Moreover, both the $\overline{N}_{\rm H,w}$ and the $\delta N_{\rm H,w}$ parameters show significant short-term variability (Fig. 9). El Mellah et al. (2020) presented a model that explores the impact of the wind clumpiness on the rapid time-variability of its column density, and describes how this observed variability can be linked to the clumps' physical properties, such as the size and the mass. In the clumpy wind model of El Mellah et al. (2020), the clumps' properties are directly linked to the scatter parameter, $\delta N_{\rm H,w}$. This parameter is expected to be maximum around superior conjunction, as indeed observed (Table 1 and Fig. 9, lower panel). The model also predicts the scatter to be an excellent tracer of the ratio $\sqrt{m_{\rm cl}}/R_{\rm cl}$, where $m_{\rm cl}$ and $R_{\rm cl}$ are the mass and radius of the clumps. In Lai et al. (2022) we measured the minimum timescales at which the passage of clumps causes significant excess X-ray variability in the power spectra of the source at superior conjunction. These timescales correspond to a clump radial size of $\gtrsim 10^{-4}\,R_*$. Using this as the smallest clumps radius, the measured maximum value of $\delta N_{\rm H,w}$ (Table 1), and the above estimate for the mass loss rate, and plugging them into Equation (18) of El Mellah et al. (2020) we infer $m_{\rm cl} \sim 10^{17}$ g as an estimate for the characteristic mass of the clumps (for $v_\infty = 2100$ km/s and $R_* = 22\,R_\odot$). This estimate is in agreement with values inferred from radiative hydrodynamical simulations (Sundqvist et al. 2018), as well as from spectral analysis (Härer et al. 2023).

5.3. The soft colour tail

The observed colour-colour tracks also reveal the presence of a soft emission component in the most absorbed stages, producing an extended soft-colour tail. We verified (Sect. 4) that the analysed *XMM-Newton* data are consistent with the presence of significant contribution from a dust-scattering halo (e.g. Jin et al. 2017). However, additional contribution from the wind itself in the form of emission lines from the diffused photoionised gas around the BH is also expected (Hirsch et al. 2019). Indeed, both contributions from a dust-scattering halo and the photoionised emitting plasma are expected to be prominent when the primary radiation from the X-ray source is significantly blocked by the absorbing material; that is, in the deepest stages of the dips. A more complex modelling (which is beyond the scope of this paper) would allow the two contributions to be disentangled, and might in turn be useful to put stronger constraints on wind models, as well as probing the interstellar dust towards Cyg X-1.

As a final remark, it is worth noting that additional spectral changes extrinsic to the wind – intrinsic changes of the hard X-ray source properties – have been neglected in our fits of time-resolved colour-colour diagrams. For example, such changes might be due to time variations in the spectral index, Γ, as induced by variations in either the temperature or the optical depth of the Comptonising gas (or both). In Sect. 4, we verified that the scattering in hard and soft colours in the uppermost, unabsorbed part of the diagrams are most likely driven by spectral variability. Mastroserio et al. (2021) showed that Γ variability of a few percents is plausible in hard state sources with ~10% fractional rms variability, typical of the hard state. Skipper et al. (2013) found a larger Γ variability (~10%) for the same source and spectral state but how these spectral changes influence the most absorbed parts of the diagram is yet to be tested.

6. Conclusions

In this paper, we have presented an analysis of the colour-colour diagrams of Cyg X-1 during a passage at superior conjunction, with the aim of constraining the properties of the stellar wind. We employed the KDE method to select wind models that best fit the colour-colour diagrams. This allowed us to overcome the problem of dealing with data that are not normally distributed. The main results are the following:

- We found that the model that best describes the characteristic 'pointy' or 'nose-like' shape of the diagrams implies the wind to be partially ionised.
- We revealed a strong temporal evolution of the colour-colour diagrams around superior conjunction. Our fits suggest this evolution to be strongly influenced by concurrent variations in the column density and covering factor of the wind.
- Both the column density and the covering factor are maximum at superior conjunction (since the LOS crosses deeper wind layers), but their overall variations follow different trends.
- We report a one-to-one scaling between long-term (>11 ks) and rapid (between 10 s and 11 ks) variations in the column density (reminiscent of the 'rms-flux' relation characterising stochastic variability in compact objects). The existence of such a correlation might have implications for the distribution of the wind clumps and for the way they combine into bigger clumps.
- Using the clumpy wind model proposed in El Mellah et al. (2020), we have estimated a wind mass loss rate of $\dot{M}_* \sim 7 \times 10^{-6}\,{\rm M}_\odot\,{\rm yr}^{-1}$ and a characteristic clump mass of $m_{\rm cl} \sim 10^{17}$ g.

Future applications of this analysis approach will require the sampling of longer stretches of the binary orbit and the application of more complex models (e.g. including contribution from scattered wind emission and from the dust-scattering halo).

Acknowledgements. This work is based on observations obtained with *XMM-Newton*, an ESA science mission instrument, and contributions directly funded by ESA Member States and NASA. EVL thanks Alex Markowitz, Piotr Życki and Maura Pilia for useful discussions. The research leading to these results has received funding from the European Union's Horizon 2020 Programme under the AHEAD2020 project (grant agreement n. 871158). EVL and AR were partially supported by the Polish National Science Centre, grant no. 2021/41/B/ST9/04110. EVL and MB are supported by the Italian Research Center on High Performance Computing Big Data and Quantum Computing (ICSC), project funded by European Union – NextGenerationEU – and National Recovery and Resilience Plan (NRRP) – Mission 4 Component 2 within the activities of Spoke 3 (Astrophysics and Cosmos Observations). BDM acknowledges support via Ramón y Cajal Fellowship (RYC2018-025950-I) and the Spanish MINECO grant PID2022-136828NB-C44. YC acknowledges support from the grant RYC2021-032718-I, financed by MCIN/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR. BDM, YC, GS, and JJ thank the Spanish MINECO grant PID2020-117252GB-I00 and the AGAUR/Generalitat de Catalunya grant SGR-386/2021. MC acknowledges support from the "Universitas Copernicana Thoruniensis In Futuro" project nr POWR.03.05.00-00-Z302/17 and from the MNiSW grant DIR/WK/2018/12. AAZ acknowledges support from the Polish National Science Center grants 2019/35/B/ST9/03944 and 2023/48/Q/ST9/00138, and from the Copernicus Academy grant CBMK/01/24. MB was funded in part by PRIN TEC INAF 2019 "SpecTemPolar! – Timing analysis in the era of high-throughput photon detectors". This research has made use of NASA's Astrophysics Data System Bibliographic Service (ADS) and of ISIS functions (`isisscripts` http://www.sternwarte.uni-erlangen.de/isis/) provided by ECAP/Remeis observatory and MIT. The authors thank Mirjam Oertel, developer of the first versions of some of the scripts calculating the colour-colour tracks. They also thank John E. Davis for the development of the slxfig (http://www.jedsoft.org/fun/slxfig/) module used to prepare some of the figures in this work. Colour schemes used in the first part of this paper were based on Paul Tol's color-blindness-friendly palettes and templates (https://personal.sron.nl/~pault/).

References

Arnaud, K. A. 1996, ASP Conf. Ser., 101, 17

Bachetti, M., Huppenkothen, D., Khan, U., et al. 2023, https://doi.org/10.5281/zenodo.7970570
Bałucińska-Church, M., Church, M. J., Charles, P. A., et al. 2000, MNRAS, 311, 861
Basak, R., Zdziarski, A. A., Parker, M., & Islam, N. 2017, MNRAS, 472, 4220
Belczynski, K., Bulik, T., & Bailyn, C. 2011, ApJ, 742, L2
Bondi, H., & Hoyle, F. 1944, MNRAS, 104, 273
Bulik, T., Belczynski, K., & Prestwich, A. 2011, ApJ, 730, 140
Castor, J. I., Abbott, D. C., & Klein, R. I. 1975, ApJ, 195, 157
Cavecchi, Y., & Patruno, A. 2022, MNRAS, 510, 1431
Corrales, L. R., García, J., Wilms, J., & Baganoff, F. 2016, MNRAS, 458, 1345
Diez, C. M., Grinberg, V., Fürst, F., et al. 2023, A&A, 674, A147
El Mellah, I., Sundqvist, J. O., & Keppens, R. 2018, MNRAS, 475, 3240
El Mellah, I., Grinberg, V., Sundqvist, J. O., Driessen, F. A., & Leutenegger, M. A. 2020, A&A, 643, A9
Feldmeier, A., Puls, J., & Pauldrach, A. W. A. 1997, A&A, 322, 878
Fornasini, F. M., Tomsick, J. A., Bachetti, M., et al. 2017, ApJ, 841, 35
Fürst, F., Pottschmidt, K., Wilms, J., et al. 2014, ApJ, 780, 133
Gies, D. R., Bolton, C. T., Thomson, J. R., et al. 2003, ApJ, 583, 424
Gilfanov, M. 2010, Lect. Notes Phys., 794, 17
Grinberg, V., Leutenegger, M. A., Hell, N., et al. 2015, A&A, 576, A117
Grinberg, V., Nowak, M. A., & Hell, N. 2020, A&A, 643, A109
Hanke, M., Wilms, J., Nowak, M. A., et al. 2008, Proceedings of the VII Microquasar Workshop: Microquasars and Beyond, 29
Hanke, M., Wilms, J., Nowak, M. A., et al. 2009, ApJ, 690, 330
Härer, L. K., Parker, M. L., El Mellah, I., et al. 2023, A&A, 680, A72
Herrero, A., Kudritzki, R. P., Gabler, R., Vilchez, J. M., & Gabler, A. 1995, A&A, 297, 556
HI4PI Collaboration (Ben Bekhti, N., et al.) 2016, A&A, 594, A116
Hill, P. D. 1985, Commun. Stat. Theory Meth., 14, 605
Hirsch, M., Hell, N., Grinberg, V., et al. 2019, A&A, 626, A64
Houck, J. C. 2002, in High Resolution X-ray Spectroscopy with XMM-Newton and Chandra, ed. G. Branduardi-Raymont, 17
Houck, J. C., & Denicola, L. A. 2000, ASP Conf. Ser., 216, 591
Hoyle, F., & Lyttleton, R. A. 1939, Proc. Camb. Philos. Soc., 35, 405
Huppenkothen, D., Bachetti, M., Stevens, A. L., et al. 2019a, ApJ, 881, 39
Huppenkothen, D., Bachetti, M., Stevens, A., et al. 2019b, J. Open Source Softw., 4, 1393
Jin, C., Ponti, G., Haberl, F., & Smith, R. 2017, MNRAS, 468, 2532
Jin, C., Ponti, G., Haberl, F., Smith, R., & Valencic, L. 2018, MNRAS, 477, 3480
Joinet, A., Kalemci, E., & Senziani, F. 2008, ApJ, 679, 655
Kallman, T. R., Bautista, M. A., Goriely, S., et al. 2009, ApJ, 701, 865
Krawczynski, H., Muleri, F., Dovčiak, M., et al. 2022, Science, 378, 650
Lai, E. V., De Marco, B., Zdziarski, A. A., et al. 2022, MNRAS, 512, 2671
Li, F. K., & Clark, G. W. 1974, ApJ, 191, L27
Lomaeva, M., Grinberg, V., Guainazzi, M., et al. 2020, A&A, 641, A144
Maeda, Y., Koyama, K., Sakano, M., Takeshima, T., & Yamauchi, S. 1996, PASJ, 48, 417
Martínez-Núñez, S., Kretschmar, P., Bozzo, E., et al. 2017, Space Sci. Rev., 212, 59
Mastroserio, G., Ingram, A., Wang, J., et al. 2021, MNRAS, 507, 55
Miller-Jones, J. C. A., Bahramian, A., Orosz, J. A., et al. 2021, Science, 371, 1046
Miškovičová, I., Hell, N., Hanke, M., et al. 2016, A&A, 590, A114
Motta, S., Belloni, T., & Homan, J. 2009, MNRAS, 400, 1603
Neijssel, C. J., Vinciguerra, S., Vigna-Gómez, A., et al. 2021, ApJ, 908, 118
Ng, C., Díaz Trigo, M., Cadolle Bel, M., & Migliari, S. 2010, A&A, 522, A96
Noble, M. S., & Nowak, M. A. 2008, PASP, 120, 821
Nowak, M. A., Hanke, M., Trowbridge, S. N., et al. 2011, ApJ, 728, 13
Oskinova, L. M., Feldmeier, A., & Kretschmar, P. 2012, MNRAS, 421, 2820
Owocki, S. P., Castor, J. I., & Rybicki, G. B. 1988, ApJ, 335, 914
Paczynski, B. 1976, Proc. Symp., 73, 75
Poutanen, J., Zdziarski, A. A., & Ibragimov, A. 2008, MNRAS, 389, 1427
Poutanen, J., Veledina, A., & Beloborodov, A. M. 2023, ArXiv e-prints [arXiv:2302.11674]
Puls, J., Vink, J. S., & Najarro, F. 2008, A&ARv, 16, 209
Rubio-Díez, M. M., Sundqvist, J. O., Najarro, F., et al. 2022, A&A, 658, A61
Saladino, M. I., Pols, O. R., van der Helm, E., Pelupessy, I., & Portegies Zwart, S. 2018, A&A, 618, A50
Scaringi, S., Körding, E., Uttley, P., et al. 2012, MNRAS, 421, 2854
Skipper, C. J., McHardy, I. M., & Maccarone, T. J. 2013, MNRAS, 434, 574
Strüder, L., Briel, U., Dennerl, K., et al. 2001, A&A, 365, L18
Sundqvist, J. O., & Owocki, S. P. 2013, MNRAS, 428, 1837
Sundqvist, J. O., & Puls, J. 2018, A&A, 619, A59
Sundqvist, J. O., Owocki, S. P., & Puls, J. 2018, A&A, 611, A17
Tarter, C. B., Tucker, W. H., & Salpeter, E. E. 1969, ApJ, 156, 943
Tomsick, J. A., Parker, M. L., García, J. A., et al. 2018, ApJ, 855, 3
Uttley, P., McHardy, I. M., & Vaughan, S. 2005, MNRAS, 359, 345
Verner, D. A., Ferland, G. J., Korista, K. T., & Yakovlev, D. G. 1996, ApJ, 465, 487
Wilkinson, T., & Uttley, P. 2009, MNRAS, 397, 666
Wilms, J., Allen, A., & McCray, R. 2000, ApJ, 542, 914
Zhou, M., Grinberg, V., Bu, Q. C., et al. 2022, A&A, 666, A172

Appendix A:

The model used to describe the broad band primary continuum for simulations reported in Sects. 3.2 and 3.3 includes a disc black body component, a soft Comptonisation component, and a hard Comptonisation component and its associated reflection, that is, `TBabs × [diskbb + nthComp + relxillCp]` in `XSPEC`. We froze the BH spin at the maximum value $a = 0.998$ (which allows the inner disc truncation radius to span the largest range of values), the inclination of the reflector i at 27° (Miller-Jones et al. 2021, see also recent results by Krawczynski et al. 2022 and Poutanen et al. 2023) and the high-energy cut-off of the hard Comptonisation component at 100 keV (Basak et al. 2017). The `nthComp` component is included in the model to account for the observed "soft excess", as done in e.g. Basak et al. (2017). The seed photon temperature of the soft Comptonisation component is tied to the best-fit inner disc temperature of the `diskbb` component.

This model was fit to the spectrum of the *XMM-Newton* observation 201 of Cyg X-1 after filtering out strong dips (the NWA dataset selected in Lai et al. 2022). After fitting this model, the data still show narrow residuals, mainly between ~0.5−2.5 keV. Since these residuals appear close to the absorption edges of the response matrix, i.e. at $E \approx 0.528$ keV, and between $E \approx 1.83-1.87$ keV, they are likely due to calibration problems. In order to account for these residuals, we modified the response matrix using the `gain` function in `XSPEC`. This command permits to change the slope or the intercept of the effective area curve, shifting the energies at which the response matrix is defined. Therefore, we introduced a linear gain shift (`intercept` parameter in `XSPEC`) of 0.01 keV (`slope` parameter of `gain` fixed to 0). As a result, the best-fit strongly improves (from χ^2/dof = 4404.17/1891 to χ^2/dof = 3923.7/1891, see Fig. A.1).

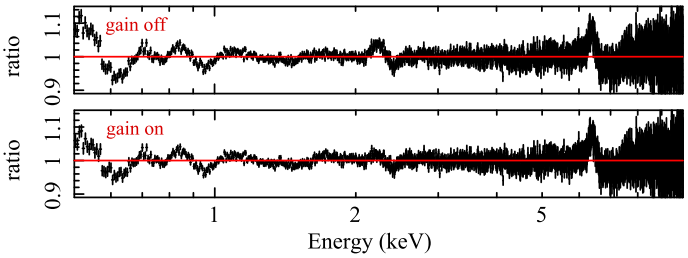

Fig. A.1. Comparison of the data to model ratios before (upper panel) and after (bottom panel) using the `gain` correction (`intercept` parameter set to 0.01 keV and `slope` parameter fixed to 0).

Nonetheless, some residuals were still present, including a strong narrow excess in the Fe Kα region. To account for these features, we added narrow ($\sigma < 0.1$ keV) emission and absorption gaussian components (see Tab. A.1). We did not investigate the nature of these features, but we point out that they might be due to incorrect calibration or residual wind absorption after the selection process of the NWA GTIs (Lai et al. 2022). On the other hand, the narrow feature observed in the Fe Kα region might indicate the need to include a second reflection component (such as from the outer disc) to obtain a better description of the broad band continuum. Nonetheless, we stress that our analysis, which focusses on testing the effects of variable absorption on the colour-colour tracks, is not strongly influenced by the details of the continuum model as long as it provides a good description of the broad-band spectrum.

Table A.1. Best-fit parameters of the continuum model.

	Component	Parameter	Value
ISM absorption	`TBabs`	$N_{\rm H}$ (10^{22}cm^{-2})	$0.79^{+0.01}_{-0.02}$
`continuum`	`diskbb`	kT_{bb} (keV)	< 0.14
	`nthComp`	Γ	< 5.9
		kT_e (keV)	> 0.67
	`relxillCp`	R_{in} (R_g)	13.24 ± 1.26
		Γ	1.59 ± 0.01
		$\log_{10}\xi$	3.32 ± 0.01
		A_{Fe}	$1.01^{+0.05}_{-0.02}$
		$\mathcal{R}$	$0.52^{+0.01}_{-0.03}$
`absorption features`	`gau1`	E_l (keV)	1.44 ± 0.01
	`gau2`	E_l (keV)	2.40 ± 0.01
`emission features`	`gau3`	E_l (keV)	6.36 ± 0.01
	`gau4`	E_l (keV)	0.52 ± 0.01
	`gau5`	E_l (keV)	0.86 ± 0.01
	`gau6`	E_l (keV)	1.05 ± 0.01

Notes. Best-fit parameters were obtained from the fit of the NWA spectrum of observation 201. The errors are reported at the 90% confidence level, i.e. $\Delta\chi^2 = 2.71$.

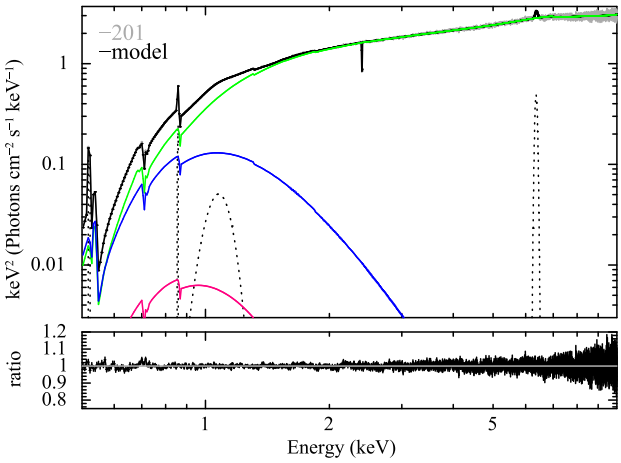

Fig. A.2. The best-fit model of the NWA time-averaged spectrum of observation 201. The complex best-fit model is overplotted in black, while the single components are shown in different colours: in magenta the disc blackbody, in blue the soft excess, in green the hard Comptonisation component and its relative reflection and in dotted grey the additional gaussians. Ratios of the data to the best-fitting model are shown in the bottom panel.

The final best-fit model yields χ^2/dof = 2263.76/1873. Fig. A.2 shows the spectrum of the NWA time-averaged spectrum of observation 201 and its best-fit model. The most relevant parameters are listed in Tab. A.1.

Appendix B: Time-resolved colour-colour diagrams: best-fit models

In Fig. B.1 we show the time-resolved colour-colour diagrams of observation 201 (colour-coded as in Fig. 7), and the best-fit models obtained from the fits discussed in Sect. 4.

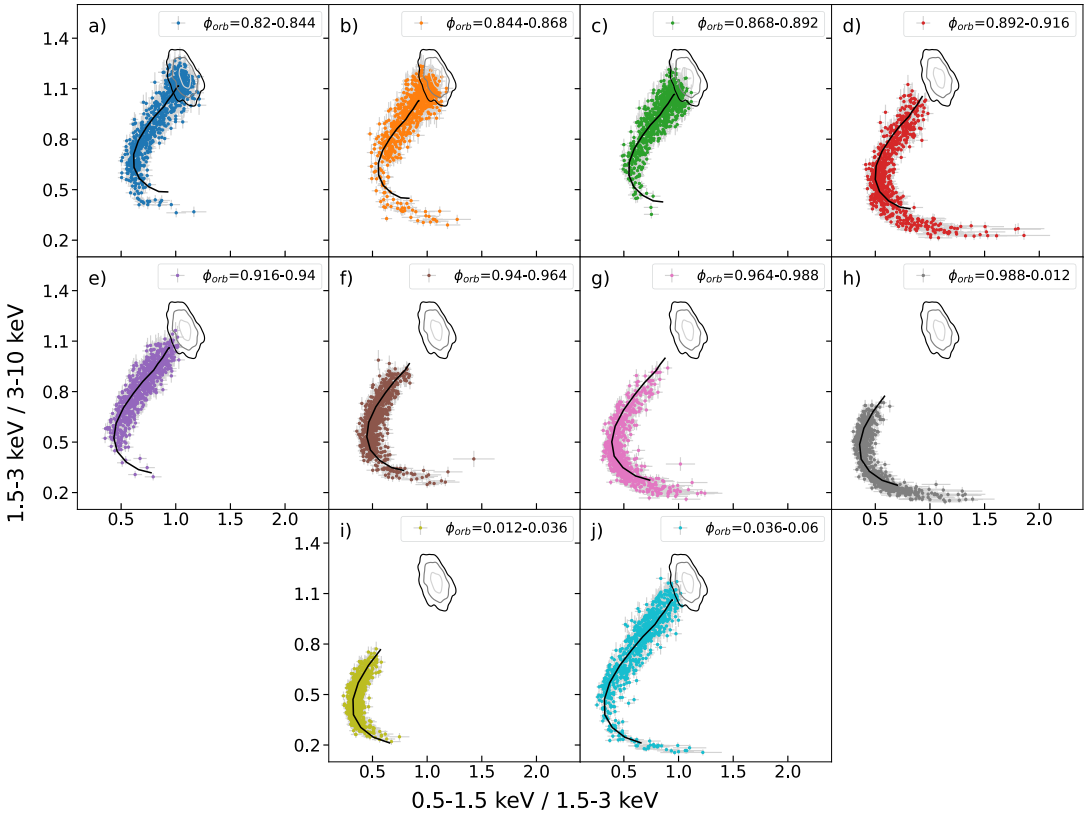

Fig. B.1. Time-resolved colour-colour diagrams of observation 201. Overplotted (black tracks) are the best-fit models obtained assuming $\log \xi = \log \frac{10+[N_{\rm H,w}/10^{22}\,{\rm cm}^{-2}]}{[N_{\rm H,w}/10^{22}\,{\rm cm}^{-2}]} + 1$. The contour plots represent the 99.7% (in black), 95% (in grey) and 68% (in light grey) confidence levels of the data distribution during the least absorbed stages of the orbit ($\phi_{\rm orb} = 0.43 - 0.46$, see Appendix C).

Appendix C: Probability distribution close to inferior conjunction

We show the colour-colour diagram calculated from orbital phases $\phi_{\rm orb} = 0.43{-}0.46$, covered by observation 501 (Lai et al. 2022). These phases are selected to be the closest to inferior conjunction (i.e. $\phi_{\rm orb} \sim 0.5$), and thus the least affected by wind absorption. Using the KDE method, we computed the probability distribution for this dataset (Sect. 4). Fig. C.1 shows the resulting 99.7%, 95% and 68% confidence contours overplotted to the data.

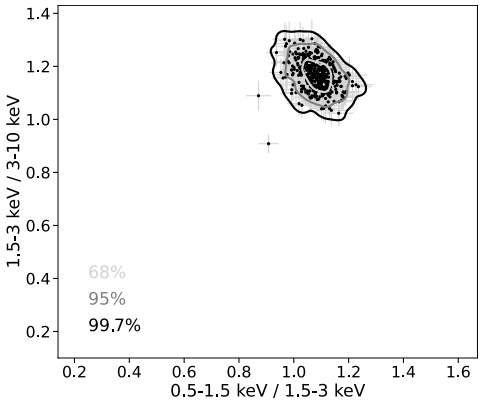

Fig. C.1. Colour-colour diagram and probability distribution map for orbital phases $\phi_{\rm orb} = 0.43{-}0.46$. The curves correspond to the 99.7% (in black), 95% (in grey) and 68% (in light grey) confidence regions.