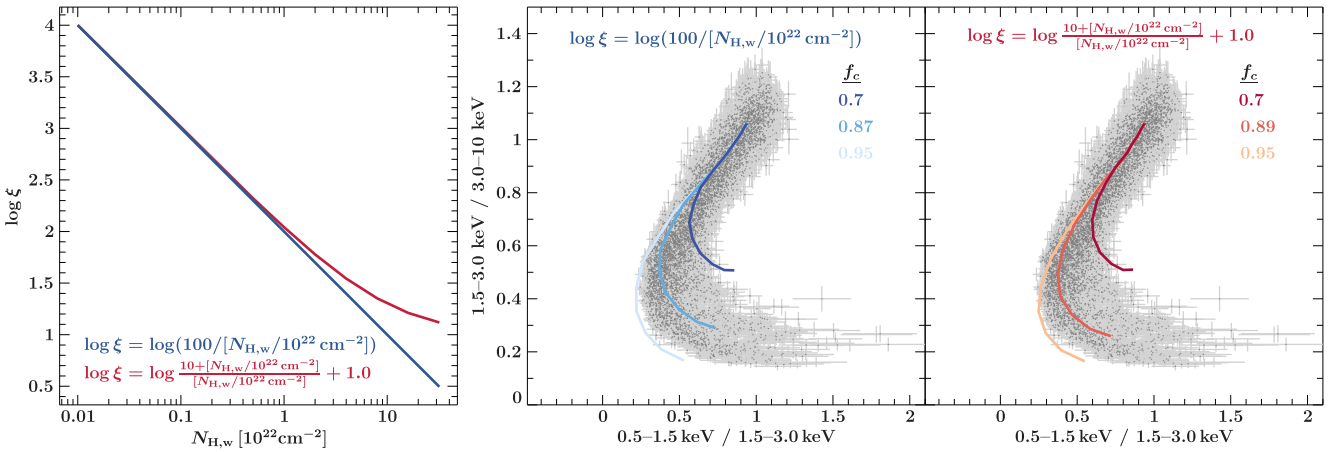

Fig. 6. Empirical functions describing the dependence of the ionisation parameter on the column density of the absorbing gas. Left panel: Assumed dependencies of the ionisation parameter, $\log\xi$, on the column density, $N_{H,w}$. Middle and right panel: Simulated tracks for a warm absorbing gas, assuming $\log\xi = \log(100/[N_{H,w}/10^{22}\,\mathrm{cm}^{-2}])$ (middle panel), and $\log\xi = \log\frac{10+[N_{H,w}/10^{22}\,\mathrm{cm}^{-2}]}{[N_{H,w}/10^{22}\,\mathrm{cm}^{-2}]} + 1$ (right panel). The different shades correspond to different values of covering factor as reported in the labels.

in part due to intrinsic variations in the covering factor between ∼0.7–0.95. We investigate this problem further in the following section.

4. Time-resolved colour-colour diagrams

To study the phase-dependent variability of the absorbing column density and covering fraction of the wind, we investigated the temporal evolution of the shape of the colour-colour diagram. We divided observation 201 in 10 segments of the same duration of ∼11 ks each, corresponding to segments of ∼0.024 in orbital phase, and produced colour-colour diagrams for each segment. The corresponding tracks are all shown in Fig. 7, colour-coded based on the range of orbital phases. The plot shows a strong temporal evolution of the tracks. The same data are shown in Fig. B.1, split in separate time-resolved colour-colour diagrams. Using the KDE method (see Sect. 3.3), we inferred the PDF of the data in each diagram. The resulting probability distribution maps are shown in Fig. 8.

The passage at superior conjunction happens at $t \sim 86.36$ ks after the starting time of the observation (see Fig. 2). Therefore, it is contained in panel "h" corresponding to $\phi_{orb} = 0.988–0.012$. The fact that the LOS passes through the densest regions of the stellar wind in this phase of the orbit is clearly seen from the shape of the corresponding colour-colour track, whereby 'only' the lower region of the diagram is populated (i.e. at hard colours $\lesssim$0.75). A similar distribution is observed both at earlier and later phases of the orbit from panel "f" to panel "i" due to strong dipping events occurring several hours before and after superior conjunction (see Fig. 2).

We built simulated tracks for each time-resolved colour-colour diagram in order to constrain the changes in the wind parameters, which can cause the observed changes in the colour-colour tracks. To this aim, we used the complex continuum model fitted to the spectrum of each time-resolved epoch described in Sect. 3.2 and the variable ionisation absorption component described by the empirical function $\log\xi = \log\frac{10+[N_{H,w}/10^{22}\,\mathrm{cm}^{-2}]}{[N_{H,w}/10^{22}\,\mathrm{cm}^{-2}]} + 1$, introduced in Sect. 3.3 (red curve in the left panel of Fig. 6). Although this function cannot be statistically preferred to the other one tested in Sect. 3.3 (blue curve in the left panel of Fig. 6), our choice is dictated by the willingness

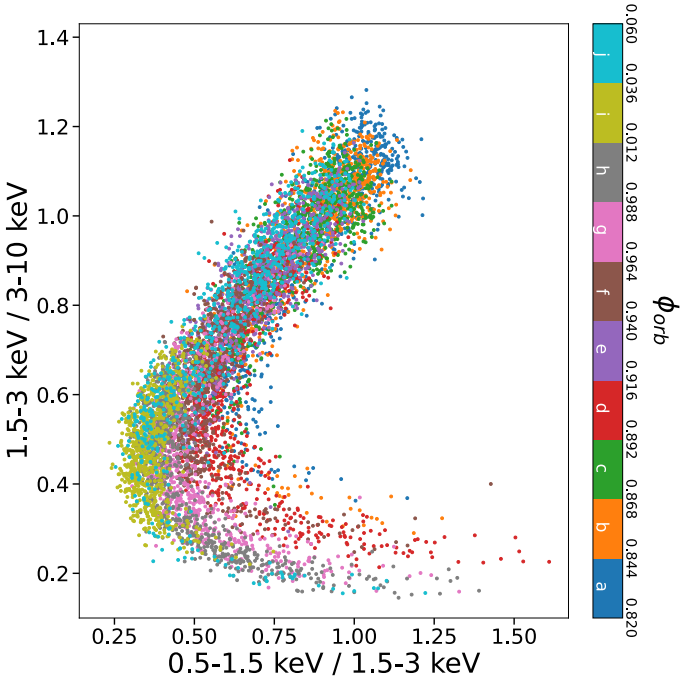

Fig. 7. Time-resolved colour-colour diagrams of observation 201, extracted from segments with a duration of 11 ks (corresponding to ∼0.024 in orbital phase) and time resolution of 10 s. The resulting ten colour-colour diagrams are overplotted using different colours that correspond to different phase intervals, according to the colour map reported on the right.

to include the effects of additional sources of ionisation becoming important in denser environments (e.g. Feldmeier et al. 1997; Sundqvist & Puls 2018), such as the base of the wind, intercepted by the LOS at superior conjunction. We selected as best-fit tracks those with the highest combined likelihood, as described in Sect. 3.3. The best-fit tracks are shown in Fig. 8 (as well as in Fig. B.1), and the corresponding best-fit covering factors for each time-resolved diagram are listed in Table 1.

We notice that the best-fit tracks do not always extend to high hard colours as much as the data do. This particularly happens

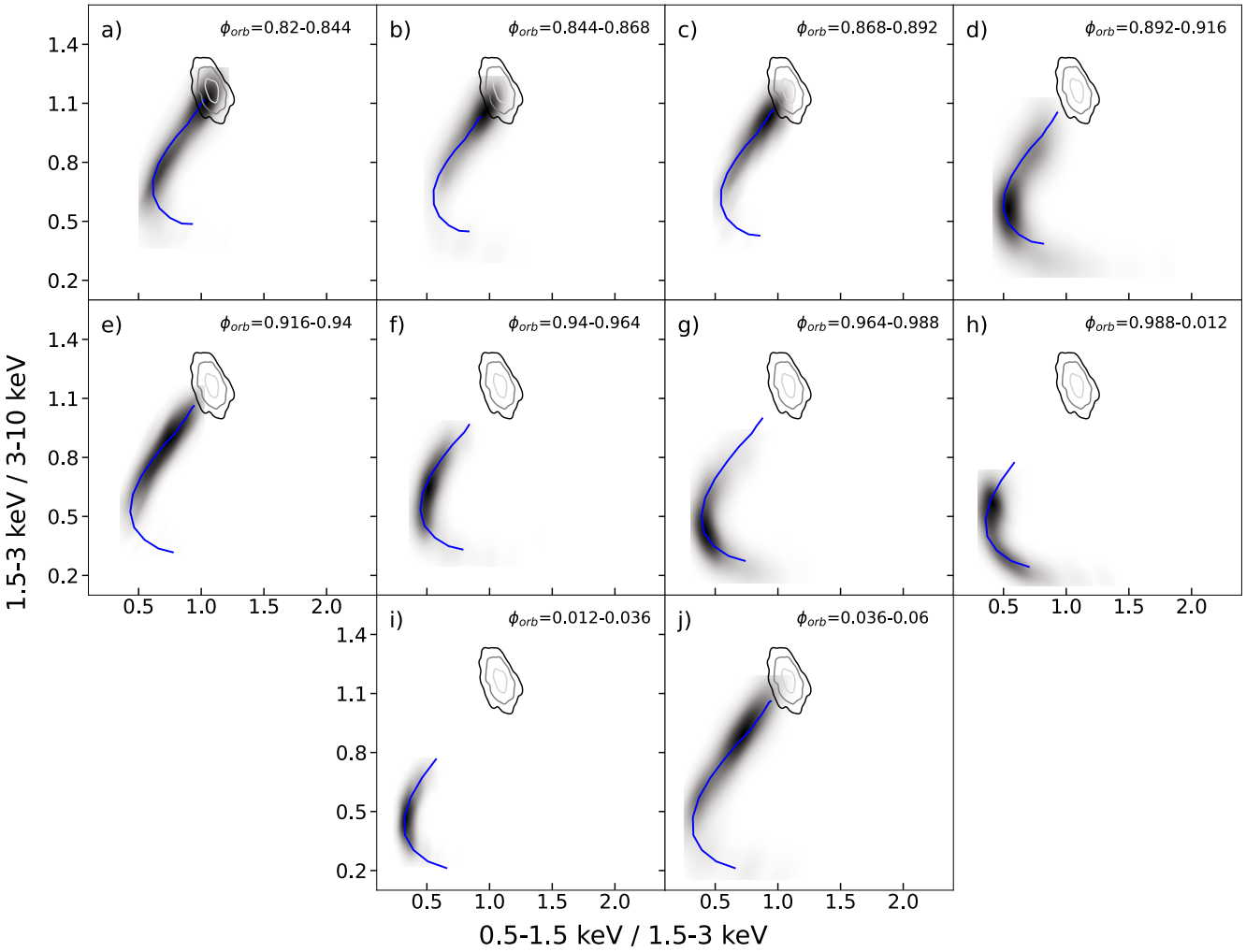

Fig. 8. Probability distribution maps of each time-resolved colour-colour diagram obtained using the KDE method. In blue, the best-fit simulated track. The grey-shaded closed curves represent the probability distribution map for orbital phases $\phi_{\rm orb}$ = 0.43−0.46 at the 99.7% (in black), 95% (in grey), and 68% (in light grey) confidence levels.

in the less absorbed phases of the orbit when the source is out of superior conjunction (i.e. panels a, b, c, and j). This is because the model reaches the maximum allowed value for the ionisation parameter (in our model, the lowest values of $N_{\rm H,w}$ correspond to the highest values of ξ, Fig. 6, left panel). The upper portion of the diagram not covered by the wind model track is populated by the least absorbed data bins. Therefore, the scattering in the measured hard and soft colours within this region might be completely driven by the intrinsic spectral variability of the source. To verify this, we calculated a colour-colour diagram from observation 501, which is one of the least affected by the wind as it was carried out right before inferior conjunction ($\phi_{\rm orb}$ = 0.5; see Lai et al. 2022). We note that the CHOCBOX campaign did not catch the passage at inferior conjunction, and therefore from this observation we selected the closest orbital phases to it ($\phi_{\rm orb}$ = 0.43−0.46). Observation 501 is also consecutive to observation 201. This allows us to rely on the implicit assumption that the intrinsic continuum spectrum of the X-ray source has not changed significantly between the two observations, and that any scatter in the colour-colour diagram would only be due to fluctuations in the parameters of the same continuum model. Using the KDE method (Sect. 3.3), we calculated the probability distribution of soft and hard colours in the diagram. The colour-colour diagram of observation 501 and the corresponding contour plots are reported in Appendix C and Fig. C.1. As it can be seen, the data are quite scattered, suggesting intrinsic short-term spectral variability of the continuum. In Figs. 8 and B.1 we overplot the 99.7%, 95% and 68% confidence contours of observation 501, to mark the area in the diagram where the intrinsic emission from the X-ray source dominates. In other words, in this region the probability for the source to be unabsorbed by the stellar wind is the highest. The data points not reached by the wind model tracks are thus fully consistent with being free from wind absorption.

From the fit of the time-resolved colour-colour diagrams we infer a steady and significant increase in the covering factor, by a factor of ∼1.2 between $\phi_{\rm orb}$ = 0.820−0.844 and $\phi_{\rm orb}$ = 0.036−0.06 (see Table 1). The observed trend is plotted in Fig. 9 in the upper panel. The maximum value of the covering factor is registered at superior conjunction, but the data suggest this parameter remains high up to at least $\phi_{\rm orb}$ = 0.06. From the PDFs of Fig. 8 and our best-fit tracks we extracted the average column density $\overline{N}_{\rm H,w}$ and its scatter $\delta N_{\rm H,w}$, as this parameter is an indicator of the amount of short-term variability of $N_{\rm H,w}$ within each epoch. These values are reported in Table 1 and plotted in Fig. 9 (middle and lower panels) for each epoch. Our procedure is this: given the best-fit track, we extract the probability of each point from the KDE, which is thus a function of $N_{\rm H,w}$. This probability

Table 1. Time-resolved colour-colour diagrams best-fit parameters.

$\phi_{\rm orb}$	$f_{\rm c}$	$\overline{N}_{\rm H,w}$	$\delta N_{\rm H,w}$
0.820–0.844	0.74 ± 0.01	6.88 ± 0.02	4.93 ± 0.03
0.844–0.868	0.74 ± 0.01	6.22 ± 0.27	6.70 ± 0.53
0.868–0.892	0.76 ± 0.01	6.00 ± 0.15	6.6 ± 0.3
0.892–0.916	0.80 ± 0.01	11.78 ± 0.48	11.84 ± 0.97
0.916–0.940	0.85 ± 0.01	5.48 ± 0.23	7.12 ± 0.47
0.940–0.964	0.84 ± 0.01	9.76 ± 0.46	9.51 ± 0.91
0.964–0.988	0.88 ± 0.01	13.55 ± 0.56	10.92 ± 1.11
0.988–0.012	0.90 ± 0.01	16.09 ± 1.31	16.15 ± 2.62
0.012–0.036	0.92 ± 0.01	10.80 ± 0.28	7.02 ± 0.55
0.036–0.060	0.92 ± 0.01	6.52 ± 0.06	8.38 ± 0.11

Notes. The table reports the ranges of orbital phases corresponding to each time-resolved diagram of observation 201. The average column density $\overline{N}_{\rm H,w}$ and its scatter $\delta N_{\rm H,w}$ are in units of 10^{22} cm^{-2}.

is a proxy for the number of data points with a given $N_{\rm H,w}$. Then, we consider the range of $N_{\rm H,w}$ that encloses 68% of the total (renormalised) probability for the corresponding track to be a proxy for the total number of data points. The width of the range gives us $\delta N_{\rm H,w}$ and its average value gives us $\overline{N}_{\rm H,w}$. We reckon this approach to be more informative than, for example, extracting the $N_{\rm H,w}$ with the highest probability, since the latter method would result in a loss of information, especially for the distributions that show more than one peak. For example, the fit of the colour-colour diagram of panel "a" of Fig. 8 would return a low $N_{\rm H,w}$, thus missing the information associated with the second peak of the PDF. Our procedure also provides a straightforward way to obtain estimates of the confidence intervals for each pair of $\overline{N}_{\rm H,w}$ and $\delta N_{\rm H,w}$ values. We now consider the two tracks corresponding to the extrema of the 68% confidence level of $f_{\rm c}$ and repeat the procedure obtaining new confidence level of $N_{\rm H,w}$. To be conservative, we take the overall maximal and minimal confidence level and from these obtain the intervals of $\overline{N}_{\rm H,w}$ and $\delta N_{\rm H,w}$.

We observe significant variations in both the $\overline{N}_{\rm H,w}$ and the $\delta N_{\rm H,w}$ parameters. While the two show the same trend of variations (when the average column density increases, its scatter also increases), their trend differs from that of $f_{\rm c}$. In particular, both the $\overline{N}_{\rm H,w}$ and the $\delta N_{\rm H,w}$ show two peaks, the second one happening at superior conjunction, when also $f_{\rm c}$ reaches its maximum value. However, when $f_{\rm c}$ is still at its maximum, $\overline{N}_{\rm H,w}$ and $\delta N_{\rm H,w}$ drop by a factor of ∼2.5 and ∼2, respectively. This is readily visible in panel "j" of Figs. 8 and B.1, where the data populate again the region of high hard colours of the colour-colour diagram (as in panels "a–e") while also stretching down to the lowest registered hard colours. Finally, it is worth noticing that the measured values of $\overline{N}_{\rm H,w}$ (between $\sim 6 \times 10^{22}$ cm^{-2} and $\sim 16 \times 10^{22}$ cm^{-2}) all fall in the regime where our model predicts additional ionisation in high density environments to become significant (see Fig. 6, left panel). In this regime, strong variations in $\overline{N}_{\rm H,w}$ along the dips are not mirrored by equally strong variations in the ionisation parameter, which remains constrained between $\log \xi \sim 1.2$–1.4.

On the soft tail of colour-colour tracks

Before discussing our results, we note that none of the models used to fit the colour-colour diagrams (Sects. 3 and 4) produces tracks that can explain the extended tail at high values of soft colours characterising the most absorbed stages of the dips (lower part of the colour-colour diagram, see e.g. Figs. 6 and B.1).

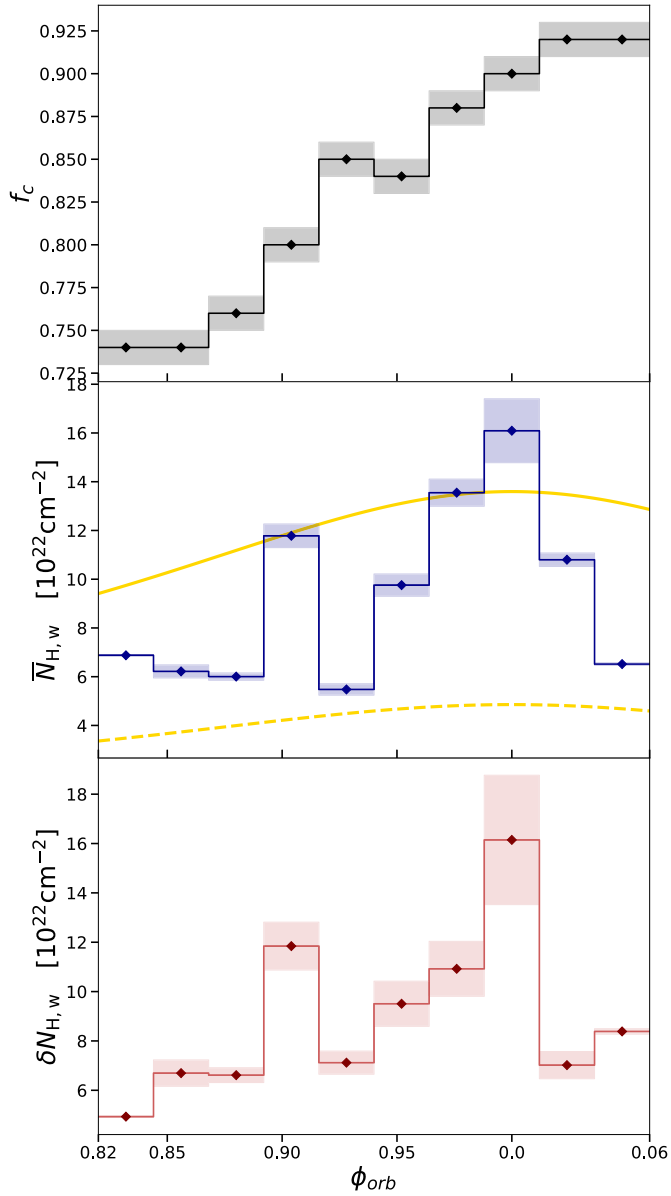

Fig. 9. Stellar wind parameters as obtained from the fit of the time-resolved colour-colour diagrams of observation 201, plotted as a function of the orbital phase (see also Table 1). The parameters are: the covering factor, $f_{\rm c}$ (upper panel), the mean column density, $\overline{N}_{\rm H,w}$ (middle panel), and the scatter in column density parameter, $\delta N_{\rm H,w}$ (lower panel). The shaded areas represent the 68% confidence contours. The yellow curves are the neutral column density profiles produced by the smooth wind model (El Mellah et al. 2020) at high (solid) and low (dashed) stellar mass loss rates.

Such high-value soft colours indicate the presence of an additional soft component that becomes significant in the deeper parts of the dips. The most likely candidates for explaining this behaviour are the contribution from emission lines produced by the wind material, and/or from a dust-scattering halo (e.g. Nowak et al. 2011). The former possibility is supported by results presented in Hirsch et al. (2019), who revealed the pres-

ence of an emission line spectrum from the photoionised plasma around the BH, emerging only when the X-ray source is highly absorbed (down in the dips). Additional contribution from a dust-scattering halo cannot be excluded. Indeed, source photons intercepted by foreground interstellar dust on the LOS are scattered away, thus dimming the source, but interstellar dust grains outside the LOS can redirect photons back to the observer, producing a halo around the source. Given the energy dependence of the dust-scattering cross-section (Corrales et al. 2016) soft X-ray photons will have more chance to be scattered back, producing an excess of soft X-ray flux in the halo (Maeda et al. 1996).

In order to check for the presence of a dust-scattering halo component in the *XMM-Newton* EPIC-pn data of Cyg X-1, we verified whether the spectrum significantly softens when extracting photons from increasingly more external regions, as described in Jin et al. (2017). It is worth noticing that the timing mode does not allow us to select regions totally unaffected by a dust-scattering halo since all pixels in the same column lose their spatial information along RAWY. Nonetheless, if a scattering halo is present and the X-ray source is bright, the inner columns will still be harder than the external columns due to the X-ray source dominating the central columns and the halo dominating the external ones.

We extracted spectra in the 2−10 keV band during the time interval ∼82−94 ks that comprises the passage at superior conjunction (see Fig. 10). Indeed, the emission from the halo is expected to be more significant when the X-ray source is highly absorbed (e.g. Jin et al. 2018), and thus in the deepest stages of the dips. We selected three detector regions at increasingly higher distance from the central pixels (and discarding the central RAWX = [36:39] to avoid pile-up effects, Sect. 2): RAWX = [32:35]−[40:43] covering an angular diameter $D = 8''-16''$ from the central pixel, RAWX = [30:32]−[43:46] ($D = 16''-32''$) and RAWX = [27:30]−[46:49] ($D = 32''-44''$). A gradual softening of the Γ parameter is observed, up to ∼10% in the most external region. In particular, $\Gamma = 1.42 \pm 0.01$ for RAWX = [32:35]−[40:43], $\Gamma = 1.54 \pm 0.01$ for RAWX = [30:32]−[43:46], and $\Gamma = 1.58 \pm 0.01$ for RAWX = [27:30]−[46:49]. This gradual softening hints at the presence of a dust-scattering halo component. We conclude that the dust-scattering halo can potentially contribute to the soft tail observed in the colour-colour tracks. Nonetheless, we notice that both a dust-scattering halo and the emission line component from the wind would contribute the most when the source is highly absorbed. Thus, we expect a soft tail to be particularly prominent when the uppermost part of the colour-colour diagram is not populated. However, all the observed time-resolved colour-colour diagrams follow these expectations (more visible in Fig. B.1), except the one at $\phi_{\rm orb} = 0.012-0.036$ (see panel "i" of Fig. B.1), which does not show a prominent soft tail despite being highly absorbed ($\overline{N}_{\rm H,w} \sim 10.80\times10^{22}$ cm^{-2} and $f_{\rm c} \sim 0.92$). We suggest this might be due to a delayed response of the illuminated dust halo or emission line region to a change of the irradiating flux from the X-ray source (see Fig. 2).

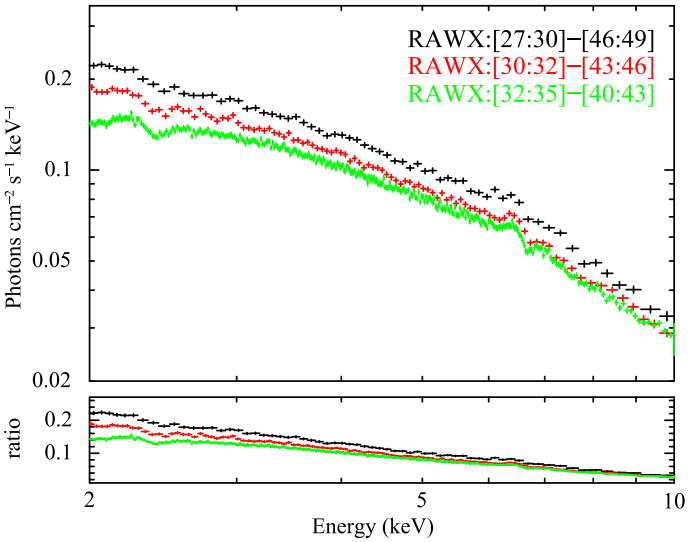

Fig. 10. Unfolded (against a constant model fixed at 1) 2−10 keV EPIC-pn spectra (upper panel) extracted in the time interval $t \sim 82-94$ ks, catching the passage at superior conjunction. The chosen RAWX selections for each spectrum are reported in the labels. The softening of the spectra suggests that there is contribution from a dust-scattering halo, easily noticeable in the ratios of the unfolded spectra (bottom panel).

5. Discussion

Several studies have demonstrated the complexity of the stellar wind in Cyg X-1 (e.g. Grinberg et al. 2015; Miškovičová et al. 2016; Hirsch et al. 2019; Grinberg et al. 2020). The wind appears to be formed of over dense clumps (e.g. Owocki et al. 1988; Feldmeier et al. 1997; Oskinova et al. 2012; Sundqvist & Owocki 2013) crossing our LOS to the X-ray source, producing stochastic variability that typically manifests itself as absorption dips in the X-ray light curves, and adds up to the intrinsic variability produced in the inner accretion flow (Lai et al. 2022). Absorption dips occurrence and concurrent absorption variability are more prominent near superior conjunction (e.g. Bałucińska-Church et al. 2000; Lai et al. 2022), when the companion star is between the observer and the BH ($\phi_{\rm orb} = 0$), as in these phases the LOS crosses deeper wind layers.

In this work, we have presented a characterisation of the stellar wind properties in Cyg X-1, through the analysis and modelling of the colour-colour diagrams of the source during a passage at superior conjunction (between orbital phases 0.82 and 0.06, see Figs. 1 and 2), with the X-ray source in its hard state. The prominence of absorption dips in this specific state, rather than in softer states, are explained as due to the X-ray source not being strong enough to fully ionise the wind (e.g. Miškovičová et al. 2016; Grinberg et al. 2015). In addition to previous studies, we employed a non-parametric fitting method based on the kernel density estimation analysis (Hill 1985), which allowed us to select the model that best describes the observed colour-colour diagrams and constrain the physical parameters of the stellar wind (Sects. 3.3 and 4).

5.1. Modelling of colour-colour diagrams and evolution of stellar wind parameters

We confirmed (see also Grinberg et al. 2020 who analysed a *Chandra* 2004 observation of the source covering similar orbital phases) that the characteristic 'pointy' or 'nose-like' shape of the colour-colour tracks requires the absorbing gas to be partially ionised (Sect. 3.2), in agreement with independent results from high-resolution spectroscopy (e.g. Hanke et al. 2009; Miškovičová et al. 2016). In particular, Hirsch et al. (2019) revealed different levels of ionisation at different absorption stages. This suggests an (unknown) functional dependence of $\log\xi$ on $N_{\rm H,w}$ (Grinberg et al. 2020). We tested two empirical functions (Sect. 3.3), one assuming a simple lin-

ear scaling and the other allowing for additional sources of ionisation at high column densities (Fig. 6). Although the data do not allow us to statistically prefer either of the two models, we notice that simulations show that collisions in the ambient stellar wind occurring in denser regions (high $N_{\rm H,w}$ regime) are expected to produce significant X-ray emission (Feldmeier et al. 1997; Sundqvist & Puls 2018). This emission can contribute to further ionise the gas.

In order to characterise the evolution of the wind as a function of the orbital phase around superior conjunction, we extracted time-resolved colour-colour diagrams, and fit them with a continuum plus variable ionisation model (see Sect. 4). The fits show a steady increase in the covering factor, $f_{\rm c}$, by a factor of ∼1.2 between $\phi_{\rm orb} \sim 0.8$ and $\phi_{\rm orb} \sim 0$, which remains high also after superior conjunction (up to at least $\phi_{\rm orb} \sim 0.06$, Fig. 9 top panel and Table 1). On the other hand, the mean column density $\overline{N}_{\rm H,w}$ shows two peaks, one at $\phi_{\rm orb} \sim 0.9$ and the other at superior conjunction. After superior conjunction, while $f_{\rm c}$ is still high, $\overline{N}_{\rm H,w}$ drops to a minimum (Fig. 9 middle and Table 1). We also measured the scatter parameter, $\delta N_{\rm H,w}$, which quantifies the amount of variability of $N_{\rm H,w}$ within each time-resolved colour-colour diagram (Fig. 9 bottom panel and Table 1). This parameter follows the same trend displayed by $\overline{N}_{\rm H,w}$ (see also Fig. 11). Given the sampling adopted for the extraction of time-resolved colour-colour diagrams (10 segments of ∼11 ks), while the inferred $\overline{N}_{\rm H,w}$ parameter probes relatively long-term modulations, the $\delta N_{\rm H,w}$ parameter probes more rapid variability (between 10 s, the time bins used to build the colour-colour diagrams, and ∼11 ks) likely driven by the smallest-scale inhomogeneities in the wind. Our study hints at a one-to-one relation between the amount of absorption at a given phase and its associated rapid variability: the higher the $\overline{N}_{\rm H,w}$, the higher its scatter $\delta N_{\rm H,w}$ (see Fig. 11). This is reminiscent of the "rms-flux" relation characterising the stochastic flux variability in compact object systems (e.g. Uttley et al. 2005; Scaringi et al. 2012), which implies a log-normal distribution of the flux and requires variability on different timescales to combine multiplicatively (Uttley et al. 2005). Since the size of the clumps influences the observed column density, confirmation of such correlation might have implications on the distribution of clump sizes and on the way they combine to form bigger clumps.

5.2. Stellar wind mass loss rate and clump mass estimations

The variations in the absorbing column density stem from two aspects. First, we expect a periodic variation due to the orbital motion of the BH around its stellar companion. This smooth component can be described by the absorption from a smooth wind, but it falls short at capturing the stochastic variability on timescales much shorter than the orbital period. The latter comes from clumps intercepting the LOS between the observer and the BH (El Mellah et al. 2020). The average column density time series match the smooth wind profile of same mass loss rate, while the dispersion above and below tell us about the clumps' properties.

The data considered in this paper cover only a small portion of the orbital period, so they cannot be used to constrain the periodic variations in column density. Instead, we plotted in the middle panel in Fig. 9, for two different stellar mass loss rates, the periodic component of the column density due to the orbital motion (yellow lines). It was computed based on the smooth and spherically symmetric stellar wind model of El Mellah et al. (2020): it accounts for all the material along the LOS to the BH, regardless of its ionisation state. Therefore, it is an estimate of the neutral column density, necessarily higher than the measurable ionised column density. If we neglect the orbital eccentricity and assume a β-law for the wind velocity profile, the shape of the neutral column density curve only depends on the orbital inclination i, on the ratio of the orbital separation d to the stellar radius R_* and on the β exponent of the velocity profile. We take $i = 27°$ and $d/R_* = 2.3$, which correspond to the values given by Miller-Jones et al. (2021), and $\beta = 1.5$, a representative value for UV-line driven winds of O supergiants (Rubio-Díez et al. 2022). Owing to the low inclination of Cyg X-1, the peak-to-peak variation remains modest across the orbit and no significant changes are expected for realistic values of d/R_* and β. On the other hand, the scale of the neutral column density curve is set by $\dot{M}_*/(m_{\rm p} v_\infty R_*)$, where $\dot{M}_*$ is the stellar mass loss rate and $m_{\rm p}$ is the proton mass. We use the values reported by Miller-Jones et al. (2021) to set $v_\infty = 2100$ km/s and $R_* = 22\,R_\odot$. Through Hα diagnostics, Gies et al. (2003) reported a stellar mass loss rate $\dot{M}_* \sim 2.6 \times 10^{-6}\,{\rm M}_\odot\,{\rm yr}^{-1}$ in the hard state. However, they do not take ionisation into account, and they rely on the stellar parameters used by Herrero et al. (1995), which do not correspond to the updated values. Therefore, there is a risk that this mass-loss rate is underestimated. And indeed, when plotted over our data points, the neutral column density profile with $\dot{M}_* \sim 2.6 \times 10^{-6}\,{\rm M}_\odot\,{\rm yr}^{-1}$ (dashed yellow line) lies below the data points. However, since the profiles, in Fig. 9, produced by the smooth wind model of El Mellah et al. (2020) represent the neutral column density while the data points stand for the absorbing column density (i.e. only the material at ionisation levels low enough to contribute to absorption), the smooth profile should be an upper limit. Therefore, we think a higher stellar mass loss rate, for instance $\dot{M}_* \sim 7 \times 10^{-6}\,{\rm M}_\odot\,{\rm yr}^{-1}$ (solid yellow line), is more compatible with the column density we measure at superior conjunction. Measurements of $\overline{N}_{\rm H,w}$ over multiple consecutive orbits are needed to obtain better constraints.

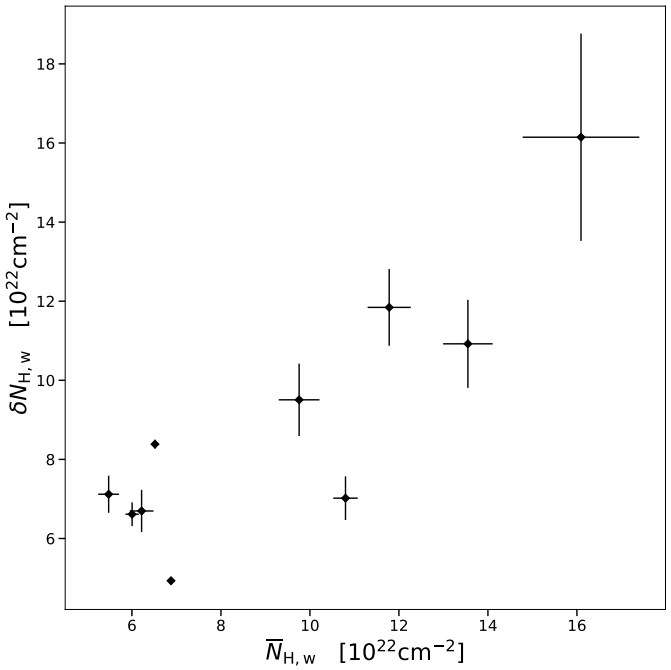

Fig. 11. Relation between the measured $\overline{N}_{\rm H,w}$ and $\delta N_{\rm H,w}$ parameters, respectively sampling long-term and rapid variations in the column density parameter of the stellar wind (the values are listed in Table 1).