is a truncated disc/hot inner flow, as shown in Fig. 2(a) with emissivity at each radius set by the Shakura–Sunyaev thin disc approximation (Fig. 2b). These two assumptions alone are enough to roughly set the transition radius from the inner edge of the disc to the outer edge of the hot flow by energetics to  $\sim 45R_g$ , though instead we set this from the QPO frequency, assuming the Lense–Thirring precession (Ingram, Done & Fragile 2009). The similarity of the two estimates gives support to the Lense–Thirring interpretation. In all the following, we use the convention that  $R = rR_g$ , where  $R_g = GM_{\text{BH}}/c^2$ .

Spectrally, we assume that the flow emits a single component at each radius. This might be unique to each radius, with e.g. each radius in the disc emitting a blackbody with temperature  $T(R)$ , while the hot flow emits a Comptonized spectrum whose parameters (electron temperature, and optical depth) scale smoothly with radius. However, spectral models are quite degenerate so we approximate the emission from the truncated disc region ( $r_{\text{out}}-r_{\text{ds}}$ ) as a disc blackbody, and we approximate the hot flow as two zones as physically we do expect that there are two main regions in the flow. Close to the disc, seed photons for Comptonization are predominantly from the disc. However, it is quite easy for this Comptonization to become optically thick along the equatorial direction, shielding the inner regions from the disc photons so that seed photons are predominantly from cyclo-synchrotron (Poutanen & Veledina 2014). Thus we assume that radii from  $r_{\text{ds}}-r_{\text{sh}}$  emit soft Comptonization, while radii from  $r_{\text{sh}}-r_{\text{in}}$  emit hard Comptonization ( $r_{\text{in}} < r_{\text{sh}} < r_{\text{ds}} < r_{\text{out}}$ ; Fig. 2a and d). Both Comptonization components illuminate the disc to produce reflection, while the energy not reflected is (mostly) thermalized, enhancing the cool disc emission. K22 show that these assumptions give a good fit to the energy spectrum from 0.5 to 80 keV.

The variability is also assumed to be radially stratified such that each radius generates fluctuations with a characteristic frequency,  $f_{\text{gen}}(r)$ . This is assumed to have a power-law form as a function of radius in the hot flow, so  $f_{\text{gen}}(r) = Br^{-m}f_K(r)$ , where  $f_K(r) = (1/2\pi)r^{-3/2}$  in units of  $c/R_g$ . The power-law scaling parameters are allowed to be different between the hot flow ( $B_f, m_f$ ) and disc ( $B_d, m_d$ ),

giving a different time-scale to reflect the different scale heights of the two flows (Rapisarda et al. 2016). Thus, the generator frequency is modelled with

$$f_{\text{gen}}(r) = \begin{cases} B_f r^{-m_f} f_K(r) & (r_{\text{in}} \leq r < r_{\text{ds}}), \\ B_d r^{-m_d} f_K(r) & (r_{\text{ds}} \leq r < r_{\text{out}}), \end{cases} \quad (1)$$

as shown in Fig. 2(c). We assume each logarithmic radial interval generates the same amplitude of variability. The fluctuations generated at each radius propagate down without losses to produce a fluctuating mass accretion rate at each radius that modulates the emitted luminosity.

As noted above, we assume the QPO is set by the Lense–Thirring precession of the entire hot flow, and the first bump in the power spectrum is set by the turbulent disc (Ingram & Done 2011). This sets  $(B_d, m_d) = (0.03, 0.5)$  and  $r_{\text{out}} = 45$ .

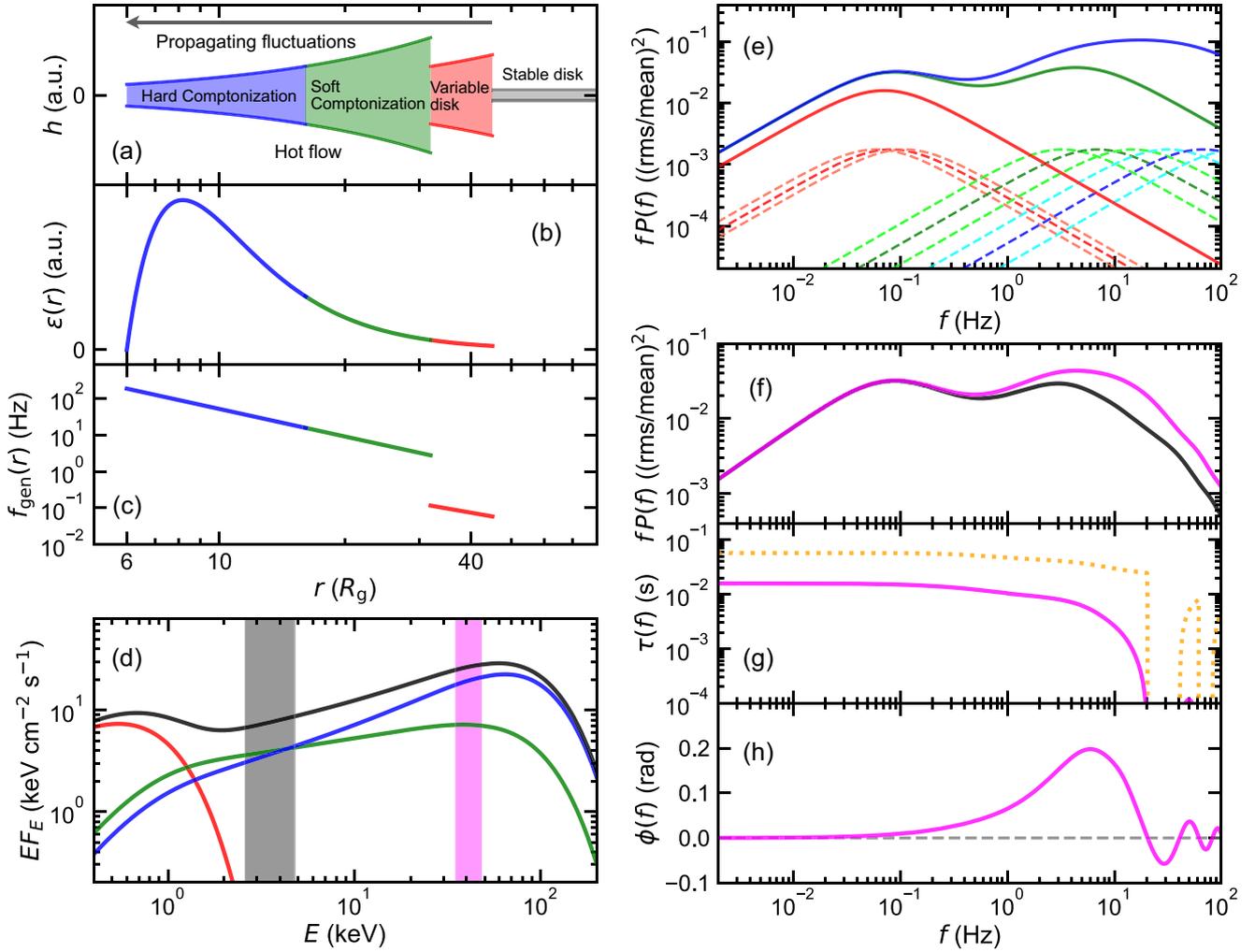
The dashed lines in Fig. 2(e) show three sample power spectra for the generated variability of the local mass accretion rate in each region (variable disc: red; soft Comptonization: green; hard Comptonization: blue; the middle ring of each highlighted in a darker colour). The functional form is a zero-centred Lorentzian with the cut-off frequency corresponding to the local generator frequency  $f_{\text{gen}}(r)$ , which yields the peak at  $f_{\text{gen}}(r)$  in the  $fP(f)$  representation.

K22 assumed that the propagation time-scale was the same as that on which the fluctuations were generated and called this the viscous time-scale (Lyubarskii 1997; Arévalo & Uttley 2006; Ingram et al. 2009). This assumption sets the propagation speed at any radius  $v_p(r) = rf_{\text{gen}}(r)$  in units of  $c$ . However, here we will revisit this assumption, so to avoid confusion, we do not use the term ‘viscous time-scale’ but use ‘generator time-scale’ and ‘propagation time-scale’ in this paper to make it clear which one we mean. K22 also assumed that the fluctuating energy release only changed the normalization of the spectral component emitted at that radius, not its shape.

The solid lines in Fig. 2(e) show the propagated (total) power spectra from each region (disc: red; soft Comptonization: green; and hard Comptonization: blue). This is not the same as the power spectrum in any given energy band as Fig. 2(d) shows that each energy band contains a mix of components.

Fig. 2(f) shows the power spectra for the two chosen energy bands, low (2.6–4.8 keV: black) and high (35–48 keV: magenta), also highlighted in the same colours in Fig. 2(d). While neither band contains the disc emission component, both bands contain the propagated disc variability. The low-energy band also contains both soft and hard Comptonization, so has the generated/propagated power in the soft Comptonization region, plus some of the highest frequency power generated in the hard Comptonization region. The high-energy band is dominated by the hard Comptonization emission, with variability that is propagated down from both the disc and soft Comptonization region, plus the highest frequency power generated/propagated through the hard Comptonization region. Thus the power spectra are almost identical for below  $\sim 1$  Hz, indicating that variability on these slow time-scales is propagated from the outer regions rather than generated at their emission regions, while they diverge at the highest frequencies where the low-energy band does not include as much of the hard Comptonization component as the high-energy band.

The propagation time-scale is explicitly seen in the time lag (Fig. 2g). The luminosity weighted mean radius of the soft Comptonization band is  $22R_g$ , whereas that for the hard Comptonization band is  $11R_g$ . This lag time (integrating  $1/(rf_{\text{prop}}(r))$ ) is 51 ms, as seen as the value of the approximately constant time lag at



**Figure 2.** Model set-up and predictions from K22. Here we show only the intrinsic components, rather than including reflection/reverberation, to focus on the physics of the propagation. (a) Assumed accretion flow geometry (height as a function of radius). There is an outer stable disc (grey) that is highly turbulent on its inner edge, forming the variable disc region. Inwards of this is the turbulent hot flow. (b) Radial emissivity, assumed to be similar to that of a thin disc. (c) Frequency at which variability is generated at each radius. There is a discontinuity between the variable disc and hot flow as their scale heights are different, so their characteristic time-scales should be different. Sample parameter values are  $(B_f, m_f) = (4, 1)$  for the hot flow and  $(B_d, m_d) = (0.03, 0.5)$  for the variable disc. (d) Time-averaged energy spectra modelled (black) assuming a disc for the stable and variable disc emission (red), while the hot flow is assumed to be approximated by two Comptonization components, soft (green) and hard (blue). The relative luminosity in each component, together with the emissivity in (b), roughly sets the size scale of each region, so that the hard Comptonization is for  $r_{in}-r_{sh} = 6-16$ , the soft Comptonization for  $r_{sh}-r_{ds} = 16-32$ , and the variable disc for  $r_{ds}-r_{out} = 32-45$ . (e) Sample power spectra of the local mass accretion rate in each spectral region. The dashed lines represent the variability generated at each radius, with  $r = 45, 38, 33$  (in the variable disc: red),  $30, 22, 16$  (in the soft Comptonization: green), and  $12, 9, 6$  (in the hard Comptonization: blue) from left to right. The frequency, at which each radius  $fP(f)$  has its peak, corresponds to the local generator frequency  $f_{gen}(r)$ . The solid lines show the total (generated plus propagated) variability at  $r = 38$  (red),  $22$  (green), and  $9$  (blue). (f) Power spectra for two energy bands highlighted in (d). Any given energy band is not just a single component. The low-energy band (black) contains roughly equal amounts of soft and hard Comptonization, while the high-energy band (magenta) has mostly hard Comptonization, but with some contribution from the soft Comptonization as well. Thus the power spectra of the light curves in the low- and high-energy bands are more similar than those of the soft and hard Comptonization components in (e). None the less, there is still more high-frequency power in the high-energy band than in the low-energy band, but at lower frequencies, the power spectra are identical as both contain the same propagated power. (g) Time-lag spectra for the low- and high-energy band light curves. The orange dotted line shows the intrinsic lag of the soft Comptonization light curve compared to the hard Comptonization light curve. This is  $\sim 50$  ms, which is longer than the measured lag of the high-energy band light curve versus the low-energy band due to the mixture of spectral components in each band. (h) Lag between high- and low-energy bands as a phase lag rather than a time lag (related by  $\phi(f) = 2\pi f\tau(f)$ ). This peaks around the frequency of the high-energy peak in the power spectrum.

low frequencies (orange dashed line). The time lag drops when the variability starts to be dominated by the generated variability (which is not correlated) rather than the propagated variability, i.e. at approximately the generator time-scale of the hard Comptonization region. At the highest frequencies, the lag also produces oscillatory structure at high frequencies, by the interference generated by the

mass accretion rate in the hard Comptonization region being the same as in the soft Comptonization region, but lagged by the mean propagation time-scale. This oscillatory structure has a period of 50 ms. However, the intrinsic lag from the spectral components is diluted by a factor of  $\sim 3$  when looking at the low- and high-energy bands rather than the soft and hard Comptonization components.

This is because the low-energy band includes both soft and hard Comptonization, while the high-energy band is dominated by the hard Comptonization. While the value of the measured lag time changes, the oscillatory structure period remains at the propagation time-scale.

This oscillation can be seen more clearly in Fig. 2(h), which shows the phase lag rather than time lag, so each time-scale is multiplied by the factor  $2\pi f$ . In this representation, the phase lag peaks at a frequency between the high-frequency peak in the power spectra at low and high energy. By comparison to the panel above, it is also clear that the oscillatory structure at high frequencies has the same period as in the undiluted (orange) time lags.

All model parameters are summarized in Table 1 and we give a more quantitative model summary in Appendix A.

The fluctuating soft and hard Comptonization regions illuminate the outer disc and produce a reflected/reprocessed signal that lags behind the generated and propagated flow variability by the light travel time to the disc. The reflected emission itself is not strong, though this real reverberation signal will add to the soft lags produced by interference in the propagation (hard) lags seen in Fig. 2(g). However, at lower energies, the photons that are not reflected heat the disc, giving a thermal reverberation signal that is strong at energies close to that of the disc emission ( $\lesssim 2$  keV; Kara et al. 2019). This reverberation signal gives an independent check on the assumption of the disc truncation radius, and the fact that it is consistent (De Marco et al. 2021) gives strong supporting evidence for the underlying assumption that the QPO mechanism is the Lense–Thirring precession. Our previous model includes the reverberation, along with the propagating fluctuations (K22).

K22 showed that this model gave a fairly good fit to the energy dependence of the power spectrum across the *NICER* energy band (0.5–10 keV) and to the lags between the same fluctuations in different energy bands as a function of frequency. However, while the spectral components were built from *NuSTAR* data, which extended above 10 keV, this instrument does not have a sufficient area to do high-time-resolution studies, so the model prediction at higher energies could not be tested.

This outburst of MAXI J1820+070 was also monitored by *Insight-HXMT* (Ma et al. 2021; You et al. 2021; Yang et al. 2022), which does have a sufficient effective area at high energies. As mentioned in the previous section, the *Insight-HXMT* data are not absolutely simultaneous with the *NICER/NuSTAR* data set we used in K22, but they are very close in time, and spectral-timing properties are nearly constant during these periods. Hence we take the spectral-timing model of K22, use it to predict the higher energy behaviour, and compare it to the *Insight-HXMT* data.

### 3.2 Comparison of our previous model to *Insight-HXMT* data

We compare the predictions of our previous model to the power spectra for the 2.6–4.8 and 35–48 keV bands and the phase-lag spectrum between these bands calculated from the *Insight-HXMT* observation data in Fig. 3 (left). Model parameter values are summarized in Table 2, which also contains those for the rest of the columns in Fig. 3. The lower energy band is well reproduced by our previous model from K22, as expected, as it is within the *NICER* energy range over which K22 got good fits. However, the power spectrum at the higher energy band is clearly overestimated, and the phase lag between the two is completely wrong, peaking at too high a frequency with a lag that is too short to match the data.

## 4 TIMING FITS: EXPLORING THE ADDITIONAL PROCESSES REQUIRED TO MATCH THE HIGH-ENERGY VARIABILITY

The specific spectral-timing model of K22 plainly cannot fit the data, so instead here we explore whether the model is capable of reproducing the observed timing properties. We give maximal flexibility by ignoring the time-averaged spectrum (but we will come back to joint spectral-timing modelling in Section 5) and attempt to minimize the sum of  $\chi^2$  values for the power spectra and phase-lag spectrum:

$$\sum_k \left\{ \left( \frac{P_{\text{data}}(E_1, f_k) - P_{\text{model}}(E_1, f_k)}{\Delta P_{\text{data}}(E_1, f_k)} \right)^2 + \left( \frac{P_{\text{data}}(E_2, f_k) - P_{\text{model}}(E_2, f_k)}{\Delta P_{\text{data}}(E_2, f_k)} \right)^2 + \left( \frac{\phi_{\text{data}}(E_1, E_2, f_k) - \phi_{\text{model}}(E_1, E_2, f_k)}{\Delta \phi_{\text{data}}(E_1, E_2, f_k)} \right)^2 \right\}, \quad (2)$$

where  $P_{\text{data}}(E, f)$  and  $\Delta P_{\text{data}}(E, f)$  are the observed power spectrum and its  $1\sigma$  error at frequency  $f$  for energy  $E$ ,  $\phi_{\text{data}}(E, E', f)$ , and  $\Delta \phi_{\text{data}}(E, E', f)$  are the equivalents for the phase-lag spectrum between energy  $E$  and  $E'$ ,  $P_{\text{model}}(E, f)$  and  $\phi(E, E', f)$  are the modelled power spectrum and phase-lag spectrum,  $E_1 = 2.6\text{--}4.8$  keV,  $E_2 = 35\text{--}48$  keV, and  $f_k$  the sampled Fourier frequency.

Our model requires the fraction of each spectral component to calculate the power spectra and cross-spectra. We express this fraction as  $S_d(E)$ ,  $S_s(E)$ ,  $S_s^{(r)}(E)$ ,  $S_h(E)$ , and  $S_h^{(r)}(E)$  for the variable disc, soft Comptonization and its reflection, and hard Comptonization and its reflection, respectively. Since X-ray energy spectra in the hard state are almost fully occupied by these five components at most (Zdziarski et al. 2021a,b, 2022), we require the sum of these fractions to correspond to unity:

$$S_d(E) + S_s(E) + S_s^{(r)}(E) + S_h(E) + S_h^{(r)}(E) = 1. \quad (3)$$

Whereas in Fig. 3 (left), these fractions were calculated from the result of spectral fit in K22 for the self-consistent spectral-timing modelling, here we let them be independent of the time-averaged spectrum in order to focus on the variability properties. We fix  $S_d(E) = 0$  because the disc emission is negligible above 2.6 keV. We also ignore the reverberation, i.e.  $S_s^{(r)}(E) = S_h^{(r)}(E) = 0$ , to simplify the model. This is not a bad approximation for our purpose, as we want to capture the broad-band power spectra and hard lags from propagation. Reverberation makes only small changes to the variability properties on the energy and variability range of interest here.

Finally, we only have the soft and hard Comptonization components with the constraints of  $S_s(E) + S_h(E) = 1$ . We do not include any models for the QPO features for simplicity.

We keep the black hole mass of  $M_{\text{BH}} = 8 M_{\odot}$  (Torres et al. 2020) and emissivity profile, i.e.  $\gamma = 3$  and  $b(r) = 1 - \sqrt{r_{\text{in}}/r}$  (Novikov & Thorne 1973; Shakura & Sunyaev 1973), where we assume that radiation energy from the annulus ranging from  $r$  to  $r + \Delta r$  is proportional to  $r^{-\gamma} b(r) 2\pi r \Delta r$ . In K22, the transition radii  $r_{\text{sh}}$ ,  $r_{\text{ds}}$  were calculated from the emissivity profile and spectral decomposition. However, we lack spectral decomposition. In addition, it turned out that model calculations are hardly affected by small changes in these parameters. Thus, we simply fix these transition radii to typical values,  $r_{\text{sh}} = 16$  and 32.

We show the result of the joint fit to the power spectra for 2.6–4.8 and 35–48 keV and the phase-lag spectrum between these energy bands in Fig. 3 (mid-left). The fit is not qualitatively improved even

**Table 1.** Summary of our model parameters. The variable flow is spectrally composed of the variable disc region and soft (outer) and hard (inner) Comptonization regions. We call the entire Comptonization region the hot flow.

Symbol	Meaning	Units	Default
$M_{\text{BH}}$	Black hole mass	$M_{\odot}$	8
$N_r$	Number of rings splitting the variable flow		40
$r_{\text{in}}$	Inner radius of the hard Comptonization	$R_g$	6
$r_{\text{sh}}$	Transition radius between the hard Comptonization and soft Comptonization	$R_g$	16
$r_{\text{ds}}$	Transition radius between the disc and soft Comptonization	$R_g$	32
$r_{\text{out}}$	Outer radius of the variable disc	$R_g$	45
$F_{\text{var},f}$ ( $F_{\text{var},d}$ )	Fractional intrinsic variability per radial decade in the hot flow (variable disc)		0.8
$D$	Damping factor		0
$B_f$ ( $B_d$ )	Coefficient of the generator frequency in the hot flow (variable disc)		0.03
$m_f$ ( $m_d$ )	Power-law index of the generator frequency in the hot flow (variable disc)		0.5
$B_f^{(p)}$ ( $B_d^{(p)}$ )	Coefficient of the propagation frequency in the hot flow (variable disc)		0.03
$m_f^{(p)}$ ( $m_d^{(p)}$ )	Power-law index of the propagation frequency in the hot flow (variable disc)		0.5
$\gamma$	Power-law index of the emissivity		3
$b(r)$	Inner boundary condition of the emissivity		$1 - \sqrt{r_{\text{in}}/r}$
$t_{0,h}$ ( $t_{0,s}$ )	Time delay of the top hat impulse response of reverberation for the hard (soft) Comptonization <sup>a</sup>	s	$5.5 \times 10^{-3}$
$\Delta t_{0,h}$ ( $\Delta t_{0,s}$ )	Time duration of the top hat impulse response of reverberation for the hard (soft) Comptonization <sup>a</sup>	s	$10 \times 10^{-3}$
$S_0(E)$	Fractional contribution of spectral components to the flux <sup>b,c,d</sup>		0.5
$\eta_{0,h}$ ( $\eta_{0,s}$ )	Constant term of the sensitivity of the hard (soft) Comptonization to change in mass accretion rate <sup>e</sup>		1
$\eta_{1,h}$ ( $\eta_{1,s}$ )	Gradient term of the sensitivity of the hard (soft) Comptonization to change in mass accretion rate <sup>e</sup>		0

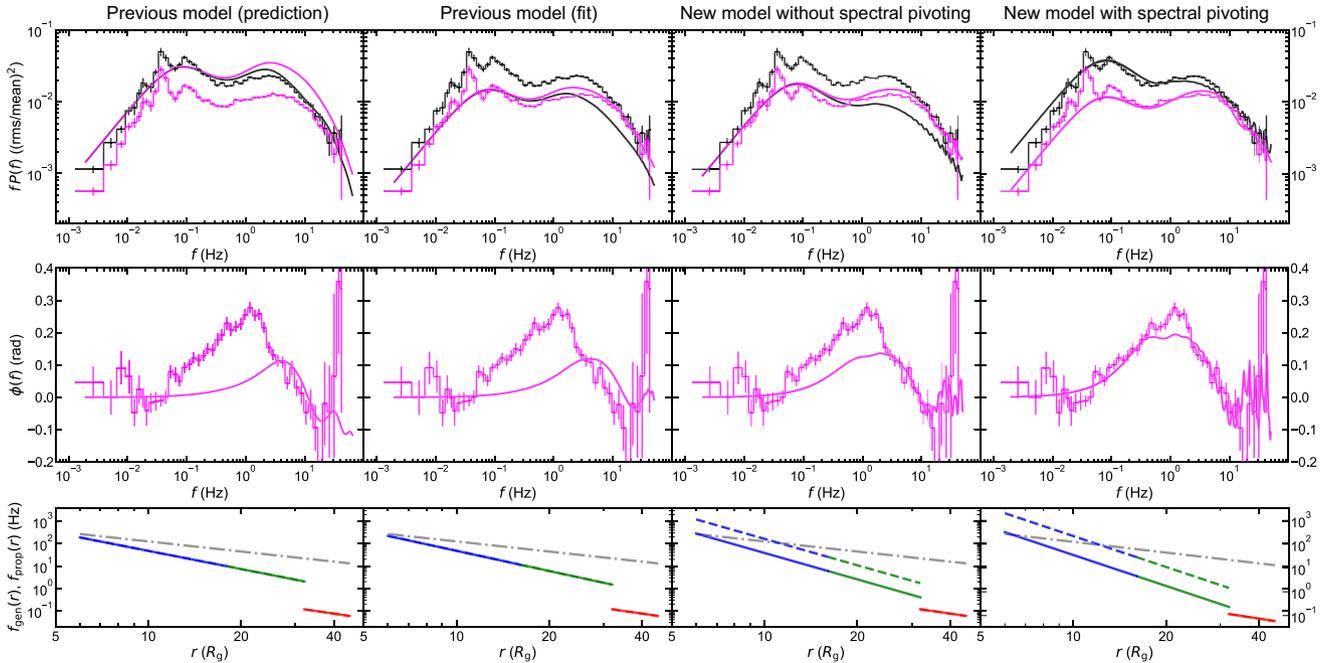
<sup>a</sup>Parameters are required when the reverberation is considered.

<sup>b</sup>Each spectral component has its own parameter:  $S_0(E)$  consists of  $S_d(E)$ ,  $S_s(E)$ ,  $S_s^{(r)}(E)$ ,  $S_h(E)$ , and  $S_h^{(r)}(E)$  (see Section 3).

<sup>c</sup> $S_0(E)$  is replaced by  $\eta(E)S_0(E)$  for timing fits when the spectral pivoting is included (Section 4.2).

<sup>d</sup> $S_0(E)$  is calculated from spectral models for spectral-timing fits (Section 5).

<sup>e</sup>Parameters are required for spectral-timing fits (Section 5).



**Figure 3.** The effect of the model updates on the timing properties. Each figure shows the low (2.4–4.8 keV; black) and high (35–48 keV; magenta) energy band power spectra (upper) and phase-lag spectra (middle), with data shown as the stepped line with errors and the model as the smooth curve. The lower panel shows the propagation frequency (solid) and generator frequency (dashed) used in the model calculations, with the Keplerian frequency (dash-dotted) for reference. Left: predictions from the previous model from K22 (Section 3) built from a full spectral-timing fit to the 0.5–10 keV data. Mid-left: fitting with the previous model (Section 3), ignoring the time-averaged spectrum. Mid-right: extending the model to include a different propagation and generator time-scale (Rapisarda et al. 2017). This shifts the frequency of the phase-lag peak but does not change the power spectra. We also gave the model the freedom to include damping (Mahmoud & Done 2018b), but the best-fitting value was close to zero, so this is not shown (Section 4.1). Right: including spectral pivoting and a difference in generator and propagation time-scale (Section 4.2). This allows the power spectral normalization of the high-energy band to be lower than at low energies, giving a significant improvement in the consistency of the model calculations.

**Table 2.** Model parameter values used in Fig. 3. Common parameter values used in all fitting are  $M_{\text{BH}} = 8$ ,  $N_r = 40$ ,  $r_{\text{in}} = 6$ ,  $r_{\text{out}} = 45$ ,  $B_{\text{d}} = B_{\text{d}}^{(\text{p})} = 0.03$ ,  $m_{\text{d}} = m_{\text{d}}^{(\text{p})} = 0.5$ ,  $\gamma = 3$ , and  $b(r) = 1 - \sqrt{r_{\text{in}}/r}$ . The transition radii are fixed to  $r_{\text{sh}} = 17.8$ ,  $r_{\text{ds}} = 32.1$  for the left-hand column and  $r_{\text{sh}} = 16$ ,  $r_{\text{ds}} = 32$  for the other columns, although these differences are too subtle to become important. Other constraints are  $F_{\text{var,f}} = F_{\text{var,d}}$  and  $D = 0$ . The mark ‘(f)’ means that a value of the corresponding parameter is fixed.

Symbol	Left	Mid-left	Mid-right	Right <sup>a</sup>
$F_{\text{var,d}}$	0.8 (f)	0.53	0.59	0.8 (f)
$B_{\text{f}}$	6 (f)	11.9	342	560
$m_{\text{f}}$	1.2 (f)	1.50	2.42	2.73
$B_{\text{f}}^{(\text{p})}$	$=B_{\text{f}}$	$=B_{\text{f}}$	80.1	166
$S_{\text{d}}(2.6\text{--}4.8\text{ keV})$	0.001 (f)	0 (f)	0 (f)	0 (f)
$S_{\text{s}}(2.6\text{--}4.8\text{ keV})$	0.356 (f)	0.505	0.471	0.305
$S_{\text{h}}(2.6\text{--}4.8\text{ keV})$	0.330 (f)	$=1 - S_{\text{s}}(E)$	$=1 - S_{\text{s}}(E)$	0.571
$S_{\text{s}}^{(\text{r})}(2.6\text{--}4.8\text{ keV})$	0.307 (f)	0 (f)	0 (f)	0 (f)
$S_{\text{h}}^{(\text{r})}(2.6\text{--}4.8\text{ keV})$	0.006 (f)	0 (f)	0 (f)	0 (f)
$S_{\text{d}}(35\text{--}48\text{ keV})$	0 (f)	0 (f)	0 (f)	0 (f)
$S_{\text{s}}(35\text{--}48\text{ keV})$	0.213 (f)	0.348	0.277	-0.056
$S_{\text{h}}(35\text{--}48\text{ keV})$	0.474 (f)	$=1 - S_{\text{s}}(E)$	$=1 - S_{\text{s}}(E)$	0.476
$S_{\text{s}}^{(\text{r})}(35\text{--}48\text{ keV})$	0.134 (f)	0 (f)	0 (f)	0 (f)
$S_{\text{h}}^{(\text{r})}(35\text{--}48\text{ keV})$	0.179 (f)	0 (f)	0 (f)	0 (f)

<sup>a</sup>  $S_0(E)$  means  $\eta(E)S_0(E)$  in this column.

giving the K22 model maximal freedom to fit without constraints from the time-averaged energy spectrum. The K22 model always has a high-energy-band power spectrum similar to that in the low-energy band everywhere except at the highest frequencies. Yet the data have very different power spectral normalizations even at low frequencies where propagation should dominate.

Plainly, while the previous model from K22 was designed to fit the data below 10 keV, it does not extrapolate to the higher energies, so does not adequately describe the physics of the propagation of fluctuations through the flow. This is important as K22 show that the propagation speed is a key determinant of the nature of the hot flow, which can allow large-scale magnetically dominated flows (magnetically arrested disc – MAD) to be distinguished from those with turbulent dynamo (standard and normal evolution – SANE) models. The poor applicability of our previous model to higher energy bands motivates our study to improve it.

#### 4.1 Suppressing variability at high energies with a constant spectral shape

The major feature missing in the previous model for the power spectra is the strong suppression of fractional variability at high energies. The generation/propagation of fluctuations in the model, where slower fluctuations generated outer regions propagate down through the flow, always leads to an increase in variability with energy, as long as the spectrum hardens inwards. In contrast, the *Insight-HXMT* observation data show that plainly the high-energy broad-band power spectrum is a factor of  $\sim 3$  lower than the low-energy power spectrum at all frequencies (Fig. 1, middle). This decrease in fractional variability with energy was not seen in the *NICER* energy band ( $\lesssim 10$  keV; K22). But it has been seen before, in e.g. the *RXTE* data of other black hole binary low/hard states (e.g. Nowak et al. 1999; Axelsson & Done 2018 for Cyg X-1; Malzac et al. 2003 for XTE J1118+480). In the context of other propagating fluctuations models, it was modelled by the damping of high-frequency fluctuations as they propagate inwards (Arévalo & Uttley 2006; Rapisarda et al. 2017), and by decreasing the intrinsic

variability power generated in the inner regions (Mahmoud & Done 2018b). To implement these effects in our model, we introduce two new parameters. One is a damping parameter  $D$ , which suppresses high-frequency variability by  $\exp(-Df\Delta t)$ , where  $\Delta t$  is the propagation time. The damping effect is ignored if  $D = 0$ . We also allow the intrinsic variability amplitude to be different between the hot flow  $F_{\text{var,h}}$  and disc  $F_{\text{var,d}}$  (the previous model from K22 has  $F_{\text{var,f}} = F_{\text{var,d}}$ ).

Another observational feature that our previous model fails to capture is the discrepancy in the frequency at which the power spectra (in the  $fP(f)$  representation) and phase-lag spectra peak. The power spectra and phase-lag spectra calculated by the K22 model have a similar peak frequency. This observational property is also seen in the *RXTE* data (e.g. XTE J1550–564; Rapisarda et al. 2017), where the proposed solution was to allow the propagation time-scale to be different from the generator time-scale on which the fluctuations are generated. Following this, we separate these time-scales and define the propagation frequency with

$$f_{\text{prop}}(r) = \begin{cases} B_{\text{f}}^{(\text{p})} r^{-m_{\text{f}}^{(\text{p})}} f_{\text{K}}(r) & (r_{\text{in}} \leq r_n < r_{\text{ds}}), \\ B_{\text{d}}^{(\text{p})} r^{-m_{\text{d}}^{(\text{p})}} f_{\text{K}}(r) & (r_{\text{ds}} \leq r_n < r_{\text{out}}), \end{cases} \quad (4)$$

such that the propagation speed is provided by  $v_{\text{p}}(r) = rf_{\text{prop}}(r)$ . It is difficult to constrain disc parameters as the disc emission has negligible contributions to the energy range of interest. Thus we keep employing  $f_{\text{gen}}(r) = f_{\text{prop}}(r)$  in the variable disc region with  $(B_{\text{d}}^{(\text{p})}, m_{\text{d}}^{(\text{p})}) = (B_{\text{d}}, m_{\text{d}}) = (0.03, 0.5)$ . To reduce the number of free parameters, we assume that  $f_{\text{prop}}(r)$  has the same radial dependence as  $f_{\text{gen}}(r)$ , i.e.  $m_{\text{f}}^{(\text{p})} = m_{\text{f}}$ . Eventually, we have only one additional parameter  $B_{\text{f}}^{(\text{p})}$ .

The modified model formalism due to the damping effect is given in Appendix B. Other additional effects,  $F_{\text{var,d}} \neq F_{\text{var,f}}$  and  $f_{\text{gen}}(r) \neq f_{\text{prop}}(r)$ , just alter the power spectrum of intrinsic mass accretion rate variability at  $n$ th ring  $|A(r_n, f)|^2$  (equation A1) and the propagation time from the outer  $k$ th ring to the inner  $n$ th ring  $\Delta t_{k,n}$  (equation A3), respectively, without affecting any other equations containing  $|A(r_n, f)|^2$  and  $\Delta t_{k,n}$ . As in the last part of the previous section, we attempt to reproduce only variability properties based on the propagating fluctuations process rather than full spectral timing. We keep those parameters fixed that are fixed in the previous fit.

Even with all these additional effects, the model is still not capable of matching the observation data. The damping parameter  $D$  is pegged to its lower bound of zero, indicating that the damping described above is ineffective in improving the fit (Mahmoud et al. 2019). This is because our model assumes that the intrinsic variability has a cut-off at the local generator frequency  $f_{\text{gen}}(r)$ , as shown in Fig. 2(c). This assumption already includes some aspects of damping. The MRI (Balbus & Hawley 1991, 1998) is expected to produce variability up to quite fast time-scales. However, only variability slower than the local propagation time can propagate inwards as the faster variability is viscously damped out (Churazov, Gilfanov & Revnivtsev 2001; Cowperthwaite & Reynolds 2014; Hogg & Reynolds 2016; Ingram 2016; Bollimpalli et al. 2020; Turner & Reynolds 2021). Our assumptions about the intrinsic variability are an approximation of this physical picture. The damping parameter being pegged to zero indicates there is no need for additional damping effects. We did not find an improvement in the fits using separate variability amplitude between the variable disc region and hot flow region, either.

Fig. 3 (mid-right) shows the results of a joint fit to the power spectra for 2.6–4.8 and 35–48 keV and the phase-lag spectrum between these energy bands by allowing  $f_{\text{gen}}(r) \neq f_{\text{prop}}(r)$  for the hot flow. For

clarity, we removed the other additional effects that did not make a difference, i.e.  $F_{\text{var,d}} = F_{\text{var,f}}$  and  $D = 0$ .

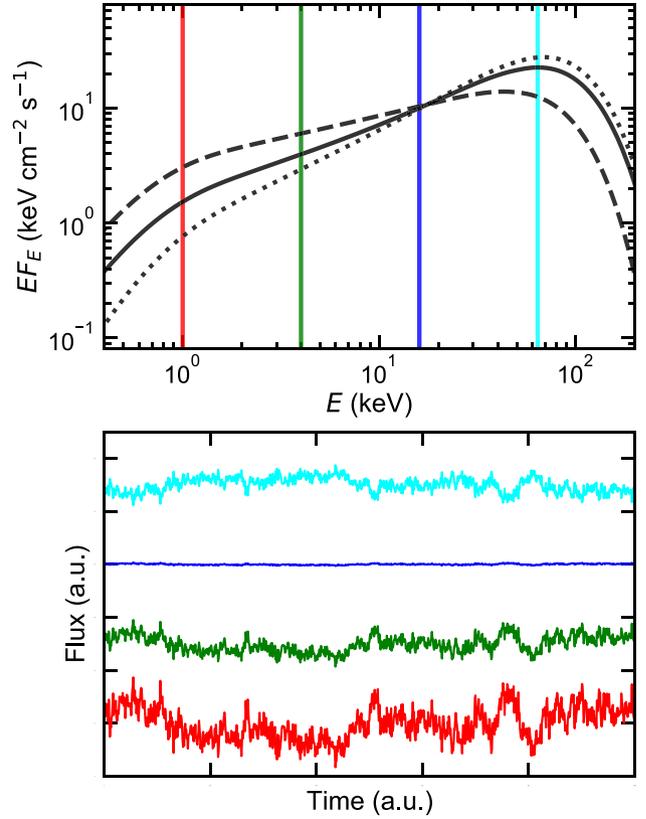
Although we see a slightly better match to the observed peak frequency of the phase-lag spectrum than the previous fit, the model still underestimates its amplitude. More importantly, we still do not solve the essential issue: the model calculates similar or larger variability for higher energy bands, inconsistent with the observation that the power spectral amplitude is larger for the lower energy band. Mahmoud & Done (2018b) and Mahmoud et al. (2019) introduce more complex radial dependence for the intrinsic variability, emissivity, and damping to capture energy-dependent variability properties. However, some assumptions involved with these complications remain to be tested. We do not explore the complex radial structure further and conclude that those additional effects implemented here are less effective than required by the high signal-to-noise ratio data obtained by *Insight-HXMT*.

We note that the difficulty in reproducing the observation data here lies in joint fitting to the power spectra and phase-lag spectrum. It is possible to reproduce power spectra for these energy bands with the current model fairly well, ignoring the phase-lag spectrum. In this case, however, the lower energy photons would come from inner regions, because inner regions are more variable than outer regions, and predict soft lags, which is completely in disagreement with the observed hard lags. This points to the importance of modelling cross-spectra and power spectra.

## 4.2 Spectral pivoting

So far, we have assumed that the spectral shape of each component does not vary in time. However, this is unphysical because mass accretion rate fluctuations make spectral parameters, e.g. the optical depth and electron temperature, vary on short time-scales (Malzac et al. 2003; Gandhi et al. 2008; Yamada et al. 2013; Bhargava et al. 2022). This oversimplification limits the model's flexibility to reproduce energy-dependent variability data. Hence we now allow the spectral shapes to fluctuate (Veledina 2016, 2018; Mastroserio et al. 2018, 2019, 2021), along with their amplitude. The schematic picture of the spectral pivoting is shown in Fig. 4 (top).

Here, we give concise explanations of how the spectral pivoting is implemented and what the model gets to be able to handle with this update. More detailed formalism is found in Appendix C. A constant spectral shape means that the spectrum at every energy reacts to mass accretion fluctuations in the same way. We consider the mass accretion rate and energy spectrum at a certain radius. By defining the average and difference from the average as  $\dot{m}_0$  and  $\Delta\dot{m}(t)$  for the mass accretion rate and as  $S_0(E)$  and  $\Delta S(E, t)$  for the spectrum, the constant spectral shape is equivalent to  $\Delta S(E, t)/S_0(E) = \Delta\dot{m}(t)/\dot{m}_0$ , which is independent of energy  $E$ . To let the spectral shape vary in time, we give the spectrum sensitivity to  $\Delta\dot{m}(t)$  as a function of energy,  $\eta(E)$ , and redefine  $\Delta S(E, t)/S_0(E) = \eta(E)\Delta\dot{m}(t)/\dot{m}_0$ , which now depends on energy. The amplitude of sensitivity parameter  $|\eta(E)|$  regulates how sensitive the spectrum is to a change in the mass accretion rate from its average, while its sign determines whether the spectrum reacts positively or negatively. The spectrum gets higher (lower) with an increase in mass accretion rate if  $\eta(E) > 0$  ( $< 0$ ). The energy at which  $\eta(E)$  crosses zero, called the pivoting point, does not react to a change in mass accretion rate. Light curves of local flux for different energies are illustrated in Fig. 4 (bottom). We note that we do not simulate light curves in the model calculations. The decrease in  $\eta(E)$  with energy, i.e. the spectrum being less sensitive to  $\Delta\dot{m}(t)$  for higher energies, could let the power spectrum decrease with energy, as observed for MAXI J1820+070, even if the mass



**Figure 4.** Schematic picture of the effect of spectral pivoting as implemented here. Top: instant local energy spectra when the mass accretion rate at the corresponding radius is higher than, equal to, and lower than the average (dashed, solid, and dotted, respectively). Bottom: light curves of the local flux for different energies, 1 keV (red), 4 keV (green), 16 keV (blue), 64 keV (cyan) as marked in the top panel. Each light curve is normalized by its average and offset for clarity. The light curves in the lower energy bands (red and green) are positively correlated with the local mass accretion rate, but the fluctuations have lower amplitude as the energy increases, going to zero at the pivot point at 16 keV, and then switching to negative correlation at higher energies (cyan).

accretion rate is more variable for central regions emitting higher energy photons. In our implementation, there arises no lag between different energies from the spectral pivoting itself except for the phase lag of  $\pi$  when  $\eta(E_1)\eta(E_2) < 0$ . Our new model shares this feature of spectral pivoting with the model developed by Veledina (2016, 2018). The new model returns to the previous one by setting  $\eta(E) = 1$ .

Each spectral component is expected to show its own sensitivity pattern. We give the sensitivity parameter to each spectral component,  $\eta_Y(E)$  ( $Y = \text{d, s, h}$ ), where the subscripts stand the variable disc, soft Comptonization, and hard Comptonization, respectively. With the implementation of spectral pivoting, all  $S_Y(E)$  ( $Y = \text{d, s, h}$ ) contained in the analytic expressions of power spectra and cross-spectra is replaced by  $\eta_Y(E)S_Y(E)$  (see Appendix C for the derivation). This means that the model's flexibility is not bound by the constraint (3) anymore because time-averaged spectra always appear as the product with their sensitivity. In addition,  $\eta_Y(E)S_Y(E)$  can be negative in contrast to  $0 \leq S_Y(E) \leq 1$ . The spectral pivoting gives freedom to the model in this way.

We attempt to fit the variability properties with the new model. We have  $\eta_Y(E)S_Y(E)$  as model parameters, instead of  $S_Y(E)$ . The

negligible disc emission  $S_d(E) = 0$  results in  $\eta_d(E)S_d(E) = 0$ . We fix  $D = 0$ , in which all intrinsic variability propagates inwards without any loss. We also fix  $F_{\text{var,d}} = F_{\text{var,f}}$  to the typical value of 0.8 because the sensitivity parameter  $\eta(E)$  can regulate the variability amplitude.

The simultaneous fit to the power spectra for 2.6–4.8 and 35–48 keV and the phase-lag spectrum between these energy bands with the new model is shown in Fig. 3 (right). We see significant improvement in variability modelling by allowing the spectral shapes to vary in time. Our new model captures the energy-dependent variability, pointing to the importance of spectral pivoting in modelling variability at high energies.

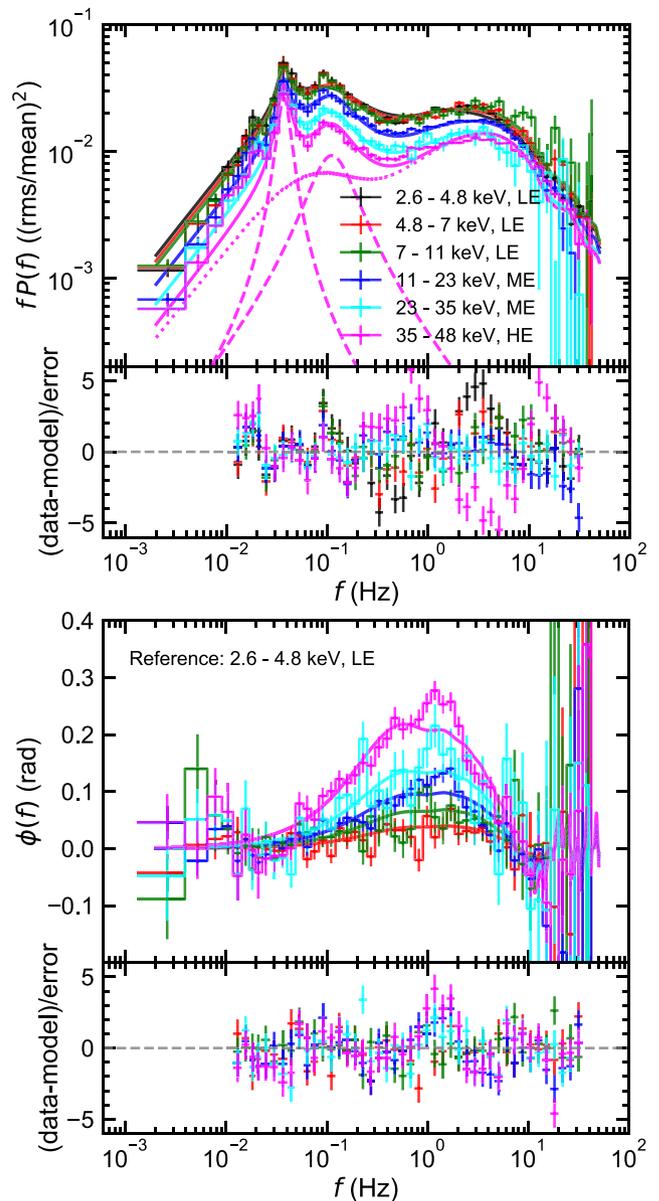
To study the variability for a continuous energy range, we split energy between 2.6–4.8 keV (LE) and 35–48 keV (HE) into four bands, i.e. 4.8–7 keV (LE), 7–11 keV (LE), 11–23 keV (ME), and 23–35 keV (ME), where the telescopes used are specified in parenthesis, and attempt to reproduce power spectra for these six energy bands and phase-lag spectra with respect to the lowest energy band for the rest of five energy bands. We minimize

$$\sum_{j,k} \left( \frac{P_{\text{data}}(E_j, f_k) - P_{\text{model}}(E_j, f_k)}{\Delta P_{\text{data}}(E_j, f_k)} \right)^2 + \sum_{\substack{j,k \\ (E_j \neq E_r)}} \left( \frac{\phi_{\text{data}}(E_r, E_j, f_k) - \phi_{\text{model}}(E_r, E_j, f_k)}{\Delta \phi_{\text{data}}(E_r, E_j, f_k)} \right)^2 \quad (5)$$

through the fit, where  $E_j$  is each energy band and  $E_r = 2.6\text{--}4.8$  keV the reference band. For more complete modelling, we add two Lorentzian functions to model the QPOs in power spectra by using the XSPEC model `lorentz`. To fix the centroid frequency and width of the QPO models, we perform a phenomenological fit to power spectra with the sum of four Lorentzian functions, where two of them are used to model each bump of the broad-band variability. We then extract the centroid and width of  $(3.66 \times 10^{-2} \text{ Hz}, 1.20 \times 10^{-2} \text{ Hz})$  for the QPO fundamental, and  $(9.44 \times 10^{-2} \text{ Hz}, 1.16 \times 10^{-1} \text{ Hz})$  for the second harmonic as typical values. Thus, the `lorentz` model has only one free parameter, the normalization. On the other hand, we do not use any additional models in phase-lag spectra due to the relatively small QPO features. The results of the joint fit to six power spectra and five phase-lag spectra are shown in Fig. 5. Each component forming power spectra is explicitly plotted with dashed (QPOs) and dotted (broad-band) lines only for the highest energy band of 35–48 keV. Model parameter values are summarized in Table 3.

We find that the new model matches observations well for all energy bands whilst keeping parameter values similar to those found in the joint fitting for 2.4–4.8 and 35–48 keV only (Fig. 3, right). It is interesting to note that the spectral parameter for the soft Comptonization component  $\mu_s(E)S_s(E)$  decreases with energy and finally reaches a negative value at the highest energy band of 35–48 keV. This means that the soft Comptonization component increases for an increase in mass accretion rate at low energies ( $\lesssim 35$  keV), whereas it decreases at high energies ( $\gtrsim 35$  keV), showing the pivoting point of  $\sim 35$  keV.

Although the broad-band variability has been studied with *Insight-HXMT* observations (e.g. Wang et al. 2020; Yang et al. 2022), we succeeded in reproducing it with a physically motivated model for the first time. In addition, while propagating fluctuations models have been applied up to  $\sim 35$  keV (Mahmoud & Done 2018a,b) with *RXTE* observations, we extend the energy range up to 48 keV using *Insight-HXMT* observations with significantly improved residuals. Our successful modelling shows the propagating fluctuations scenario



**Figure 5.** Joint fit to six power spectra (top) and five phase-lag spectra (bottom) across 2.6–48 keV with our new model including the spectral pivoting. We include two Lorentzian functions for the QPO (dashed) and harmonic (dotted). In the calculation of phase-lag spectra, the lowest band of 2.6–4.8 keV is chosen as the reference band. The lower plot for each panel is the difference between data and model divided by  $1\sigma$  errors. The new model including the spectral pivot successfully reproduces all the timing data across this bandpass.

holds good up to high-energy bands, keeping it the most plausible explanation for the aperiodic variability.

## 5 JOINT SPECTRAL-TIMING FIT WITH SPECTRAL PIVOTING

We come back to implement a full spectral-timing analysis, rather than just the series of timing analyses above. We attempt to fit the energy-dependent timing properties from 2.6–48 keV along with the time-average energy spectrum at the corresponding energy range by

**Table 3.** Model parameter values used in Fig. 5. Fixed parameter values are the same as in Table 2.

Symbol	Value
$F_{\text{var,d}}$	0.8 (f)
$B_f$	401
$m_f$	2.59
$B_f^{(p)}$	108
$\eta_s S_s(2.6\text{--}4.8\text{ keV})$	0.315
$\eta_h S_h(2.6\text{--}4.8\text{ keV})$	0.588
$\eta_s S_s(4.8\text{--}7\text{ keV})$	0.242
$\eta_h S_h(4.8\text{--}7\text{ keV})$	0.601
$\eta_s S_s(7\text{--}11\text{ keV})$	0.189
$\eta_h S_h(7\text{--}11\text{ keV})$	0.617
$\eta_s S_s(11\text{--}23\text{ keV})$	0.116
$\eta_h S_h(11\text{--}23\text{ keV})$	0.549
$\eta_s S_s(23\text{--}35\text{ keV})$	0.049
$\eta_h S_h(23\text{--}35\text{ keV})$	0.498
$\eta_s S_s(35\text{--}48\text{ keV})$	−0.052
$\eta_h S_h(35\text{--}48\text{ keV})$	0.484
$\chi^2/\text{degrees of freedom}$	1095.0/457

minimizing

$$\begin{aligned} & \sum_i \left( \frac{S_{\text{data}}(E_i) - S_{\text{model}}(E_i)}{\Delta S_{\text{data}}(E_i)} \right)^2 \\ & + \sum_{j,k} \left( \frac{P_{\text{data}}(E_j, f_k) - P_{\text{model}}(E_j, f_k)}{\Delta P_{\text{data}}(E_j, f_k)} \right)^2 \\ & + \sum_{\substack{j,k \\ (E_j \neq E_i)}} \left( \frac{\phi_{\text{data}}(E_r, E_j, f_k) - \phi_{\text{model}}(E_r, E_j, f_k)}{\Delta \phi_{\text{data}}(E_r, E_j, f_k)} \right)^2, \end{aligned} \quad (6)$$

where  $S_{\text{data}}(E)$  and  $\Delta S_{\text{data}}(E)$  are the observed time-averaged spectrum and its  $1\sigma$  error,  $S_{\text{model}}(E)$  the modelled time-averaged spectrum, and  $E_i$  each energy bin in the time-averaged spectrum. We remove clear calibration features seen in the ME spectrum for 20–24 keV (light green regions in Fig. 1, top) from the spectral modelling.

To model the energy spectrum, we account for not only the soft and hard Comptonization components but their disc reflection. We ignore emission from the turbulent disc due to its negligible contribution above the lowest energy of 2.6 keV (few per cent at 2.6 keV in the spectral fit found in K22). We also ignore the negligible effect of galactic absorption. We use the XSPEC model `nthcomp` (Zdziarski, Johnson & Magdziarz 1996; Życki, Done & Smith 1999) for the Comptonization components, and `relxillCp` provided in RELXILL version 2.0 (García et al. 2014; Dauser et al. 2022) for the reflected components. Finally, we use

$$(\text{nthcomp} + \text{relxillCp}) + (\text{nthcomp} + \text{relxillCp}), \quad (7)$$

where each bracket corresponds to the soft Comptonization/reflection and hard Comptonization/reflection, respectively.

To connect the time-averaged spectrum and variability consistently, we take reverberation into account in our timing model. Its implementation is updated from that in K22 mainly due to the inclusion of spectral pivoting. We summarize how the reverberation behaves in our new model here, while the detailed formalism is described in Appendix D.

The illuminating Comptonization spectrum changing its shape with time results in the reflected spectrum also changing its shape

with time. As in the previous section, we consider a certain radius. Along with the mass accretion rate and direct emission, we account for the reflected emission associated with the direct emission. Defining the average and difference from it as  $S_0^{(r)}(E)$  and  $\Delta S^{(r)}(E, t)$ , we assume  $\Delta S^{(r)}(E, t)/S_0^{(r)}(E) = (\Delta S(E, t)/S_0(E)) \otimes h(t) = \eta(E)(\Delta \dot{m}_0(t)/\dot{m}_0) \otimes h(t)$ . We use the superscript ‘(r)’ to stand for the reflected emission. The convolution in time is denoted by  $\otimes$ , and  $h(t)$  is called the impulse response, which is the time evolution of reflected emission for a flash of illumination. All information as to the disc response, such as the delay for the direct emission due to an additional light crossing path and the duration due to the different delay times for different locations of reflection, are encoded in  $h(t)$ .

The relation of spectral variation between the direct and reflected emission means that the reflected emission follows variations of the direct emission at the corresponding energy with some time delay, as long as the variability is slow enough not to be washed out via reprocessing, i.e. via the operation of the convolution. In the simple case of  $h(t) = \delta(t - \tau)$ , variations of the reflected emission exactly lag behind those of the direct emission with the time delay of  $\tau$ :  $\Delta S^{(r)}(E, t)/S_0^{(r)}(E) = \Delta S(E, t - \tau)/S_0(E)$ .

Each reflected component has its own impulse response as each Comptonization component illuminates different parts of the accretion flow (Zdziarski et al. 2021a). We define the impulse response with a top-hat function:

$$h_Y(t) = \begin{cases} 1/\Delta t_{0,Y} & (|t - t_{0,Y}| < \Delta t_{0,Y}/2), \\ 0 & (\text{otherwise}), \end{cases} \quad (8)$$

where  $Y = s, h$  are associated with the soft and hard Comptonization components, respectively. The parameters  $t_{0,Y}$  and  $\Delta t_{0,Y}$  characterize the delay and duration, respectively. More realistic impulse responses are required, especially for low-energy bands  $E \lesssim 2$  keV, where the quasi-thermal emission due to the reprocessing dominates high-frequency variability ( $\gtrsim 1$  Hz). However, the top-hat function appears to be a good approximation for high energies, where Comptonization largely determines variability properties.

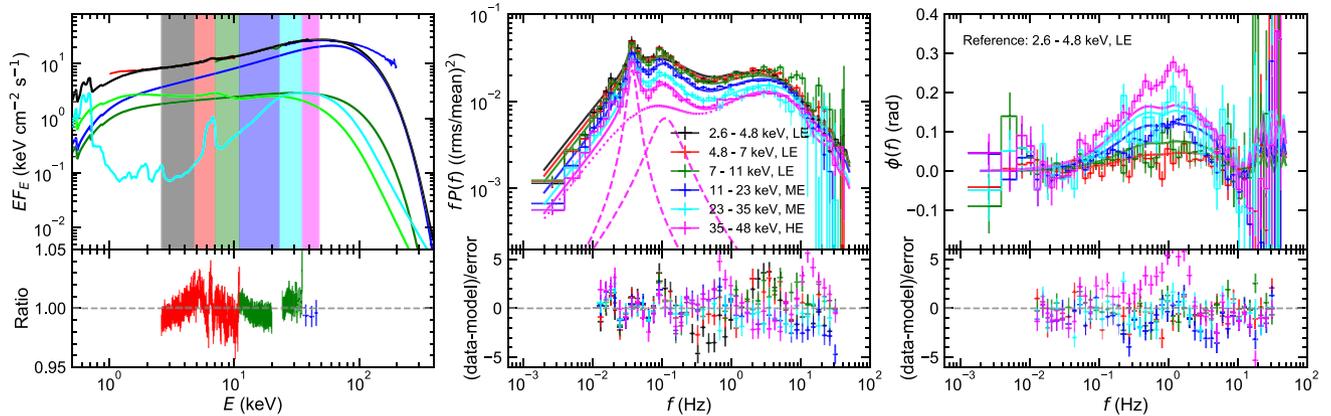
For the consistency between the spectral modelling and variability modelling, we calculate the fractional time-averaged spectra required in our timing model,  $S_d(E)(= 0)$ ,  $S_s(E)$ ,  $S_h(E)$ ,  $S_s^{(r)}(E)$ ,  $S_h^{(r)}(E)$ , from the spectral models, `nthcomp` and `relxillCp`. To obtain  $\eta_s(E)S_s(E)$ ,  $\eta_h(E)S_h(E)$ ,  $\eta_s(E)S_s^{(r)}(E)$ , and  $\eta_h(E)S_h^{(r)}(E)$ , it is simple to assume a function for the sensitivity parameters  $\eta_Y(E)$  ( $Y = s, h$ ). We note that we do not need  $\eta_d(E)$  due to  $S_d(E) = 0$ . Given that  $\eta_Y(E)$  can switch its sign at the pivoting point and that the fractional rms of the broad-band variability roughly changes with energy logarithmically (Gierliński & Zdziarski 2005; Yang et al. 2022), it is fair to make a phenomenological assumption of

$$\eta_Y(E) = \eta_{0,Y} + \eta_{1,Y} \log_{10}(E [\text{keV}]). \quad (9)$$

The model parameter  $\eta_{0,Y}$  is the sensitivity at 1 keV, while  $\eta_{1,Y}$  determines its gradient to energy.

We note the difference in the model calculations between the timing fits (Section 4.2) and spectral-timing fits. In the timing fits,  $\eta(E)S_0(E)$  is a model parameter, and it is impossible to disentangle this product. On the other hand,  $S_0(E)$  and  $\eta(E)$  are separately modelled in the spectral-timing fits. The former is calculated from spectral models, the latter is from equation (9).

In the joint spectral-timing fit, we fix the seed photon temperature of Comptonization components to the typical disc temperature in this state,  $kT_{\text{seed},s} = kT_{\text{seed},h} = 0.2$  keV (De Marco et al. 2021; K22). Since the electron temperature is difficult to constrain from the energy band of interest, we fix it to  $kT_{e,s} = kT_{e,h} = 23$  keV, as in K22. While we allow the inner radius of the reflector for the hard Comptonization



**Figure 6.** Spectral-timing fit to the time-averaged energy spectrum (left), six power spectra (middle), and five phase-lag spectra (right) across 2.6–48 keV with our new model including spectral pivoting. In the left-hand panel, the soft Comptonization and its associated reflection are plotted with the green and light green lines, while the hard Comptonization and its associated reflection are plotted with the blue and light blue lines. The black line shows their sum. We include the effect of galactic absorption of  $N_{\text{H}} = 1.4 \times 10^{21} \text{ cm}^{-2}$  on the spectral components. The colours of the shaded regions in the energy spectral plot show the energy band used in the power spectra and phase-lag spectra. In the middle panel, the power spectrum at the highest energy band includes the QPO and harmonic, as in Fig. 5. The bottom panels show residuals. The data-to-model ratio is used for the energy spectrum, while the difference between data and model divided by  $1\sigma$  errors is used for power spectra and phase-lag spectra. We successfully fit all the data in this bandpass with our updated model including the spectral pivoting, except for a slight underestimate of the phase-lag spectrum around the peak for the highest energy band.

component to be free, we fix that for the soft Comptonization component to  $R_{\text{in},s} = 45R_{\text{g}}$  corresponding to the outer edge of the variable flow located at  $r_{\text{out}} = 45$ . Following K22, we fix the inclination angle to  $i = 66^\circ$  (Torres et al. 2020) and Fe abundance to  $Z_{\text{Fe}} = 1.1$ . We also set the black hole spin to  $a^* = 0$ , consistent with  $r_{\text{in}} = 6$  in the timing model, and use the high electron density of  $N_e = 10^{20} \text{ cm}^{-3}$  (García et al. 2016; Mastroserio et al. 2021). The delay and duration of the impulse response are, in principle, derived from the location and geometry of illuminating source and reflector. However, in the geometry assumed, the time-scales of reverberation  $\lesssim 10$  ms (corresponding to the light crossing of  $\lesssim 250R_{\text{g}}$  for  $M_{\text{BH}} = 8 M_{\odot}$ ) are shorter than variability time-scales of interest (20 ms–100 s). In addition, reverberation signatures are unclear across the energy bands of interest (2.6–48 keV), and small alterations of the impulse response due to small changes of the accretion flow geometry do not significantly affect the variability properties. Thus, we simply fix  $t_{0,s} = t_{0,h} = 6$  ms and  $\Delta t_{0,s} = \Delta t_{0,h} = 10$  ms as typical values. The top-hat impulse response with these values appears to be good approximations of more realistic ones (K22).

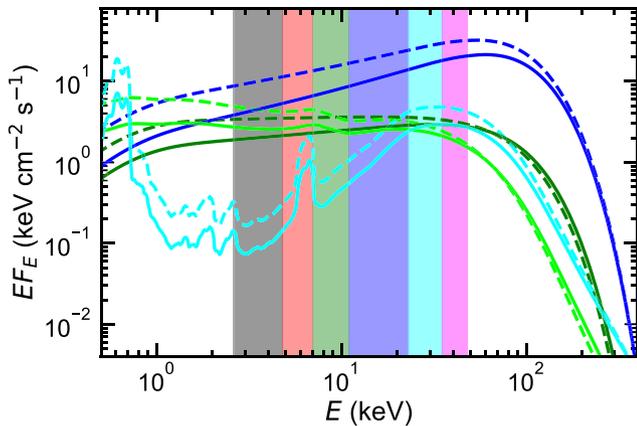
The results of simultaneous modelling of the energy spectrum, six power spectra, and five phase-lag spectra are shown in Fig. 6. The comparison between the data and model is also plotted as the ratio for the energy spectrum and the difference divided by  $1\sigma$  errors for the variability. Model parameter values are found in Table 4. Overall, our new model successfully reproduces both time-averaged and variability properties, although the discrepancies are seen in the phase-lag spectrum between 35–48 and 2.6–4.8 keV (magenta), which is discussed in Section 6.3. This modelling is the first simultaneous fit to spectrum and variability using our model. The uncertainties of the derived parameter values are evaluated with a Markov chain Monte Carlo (MCMC) analysis in Appendix E.

The spectral variation derived from the fit is shown in Fig. 7. The spectra for the mass accretion rate being its average and double the average are plotted with solid and dashed lines. For illustration purposes, we ignore all effects from the impulse response for reverberation, such as time delay, i.e. we assume  $h(t) = \delta(t)$ . This means that the Comptonization and its associated reflection behave completely in the same way,  $\Delta S(E, t)/S_0(E) = \Delta S^{(r)}(E, t)/S_0^{(r)}(E)$ .

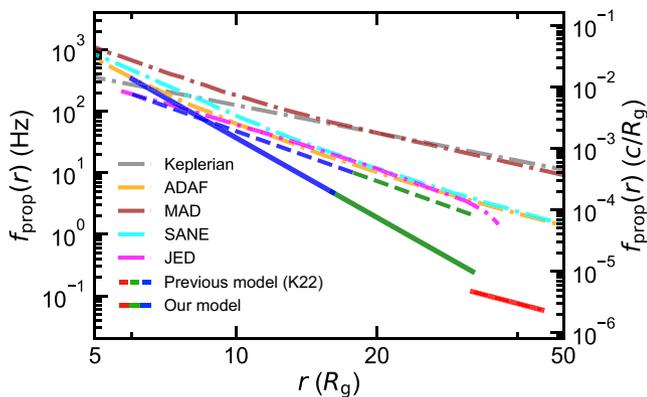
**Table 4.** Model parameter values derived from the joint spectral-timing fit in Fig. 6. Fixed parameters related to the spectrum are the seed photon temperature  $kT_{\text{seed},s} = kT_{\text{seed},h} = 0.2$  keV, electron temperature  $kT_{e,s} = kT_{e,h} = 23$  keV, Fe abundance  $Z_{\text{Fe}} = 1.1$ , inclination angle  $i = 66^\circ$ , black hole spin  $a^* = 0$ , electron density  $N_e = 10^{20} \text{ cm}^{-3}$ , inner radius of reflection region  $R_{\text{in},s} = 45R_{\text{g}}$ , and outer radii of reflection region  $R_{\text{out},s} = R_{\text{out},h} = 1000R_{\text{g}}$ . The subscript ‘s’ (‘h’) denotes the soft (hard) Comptonization or its associated reflection component. Fixed parameters related to variability are the same as in Table 2 in addition to the extra parameters about reverberation,  $t_{0,s} = t_{0,h} = 6$  ms and  $\Delta t_{0,s} = \Delta t_{0,h} = 10$  ms.

Component	Model	Symbol	Value
Spectral parameters			
Soft Comptonization and reflection	nthcomp	$\Gamma_s$	1.81
		$\text{norm}_s$	1.38
	relxillCp	$\log_{10}\xi_s$	3.44
		$\text{norm}_s^{(r)}$	0.0395
Hard Comptonization and reflection	nthcomp	$\Gamma_h$	1.50
		$\text{norm}_h$	2.11
	relxillCp	$R_{\text{in},h}$	78
		$\log_{10}\xi_h$	1.70
		$\text{norm}_h^{(r)}$	0.0328
Variability parameters			
Broad-band	Our model	$B_f$	862
		$m_f$	2.81
		$B_f^{(p)}$	189
		$\eta_{0,s}$	1.023
		$\eta_{1,s}$	−0.568
		$\eta_{0,h}$	1.527
		$\eta_{1,h}$	−0.580
		$\chi^2/\text{d.o.f.}$	1993.1/1795

Generally, all spectra are less sensitive for higher energies to mass accretion rate fluctuations, which results in a decrease in the power spectrum with energy. We see the pivoting point at  $\sim 50$  keV for the



**Figure 7.** Changes in the shape of the energy spectral components due to changes in mass accretion rate derived from the spectral-timing fit in Fig. 6. The average energy spectrum for the soft (green) and hard (blue) Comptonization components and their reflected emission (pale green and cyan, respectively) are shown as solid lines, while the dashed lines show the effect of doubling the mass accretion rate.



**Figure 8.** Propagation frequency (solid) as a function of radius derived from the spectral-timing fit. Red, green, and blue colours denote the variable disc, soft Comptonization, and hard Comptonization regions, respectively. The propagation frequency derived from our previous work and predicted from theoretical models are plotted with a dashed line and dash-dotted lines (see K22 for details). A black hole mass of  $M_{\text{BH}} = 8 M_{\odot}$  is assumed. The Keplerian frequency is also plotted with a grey dash-dotted line for reference. We caution that none of these models except the JED-SAD has an explicit transition from the flow to the disc.

soft Comptonization and its reflection, which roughly agrees with that at  $\sim 35$  keV derived from the fit only to the timing properties in the previous section.

## 6 DISCUSSION

### 6.1 Generator time-scale and propagation time-scale

The characteristic time-scales on which the fluctuations are propagated at each radius are derived from the spectral-timing fit (Fig. 6). We compare the propagation frequency derived (solid) to those predicted by different hot flow models (dash-dotted) in Fig. 8. The theoretical propagation frequencies for the advection-dominated accretion flow (ADAF; Narayan, Kato & Honma 1997), standard and normal evolution (SANE; Narayan et al. 2012), magnetically arrested disc (MAD; Narayan et al. 2012), and jet emitting disc (JED; Marcel

et al. 2018) are calculated in a standard way by dividing the accretion speed by radius assuming  $M_{\text{BH}} = 8 M_{\odot}$  (see K22 for details). The propagation frequency derived from our previous model from K22 is also plotted (dashed).

In K22, we found fairly good agreement of the derived propagation time-scales with those in ADAF, SANE, and JED rather than MAD. The propagation time-scale derived from our new model is now not very similar to any theoretical predictions. Here, allowing the generator and propagation frequencies to be different makes their radial dependence steeper, i.e. from  $f_{\text{prop}}(r) \propto r^{-2.7}$  to  $f_{\text{prop}}(r) \propto r^{-4.31}$ . This steep radial dependence is required to reproduce both the observed large phase lags and high-frequency broad-band variability simultaneously. Indeed, the propagation time-scale derived is robust against uncertainties of the relationship between the generator time-scale and propagation time-scale. Using our new model including the spectral pivoting, the assumption of  $f_{\text{gen}}(r) = f_{\text{prop}}(r)$  also gives a similar propagation time-scale, although the fit is not as good as that obtained in the previous section with  $f_{\text{gen}}(r) \neq f_{\text{prop}}(r)$ . The key feature that our new model requires is a slow propagation speed enough to reproduce observed phase lags. The propagation time-scale of MAD (brown dash-dotted line) is too short to explain the observed phase lags. Thus, our results still prefer SANE rather than MAD, although MAXI J1820+070 displays a powerful jet (Bright et al. 2020).

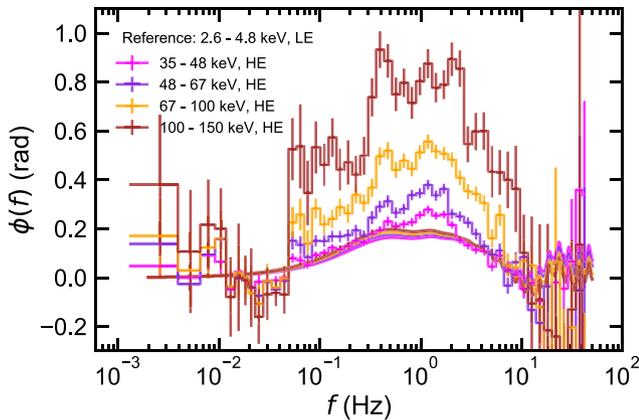
### 6.2 Origin of QPOs

Our full spectral-timing modelling accounts for all the X-ray spectrum and rapid variability except for the QPOs. We model these QPOs by adding peaked Lorentzian components in the power spectra. No extra component is added to the phase-lag spectra simply because the QPO features are not very clear across 2.6–48 keV. We did not add any other spectral components for the QPOs, implicitly assuming that the QPO is a modulation of the spectral components already included in the model (multiplicative) rather than being associated with an additional spectral component (additive).

Our successful modelling does not give much room for an additional emission component only related to the QPOs (Fig. 6), supporting the assumption above. This result is consistent with a QPO produced predominantly from a global mode of flow rather than an intrinsic change in intensity with QPO frequency. We specifically have in mind the Lense–Thirring (vertical) precession of the entire hot flow, where the observed luminosity of the Comptonization component(s), including all their stochastic variability, are modulated by the changing projected area of the translucent hot flow as the viewing angle changes with QPO phase (Fragile et al. 2007; Ingram et al. 2009; Ingram & Done 2011, 2012). This picture agrees with the new polarization results for Cyg X-1 in the low/hard state, which requires the hot X-ray emitting plasma to be radially extended (Krawczynski et al. 2022). Conversely, the alternative model of a precessing jet suggested by Ma et al. (2021) is challenged by the polarization results because it requires the hot X-ray emitting region to align with the jet.

### 6.3 Limitations of our new model

Our new model gives a poor fit to both the energy spectrum and phase lags beyond  $\sim 40$ – $50$  keV. From Fig. 6 (left), the spectral model clearly underestimates the data above  $\sim 100$  keV. The hard Comptonization spectrum rolls over too fast to match the observed data. The comparison of the phase-lag spectrum between the model and data at high energies is shown in Fig. 9. The model phase lags increase smoothly with energy to  $\sim 40$  keV (Fig. 6, right) but then



**Figure 9.** The predicted high-energy phase lags (lines) versus the data (points), calculated using the standard 2.6–5.8 keV reference band. The lowest energy band on this plot is 35–48 keV (magenta), which is the highest energy band included in the spectral-timing fits. The predicted phase lags at higher energies saturate beyond 40 keV, unlike the data, which show a clear increase in phase lag with energy, pointing out the limitations of the model.

saturates to a constant value rather than continually increasing as in the data. The lag behaviour arises as the fraction of the total spectrum that is made of the hard (and long lagged) Comptonization spectrum increases up to around  $\sim 40$  keV, but after this point, the hard Comptonization dominates, leading to the saturation of lag.

The spectral mismatch could be fixed if there is additional stratification of the energy spectrum of the hot flow, so the very innermost regions produce an even hotter/harder Comptonization component. In many ways, this is quite natural. The two Comptonization components used here for the spectral decomposition are only an approximation to a continuous flow with (presumably) continuous stratification, even if we do expect there physically to be two main regions. Close to the disc, seed photons for Comptonization are predominantly from the disc. However, it is quite easy for this Comptonization to become optically thick along the equatorial direction, shielding the inner regions from the disc photons so that seed photons are predominantly from cyclo-synchrotron (Poutanen & Veledina 2014). None the less, there could still be some radial temperature/spectral hardness gradients in this second region that could produce additional emissions at the hardest energies (Poutanen & Veledina 2014). We note that the JED models (e.g. Marcel et al. 2018) also predict a continuously increasing temperature/harder spectrum with radius in their hot flow.

However, including the additional harder Comptonization component probably does not fully solve issues with the phase lag, as the amount of increased lag should be rather small as the propagation speed is already high. Yet the data show a large increase in lag between high-energy bands. It seems more likely that there are other factors at work affecting the lags, potentially related to changing temperature in the flow.

There is another feature that is lacking from our new model. It is a physical description of the spectral pivoting from the Comptonization process. Currently, the model assumes that the spectra pivot in a synchronous way, i.e. the local spectrum at every energy responds to fluctuations of the local mass accretion rate simultaneously. Although the magnitude of the response can be different between different energies, as seen in Fig. 7, there is no causal connection between them. There can be 0 or  $\pi$  of phase lags arising from the spectral pivoting itself (the phase lag of  $\pi$  arises if one energy band is above the pivoting point and the other band is below it, i.e.  $\eta(E_1)\eta(E_2) <$

0). Thus, in our new model, the lags between different energies are still due to the propagating fluctuation process, as is the case for the model developed by Veledina (2016, 2018). The spectral pivoting implemented in our new model can strongly affect power spectra but has only a relatively mild effect on phase-lag spectra. Indeed, our new model is able to reproduce energy-dependent power spectra fairly well up to  $\sim 100$  keV, even though it fails to match the phase-lag spectra.

However, the physical picture of Comptonization described above should give a characteristic spectral pivoting pattern. A fluctuation from the edge of the truncated disc first gives a change in seed photons to the soft Comptonization. Assuming an increase in seed photons, it increases the Compton cooling on the light travel time without any change in electron heating, so the spectrum softens. Then, after the accretion time-scale (propagation time-scale), the same fluctuation modulates the soft Comptonization by increasing the electron density, increasing the heating rate, and causing the spectrum to harden. This gives an asynchronous rocking of the soft Comptonization, where two mutually correlated but lagged variability sources form its time-dependent behaviour. By contrast, in the hard Comptonization region, the fluctuation gives a synchronous change in seed photons and electron heating as both are produced together around its outer edge. The synchronous pivoting implemented in our model may be limiting its ability to properly model the data, as it is suppressing a real lag that occurs from the two time-scales propagation mechanism in the soft Comptonization.

It is worth noting that our implementation of the spectral pivoting is different from that in the RELTRANS model (e.g. Mastroserio et al. 2018; Ingram et al. 2019). Mastroserio et al. (2018, 2021) and Mastroserio, Ingram & van der Klis (2019) consider the non-linear effects in the time-varying continuum spectrum and have two variable terms in its expression to allow lags to arise from the spectral pivoting itself (Kotov et al. 2001). However, not specifying the underlying process causing the spectral pivoting may make RELTRANS too flexible in producing the observed hard-lag data. On the other hand, our new model has only one variable term, i.e. the local mass accretion rate, in the expression of the local spectrum. The local spectrum varies linearly to this term, which does not produce lags except for  $\pi$ . As mentioned above, our model relies on the hard lags caused by the combination of the propagating fluctuations process and energy-dependent emission profile (Veledina 2016, 2018).

We suspect that this lack of a physical spectral pivoting model, including the light crossing time spectral softening and the propagation time, is the major reason our new model fails to fit the phase-lag spectra from 50 to 150 keV. This more physical model for spectral pivoting is beyond the scope of this paper but will be considered in future work.

## 7 CONCLUSIONS

We have studied X-ray spectral-timing properties of the black hole binary MAXI J1820+070 in the bright low/hard state using *Insight-HXMT* observation data. Particularly, we have focused on the energy-dependent broad-band variability on time-scales from milliseconds to seconds.

We started with testing our previous model from K22, which included the propagating fluctuations process and reverberation, and successfully explained soft X-ray timing properties ( $< 10$  keV), and found that it cannot be applied to higher energy bands. The key variability feature that our previous model missed was the decrease in fractional power spectrum with energy above  $\sim 10$  keV, which is difficult to explain with the simple propagating fluctuations picture

but is typically observed (Nowak et al. 1999; Malzac et al. 2003; Axelsson & Done 2018). We have seen that additional effects proposed in the literature, such as the damping (Rapisarda et al. 2017; Mahmoud & Done 2018b), are not very effective in reproducing both observed power spectra and phase-lag spectrum simultaneously.

We updated our model by implementing spectral pivoting. This is physically expected in a Comptonization model with fluctuating power (Malzac et al. 2003; Gandhi et al. 2008; Veledina 2016; Mastroserio et al. 2018). This, plus allowing the propagation speed to be different from the time-scale on which fluctuations are generated, allows us to reproduce both power spectra and phase-lag spectra across the 2.6–48 keV band. The pivoting reduces the amplitude of the response of the spectrum at high energies compared to low energies, suppressing the power spectral normalization at high energies but keeping its shape. We are finally able to do a full spectral-timing fit in the 2–50 keV bandpass, demonstrating that our timing model can be self-consistently combined with spectral models.

The propagation derived from the spectral-timing fit favours SANE over MAD, as the propagation speed is too fast to explain the observed lags for MAD. Our spectral model for the accretion flow consists of emission from the turbulent disc, plus soft and hard Comptonization regions and their associated disc reflection. These spectral components fit all the emissions in our bandpass, and their timing components fit all of the power spectra and phase lags apart from the QPO. Thus there is very little room for any additional spectral component to make the QPO, supporting models where the QPOs originate from a global modulation (multiplicative process) of the existing hot flow such as the Lense–Thirring precession (Ingram et al. 2009) rather than e.g. an additional component from the jet (Ma et al. 2021).

None the less, our new model still has some limitations, with clear discrepancies with data above  $\sim 40$  keV. The observed phase lags keep on increasing up to  $\sim 150$  keV while those in our model saturate above  $\sim 40$  keV. These may point to a more complex description of spectral pivoting and/or additional spectral stratification of the inner parts of the hot flow, but clearly, they show that the data are better than the best current models of the hot flow. We stress that this is a motivation for better physical modelling of the flow so that we can robustly use the observed spectral and timing data to explore the underlying nature and geometry of the accretion flow in the region where the jet is launched.

## ACKNOWLEDGEMENTS

We thank Zi-Xu Yang, Liang Zhang, and Xiang Ma for help with *Insight-HXMT* data reduction. We acknowledge the anonymous referee for their valuable comments. This work was supported by JSPS KAKENHI, grant numbers 18H05463 and 20H00153, and WPI MEXT. TK acknowledges support from JST SPRING, grant number JPMJFS2108 and JSPS Overseas Challenge Program for Young Researchers. CD acknowledges support from the STFC consolidated grant ST/T000244/1. Our work also benefited from discussions during the workshop at the International Space Science Institute (Bern).

We dedicate this work to the memory of Magnus Axelsson, who died suddenly as we were finishing the paper.

## DATA AVAILABILITY

The observational data underlying this paper are available at <http://hxmten.ihep.ac.cn/>. The model will be shared upon reasonable request to the corresponding author.

## REFERENCES

- Arévalo P., Uttley P., 2006, *MNRAS*, 367, 801  
 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  
 Axelsson M., Done C., 2018, *MNRAS*, 480, 751  
 Axelsson M., Veledina A., 2021, *MNRAS*, 507, 2744  
 Axelsson M., Borgonovo L., Larsson S., 2005, *A&A*, 438, 999  
 Balbus S. A., Hawley J. F., 1991, *ApJ*, 376, 214  
 Balbus S. A., Hawley J. F., 1998, *Rev. Mod. Phys.*, 70, 1  
 Belloni T., Psaltis D., van der Klis M., 2002, *ApJ*, 572, 392  
 Bhargava Y., Hazra N., Rao A. R., Misra R., Bhattacharya D., Roy J., Alam M. S., 2022, *MNRAS*, 512, 6067  
 Bollimpalli D. A., Mahmoud R., Done C., Fragile P. C., Kluźniak W., Narayan R., White C. J., 2020, *MNRAS*, 496, 3808  
 Bright J. S. et al., 2020, *Nat. Astron.*, 4, 697  
 Buisson D. J. K. et al., 2019, *MNRAS*, 490, 1350  
 Churazov E., Gilfanov M., Revnivtsev M., 2001, *MNRAS*, 321, 759  
 Cowperthwaite P. S., Reynolds C. S., 2014, *ApJ*, 791, 126  
 Dauser T., García J. A., Joyce A., Lickleder S., Connors R. M. T., Ingram A., Reynolds C. S., Wilms J., 2022, *MNRAS*, 514, 3965  
 De Marco B., Ponti G., Muñoz-Darias T., Nandra K., 2015, *ApJ*, 814, 50  
 De Marco B., Zdziarski A. A., Ponti G., Migliori G., Belloni T. M., Segovia Otero A., Dziełek M. A., Lai E. V., 2021, *A&A*, 654, A14  
 Done C., Gierliński M., Kubota A., 2007, *A&AR*, 15, 1  
 Foreman-Mackey D., 2016, *J. Open Source Softw.*, 1, 24  
 Fragile P. C., Blaes O. M., Anninos P., Salmonson J. D., 2007, *ApJ*, 668, 417  
 Gandhi P. et al., 2008, *MNRAS*, 390, L29  
 García J. et al., 2014, *ApJ*, 782, 76  
 García J. A., Fabian A. C., Kallman T. R., Dauser T., Parker M. L., McClintock J. E., Steiner J. F., Wilms J., 2016, *MNRAS*, 462, 751  
 Gierliński M., Zdziarski A. A., 2005, *MNRAS*, 363, 1349  
 Grinberg V. et al., 2014, *A&A*, 565, A1  
 Hogg J. D., Reynolds C. S., 2016, *ApJ*, 826, 40  
 Homan J. et al., 2020, *ApJ*, 891, L29  
 Ingram A. R., 2016, *Astron. Nachr.*, 337, 385  
 Ingram A., 2019, *MNRAS*, 489, 3927  
 Ingram A., Done C., 2011, *MNRAS*, 415, 2323  
 Ingram A., Done C., 2012, *MNRAS*, 419, 2369  
 Ingram A., van der Klis M., 2013, *MNRAS*, 434, 1476  
 Ingram A., Done C., Fragile P. C., 2009, *MNRAS*, 397, L101  
 Ingram A., Mastroserio G., Dauser T., Hovenkamp P., van der Klis M., García J. A., 2019, *MNRAS*, 488, 324  
 Kara E. et al., 2019, *Nature*, 565, 198  
 Kawamura T., Axelsson M., Done C., Takahashi T., 2022, *MNRAS*, 511, 536(K22)  
 Kawamura T. et al., 2018, *Astron. Telegram*, 11399, 1  
 Kotov O., Churazov E., Gilfanov M., 2001, *MNRAS*, 327, 799  
 Krawczynski H. et al., 2022, *Science*, 378, 650  
 Li X. et al., 2020, *J. High Energy Astrophys.*, 27, 64  
 Lyubarskii Y. E., 1997, *MNRAS*, 292, 679  
 Ma X. et al., 2021, *Nat. Astron.*, 5, 94  
 Mahmoud R. D., Done C., 2018a, *MNRAS*, 473, 2084  
 Mahmoud R. D., Done C., 2018b, *MNRAS*, 480, 4040  
 Mahmoud R. D., Done C., De Marco B., 2019, *MNRAS*, 486, 2137  
 Malzac J., Belloni T., Spruit H. C., Kanbach G., 2003, *A&A*, 407, 335  
 Marcel G. et al., 2018, *A&A*, 615, A57  
 Mastroserio G., Ingram A., van der Klis M., 2018, *MNRAS*, 475, 4027  
 Mastroserio G., Ingram A., van der Klis M., 2019, *MNRAS*, 488, 348  
 Mastroserio G. et al., 2021, *MNRAS*, 507, 55  
 Miyamoto S., Kitamoto S., Mitsuda K., Dotani T., 1988, *Nature*, 336, 450  
 Miyamoto S., Kimura K., Kitamoto S., Dotani T., Ebisawa K., 1991, *ApJ*, 383, 784  
 Narayan R., Kato S., Honma F., 1997, *ApJ*, 476, 49  
 Narayan R., Sądowski A., Penna R. F., Kulkarni A. K., 2012, *MNRAS*, 426, 3241

- Novikov I. D., Thorne K. S., 1973, in DeWitt C., DeWitt B., eds, *Black Holes (Les Astres Occlus)*. Gordon & Breach, New York, p. 343
- Nowak M. A., Wilms J., Dove J. B., 1999, *ApJ*, 517, 355
- Pottschmidt K. et al., 2003, *A&A*, 407, 1039
- Poutanen J., Veledina A., 2014, *Space Sci. Rev.*, 183, 61
- Poutanen J., Veledina A., Zdziarski A. A., 2018, *A&A*, 614, A79
- Prabhakar G., Mandal S., Athulya M. P., Nandi A., 2022, *MNRAS*, 514, 6102
- Rapisarda S., Ingram A., Kalamkar M., van der Klis M., 2016, *MNRAS*, 462, 4078
- Rapisarda S., Ingram A., van der Klis M., 2017, *MNRAS*, 469, 2011
- Shakura N. I., Sunyaev R. A., 1973, *A&A*, 24, 337
- Shidatsu M., Nakahira S., Murata K. L., Adachi R., Kawai N., Ueda Y., Negoro H., 2019, *ApJ*, 874, 183
- Tetarenko A. J. et al., 2021, *MNRAS*, 504, 3862
- Torres M. A. P., Casares J., Jiménez-Ibarra F., Álvarez-Hernández A., Muñoz-Darias T., Armas Padilla M., Jonker P. G., Heida M., 2020, *ApJ*, 893, L37
- Tucker M. A. et al., 2018, *ApJ*, 867, L9
- Turner S. G. D., Reynolds C. S., 2021, *MNRAS*, 504, 469
- Uttley P., McHardy I. M., Vaughan S., 2005, *MNRAS*, 359, 345
- Uttley P., Wilkinson T., Cassatella P., Wilms J., Pottschmidt K., Hanke M., Böck M., 2011, *MNRAS*, 414, L60
- Uttley P., Cackett E. M., Fabian A. C., Kara E., Wilkins D. R., 2014, *A&AR*, 22, 72
- Uttley P. et al., 2018, *Astron. Telegram*, 11423, 1
- Vaughan S., Edelson R., Warwick R. S., Uttley P., 2003, *MNRAS*, 345, 1271
- Veledina A., 2016, *ApJ*, 832, 181
- Veledina A., 2018, *MNRAS*, 481, 4236
- Wang Y. et al., 2020, *ApJ*, 896, 33
- Wang J. et al., 2021, *ApJ*, 910, L3
- Wang J. et al., 2022, *ApJ*, 930, 18
- Wilkinson T., Uttley P., 2009, *MNRAS*, 397, 666
- Yamada S., Negoro H., Torii S., Noda H., Mineshige S., Makishima K., 2013, *ApJ*, 767, L34
- Yang Z.-X. et al., 2022, *ApJ*, 932, 7
- You B. et al., 2021, *Nat. Commun.*, 12, 1025
- Zdziarski A. A., Johnson W. N., Magdziarz P., 1996, *MNRAS*, 283, 193
- Zdziarski A. A., Dzielak M. A., De Marco B., Szanecki M., Niedźwiecki A., 2021a, *ApJ*, 909, L9
- Zdziarski A. A. et al., 2021b, *ApJ*, 914, L5
- Zdziarski A. A., You B., Szanecki M., Li X.-B., Ge M., 2022, *ApJ*, 928, 11
- Zhang S.-N. et al., 2020, *Sci. China Phys. Mech. Astron.*, 63, 249502
- Życki P. T., Done C., Smith D. A., 1999, *MNRAS*, 309, 561

## APPENDIX A: SUMMARY OF OUR PREVIOUS MODEL

We give the detailed formalism of our model in the whole appendix. We start with the summary of our previous model presented in K22 before formulating model updates made in this work in the subsequent appendices. We note that while we assume three radially stratified spectral components as the variable flow, i.e.  $N_s = 3$  (see Fig. 2a), the formalism can be straightforwardly applied to an arbitrary number of spectral components. Needless to say, the formalism is the simplest for the single spectral component,  $N_s = 1$ .

We split the variable flow ranging from  $r_{in}$  to  $r_{out}$  into  $N_r$  rings logarithmically, such that the central radius of the  $n$ th ring,  $r_n$ , and the distance between neighbouring central radii,  $\Delta r_n = r_n - r_{n+1}$ , follow  $\Delta r_n/r_n = \text{constant}$  ( $n = 1, 2, \dots, N_r$  from outer to inner rings). We use the radius  $r_n$  in units of gravitational radii  $R_g = GM_{BH}/c^2$ , where  $G$ ,  $M_{BH}$ , and  $c$  are the gravitational constant, black hole mass, and speed of light in vacuum, respectively. The number of rings  $N_r$  determines the resolution of the model calculation (Ingram & van der Klis 2013). The larger number of  $N_r$  allows more accurate calculations of power spectra and cross-spectra for high frequencies at the expense of computation efficiency.

The mass accretion rate varies stochastically everywhere in the variable flow. We define this intrinsically fluctuating mass accretion rate for each ring as the product of its mean  $m_0$  and time-variable term  $a(r_n, t)$  having the mean of unity,  $\mu = \langle a(r_n, t) \rangle_t = 1$ , where  $t$  denotes time. Employing a dimensionless mass accretion rate prescription, we set  $\dot{m}_0 = 1$ . We assume that the power spectrum of  $a(r_n, t)$ ,  $|A(r_n, f)|^2$ , is provided by a zero-centred Lorentzian function with the cut-off frequency equal to the generator frequency  $f_{gen}(r_n)$ ,

$$|A(r_n, f)|^2 = \frac{2\sigma^2}{\pi\mu^2} \frac{f_{gen}(r_n)}{f^2 + (f_{gen}(r_n))^2}, \quad (\text{A1})$$

where  $f$  is the Fourier frequency and  $\sigma^2$  the variance of  $a(r_n, t)$ . In the expression (A1), we employ the normalized power spectra such that their integral over positive frequency corresponds to  $(\sigma/\mu)^2$ . We use the parameter  $F_{var}$  to set the variance through  $\sigma/\mu = F_{var}/\sqrt{N_{dec}}$ , where  $N_{dec}$  is the number of rings per radial decade and thus,  $F_{var}$  is the fractional variability per radial decade. The generator frequency is defined as equation (1) (see also Fig. 2c). Sample power spectra are shown in Fig. 2(e) with dashed lines.

While mass accretion rate fluctuations are generated, they propagate towards the central object at the same time. The propagation is expected to happen in a multiplicative manner (Uttley, McHardy & Vaughan 2005), which leads to assuming

$$\dot{m}(r_n, t) = \dot{m}_0 \prod_{k=1}^n a(r_k, t - \Delta t_{k,n}), \quad (\text{A2})$$

where  $\Delta t_{k,n}$  is the propagation time from the outer  $k$ th ring to inner  $n$ th ( $n \geq k$ ) ring. We note that  $\dot{m}_0 = 1$  is left in the expression to preserve generality. As the propagation speed  $v_p(r)$  is set by radial velocity, i.e.  $v_p(r) = r f_{visc}(r)$ , the propagation time is expressed as

$$\Delta t_{k,n} = \frac{\Delta r}{r} \sum_{l=k+1}^n \frac{1}{f_{visc}(r)}, \quad (\text{A3})$$

where  $\Delta r/r = \Delta r_n/r_n = \text{constant}$ . In the statistical equilibrium under the generation and propagation of mass accretion rate fluctuations, the power spectra,  $|\dot{M}(r_n, f)|^2$ , and cross-spectra,  $(\dot{M}(r_m, f))^* \dot{M}(r_n, f)$  ( $m \neq n$ ), can be calculated analytically (see appendix A in K22). The asterisk  $*$  denotes the complex conjugate. Sample power spectra are shown in Fig. 2(e) with solid lines.

The flux at energy  $E$  and time  $t$ ,  $x(E, t)$ , is the sum of contributions from each ring in the variable flow. Ignoring the effect of reverberation for simplicity (we take it into account in Appendix D), we express the flux as

$$x(E, t) = \sum_{n=1}^{N_r} \lambda(r_n) S(E, r_n, t). \quad (\text{A4})$$

The flux represents the number of photons per unit of time because the variability is studied for count rather than energy. Both treatments give the same model calculations for a single energy. However, using count rates is necessary to compare the model to observations because observation data cannot be processed for an infinitesimal energy band. The local flux emitted from  $n$ th ring,  $S(E, r_n, t)$ , is defined as

$$S(E, r_n, t) = S_0(E, r_n) \frac{\dot{m}(r_n, t)}{\dot{m}_0}, \quad (\text{A5})$$

where  $S_0(E, r_n)$  is the time-averaged spectrum at the  $n$ th ring. We note that  $\dot{m}_0 = 1$  is explicitly written in equation (A5). Since we assume three spectral components ( $N_s = 3$ ), this time-averaged spectrum is

categorized into three spectra:

$$S_0(E, r_n) = \begin{cases} S_h(E) & (r_{\text{in}} \leq r_n < r_{\text{sh}}), \\ S_s(E) & (r_{\text{sh}} \leq r_n < r_{\text{ds}}), \\ S_d(E) & (r_{\text{ds}} \leq r_n < r_{\text{out}}), \end{cases} \quad (\text{A6})$$

depending on to which spectral regions the  $n$ th ring belongs.  $r_{\text{in}}$  is the inner edge of the hard Comptonization component,  $r_{\text{sh}}$  is the transition radius between the hard Comptonization and soft Comptonization components,  $r_{\text{ds}}$  is the transition radius between the soft Comptonization and variable disc components, and  $r_{\text{out}}$  is the outer edge of the variable disc component.  $S_h(E)$ ,  $S_s(E)$ , and  $S_d(E)$  are the time-averaged spectra for the hard Comptonization, soft Comptonization, and disc components, as shown in Fig. 2(d).

To let equations be concise, we define the collection of radii affiliated with the variable disc, soft Comptonization, and hard Comptonization regions as  $r_d$ ,  $r_s$ , and  $r_h$ , respectively. Then we express  $S_0(E)$  as

$$S_0(E, r_n) = S_Y(E) \quad (r_n \in r_Y), \quad (\text{A7})$$

for  $Y = h, s, d$ . In the model formalism, we use  $S_0(E, r_n)$  as a fraction, giving the constraint of

$$\sum_{Y=h,s,d} S_Y(E) = 1. \quad (\text{A8})$$

We note that we ignore any other spectral components here, such as reflected emission.

Along with the time-averaged spectrum  $S_0(E, r_n)$ , the contribution from each ring is also regulated by  $\lambda(r_n)$  in terms of energy dissipation:

$$\lambda(r_n) = \epsilon(r_n) 2\pi r_n \Delta r_n \bigg/ \sum_{m(r_m \in r_Y)} \epsilon(r_m) 2\pi r_m \Delta r_m, \quad (\text{A9})$$

where  $\epsilon(r)$  is the emissivity. The emissivity is defined as the product of a power-law function and inner boundary condition:  $\epsilon(r) \propto r^{-\gamma} b(r)$ , where  $b(r) = 1 - \sqrt{r_{\text{in}}/r}$  or  $b(r) = 1$  for the ‘stress-free’ or ‘stressed’ boundary condition (Ingram & Done 2012). Because of the normalization of  $\lambda(r_n)$  and mass accretion rate fluctuations, the time-averaged flux from an entire spectral region corresponds to the time-averaged spectrum:

$$\sum_{n(r_n \in r_Y)} \lambda(r_n) S(E, r_n, t) = S_Y(E), \quad (\text{A10})$$

giving rise to the time-averaged total flux being unity,  $\langle x(E, t) \rangle_t = \sum_{Y=h,s,d} S_Y(E) = 1$ .

From equations (A4) and (A5), the Fourier transform of the flux  $x(E, t)$  is proportional to the Fourier transform of the local mass accretion rate  $\dot{m}(r_n, t)$ :

$$\begin{aligned} X(E, f) &= \sum_{n=1}^{N_r} \lambda(r_n) \frac{S_0(E, r_n)}{\dot{m}_0} \dot{M}(r_n, f) \\ &= \sum_{n=1}^{N_r} w(r_n, E) \dot{M}(r_n, f), \end{aligned} \quad (\text{A11})$$

where we call the coefficient  $w(r_n, E) = \lambda(r_n) S_0(E, r_n) / \dot{m}_0$  the weight. The weight  $w(r_n, E)$  is proportional to the product of the emissivity and energy spectrum, finally determining how much  $n$ th ring contributes to the variability of the flux. Since we know the analytic forms of power spectra and cross-spectra for the local mass accretion rate,  $(\dot{M}(r_m, t))^* \dot{M}(r_n, t)$  ( $m, n = 1, 2, \dots, N_r$ ), we can calculate the power spectrum  $|X(E, f)|^2$  and cross-spectrum for two different energies  $E_1$  and  $E_2$ ,  $(X(E_1, f))^* X(E_2, f)$ , analytically (see

appendix A in K22), which can be directly compared to observation data. The sample model calculations of power spectra, time-lag spectrum, and phase-lag spectrum are shown in Figs 2(f), (g), and (h), respectively.

## APPENDIX B: INCLUDING DAMPING EFFECTS

We summarize the expressions of power spectra and cross-spectra for the local mass accretion rate and flux, where the damping effects are taken into account. The derivations are described in Rapisarda et al. (2017) appendix A, and in Mahmoud & Done (2018b) appendix A. We note that the damping effects are excluded in any results presented since they turned out to be ineffective for this work.

The local mass accretion rate is expressed as

$$\dot{m}(r_n, t) = \prod_{k=1}^n g(r_k, r_n, t) \otimes a(r_k, t), \quad (\text{B1})$$

where the symbol  $\otimes$  denotes the convolution in time  $t$ . The green function  $g(r_k, r_n, t)$  describes how the local mass accretion rate at the inner  $n$ th ring responds to that at the outer  $k$ th ring ( $n \geq k$ ). The Fourier transform of the green function,  $G(r_k, r_n, t)$ , is required in model calculations. While we employ  $G(r_k, r_n, t) = e^{-i2\pi f \Delta t_{k,n}}$  in K22 and Appendix A, which is equivalent to  $g(r_k, r_n, t) = \delta(t - \Delta t_{k,n})$ , we use

$$G(r_k, r_n, f) = e^{-Df \Delta t_{k,n}} e^{-i2\pi f \Delta t_{k,n}} \quad (\text{B2})$$

in order to include the damping effects.

From equation (B1), the following expressions for power spectra and cross-spectra for the local mass accretion rate are obtained:

$$|\dot{M}(r_n, f)|^2 = \frac{1}{N^2} |A(r_n, f)|^2 \otimes |G(r_{n-1}, r_n, f) \dot{M}(r_{n-1}, f)|^2, \quad (\text{B3})$$

$$(\dot{M}(r_k, f))^* \dot{M}(r_n, f) = \Lambda_{k,n} G(r_k, r_n, f) |\dot{M}(r_k, f)|^2 \quad (n \geq k), \quad (\text{B4})$$

where  $\otimes$  denotes the convolution in frequency  $f$ . The consecutive product of the time-averaged local mass accretion rate  $\mu_l$  ( $l = 1, \dots, N_r$ ) from  $k$ th ring to  $n$ th ring is defined as  $\Lambda_{k,n}$ , i.e.  $\Lambda_{k,n} = \prod_{l=k+1}^n \mu_l$ , which is unity. We note that we use continuous Fourier frequency  $f$  rather than discrete Fourier frequency  $f_m$  for simplicity. The coefficient  $1/N^2$  in equation (B3), where  $N$  is the number of data points used, comes from the convolution theorem

$$X(f_m) = \frac{1}{N} A(f_m) \otimes B(f_m), \quad (\text{B5})$$

where  $x(t_l) = a(t_l)b(t_l)$ . In the convolution theorem, the discrete Fourier transform is defined as

$$X(f_m) = \sum_{l=0}^{N-1} x(t_l) e^{-i2\pi f_m t_l}, \quad (\text{B6})$$

where  $f_m = m/(N\Delta t)$  ( $m = -N/2 + 1, \dots, N/2$  for even  $N$  or  $m = (N-1)/2, \dots, (N-1)/2$  for odd  $N$ ) and  $t_l = l\Delta t$  with  $\Delta t$  being the sampling interval. The discrete inverse Fourier transform is thus

$$x(t_l) = \frac{1}{N} \sum_m X(f_m) e^{+i2\pi f_m t_l}. \quad (\text{B7})$$

Combining equations (B3) and (B4) with equation (C4) yields the expressions of the power spectra for the flux:

$$|X(E, f)|^2 = \sum_{n=1}^{N_r} \left[ (w(r_n, E))^2 |\dot{M}(r_n, f)|^2 + 2 \sum_{k=1}^{n-1} \left\{ w(r_k, E) w(r_n, E) \Lambda_{k,n} \times |G(r_k, r_n, f)| \cos(\Phi(r_k, r_n, f)) |\dot{M}(r_k, f)|^2 \right\} \right], \quad (\text{B8})$$

where  $G(r_k, r_n, f) = |G(r_k, r_n, f)| e^{i\Phi(r_k, r_n, f)}$ . The cross-spectra for the flux can be calculated in the same manner:

$$(X(E_1, f))^* X(E_2, f) = \sum_{n=1}^{N_r} \left[ w(r_n, E_1) w(r_n, E_2) |\dot{M}(r_n, f)|^2 + \sum_{k=1}^{n-1} \left\{ (w(r_k, E_1) w(r_n, E_2) e^{-i\Phi(r_k, r_n, f)} + w(r_n, E_1) w(r_k, E_2) e^{+i\Phi(r_k, r_n, f)}) \times \Lambda_{k,n} |G(r_k, r_n, f)| |\dot{M}(r_k, f)|^2 \right\} \right]. \quad (\text{B9})$$

### APPENDIX C: INCLUDING SPECTRAL PIVOTING

To include the change in the spectral shape on short time-scales, we redefine the variable local energy spectrum  $S(E, r_n, t)$  as

$$S(E, r_n, t) = \left( 1 + \eta(E, r_n) \frac{\dot{m}(r_n, t) - \dot{m}_0}{\dot{m}_0} \right) S_0(E, r_n). \quad (\text{C1})$$

Again, although we assume  $\dot{m}_0 = 1$ , we leave the symbol explicitly in the formulation to preserve generality. The new energy-dependent parameter  $\eta(E, r_n)$  regulates how the spectrum at the energy  $E$  responds to mass accretion rate fluctuations. While the amplitude of  $\eta(E, r_n)$  regulates the sensitivity, its sign determines the correlation patterns (see Fig. 4). Since we have three spectral regions for the variable flow ( $N_s = 3$ ), we accordingly classify this sensitivity parameter into three:

$$\eta(E, r_n) = \eta_Y(E) \quad (r_n \in r_Y), \quad (\text{C2})$$

for  $Y = h, s, d$ . The spectrum  $S(E, r_n, t)$  is positively (negatively) correlated to mass accretion rate fluctuations if  $\eta(E, r_n) > 0 (< 0)$ , as shown in Fig. 4. Both correlations can be physically realized (Veledina 2018). The new definition (C1) is reduced to the previous definition (A5) if  $\eta(E, r_n) = 1$ . Thus, we see that equation (C1), in which the spectral shapes vary in time, is a natural extension of equation (A5), in which the spectral shapes are fixed.

Employing equation (C1), the Fourier transform of the flux is

$$X(E, f) = \sum_{n=1}^{N_r} \lambda(r_n) \frac{\eta(E, r_n) S_0(E, r_n)}{\dot{m}_0} \dot{M}(r_n, f) = \sum_{n=1}^{N_r} w(r_n, E) \dot{M}(r_n, f), \quad (\text{C3})$$

where the weight  $w(r_n, E)$  is redefined as

$$w(r_n, E) = \lambda(r_n) \frac{\eta(E, r_n) S_0(E, r_n)}{\dot{m}_0}. \quad (\text{C4})$$

The Fourier transform  $X(E, f)$  is expressed in the same manner as equation (A11) and thus linear to  $\dot{M}(r_n, f)$ . Thus, we can calculate power spectra and cross-spectra for the flux analytically.

The only difference that appeared in the expression of  $X(E, f)$  (compare equations A11 and C3) is the presence of the new parameter  $\eta(E, r_n)$ . Allowing the spectral variation is equivalent to regarding  $S_0(E)$  in equation (A11) as  $\eta(E, r_n) S_0(E, r_n)$ . Because of this replacement, the spectral pivoting releases the model from the constraint (A8) in the sense that the model calculations always require the product  $\eta(E, r_n) S(E, r_n)$ , not  $S_0(E, r_n)$ . The spectral pivoting gives much flexibility to the model in this way, as seen in the successful modelling in Fig. 3 (rightmost) compared to the failed ones in the rest of the columns.

### APPENDIX D: MODIFYING REVERBERATION

We modify the implementation of reflection due to the implementation of the spectral variation described above. While a part of the emission from the variable flow directly hits a detector, it also irradiates the outer disc, resulting in the reflected/reprocessed emission. Considering this reprocessed emission also hits the detector, the observed flux is expressed as

$$x(E, t) = x^{(d)}(E, t) + x^{(r)}(E, t), \quad (\text{D1})$$

where  $x^{(d)}(E, t)$  and  $x^{(r)}(E, t)$  are the direct and reflected components, respectively. Hereafter, we use the subscripts ‘(d)’ and ‘(r)’ to specify the direct and reflected components. Since we have only considered the direct component above, symbols that appeared in previous appendices will be used with the superscript of ‘(d)’. The expression of the direct component  $x^{(d)}(E, t)$  corresponds to equation (A4).

The variable disc component is unlikely to contribute to the reflected emission. In this regard, the variable disc component is distinct from the soft and hard Comptonization components. However, in terms of formalism, it is simple to assume that every spectral component in the variable flow, i.e. every direct component, has its associated reflected component. Therefore, we account for the reflected component for the variable disc in the formalism and remove it in parameter space.

Each ring of the variable flow illuminates the outer disc, yielding the following expression of the reflected component:

$$x^{(r)}(E, t) = \sum_{n=1}^{N_r} \int dE' \lambda(r_n) S^{(d)}(E', r_n, t) \otimes h(E, E', r_n, t), \quad (\text{D2})$$

where the symbol  $\otimes$  denotes the convolution in time  $t$ . The impulse response  $h(E, E', r_n, t)$  describes the time-evolution of reflected flux at the photon energy  $E$  for the instant incident flux at  $E'$  from the  $n$ th ring. The impulse response also encompasses the probability that a photon at energy  $E$  is generated from one at  $E'$ . Here, we assume that the form of the impulse response is common within each spectral component:

$$h(E, E', r_n, t) = h_Y(E, E', t) \quad (r_n \in r_Y), \quad (\text{D3})$$

for  $Y = h, s, d$ . The radius of the ring  $r_n$  is only needed to specify to which spectral component the ring belongs. The impulse response for each spectral component is independent of radius, i.e. the location of illuminating source. This approximation should be validated since the size of a reflector ( $\sim 10^3 R_g$ ) is expected to be much larger than the size of spectral regions in the variable flow ( $\sim 10 R_g$ ).

We also assume that the shape of the impulse response is independent of photon energies  $E$  and  $E'$  and separate its amplitude and

shape in the expression:

$$h(E, E', r_n, t) = C(E, E', r_n) \tilde{h}(r_n, t), \quad (\text{D4})$$

where we normalize the shape  $\tilde{h}(t)$  to follow

$$\int_{-\infty}^{+\infty} dt \tilde{h}(r_n, t) = 1. \quad (\text{D5})$$

Again, each term of the impulse response is common within spectral regions:

$$C(E, E', r_n) = C_Y(E, E'), \quad \tilde{h}(r_n, t) = \tilde{h}_Y(t) \quad (r_n \in r_Y), \quad (\text{D6})$$

for  $Y = h, s, d$ . Hereafter, we replace the symbol of the normalized impulse response  $\tilde{h}(r_n, t)$  with  $h(r_n, t)$ , since the symbol of the impulse response including normalization,  $h(E, E', r_n, t)$ , is not required in the formalism anymore. Substituting equations (C1) and (D4) into equation (D2) yields

$$x^{(r)}(E, t) = \sum_{n=1}^{N_r} \int dE' \lambda(r_n) \left( 1 + \eta^{(d)}(E', r_n) \frac{\dot{m}(r_n, t) - \dot{m}_0}{\dot{m}_0} \right) \times S_0^{(d)}(E', r_n) \otimes C(E, E', r_n) h(r_n, t). \quad (\text{D7})$$

We note that (d) is explicitly used here to distinguish the direct component from the reflected one, although not in the previous subsections, where the reflection is not included.

To connect equation (D7) with the time-averaged reflected spectrum,

$$S_0^{(r)}(E, r_n) = S_Y^{(r)}(E) \quad (Y = h, s, d), \quad (\text{D8})$$

we average the reflected flux  $x^{(r)}(E, t)$  in time:

$$\langle x^{(r)}(E, t) \rangle_t = \sum_{Y=h,s,d} \int dE' C_Y(E, E') S_Y^{(d)}(E'), \quad (\text{D9})$$

which needs to correspond to  $\sum_{Y=h,s,d} S_Y^{(r)}(E)$ . Each term on the right-hand side corresponds to each time-averaged reflected flux. In practical applications, we fix  $S_d^{(r)}(E) = 0$  to exclude the reflected emission arising from irradiation by the variable disc component. Equation (D7) cannot be simplified with this constraint because  $\eta^{(d)}(E', r_n)$  depends on the incident photon energy  $E'$ . The existence of  $\eta^{(d)}(E', r_n)$  in the integration in equation (D7) means that the spectral variation at every energy affects the reflected emission as long as  $C(E, E', r_n) \neq 0$ . Although this situation is physically natural, its proper treatment would require many complications. Moreover,  $\eta^{(d)}(E', r_n)$  is an ad hoc parameter introduced to include the spectral variation and not connected directly to physical parameters. Given these circumstances, we replace  $\eta^{(d)}(E', r_n)$  with  $\eta^{(r)}(E, r_n)$  related to the reflection component and let  $\eta^{(r)}(E, r_n)$  be only dependent of the output photon energy  $E$ , resulting in

$$x^{(r)}(E, t) = \sum_{n=1}^{N_r} \lambda(r_n) \left( 1 + \eta^{(r)}(E, r_n) \frac{\dot{m}(r_n, t) - \dot{m}_0}{\dot{m}_0} \right) \times S_0^{(r)}(E, r_n) \otimes h(r_n, t). \quad (\text{D10})$$

Defining

$$S^{(r)}(E, r_n, t) = \left( 1 + \eta^{(r)}(E, r_n) \frac{\dot{m}(r_n, t) - \dot{m}_0}{\dot{m}_0} \right) S^{(r)}(E, r_n) \quad (\text{D11})$$

yields an expression for the reflected component comparable to that for the direct component (A4):

$$x^{(r)}(E, t) = \sum_{n=1}^{N_r} \lambda(r_n) S^{(r)}(E, r_n, t) \otimes h(r_n, t). \quad (\text{D12})$$

The difference lies in the presence of convolution. As each direct spectral component does, each reflected spectral component shares the sensitivity parameter within its associated radii:

$$\eta^{(r)}(E, r_n) = \eta_Y(E) \quad (r_n \in r_Y), \quad (\text{D13})$$

for  $Y = h, s, d$ . In the spectral-timing fit in Section 5, we assume  $\eta_Y^{(d)}(E) = \eta_Y^{(r)}(E)$ , which means that the direct and its associated reflected components react to a deviation of the mass accretion rate from its average in the same way.

Finally, substituting equations (A4) and (D12) into equation (D1), we get the expression of the flux:

$$x(E, t) = \sum_{n=1}^{N_r} \lambda(r_n) (S^{(d)}(E, r_n, t) + S^{(r)}(E, r_n, t) \otimes h(r_n, t)). \quad (\text{D14})$$

Using equations (C1) and (D11), its Fourier transform is

$$X(E, f) = \sum_{n=1}^{N_r} (w^{(d)}(r_n, E) + w^{(r)}(r_n, E) H(r_n, f)) \dot{M}(r_n, f), \quad (\text{D15})$$

where

$$w^{(r)}(r_n, E) = \lambda(r_n) \frac{\eta^{(r)}(E, r_n) S_0^{(r)}(E, r_n)}{\dot{m}_0} \quad (\text{D16})$$

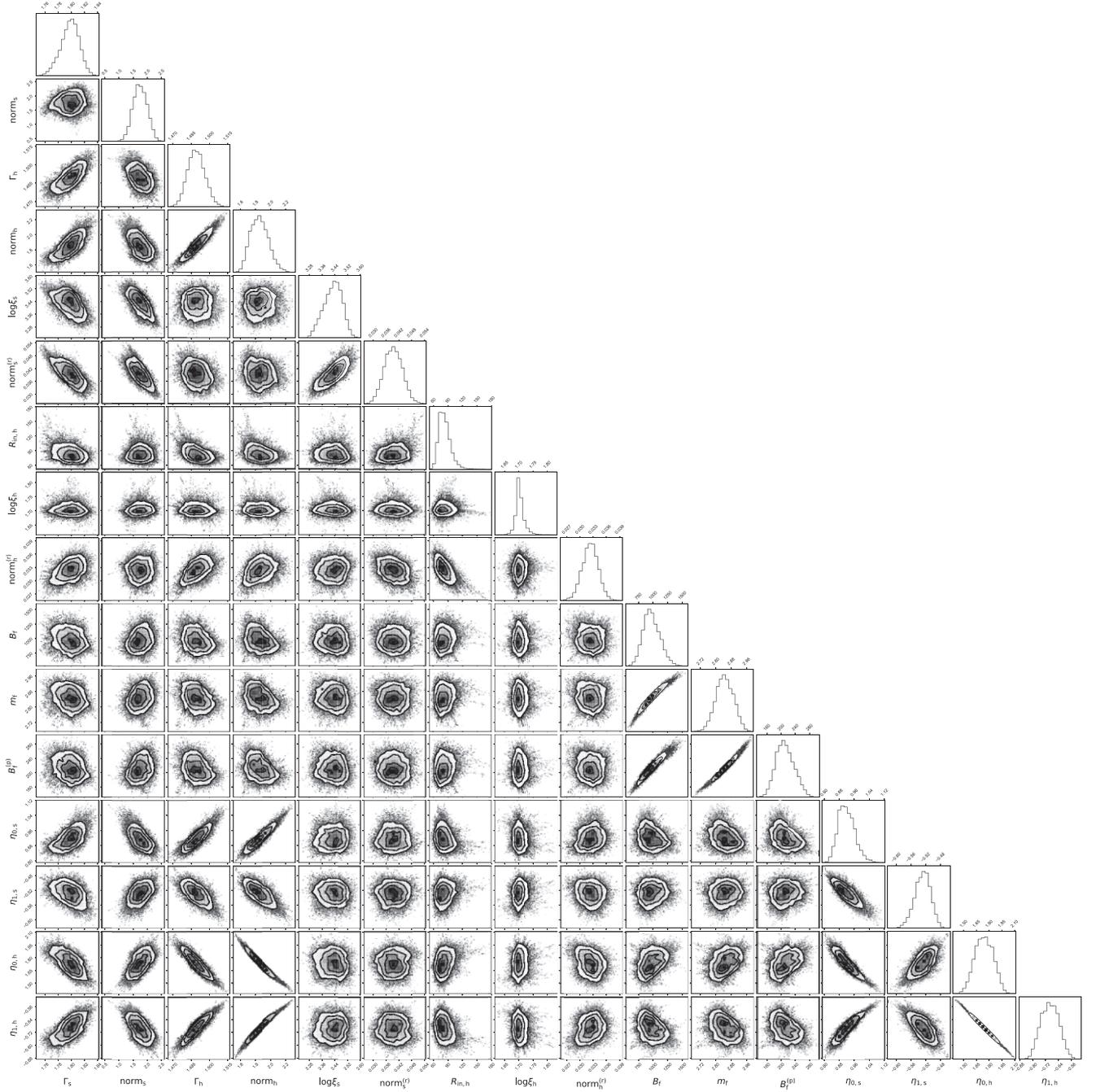
is the weight for the reflection component, and  $H(r_n, f)$  is the transfer function, the Fourier transform of the impulse response  $h(r_n, t)$ . Again, the linearity of  $X(E, f)$  to  $\dot{M}(r_n, f)$  allows analytic calculations for the power spectra and cross-spectra of the flux.

## APPENDIX E: MCMC FIT

We demand 28 free parameters in the spectral-timing fit performed in Section 5. In addition to these many free parameters, our model computations are quite expensive, even though not prohibitive. These situations make it difficult to properly estimate the errors on the derived parameter values.

Here, we perform a Markov chain Monte Carlo (MCMC) fit to evaluate the uncertainties of the model parameter values. We use the `chain` command in `XSPEC`, in which we employ the Goodman–Weare algorithm with 100 walkers and 100 000 steps in total. Whereas in Section 5, we free the normalizations of the `lorentz` model used to capture QPO features in the power spectra, we fix all of them in the MCMC fit. Fixing them decreases 12 free parameters because there are two `lorentz` models for each of the six energy bands, enabling the MCMC analysis with 16 free parameters. We burn 1000 steps because the  $\chi^2$  values converge after  $\sim 1000$  steps.

The corner plots for all 16 free parameters are shown in Fig. E1. The distribution is generally centrally peaked for every free parameter, signifying good convergence. The parameters regulating the propagation time-scale,  $B_r^{(p)}$  and  $m_r^{(p)}$ , take values among 160–240 and 2.78–2.88, respectively. These spreads give rise to the spreads of the propagation time-scale as narrow as the line width in Fig. 8. Thus, the result of our analysis that the propagation time-scale derived fairly disagrees with any theoretical predictions in Fig. 8 is robust.



**Figure E1.** Corner plots for 16 free parameters of the spectral-timing fit in Section 5. Every free parameter has its row and column. The correlation between two free parameters is shown at the intersection of the corresponding row and column. The distribution of each parameter is shown at the rightmost (top) of the corresponding row (column). We use the `corner` module (Foreman-Mackey 2016) to produce the plots. The figure is viewed clearly in the digital version.

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