

# Gaia DR2 Distances and Peculiar Velocities for Galactic Black Hole Transients

Poshak Gandhi,<sup>1</sup><sup>★</sup> Anjali Rao,<sup>1</sup> Michael A.C. Johnson,<sup>1,2</sup> John A. Paice,<sup>1,3</sup> Thomas J. Maccarone<sup>4</sup>

<sup>1</sup>Department of Physics & Astronomy, University of Southampton, Highfield, Southampton SO17 1BJ, UK

<sup>2</sup>Electronics and Computer Science, University of Southampton, Highfield, Southampton, SO17 1BJ, UK

<sup>3</sup>Inter-University Centre for Astronomy and Astrophysics, Pune, Maharashtra 411007, India

<sup>4</sup>Department of Physics & Astronomy, Box 41051, Science Building, Texas Tech University, Lubbock, TX 79409-1051, USA

Accepted 2019 Feb 05.

## ABSTRACT

We report on a first census of Galactic black hole X-ray binary (BHXR) properties with the second data release (DR2) of *Gaia*, focusing on dynamically confirmed and strong candidate black hole transients. DR2 provides five-parameter astrometric solutions including position, parallax and proper motion for 11 of a sample of 24 systems. Distance estimates are tested with parallax inversion as well as Bayesian inference. We derive an empirically motivated characteristic scale length of  $L = 2.17 \pm 0.12$  kpc for this BHXR population to infer distances based upon an exponentially decreasing space density prior. Geometric DR2 parallaxes provide new, independent distance estimates, but the faintness of this population in quiescence results in relatively large fractional distance uncertainties. Despite this, DR2 estimates generally agree with literature distances. The most discrepant case is BW Cir, for which detailed studies of the donor star have suggested a distant location at  $\gtrsim 25$  kpc. A large DR2 measured parallax and relatively high proper motion instead prefer significantly smaller distances, suggesting that the source may instead be amongst the nearest of XRBs. However, both distances create problems for interpretation of the source, and follow-up data are required to resolve its true nature. DR2 also provides a first distance estimate to one source, MAXI J1820+070, and novel proper motion estimates for 7 sources. Peculiar velocities relative to Galactic rotation exceed  $\approx 50$  km s<sup>-1</sup> for the bulk of the sample, with a median system kinetic energy of peculiar motion of  $\sim 5 \times 10^{47}$  erg. BW Cir could be a new high-velocity BHXR if its astrometry is confirmed. A putative anti-correlation between peculiar velocity and black hole mass is found, as expected in mass-dependent BH kick formation channels, but this trend remains weak in the DR2 data.

**Key words:** stars: black holes – stars: distances – parallaxes – proper motions – accretion, accretion discs

## 1 INTRODUCTION

Black hole X-ray binaries (BHXRBS) are key sources for studying accretion in the strong gravitational regime. About 60 low mass BHXRBS are known in the Galaxy, though only about 20 of these are well characterised and confirmed as hosting black holes, based upon dynamical measurements of the companion star orbiting around the primary (Casares & Jonker 2014; Corral-Santana et al. 2016; Tetarenko et al. 2016).

Distances to most XRBs remain highly uncertain. The most reliable distances come from geometric estimates based upon measurements of trigonometric parallax shifts relative to background

sources. However, such measurements require exquisitely sensitive observations over several years, which have not been generally available except in the radio for a handful of cases where very long baseline interferometric studies have been possible: e.g. Cyg X-1 (Reid et al. 2011), V404 Cyg (Miller-Jones et al. 2009) and GRS 1915+105 (Reid et al. 2014b). In the absence of geometric distances, estimates become reliant upon emission- and evolutionary-dependent model assumptions (cf., review by Jonker & Nelemans 2004). Testing the validity of such model-dependent distances is important in order to properly model source luminosities and understand the origin and evolution of these systems.

The *Gaia* mission (Gaia Collaboration et al. 2016a) has enabled a step change in the field of astrometric measurements of stars within the Galaxy. Geometric parallaxes along with highly accurate

\* E-mail: p.gandhi@soton.ac.uk

positions and proper motions for more than 1 billion stars are expected (Gaia Collaboration et al. 2016a). Among the population of Galactic sources, BHXRBS are expected to be found. The typical brightness of these systems in quiescence is close to  $V \sim 20$  mag or fainter. *Gaia* is expected to probe sources down to  $G \approx 21$  mag, where the  $G$  band covers 330–1050 nm. So the viability and utility of astrometric measurements with *Gaia* is not clear *a priori*.

A significant new data release (DR2, Gaia Collaboration et al. 2018) occurred on 2018 April 25, and we present here an analysis of the properties of Galactic BHXRBS as seen in DR2, focusing on systems that are well characterised in the optical – they are all either known to host a black hole based upon orbital dynamical measurements of the donor star, or are thought to be strong candidates to lie in this class. DR2 covers data gathered over 22 months from July 2014 to May 2016. We present astrometric measurements, including parallax and proper motion measurements where available, and examine their quality. These measurements are used to infer parallax-based distance estimates, and are compared with estimates from the literature based upon non-geometric techniques. We also present estimates of the peculiar Galactic velocities of the sources based upon three-dimensional kinematics, and carry out tests aimed at ascertaining formation pathways. This work is meant as a first look at the census of BHXRBS in this massive new parameter space of optical astrometric measurements. Future data releases can be expected to further improve upon this.

## 2 GAIA COUNTERPART SEARCH AND ASSOCIATION

We queried the DR2 public release on 2018 April 25. A search radius of  $2''$  was used for DR2 entries around each of 24 BHXRBS, though we also experimented with larger radii in individual ambiguous cases as detailed later. The targets are chosen from the BlackCAT catalogue (Corral-Santana et al. 2016) as having dynamically confirmed mass measurements or other strong evidence of hosting a black hole. In addition, we also included a new, bright BHXRBS transient candidate MAXI J1820+070 Kawamuro et al. (2018). Finally, our list includes Cygnus X-1. Though a fairly persistent high mass X-ray binary, this source is amongst the best studied of black hole systems, and serves as a useful comparison source in terms of demonstrating the potential of *Gaia* for X-ray binary studies at the bright end. In order to assign *Gaia* counterparts, we used coordinates from the literature which are based upon optical, infrared or radio identifications. We stress that we did not use primary source coordinates listed in the SIMBAD database<sup>1</sup>, as *Gaia* associations therein could be spurious; this point is discussed further below.

In Table 1, we list all BHXRBS together with the associated *Gaia* counterpart. Literature J2000 coordinates are listed together with the *Gaia* coordinates for epoch 2015.5 and the counterpart  $G$  mag, where available. No conversion between epochs is carried out here using the DR2 proper motions discussed later for some sources. We present the *Gaia* coordinates as reported in DR2 for ease of comparison with the archive. Maximal coordinate shifts expected based upon the proper motions are  $< 0.''15$  between the two epochs, small enough not to affect the counterpart identification discussed here.

<sup>1</sup> <http://simbad.u-strasbg.fr/simbad>

## 3 DISTANCE ESTIMATION WITH GAIA

For a measured geometric parallax  $\pi$  in milliarcsec (mas), distance can be obtained by parallax-inversion as  $r_{\text{inv}} = \frac{1}{\pi}$  kpc. However, parallax inversion serves as a poor estimator, especially when uncertainties on  $\pi$  are large ( $\gtrsim 20\%$ ; cf. Bailer-Jones 2015). This becomes increasingly relevant in the faint regime for the quiescent BHXRBS studied here, which tend to have magnitudes of  $G > 19$ .

Treating the problem as a case of Bayesian inference can help improve the estimate, if an appropriate choice of prior distribution of expected distances is possible. Several works have attempted various priors for Galactic stellar populations (e.g., Bailer-Jones 2015, Astraatmadja & Bailer-Jones 2016a, Astraatmadja & Bailer-Jones 2016b, Bailer-Jones et al. 2018). But the validity of stellar-based priors is uncertain when it comes to the BHXRBS population for two reasons. Firstly, there is the possibility of natal kicks scrambling the spatial distribution of XRBs relative to the underlying stellar population. Secondly, the selection of these sources differs markedly from stars: they are discovered in X-rays, which have a longer mean free path through Galactic gas and dust than the optical. Discovered sources are then often subject to targeted multiwavelength follow-up, which is typically much more sensitive than field stars. Intuitively, this would argue for a larger characteristic scale length for BHXRBS.

In the following section, we set up two possible priors for our sample: a simple empirical exponential model based upon the known BHXRBS population, as well as a three-dimensional model of the space distribution of XRBs in X-rays.

The prior  $P(r)$  is combined with a likelihood function  $P(\pi|r)$  to produce the final posterior estimate of  $r$ ,  $P(r|\pi)$ . The likelihood  $P(\pi|r)$  is assumed to be normally distributed with mean of  $\frac{1}{r}$  and standard deviation  $\sigma_\pi$ , as described in the above works. We quote the mode ( $r_{\text{mode}}$ ) of the posterior as the final distance estimate and the uncertainties represent the highest density confidence interval, corresponding to a  $1-\sigma$  interval for a Gaussian distribution. As described in Bailer-Jones et al. (2018), this is equivalent to lowering down a horizontal line from  $r_{\text{mode}}$ , until the enclosed probability encompasses 68.3 % of the full posterior probability density.

We have not considered systematic uncertainties in our calculation, as recommended in Gaia Collaboration et al. (2016b). Systematic astrometric uncertainties have not been accounted for individually in DR2, though generic corrections are included (Gaia Collaboration et al. 2018; Lindegren et al. 2018). Residual zero-point systematics are expected to be only  $\approx 0.03$  mas, much smaller than typical statistical uncertainties in the faint regime (Ibid.). Global systematics may be as large as 0.1 mas, but the Collaboration recommendation at present is to not combine these with statistical uncertainties, but rather keep them in mind for the interpretation (Ibid.).

Table 2 summarises astrometric parameters for the sources where available. Values of parallax and proper motion are stated, together with several DR2 pipeline flags relevant to the fits. These flags are described in the *Gaia* Data Model<sup>2</sup>, and will be discussed in detail in Section 7.

<sup>2</sup> *Gaia* Data model:

[https://gea.esac.esa.int/archive/documentation/GDR2/Gaia\\_archive/chap\\_datamodel](https://gea.esac.esa.int/archive/documentation/GDR2/Gaia_archive/chap_datamodel)

## 4 DISTANCE PRIORS FOR THE BHXR POPULATION

### 4.1 Exponential Model

This prior adopts an exponentially decreasing space density with distance  $r$  from the Sun, as in several recent works focusing on distributions of stars in the Galaxy (Astraatmadja & Bailer-Jones 2016a, Astraatmadja & Bailer-Jones 2016b, Bailer-Jones et al. 2018). However, we determine a scale length more appropriate to dynamically confirmed BHXRBS as follows.

We first collated previous estimates ( $r_{\text{lit}}$ ) of distances to our BHXR sample from the literature. These distances come from a number of published works, and are based upon a variety of techniques (sometimes used in combination) including photometric and spectral classification of the companion star, radio parallax measurements, and constraints based upon line-of-sight reddening columns. These  $r_{\text{lit}}$  estimates, together with uncertainties, are listed in Table 3. We then fitted these distances with the following exponential prior probability model:

$$P(r) = \frac{1}{2L^3} r^2 e^{-r/L} \quad \text{for } r > 0, \quad (1)$$

where  $L$  is the scale length and the shell volume element component scales with  $r^2$ . Such a procedure is very similar to the position-dependent scale length fitting carried out by Bailer-Jones et al. (2018) for the Galactic stellar population, albeit with a small sample size here.

An unbinned maximum likelihood algorithm<sup>3</sup> was used for the fit. In order to determine the uncertainty on  $L$ , randomised ensembles of  $r_{\text{lit}}$  values were generated by resampling from a Normal distribution for each object, with the assumed mean and standard deviation being the published value of  $r_{\text{lit}}$  and its uncertainty, respectively. There are four sources with lower limits on  $r_{\text{lit}}$ , for which we drew random values uniformly between  $r_{\text{lit}}$  and  $r_{\text{lit}} + 5$  kpc, a threshold based upon typical upper limits suggested in the literature. We also tested that our fits are not strongly sensitive to this threshold; e.g. doubling the threshold to  $r_{\text{lit}} + 8$  kpc changes our fitted  $L$  by less than the scatter quoted below.

A total of 10,000 ensembles were randomised, resulting in a measured mean value of the characteristic scale length  $L = 2.17 \pm 0.12$  kpc, with the uncertainty here being the standard deviation amongst the randomised ensembles. The prior model and the distribution of randomised scale lengths are plotted in Fig. 1. The posterior estimator of the distance based upon this prior is denoted hereafter as  $r_{\text{exp}}$ .

### 4.2 Milky Way X-ray source distribution

Our second tested prior is based upon X-ray detections of XRBs in the Milky Way. Our model is taken from Grimm et al. (2002), who present several equations to describe the spatial density of sources separately in the Galactic thin disc, bulge and halo. Their base model, in turn, was built from early works describing the Galactic stellar mass distribution (Bahcall & Soneira 1980; Dehnen & Binney 1998). The space density of BHXRBS is taken to correspond to the Galactic low mass X-ray binary (LMXB) density as described by Eqs. (4), (5) and (6) of Grimm et al. (2002), representing the density in the bulge, disk, and halo of the Milky Way, respectively.

This is a three-dimensional model with a scale length depending upon Galactic coordinates. The full equations and relevant constants for the model can be found in Grimm et al. (2002), and are also reproduced in the Appendix. The number density prior,  $P(r)$ , is composed of the space density multiplied by  $r^2$  representing the volume element, with  $r$  representing distance from the Sun, as for the exponentially-decreasing space density prior. The posterior estimate of the distance is denoted by  $r_{\text{MW}}$ .

In Sections 6 and 7, we will compare and discuss these various distance estimates.

## 5 KINEMATICS

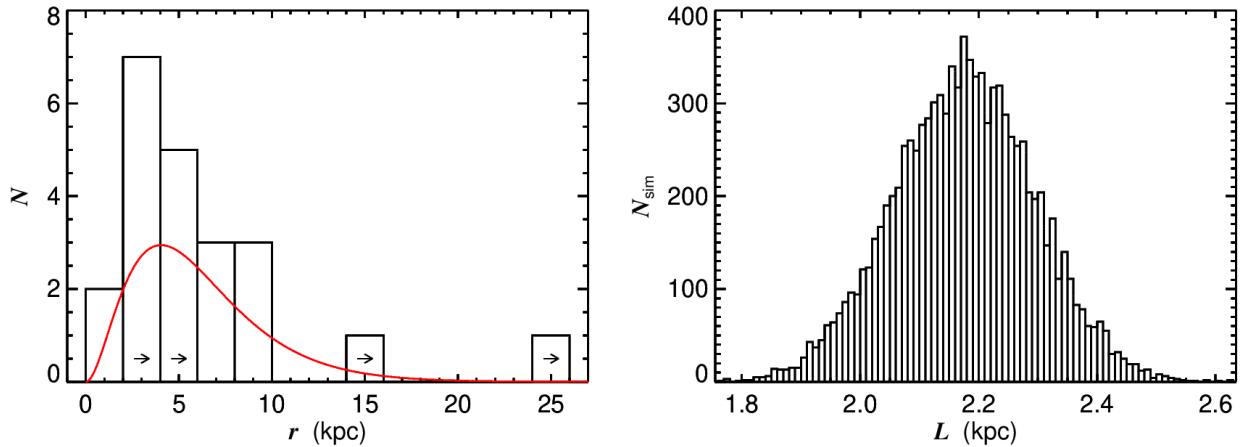
The space velocities of XRBs remain poorly known. If black holes originate as remnants of supernova explosions, they may receive natal kicks that could potentially recoil and scatter them into new trajectories. Symmetric as well as asymmetric natal kicks have been debated for binary systems (Blaauw 1961; Brandt & Podsiadlowski 1995). Studying the distribution of BHXR locations and their kinematics in excess of Galactic rotation can thus shed light on their origins, and there are various works that have attempted this in the past (van Paradijs & White 1995; Jonker & Nelemans 2004; Repetto et al. 2017).

Two black hole binaries for which there is evidence of strong dynamical scattering are XTE J1118+480 and GRO J1655-40 (Mirabel 2017). But such measurements typically require full astrometric and kinematic information: i.e. parallaxes ( $\pi$ ), proper motions ( $\mu_\alpha, \mu_\delta$  in RA, Dec) and radial velocities ( $\gamma$ ), and thus exist for only a handful of sources which have sensitive high angular resolution observations spread over time. The best of these so far come from radio very long baseline interferometric observations.

DR2 now provides measurements of  $(\mu_\alpha \cos\delta, \mu_\delta)$  in addition to  $\pi$ , in the optical. Together with previous measured values or assumed plausible ranges of  $\gamma$ , this allows us to estimate source velocities. We used the formalism of Reid et al. (2009) to transform the observables into heliocentric space velocities, and then remove Solar motion as well as Galactic rotation to determine peculiar velocities  $v_{\text{pec}}$  (i.e. relative to expected motion in the Galaxy). Relevant constants are listed in Table A1. Uncertainties were determined using Monte Carlo resampling of all parameters including the constants.

Whereas literature distance estimates are available for many of our BHXRBS, the DR2 proper motion measurements are novel for most. Furthermore, as we discuss below, typical uncertainties on these proper motions are small. So one would expect to be able to infer interesting dynamical information for most sources. However, this is hampered by the large uncertainties on  $\pi$ . In order to mitigate these uncertainties, for the systems with particularly poor DR2 parallax measurements  $|\frac{\sigma_\pi}{\pi}| > 1$ , we chose to use the literature distance estimates,  $r_{\text{lit}}$ , instead of the *Gaia* distance estimates, when inferring  $v_{\text{pec}}$  since the uncertainties on  $r_{\text{lit}}$  are typically smaller. This affected only 3 sources (XTE J1118+480, 4U 1543-475, Swift J1753.5-0127), and allows us to better focus on the novel kinematic information provided by DR2 in these cases. We verified that using the full DR2 astrometry also resulted in consistent results in these cases.

<sup>3</sup> [https://www.harrisgeospatial.com/docs/ml\\_distfit.html](https://www.harrisgeospatial.com/docs/ml_distfit.html)



**Figure 1.** Fitting of a characteristic scale length  $L$  for the spatial distribution of known BHXRBS.  $r$  represents distance from the Sun. The panel on the left shows the best fitting model to the sample in red, and the panel on the right shows the histogram of  $L$  values resulting from the resampling simulations. See section 4.1 for details.

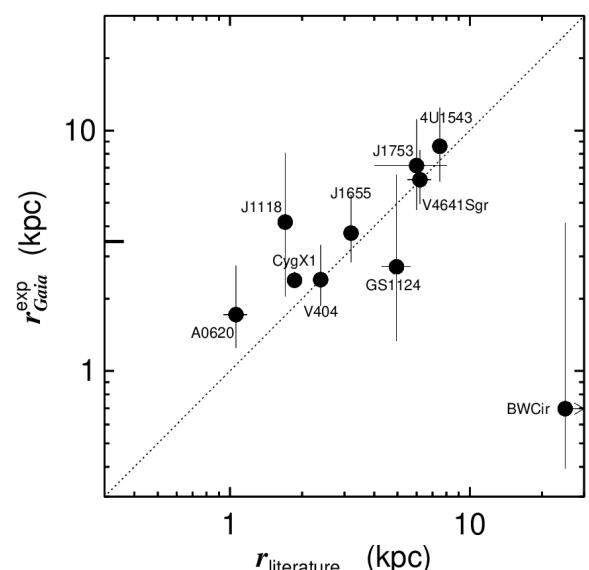
## 6 RESULTS

DR1 reported detections and  $G$  mags of several BHXRBS in our sample, but without parallaxes or proper motion estimates. DR2 is more complete in terms of sky coverage and depth.

We find potential DR2 counterparts within  $2''$  of the known positions of 21 BHXRBS, listed in Table 1. Examining these fields individually, we find that in 18 cases, the identified *Gaia* counterparts can be associated with the BHXRBS. Two sources, GX 339–4 and V404 Cyg, had multiple possible counterparts within  $2''$ , and the final association was confirmed from close examination of the fields. Finding charts for both cases are presented in the Appendix. The offset from the literature position is found to be less than  $1''$  in most cases, with only one secure counterpart (for V4641 Sgr) lying at an offset of just above  $1''$ .

All shortlisted counterparts at larger separations are, in fact, unassociated with the parent BHXRBS. Of these, we note a potential confusion issue for the particular case of GRS 1009–45. The nearest DR2 neighbour is a star with DR2 designation 5414249796406957056 lying about  $1.''2$  to the south-east of the BlackCAT position, with DR2 coordinates of RA =  $10:13:36.40$ , Dec =  $-45:04:32.52$  (epoch 2015.5). The BHXRBS itself is fainter in the optical, and is not detected in DR2. A finding chart is presented in the Appendix. We highlight this source because at the time of writing this manuscript, the SIMBAD database has associated the above (incorrect) neighbouring star with the BHXRBS.<sup>4</sup> We stress that this association is spurious.

Complete DR2 astrometric solutions consisting of position, parallax and proper motion are available for 11 of the 18 associated systems (Table 2). But uncertainties  $\sigma_\pi$  on parallax in most cases are significant, with a median  $\pi/\sigma_\pi = 3.2$  and a fractional uncertainty  $> 1$  for 3 sources (XTE J1118+480, 4U 1543–475 and Swift J1753.5–0127). The reported parallax in one case is negative (Swift J1753.5–0127). Negative parallaxes can arise in cases of high fractional uncertainty, as a result of astrometric fits to noisy data (e.g. Bailer-Jones 2015). The weak parallax constraints are not surprising when considering the faintness of the detected DR2



**Figure 2.** Comparison of *Gaia* DR2 distances (based on the exponential prior,  $r_{\text{exp}}$ ) with literature estimates ( $r_{\text{lit}}$ ) for our BHXRBS targets, where known. The dotted line represents equality between the two. The long dash at  $r_{\text{exp}} \approx 3.5$  kpc denotes the most likely value of  $r_{\text{exp}}$  for MAXI J1820+070, a new XRB that lacks  $r_{\text{lit}}$  estimates.

counterparts. The median mag of our sample is  $G = 19.35$  mag, with a standard deviation of 3.18 mag.

Distance estimates and  $1-\sigma$  confidence intervals based upon the DR2 parallaxes are reported in Table 3. We report  $r_{\text{inv}}$ ,  $r_{\text{exp}}$ , and  $r_{\text{MW}}$ , as described in Section 3. There is no case of strong disagreement between these within the uncertainties, and the Bayesian priors also yield positive distance estimates for the one case with negative  $\pi$ .

In Fig. 2, we compare one of these estimates ( $r_{\text{exp}}$ ) with the literature distances  $r_{\text{lit}}$  compiled in Table 3, for the 10 systems where both are available. The two kinds of estimates agree within uncertainty for 6 objects. There is only modest discrepancy in two cases:

<sup>4</sup> <http://simbad.u-strasbg.fr/simbad/sim-id?Ident=GRS+1009-45>

A 0620–00 and XTE J1118+480, which is not surprising given the high fractional  $\pi$  uncertainties. Two cases disagree more strongly: BW Cir and Cyg X–1. We will delve into individual systems in the Discussion section.

DR2 proper motion measurements for the 11 systems with astrometric solutions are typically much more constrained than their parallaxes, with fractional uncertainties of only  $\approx 0.05$ . The mean magnitude of the proper motions is  $7.3 \pm 0.2 \text{ mas yr}^{-1}$ , most of them being in the direction of decreasing projected RA and also in decreasing Dec. The source with the highest magnitude of proper motion is XTE J1118+480, about 2.5 times the average value above. In Fig. 3, the sources are plotted on a Hammer-Aitoff projection of the sky, along with arrows showing the current direction of their proper motion. In Galactic coordinates, there is no significant average motion in Galactic latitude ( $b$ ) after excluding the outlier XTE J1118+480. There is, however, an overall average motion towards decreasing Galactic longitude  $\mu_l \cos(b) = -3.2 \pm 0.2 \text{ mas yr}^{-1}$ , also discernible in Fig. 3, though we note the small size of this sample. Since these motions are heliocentric, Solar motion along the direction of Galactic rotation, combined with differential disc rotation, should contribute to this trend.

Table 3 also lists the peculiar velocities relative to Galactic rotation. The median value of  $v_{\text{pec}}$  for these 11 systems is  $80.9 \text{ km s}^{-1}$ , but there is also substantial scatter resulting from the large parallax uncertainties in individual cases. Further discussion on the  $v_{\text{pec}}$  values and their implications will follow in the next section.

Distance and peculiar velocity posterior distributions are presented in the Appendix for all sources.

## 7 DISCUSSION

### 7.1 Distance estimates

Table 3 shows agreement between the various DR2 distance estimators within measured uncertainties for all systems. The most likely values of  $r_{\text{exp}}$  are typically straddled by  $r_{\text{inv}}$  and  $r_{\text{MW}}$ . The high fractional uncertainties on the parallaxes, however, imply that  $r_{\text{inv}}$  is unlikely to be a reliable distance estimator for the majority of systems. On the other hand,  $r_{\text{MW}}$  is biased systematically higher than  $r_{\text{exp}}$  in most cases, with a median excess factor of 1.29. The agreement between  $r_{\text{exp}}$  and  $r_{\text{inv}}$  is much closer, with the corresponding factor being 1.05.

The  $r_{\text{MW}}$  overestimate likely arises from the fact that the Milky Way prior is constructed by considering the space distribution of X-ray sources, without accounting for effects such as their expected optical flux detectability and the impact of Galactic dust reddening. Consideration of these issues will require modelling the disc and donor star contributions for each source to the *Gaia* band, and trying to account for the patchy and uncertain nature of Galactic dust; these refinements are beyond the scope of the present work.

Given the above comparisons, we adopt  $r_{\text{exp}}$  as our primary distance estimate for the rest of this work, unless otherwise stated. Such a strategy mimics the adopted choice of an exponential prior in several other works (Astraatmadja & Bailer-Jones 2016b, Bailer-Jones et al. 2018), but now with a custom scale length  $L$  derived for the dynamically confirmed BHXR population.

Regarding comparisons with published distance estimates, the relative uncertainties on the DR2 estimates typically exceed corresponding uncertainties on  $r_{\text{lit}}$ , by a mean factor of at least 3. As

already discussed, the ill constrained nature of the geometric parallax measurements is ascribable to the faint nature of the quiescent BHXR population. Further ongoing astrometric observations should improve these constraints in future data releases.

We find agreement at a 90 % confidence level for 8 of 10 objects between  $r_{\text{lit}}$  and  $r_{\text{exp}}$  (Section 6; Fig. 2). This is not unexpected, since the scale length  $L$  has been derived from the full population of known  $r_{\text{lit}}$  values. But each source only makes a small contribution to the prior derivation in Section 4.1, so the match for individual sources is encouraging. This is a testament to the DR2 astrometric uncertainties being well characterised on the one hand. Conversely, the DR2 estimates provide a geometric test of the validity of the published distances (based upon a variety of photometric, spectroscopy and astrometric techniques). An additional graphical comparison against  $r_{\text{inv}}$  and  $r_{\text{MW}}$  can be found in the Appendix.

For A 0620–00 and XTE J1118+480, there is only modest disagreement at the 1- $\sigma$  level, and full consistency at the 90 % confidence level. The astrometric solution for XTE J1118+480 is particularly poor, with  $\pi/\sigma_\pi < 1$ , so this should be treated with caution. For A 0620–00, the adopted value of  $r_{\text{lit}}$  comes from a detailed dynamical model by Cantrell et al. (2010), which accounts for the disc contribution to the ellipsoidal orbital light curves, yielding a derived distance of  $1.06 \pm 0.12 \text{ kpc}$ . Cantrell et al. show that there is a weak dependence of this distance on the donor spectral type and reddening assumed, with mean values of 0.92–1.26 kpc possible. The upper end of this range is already consistent with the  $r_{\text{exp}}$  estimate. We also note agreement of  $r_{\text{lit}}$  with the  $r_{\text{inv}}$  estimate, so this mild disagreement could simply be reflective of some intrinsic scatter in the spatial distribution beyond that modelled by our prior assumptions.

The two cases of stronger disagreement are Cyg X–1 and BW Cir. Cyg X–1 is one of the few sources with a prior radio parallax distance,  $r_{\text{lit}} = 1.86^{+0.12}_{-0.11} \text{ kpc}$  (Reid et al. 2011). Our DR2 estimate of  $r_{\text{exp}} = 2.38^{+0.20}_{-0.17} \text{ kpc}$  disagrees with this at the 98.8 % level assuming a Normal distribution for the radio parallax uncertainties. For BW Cir, there is no radio parallax, but the best constraints suggest a distance of  $r_{\text{lit}} \gtrsim 25 \text{ kpc}$  (Casares et al. 2004, 2009). For this system, we derive  $r_{\text{exp}} = 0.70^{+3.43}_{-0.30} \text{ kpc}$ . Only 0.02 % of the  $r_{\text{exp}}$  posterior distribution lies beyond 25 kpc (see Appendix), implying disagreement between the two estimates of 99.98 %. Whereas the DR2 mode estimate is relatively close to  $r_{\text{lit}}$  for Cyg X–1 (differing by a factor of only 1.28), the difference in the case of BW Cir is apparently a factor  $> 30$ . The causes of these mismatches probably differ between the two cases, and we will explore these further below.

Fig. 4 shows our BHXR sample superposed on an schematic representation of the Galaxy. DR2  $r_{\text{exp}}$  estimates are plotted where available, together with  $r_{\text{lit}}$  values for other systems. This illustration serves to depict the ‘*Gaia* view’ of BHXRs projected on the plane of the Milky Way.

### 7.2 Astrometric fits

Table 2 lists several flags and parameters relevant for assessing the quality of the DR2 astrometric fits. A description of the flags can be found in the reference cited in footnote 2. All reported fits are based on 9 or more visibility periods (`visibility_periods_used`), i.e. distinct groups of observations separated from other groups by 4 or more days. Small values (e.g. less than 10) could be indicative of additional modelling uncertainties (Ibid.). In only one case (V4641 Sgr) is the astromet-

**Table 1.** *Gaia* DR2 astrometric and photometric data for our sample of 24 Galactic black hole X-ray binaries.

#	Source	Literature			<i>Gaia</i>			Offset (arcsec)	G mag
		(h:m:s) (1)	(d:m:s) (2)	Ref. (3)	(h:m:s) (6)	(d:m:s) (7)	(8)		
1	GRO J0422+32	04:21:42.79	+32:54:27.10	[1]	04:21:42.72	+32:54:26.94	0.86	20.85	
2	A 0620–00	06:22:44.50	-00:20:44.72	[2]	06:22:44.54	-00:20:44.37	0.72	17.52	
3	GRS 1009–45	10:13:36.34	-45:04:31.50		+10:13:36.40	-45:04:32.52	4.22	47.73	
4	XTE J1118+480	11:18:10.79	+48:02:12.42	[3]	11:18:10.77	+48:02:12.21	0.33	19.35	
5	GS 1124–684	11:26:26.65	-68:40:32.83		11:26:26.59	-68:40:32.90	0.35	19.57	
6	Swift J1357.2–0933	13:57:16.82	-09:32:38.55		13:57:16.84	-09:32:38.79	0.34	20.75	
7	BW Cir	13:58:09.70	-64:44:05.80	[4]	13:58:09.71	-64:44:05.29	0.51	20.53	
8	4U 1543–475	15:47:08.32	-47:40:10.80		15:47:08.27	-47:40:10.37	0.70	16.48	
9	XTE J1550–564	15:50:58.70	-56:28:35.20	[5]	15:50:58.65	-56:28:35.31	0.41	20.90	
10	XTE J1650–500	16:50:00.98	-49:57:43.60	[6]	+16:50:00.81	-49:57:43.76	4.67	20.02	
11	GRO J1655–40	16:54:00.14	-39:50:44.90	[7]	16:54:00.14	-39:50:44.88	0.05	16.22	
12	MAXI J1659–152	16:59:01.68	-15:15:28.73	[8]	—	—	—	—	
13	GX 339–4	17:02:49.40	-48:47:23.30	[9]	17:02:49.38	-48:47:23.16	0.23	16.47	
14	H 1705–250	17:08:14.52	-25:05:30.15	[10]	17:08:14.50	-25:05:30.29	0.30	20.83	
15	XTE J1752–223	17:52:15.10	-22:20:32.36	[11]	17:52:15.11	-22:20:31.43	0.96	20.20	
16	Swift J1753–0127	17:53:28.29	-01:27:06.22	[12]	17:53:28.29	-01:27:06.31	0.09	16.70	
17	XTE J1817–330	18:17:43.53	-33:01:07.47	[13]	+48:17:43.49	-33:01:08.80	4.41	49.86	
18	V4641 Sgr	18:19:21.58	-25:24:25.10		18:19:21.63	-25:24:25.84	1.04	13.57	
19	MAXI J1820+070	18:20:21.93	+07:11:07.08	[14]	18:20:21.94	+07:11:07.19	0.19	17.41	
20	XTE J1859+226	18:58:41.58	+22:39:29.40	[15]	—	—	—	—	
21	GRS 1915+105	19:15:11.55	+10:56:44.76	[16]	—	—	—	—	
22	GS 2000+251	20:02:49.48	+25:14:11.36		20:02:49.52	+25:14:10.64	0.92	21.22	
23	V404 Cyg	20:24:03.82	+33:52:01.90	[17]	20:24:03.82	+33:52:01.84	0.06	17.19	
24	Cyg X–1	19:58:21.67	+35:12:05.73	[18]	19:58:21.67	+35:12:05.69	0.05	08.52	

Notes: Cols. (3)–(5): ICRS coordinates (J2000.0) from the literature; Cols. (6)–(7): ICRS coordinates (J2015.5) of the nearest counterparts within 2'' from *Gaia* DR2; (8) Offset in arcsec between the two sets of coordinates. Unassociated neighbours within 2'' have been struck out. References for literature coordinates in Col. (5): Where not stated, the coordinates are taken from Corral-Santana et al. 2016; [1] Shrader et al. (1994); [2] Gallo et al. (2006). [3] Fender et al. (2001); [4] Brocksopp et al. (2001); [5] Corbel et al. (2001); [6] Tomsick et al. (2004); [7] Bailyn et al. (1995); [8] Paragi et al. (2013); [9] Corbel et al. (2000); [10] Yang et al. (2012); [11] Miller-Jones et al. (2011); [12] Fender et al. (2005); [13] Rupen et al. (2006); [14] Kennea (2018); [15] Garnavich et al. (1999); [16] Dhawan et al. (2000); [17] Miller-Jones et al. (2009); [18] Reid et al. (2011).

**Table 2.** *Gaia* DR2 astrometric data for the sample of black hole binaries.

Source	Parallax	Proper Motion			$v_{\text{radial}}$		Fit		
		$\pi$ (mas) (1)	$\mu_\alpha \cos\delta$ (mas $y^{-1}$ ) (2)	$\mu_\delta$ (mas $y^{-1}$ ) (3)	$\gamma$ (km $s^{-1}$ ) (5)	Ref. (6)	n_vis (7)	gof (8)	noise (mas) (9)
A 0620–00	$0.64 \pm 0.16$	$-0.09 \pm 0.25$	$-5.20 \pm 0.30$	$8.5 \pm 1.8$	[1]	10	3.026	0.52 [3.13]	0.30
XTE J1118+480	$0.30 \pm 0.40$	$-17.57 \pm 0.34$	$-6.98 \pm 0.43$	$-15 \pm 10$	[2]	15	4.694	1.05 [3.12]	0.44
GS 1124–684	$0.61 \pm 0.34$	$-2.44 \pm 0.61$	$-0.71 \pm 0.46$	$14.2 \pm 6.3$	[3]	16	0.352	0.0	0.54
BW Cir	$1.83 \pm 0.58$	$-9.38 \pm 2.22$	$-5.70 \pm 2.26$	$103 \pm 4$	[4]	10	1.592	0.0	2.25
4U 1543–475	$0.04 \pm 0.07$	$-7.41 \pm 0.14$	$-5.33 \pm 0.10$	$-87 \pm 3$	[5]	14	0.344	0.0	0.13
GRO J1655–40	$0.27 \pm 0.08$	$-4.20 \pm 0.13$	$-7.44 \pm 0.09$	$-142 \pm 1.5$	[6]	12	4.035	0.20 [2.21]	0.13
Swift J1753–0127	$-0.01 \pm 0.13$	$1.13 \pm 0.16$	$-3.53 \pm 0.15$	$6 \pm 6$	[7]	10	0.570	0.0	0.18
V4641 Sgr	$0.15 \pm 0.04$	$-0.73 \pm 0.07$	$0.42 \pm 0.06$	$107.4 \pm 2.9$	[8]	9	0.675	0.06 [0.56]	0.08
MAXI J1820+070	$0.31 \pm 0.11$	$-3.14 \pm 0.19$	$-5.90 \pm 0.22$	—	—	10	-0.871	0.0	0.25
V404 Cyg	$0.44 \pm 0.10$	$-5.77 \pm 0.17$	$-7.85 \pm 0.17$	$-0.4 \pm 2.2$	[9]	15	9.525	0.50 [7.12]	0.18
Cyg X–1	$0.42 \pm 0.03$	$-3.88 \pm 0.05$	$-6.17 \pm 0.05$	$-5.1 \pm 0.5$	[10]	16	3.626	0.0	0.04

Notes: Col. (6) References for the radial velocity: [1] González Hernández, J. I. & Casares, J. (2010); [2] Mirabel et al. (2001); [3] Wu et al. (2015); [4] Casares et al. (2004); [5] Orosz et al. (1998); [6] Mirabel et al. (2002); [7] Neustroev et al. (2014), though note that the authors stress potential caveats related to systematic uncertainties; [8] Orosz et al. (2001); [9] Casares & Charles (1994); [10] Gies et al. (2008). Col. (7): visibility\_periods\_used: Number of distinct visibility periods used in the DR2 astrometric solution; Col. (8) astrometric\_gof\_a1: Gaussianised  $\chi^2$  goodness-of-fit statistic; Col. (9) astrometric\_excess\_noise: Excess residuals, expressed as an angle, of the data fit to the best astrometric model. Numbers in square brackets denote the corresponding significance of this value (astrometric\_excess\_noise\_sig); Col. (10) astrometric\_sigma5d\_max: The longest principal axis, expressed as an angle, in the 5-dimensional DR2 covariance matrix.

**Table 3.** Distances and space velocity estimates based upon *Gaia* astrometry.

Source	$r_{\text{inv}}$ (kpc)	$r_{\text{exp}}$ (kpc)	$r_{\text{MW}}$ (kpc)	$v_{\text{pec}}$ (km s $^{-1}$ )	$r_{\text{lit}}$ (kpc)	Ref.
(1)	(2)	(3)	(4)	(5)	(6)	(7)
GRO J0422+32	—	—	—	—	$2.49 \pm 0.30$	[1]
A 0620–00	$1.57 \pm 0.40$	$1.72^{+1.02}_{-0.47}$	$1.70^{+0.90}_{-0.44}$	$46.5^{+38.0}_{-11.0}$	$1.06 \pm 0.12$	[2]
GRS 1009–45	—	—	—	—	$3.8 \pm 0.3$	[3]
XTE J1118+480	$3.34 \pm 4.42$	$4.16^{+3.91}_{-2.12}$	$8.22^{+9.26}_{-6.71}$	$166.5^{+47.0}_{-49.0}\dagger$	$1.7 \pm 0.1$	[4]
GS 1124–684	$1.65 \pm 0.93$	$2.71^{+3.82}_{-1.38}$	$4.01^{+3.91}_{-2.27}$	$74.5^{+49.0}_{-27.0}$	$4.95^{+0.69}_{-0.65}$	[5]
Swift J1357.2–0933	—	—	—	—	$> 2.29$	[6]
BW Cir	$0.55 \pm 0.17$	$0.70^{+3.43}_{-0.30}$	$7.01^{+3.43}_{-6.32}$	$114.5^{+62.0}_{-17.0}$	$\gtrsim 25$	[7]
4U 1543–475	$24.72 \pm 41.15$	$8.60^{+3.84}_{-2.46}$	$9.47^{+3.91}_{-2.53}$	$93.9^{+16.2}_{-11.6}\dagger$	$7.50 \pm 0.5$	[8]
XTE J1550–564	—	—	—	—	$4.38^{+0.58}_{-0.41}$	[9]
XTE J1650–500	—	—	—	—	$2.60 \pm 0.70$	[10]
GRO J1655–40	$3.66 \pm 1.01$	$3.74^{+1.65}_{-0.91}$	$7.41^{+1.49}_{-3.33}$	$150.6^{+14.8}_{-12.3}$	$3.2 \pm 0.2$	[11]
MAXI J1659–152	—	—	—	—	$8.6 \pm 3.7$	[12]
GX 339–4	—	—	—	—	$> 5.0$	[13]
H 1705–250	—	—	—	—	$8.6 \pm 2.0$	[14]
XTE J1752–223	—	—	—	—	$5.75 \pm 2.25$	[15]
Swift J1753–0127	—	$7.15^{+3.99}_{-2.48}$	$8.42^{+4.32}_{-2.85}$	$88.5^{+97.0}_{-39.0}\dagger$	$6 \pm 2$	[16]
XTE J1817–330	—	—	—	—	$5.5 \pm 4.5$	[17]
V4641 Sgr	$6.62 \pm 1.81$	$6.24^{+2.04}_{-1.30}$	$8.22^{+1.03}_{-1.18}$	$64.5^{+105.0}_{-23.0}$	$6.2 \pm 0.7$	[18]
MAXI J1820+070	$3.23 \pm 1.14$	$3.46^{+2.18}_{-1.03}$	$3.82^{+2.89}_{-1.23}$	$80.0^{+08.6}_{-23.8}a$	—	
XTE J1859+226	—	—	—	—	$\gtrsim 14$	[19]
GRS 1915+105	—	—	—	$22 \pm 24^b$	$8.6^{+2.0}_{-1.6}$	[20]
GS 2000+251	—	—	—	—	$2.7 \pm 0.7$	[21]
V404 Cyg	$2.28 \pm 0.52$	$2.40^{+0.95}_{-0.53}$	$2.54^{+1.32}_{-0.64}$	$51.5^{+18.0}_{-14.0}$	$2.39 \pm 0.14$	[22]
Cyg X–1	$2.37 \pm 0.18$	$2.38^{+0.20}_{-0.17}$	$2.39^{+0.21}_{-0.18}$	$22.8^{+0.8}_{-0.8}$	$1.86^{+0.12}_{-0.11}$	[23]

Notes: Col. (2) Parallax-inversion distance. (3) Bayesian distance assuming an exponentially-decreasing space density prior; (4) Assuming a Milky Way space density prior. (5) Peculiar velocity relative to Galactic rotation. (6) Previously published distance from literature, and (7) reference for  $r_{\text{lit}}$ .

$\dagger$  Assuming  $r_{\text{lit}}$  in the inference of  $v_{\text{pec}}$ .  $a$ assuming  $\gamma = 0$  km s $^{-1}$ .  $b$ Reid et al. (2014b).

References for  $r_{\text{lit}}$ : [1] Gelino & Harrison (2003); [2] Cantrell et al. (2010); [3] Gelino (2002); [4] Gelino et al. (2006); [5] Wu et al. (2016); [6] Mata Sánchez et al. (2015); [7] Casares et al. (2009); [8] Jonker & Nelemans (2004); [9] Orosz et al. (2011); [10] Homan et al. (2006); [11] Hjellming & Rupen (1995); [12] Kuulkers et al. (2013); [13] Heida et al. (2017); [14] Jonker & Nelemans (2004); [15] Ratti et al. (2012); [16] Cadolle Bel et al. (2007); [17] Sala et al. (2007); [18] MacDonald et al. (2014); [19] Corral-Santana et al. (2011); [20] Reid et al. (2014b); [21] Jonker & Nelemans (2004); [22] Miller-Jones et al. (2009); [23] Reid et al. (2011).

ric solution based upon 9 visibility periods, and the DR2 estimated distance in this case agrees with  $r_{\text{lit}}$ . The total number of individual observations spanning all visibility periods exceeds 100 for all sources. So none of these solutions appears to be strongly biased by sparse or irregular sampling.

The `astrometric_gof_a1` (hereafter, `gof`) values tabulated represent the ‘Gaussianised  $\chi^2$ ’ goodness-of-fit statistic with zero mean and unit standard deviation (Ibid.). While most are below a value of +3 and thus indicate a reasonable astrometric fit solution, there are a few which are notably above this, e.g. V404 Cyg ( $\text{gof} = 9.5$ ) and Cyg X–1 ( $\text{gof} = 3.6$ ). So could we explain away the significant offset between  $r_{\text{exp}}$  and  $r_{\text{lit}}$  for Cyg X–1 based upon a bad astrometric fit? This is not obviously the case, because three other sources with DR2  $r_{\text{exp}}$  values matching literature estimates also have an apparently bad astrometric `gof`: XTE J1118+480, GRO J1655–40 and V404 Cyg. Furthermore, as we will discuss shortly, the peculiar velocity inferred from the astrometric solution for Cyg X–1 matches previous estimates (Section 7.3).

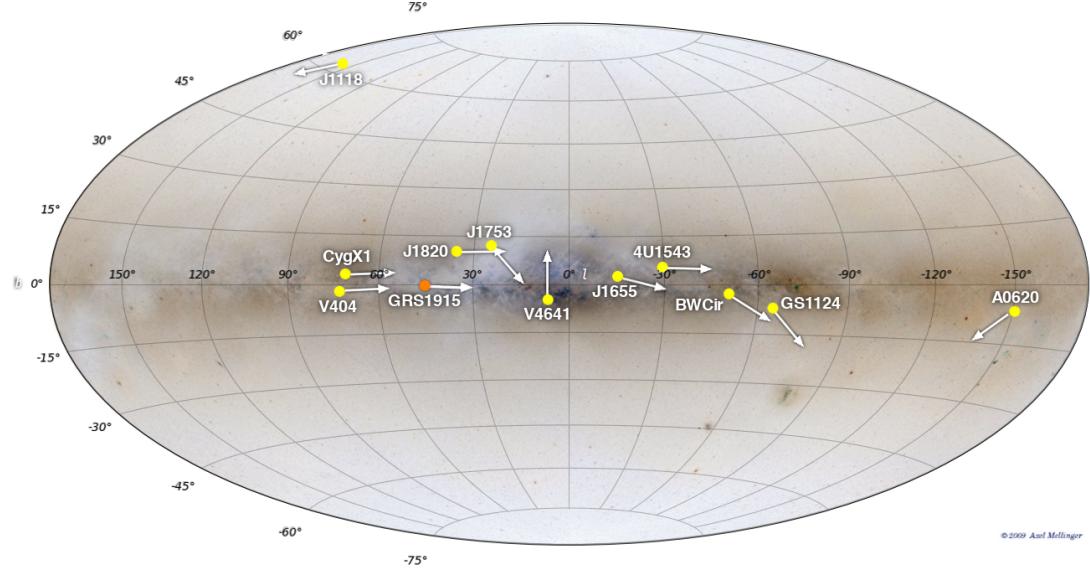
Table 2 does reveal an approximate trend of high `gof` values occurring together with significant `astrometric_excess_noise` (hereafter, `noise`), which is a measure of excess residuals in the best-fitting astrometric solution. Such residuals could result from instrumental effects such

as image centroiding errors and issues with the telescope attitude, or could be indicative of an intrinsic complexity in the nature of the system under study (Ibid.). In particular, such complexity may arise for our sample of BHXBs due to ‘orbital wobble’, when the flux-weighted centroid of emission wobbles at the orbital period of the binary system.

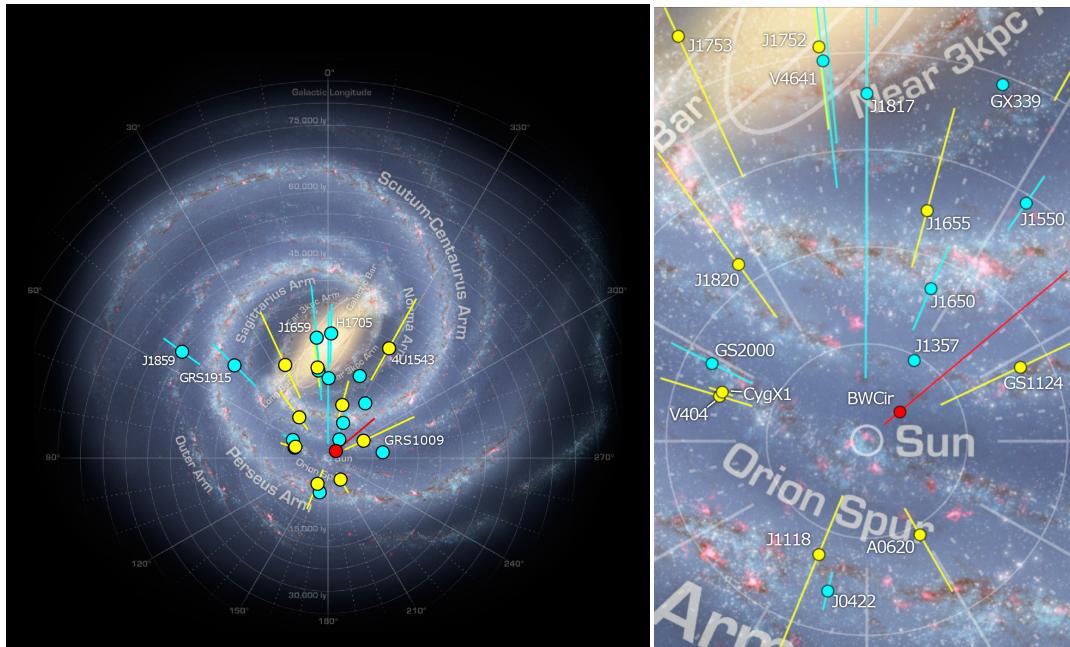
An estimate of the maximal possible astrometric wobble may be obtained by assuming that the quiescent flux of BHXBs is dominated by the companion star. In this case, the angular wobble  $w$  is simply  $w = 2 \frac{a_2}{r}$ , where  $a_2$  is the separation of the companion from the centre of mass and  $r$  is the source distance. With knowledge of the system’s orbital parameters, this can be expressed as

$$w = 2 \frac{GM_1}{K_2^2(1+q)^2 r} \sin^2 i \quad (2)$$

where  $M_1$  is the mass of the compact object,  $K_2$  the velocity amplitude of the companion,  $q$  is the mass ratio  $\frac{M_2}{M_1}$ , and  $i$  the binary inclination. Table 4 lists the known system parameters together with the expected angular wobble in mas. The distances are taken to be the  $r_{\text{lit}}$  values, since the aim here is to check the DR2 solutions. For BW Cir,  $r$  is assumed to be 25 kpc and  $i = 79$  deg, giving a robust upper limit on  $w$ . The lower limit on  $M_{\text{BH}}$  of  $7 M_{\odot}$ ,



**Figure 3.** Schematic illustration of the sky-projected motions of BHXRBS in the Milky Way. The layout is in Galactic Coordinates, using the Hammer-Aitoff projection. Yellow circles represent the locations of BHXRBS for which we have proper motions from *Gaia* DR2, while the orange circle represents radio measurements for GRS 1915+105 (Dhawan et al. 2007, Reid et al. 2014b). The white arrows show the direction of their proper motion. (Galactic sky image: Mellinger 2009).



**Figure 4.** The *Gaia* DR2 map of BHXRBS projected on the plane of the Milky Way. Yellow points represent locations calculated using *Gaia* data and  $r_{\text{exp}}$ . Cyan points represent objects for which DR2 reports no parallaxes, and their distances are taken from the literature. BW Cir is marked in red due to potential caveats on its distance; see Section 7.4 (Milky way image: NASA/JPL-Caltech, ESO, J. Hurt; concentric white circles mark distance increments of 5,000 light-years from the Sun).

however, results in a lower limit on  $w$ . For two systems of those with DR2 astrometric solutions listed in Table 2, Swift J1753.5–0127 and MAXI J1820+070, the orbital solutions are currently not well constrained enough for use here.

$w$  is found to be larger than the statistical uncertainties on the parallax,  $\sigma_\pi$ , for both V404 Cyg and Cyg X-1. This indicates that *Gaia* may have detected the orbital astrometric wobble in these two systems, and could explain the high  $\text{gof}$  statistic for both sources

as well as the significant noise for V404 Cyg. One additional systematic factor for V404 Cyg could be associated with a bright outburst that the source underwent 2015. This resulted in the disc brightening by several mags in the optical during June and December of that year (e.g. Kimura et al. 2016, Muñoz-Darias et al. 2017). This would introduce brief systematic shifts of the optical light centre from the companion star to the disc, and it will be interesting

to search for this effect in the full *Gaia* astrometric measurements once those are released.

Amongst other sources, there could be mild contributions in the case of V4641 Sgr and GRO J1655–40 ( $\frac{w}{\sigma_w} \approx 0.35$  for both). The astrometry of the remaining systems is unlikely to be affected by  $w$ , even assuming reasonable mass ranges for BW Cir.

Cyg X–1, however, shows `noise=0`. This is somewhat surprising given the brightness of this source as compared to V404 Cyg, but it may be that systematic uncertainties act to suppress the noise. A global systematic error of up to 0.1 mas in DR2 (e.g. Luri et al. 2018) would easily bring Cyg X–1 in line with the radio parallax distance. Based upon the Collaboration recommendation, we have not included this systematic uncertainty in the quoted errors (Ibid.). True offsets between the published radio and the optical DR2 measurements may also arise if the radio jet wobble differs from the optical wobble, as a result of the jet emission arising from an extended, beamed and precessing zone from one side of the jet. The companion star in this case is more massive than the compact object, so the optical wobble is expected to be smaller than any radio one-sided jet wobble. Investigation of this issue will require a detailed comparison of the radio data with the full *Gaia* measurements once those are released. Any optical emission from the jet is likely to be much weaker than the companion star in this case (Rahoui et al. 2011) and, in any case, optical jet emission should arise on much more compact scales (e.g. Gandhi et al. 2017).

For the final discrepant source, BW Cir, the `gof` and `noise` parameters are not indicative of a significant issue with the fit. In this case, one further possibility we consider is the impact of correlations between the astrometric fit parameters (e.g. Lindegren et al. 2018, Luri et al. 2018). The DR2 fit solutions include 10 pair-wise correlation coefficients for these; additionally, the DR2 flag `astrometric_sigma5d_max` denotes the longest semi-major axis of the 5-d error ellipse when including correlations – a rough single measure of the effect of the covariances. Amongst our sample of sources with five-parameter fits, BW Cir turns out to have the largest value of this flag, by far (Table 2), indicative of potential additional covariances. Lindegren et al. (2018) suggest several checks for accepting or rejecting the DR2 astrometric solutions, including a check on the `sigma5d` parameter – see their Eq. 11, where they recommend accepting a solution if  $\text{sigma5d\_max} \leq (1.2 \text{ mas}) \times \max[1, 10^{0.2(G-18)}]$ , where  $G$  is the average source mag. BW Cir passes this check.

In order to test the effect of parameter correlations, we used the Markov Chain Monte Carlo code of Bailer-Jones (2017) to infer the distance posterior when accounting for covariances between parameters. The code accounts for covariances between parallax and transverse (proper motion) speed by default. We modified this to additionally include covariances on position, incorporating the full DR2 covariance matrix. The prior assumed was our canonical exponentially decreasing space density function for distance, with  $L = 2.17 \text{ kpc}$ . Priors on other parameters were assumed to be uniform. A set of 200 walkers ran over 1500 iterations with an initial burn-in period of 500, using step sizes of 5 pc in distance, 0.05 mas  $\text{y}^{-1}$  in proper motion (both RA and Dec), and 0.05 mas in position (both RA and Dec). Even after accounting for these correlations, we find  $r_{\text{mode}} = 1.54^{+4.08}_{-1.25} \text{ kpc}$ , with a distance of 25 kpc still strongly disfavoured. This posterior distribution is also presented in the Appendix.

In summary, the astrometric flags do not reveal any obvious spurious solutions at the present time. Cases with apparently bad goodness-of-fit parameters may signal the presence of inherent

complexity such as orbital wobble, and confirmation of this will be possible after the final (full) *Gaia* data release. The most discrepant case, BW Cir, does have any astrometric flag indicative of a poor solution. This source will be discussed in more detail in Section 7.4.

### 7.3 Peculiar Velocities and System Kinetic Energies

DR2 infers  $v_{\text{pec}}$  values significantly above  $100 \text{ km s}^{-1}$  for XTE J1118+480 and GRO J1655–40 (Table 3). High runaway velocities for both were first inferred by Mirabel et al. (2001) and Mirabel et al. (2002) based upon radio and optical (*Hubble Space Telescope*) astrometry, respectively. DR2 agrees with this assessment in both cases, though our DR2 estimate is larger than that of Mirabel et al. (2002) for GRO J1655–40 by about 50% due to apparently higher values of  $\mu$ . In particular, for  $\mu_\delta$ , these authors found a value of  $-4.0 \pm 0.4 \text{ mas y}^{-1}$  whereas DR2 finds  $-7.44 \pm 0.09 \text{ mas y}^{-1}$ . Other published kinematic measurements include V404 Cyg and Cyg X–1. Our DR2 estimate for V404 Cyg agrees well within uncertainties with the measurement of  $v_{\text{pec}} = 39.9 \pm 5.5 \text{ km s}^{-1}$  by Miller-Jones et al. (2009). Similarly, for Cyg X–1, Reid et al. (2011) estimate  $v_{\text{pec}} \approx 21 \text{ km s}^{-1}$ , in agreement with our DR2 estimate.

First proper motions measurements for 7 other sources are now available from DR2 (Table 2). BW Cir is identified as a potential high velocity system, with  $v_{\text{pec}}$  of about  $\gtrsim 100 \text{ km s}^{-1}$ . At the apparently small most likely distance of  $r_{\text{exp}} \lesssim 1 \text{ kpc}$ , the peculiar velocity is dominated by the radial velocity of  $\gamma = 103 \text{ km s}^{-1}$ . Should this distance be an underestimate, the relative contribution of  $\gamma$  may drop, but the measured DR2 proper motion of  $\mu = 11.0 \pm 2.2 \text{ mas y}^{-1}$  will then push up  $v_{\text{pec}}$  (for further analysis on this source, see section 7.4).

DR2 infers modestly high likely  $v_{\text{pec}}$  values of  $\sim 50\text{--}100 \text{ km s}^{-1}$  for all other sources: A 0620–00, V4641 Sgr, MAXI J1820+070, GS 1124–684, 4U 1543–475 and Swift J1753.5–0127. The last two sources have particularly poor parallax constraints ( $| \frac{\pi}{\sigma_\pi} | < 1$ ) and we verified that these sources lie in the class of intermediate to modestly high  $v_{\text{pec}}$  values irrespective of whether we use adopt  $r_{\text{lit}}$  or  $r_{\text{exp}}$  as their distance estimator.

In Fig. 5, we present correlations of  $M_{\text{BH}}$  against  $v_{\text{pec}}$ , and against system kinetic energy,  $K_{\text{pec}} = \frac{1}{2} M_{\text{Total}} \times v_{\text{pec}}^2$  where  $M_{\text{Total}} = M_1 + M_2$ . Systems with a lack of secure black hole masses (BW Cir and MAXI J1820+070) were not included. Furthermore, as pointed out by Mirabel & Rodrigues (2003), the peculiar velocity of Cyg X–1 relative to its parent association of massive stars, Cyg OB3, is likely even smaller than the velocity relative to the Galaxy. They quote a value lower than  $9 \pm 2 \text{ km s}^{-1}$ . Therefore, we treat our value of  $v_{\text{pec}}$  for Cyg X–1 as an upper limit here. The median  $v_{\text{pec}}$  of this sample is  $65 \text{ km s}^{-1}$  and the median  $K_{\text{pec}}$  is  $5 \times 10^{47} \text{ erg}$ . Assuming typical supernova energies of  $E_{\text{SN}} \sim 10^{51} \text{ erg}$  (Kasen & Woosley 2009),  $K_{\text{pec}}$  represents 0.05 % of  $E_{\text{SN}}$ .

The figure reveals potential weak anti-correlations in both cases: Spearman’s rank correlation coefficient  $r_S = -0.53$  and  $-0.43$  for the relations against  $v_{\text{pec}}$  and against  $K_{\text{pec}}$ , respectively. But with insignificant corresponding  $p$ -values of only 0.14 and 0.24, respectively, it is likely that anti-correlation trends as steep as observed arise simply as a result of chance, within this small sample.

One can ask the alternative question of how likely is it that the correlation coefficients  $r_S$  (or equivalently, the slopes) are *negative*

**Table 4.** Estimated magnitude of orbital wobble

Source	$K_2$ km s $^{-1}$	$i$ deg	$M_1$ $M_\odot$	$q$	$r_{\text{literature}}$ kpc	$w$ mas
(1)	(2)	(3)	(4)	(5)	(6)	(7)
A 0620–00	437	51	$6.6 \pm 0.3$	0.074	1.06	0.030
XTE J1118+480	709	74	$7.5 \pm 0.7$	0.024	1.7	0.014
GS 1124–684	407	43	$11.0 \pm 1.75$	0.079	4.95	0.010
BW Cir	279	<79 $^\dagger$	>7.0 $^\dagger$	0.12	$\gtrsim 25^\dagger$	0.005
4U 1543–475	124	21	$9.4 \pm 1.0$	0.28	7.5	0.011
GRO J1655–40	226	69	$6.0 \pm 0.4$	0.42	3.2	0.028
V4641 Sgr	211	72	$6.4 \pm 0.6$	0.66	6.2	0.014
V404 Cyg	208	67	$9.0^{+0.2}_{-0.6}$	0.067	2.39	0.115
Cyg X–1	75	27	$14.8 \pm 1.0$	1.30	1.86	0.098

Orbital parameters in Cols. (2)–(5) are taken from BlackCAT, Casares & Jonker (2014), and papers listed in Tables 1 and 2. Col. (6): Literature distances; see Table 3. Col. (7) Orbital wobble  $w$  computed according to Eq. 2.  $^\dagger$ Value assumed to be the quoted limit for computing  $w$ .

(i.e. consistent with an *anti-correlation*), irrespective of their exact values. In order to answer this, we carry out resampling of the  $v_{\text{pec}}$ ,  $M_{\text{BH}}$  and  $K_{\text{pec}}$  values to generate 10,000 randomised ensembles each comprising 9 objects, by using the known system parameter uncertainties together with the  $v_{\text{pec}}$  posterior distributions for the resampling. For the case of Cyg X–1, our measured  $v_{\text{pec}}$  was treated as an upper limit (as mentioned above) for uniform sampling of smaller randomised values. Each random ensemble was analysed as before, resulting in a distribution of 10,000 corresponding randomised  $r_S$  values for both correlations. The strength of any apparent anti-correlation can then be quantified by checking the fraction of  $r_S$  values that are negative. With this test, we find that the correlation coefficients are negative for 99.9 % and 98.8 % of the simulated trends against  $v_{\text{pec}}$  and against  $K_{\text{pec}}$ , respectively. These percentages are consistent with more statistically significant anti-correlations than implied by the raw sample  $p$ -values in the preceding paragraph. This is because several of the individual source  $v_{\text{pec}}$  uncertainties in the low-to-intermediate  $M_{\text{BH}}$  regime in Fig. 5 are skewed towards high positive  $v_{\text{pec}}$  values which, in turn, would favour any trends being skewed towards negative (anti)correlation coefficients.

We also used these 10,000 randomised ensembles to fit linear regression slopes to the trends together with uncertainties. For the regression trend of  $v_{\text{pec}}$  against  $M_{\text{BH}}$ , we find a slope of  $-10.6^{+4.6}_{-6.2} \text{ km s}^{-1} M_\odot^{-1}$ , whereas for the regression of  $K_{\text{pec}}$  against  $M_{\text{BH}}$ , we find a slope of  $-0.09^{+0.10}_{-0.14} \times 10^{48} \text{ erg M}_\odot^{-1}$ . The quoted values are the modes of the ensemble slope distributions, and the uncertainties refer to the corresponding 68.3 % highest density intervals.

Similar correlations have been investigated before using smaller subsets of objects (Miller-Jones 2014; Mirabel 2017), and have the potential to constrain black hole (BH) formation scenarios. Theoretical models suggest potentially distinct formation channels of direct collapse for high mass BHs, and delayed supernova fallback onto central neutron stars for low mass BHs (e.g., Fryer & Kalogera 2001). In such cases, inverse correlations between the imparted natal kicks and system mass may be expected. Our result of putative anti-correlations is intriguing and is consistent with such formation channels. We also note that our median  $v_{\text{pec}}$  value of  $\approx 80 \text{ km s}^{-1}$  would easily exceed the kick velocities inferred for BH formation. But we caution against drawing substantive conclusions at present. Despite the exquisite DR2 constraints on  $\mu$ , the relatively small sample size and the uncertainties on  $\pi$  and  $M_{\text{BH}}$  still contribute significant scatter. In particular, only the slope for

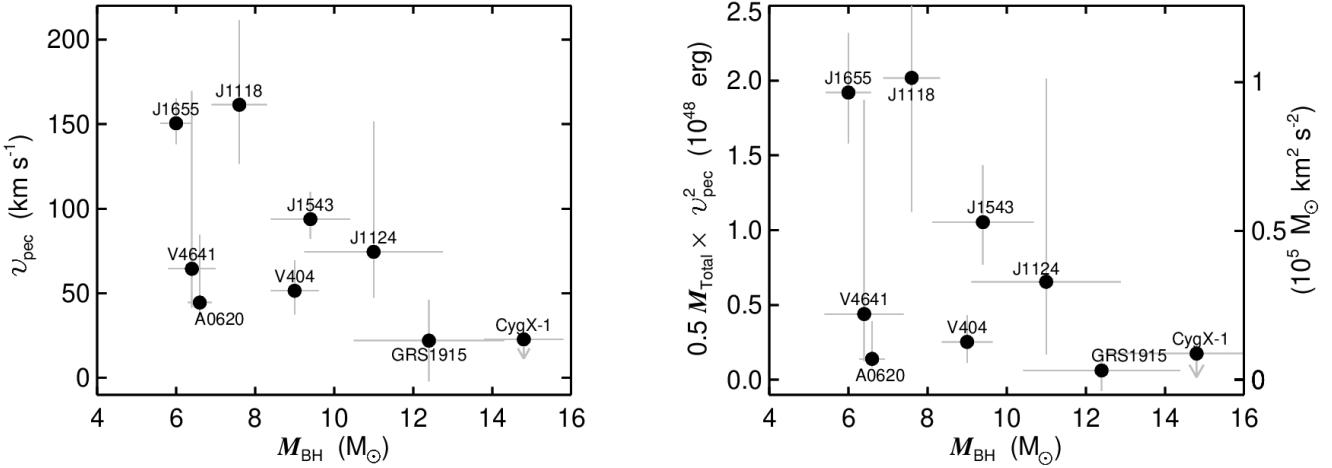
the former correlation ( $v_{\text{pec}}$  against  $M_{\text{BH}}$ ) are significantly negative using the simulations above; any anti-correlation of  $K_{\text{pec}}$  against  $M_{\text{BH}}$  remains insignificant in the present data. This could simply be because the definition of  $K_{\text{pec}}$  conflates uncertainties of system mass together with peculiar velocity. In any case, follow up of these sources to further mitigate  $M_{\text{BH}}$  as well as  $K_{\text{pec}}$  uncertainties, as well as enlarging the base sample, will be needed to confirm or rule out the validity of these trends.

#### 7.4 The curious case of BW Cir

BW Cir, or GS 1354–64, is a dynamically confirmed black hole (Casares et al. 2009). Previous estimates on its distance have placed it at a possible distance of  $\gtrsim 25$  kpc (Casares et al. 2004, 2009), on the other side of the galaxy. There are good reasons for such an inference, as discussed in the above papers, including the donor star classification and expected high luminosity given its relatively long orbital period (2.5 days), as well as the fact that its radial velocity matches with the Galactic differential speed at that distance. However, *Gaia* apparently disagrees with this estimate strongly, placing it at a likely distance of  $r_{\text{exp}} = 0.7^{+3.4}_{-0.3}$  kpc (see discussion in Section 7.1). For this source, the Milky Way prior based estimate  $r_{\text{MW}} = 7.0^{+3.4}_{-6.3}$  kpc. Although apparently less discrepant with respect to  $r_{\text{lit}}$ , note that the lower confidence interval for  $r_{\text{MW}}$  shows a significant tail towards small distances, and at the upper end, values of more than 25 kpc still only have a probability of being favoured at  $\approx 0.5$  % (we refer to the Appendix for a plot of the distance posterior distribution). As discussed in Section 7.2, there is no reason to suspect a spurious solution based upon the astrometric solution flags, so this discrepancy is puzzling.

In order to rule out any systematics related to source identification, we present in Fig. 6 the DR2 counterparts found within a larger search radius around the source, overlaid on a ground-based (VLT/FORS2) optical image of the field. There is little doubt that *Gaia* correctly identifies BW Cir and its nearest neighbours, implying no immediate identification issues.

At its inferred DR2 distance,  $v_{\text{pec}}$  is dominated by the systemic heliocentric radial velocity  $\gamma$ . However, we note that if the source identification were correct but the distance were significantly larger, the DR2 proper motions would start pushing up  $v_{\text{pec}}$  to several hundred  $\text{km s}^{-1}$  or more. This trend is displayed in Fig. 7. While this cannot be ruled out a priori, it adds additional support to a closer distance. We also note that a high proper motion was measured in the Hot Stuff for One Year Catalogue (Altmann et al.



**Figure 5.** Black hole mass ( $M_{\text{BH}}$ ) vs. peculiar velocity (left), and vs. system kinetic energy relative to that expected if the source were undergoing pure Galactic rotation ( $K_{\text{pec}}$ ; right).

2017), although that is now superseded in terms of precision by DR2.

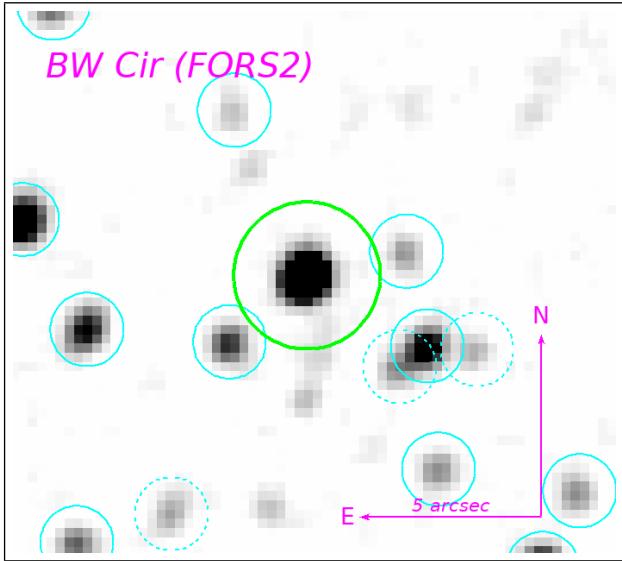
The one additional potential systematic that is difficult to quantify at present is the effect of flux variability. BW Cir underwent an outburst for several months in 2015, during which time it brightened to  $V \approx 17.6$  mag (e.g. Koljonen et al. 2016). Amongst the full DR2 data release, there are a very small number of cases where the pipeline photometric processing fails in the presence of significant intrinsic variability – one prominent case highlighted by the *Gaia* team is that of RR Lyrae, whose  $G$ -mag as well as parallax are reported to be incorrect (Gaia Collaboration et al. 2018). For this source, however, both the  $\text{gof} = 212.5$  and the  $\text{noise} = 3.60$  flags strongly indicate a bad fit, unlike BW Cir. Moreover, the reported  $G$ -mag of BW Cir is not dramatically different from that expected in quiescence (Corral-Santana et al. 2016). So any adverse impact of outburst flux variability on the DR2 astrometric solution remains unclear at this stage.

In summary, the DR2 identification for BW Cir appears reliable, and the parallax and proper motion argue for this source being a nearby system. Systematic uncertainties on the parallaxes and those related to variable source processing would not obviously result in spurious astrometry. A nearby distance would also bring the source quiescent X-ray luminosity  $L_X$  in line with the  $L_X$ – $P_{\text{orb}}$  (orbital period) trend for black hole binaries (cf. Reynolds & Miller 2011, who point out that the source is overluminous by more than an order of magnitude in X-rays as compared to other BHXBs, if at a distance of 25 kpc), and also revives the possibilities of the source being associated with Cen X-2 (cf. Casares et al. 2004, 2009).

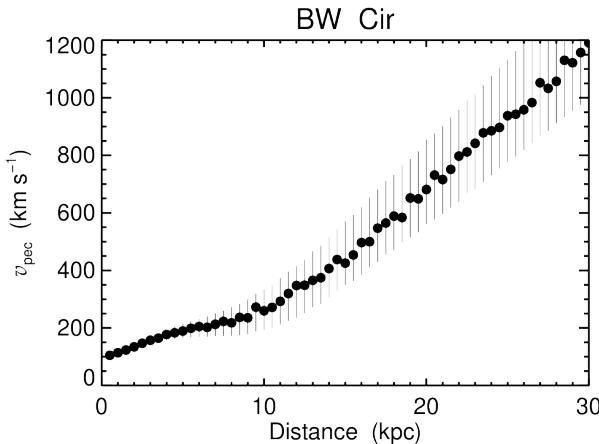
But we caution that further investigation is warranted in order to (1) carry out other complementary checks on the distance – e.g. with future radio facilities – and, (2) to understand implications for the quiescent (and outburst) source nature. In particular, taking the distance at the low end of the distribution has major implications for the amount of light that can be produced by the donor star, and, in turn the radius of the donor star. The donor star’s period is well established (Casares et al. 2004, 2009) and its spectral temperature classification is unlikely to be subject to significant revision, so changing the donor star distance changes its radius. If the donor star is then also still assumed to fill its Roche

lobe, then the mass must be  $\sim 25^3$  times smaller for a distance of 1 kpc than for a distance of 25 kpc in order to maintain the same density, and the same optical flux for the donor star. This would yield a donor mass of  $\sim 5 \times 10^{-5} M_{\odot}$ , unreasonably small. While the donor may potentially be highly bloated, and overluminous for its mass, it is unlikely to be much less than  $\sim 0.1 M_{\odot}$ . For such a star to be filling its Roche lobe, the distance could be changed by only a factor of about 2 from the 25 kpc approximate estimate of Casares et al. (2009). Some additional flexibility can be derived by allowing the fraction of the optical light coming from the donor star to be a bit smaller than the 30–50 % range for different quiescent fluxes assumed by Casares et al. (2009), but given the detectability of ellipsoidal modulations from the source in the presence of other variability, and the detectability of the stellar absorption lines, the veiling of the donor star by accretion light cannot be too extreme. Similarly, the system could potentially be wind-fed, and underfilling its Roche lobe, but it could only underfill the Roche lobe by a small amount while still having a tidally locked donor star that produces ellipsoidal modulations.

The hypothesis of a large distance creates a different problem: a space velocity that is potentially above the escape velocity from the Galaxy despite the source’s location quite close to the Galactic Plane. The posterior distributions do allow distances of order  $\sim 1$ –10 kpc with a reasonable probability and a large, but not excessive, space velocity. It could be that the source just happens to lie at such a distance, in the tail of the likelihood distribution  $P(\pi|r)$ . This would partially reconcile the various distance, stellar classification, and space velocity constraints. But such a solution would remain an unsatisfactory compromise until further data can be obtained, and is speculative at this stage. Further data releases from *Gaia* should help with improving both the statistical uncertainties, as well as understanding the potential systematics discussed above. We also suggest that other high angular resolution measurements (e.g. with future radio facilities, or pushing current sub-mm facilities to deep limits) to corroborate or revise the DR2 parallax and motion should be carried out.



**Figure 6.** FORS2/VLT V-band image of the field of BW Cir. The source counterpart is within the green  $2''$  radius circle. Neighbouring sources are highlighted within cyan apertures of radius  $1''$ . Sources with a reported parallax in DR2 are outlined by the continuous circles; dashed circles denote sources without a reported parallax. The image was taken on 2015 Jun 26 near the beginning of a source outburst.



**Figure 7.** Expected  $v_{\text{pec}}$  for BW Cir using the DR proper motions and known radial velocity, for a range of possible distances over  $0.5$ – $30$  kpc.

## 7.5 MAXI J1820+070

MAXI J1820+070 is a recently-discovered transient detected with MAXI on March 11, 2018 (Kawamuro et al. 2018) following an optical brightening reported by the ASASSN survey (Denisenko 2018, where the designation is ASASSN-18ey). The system peak brightness lay amongst the brightest sources in the extra-Solar X-ray sky, and it was observed by many observatories across the electromagnetic spectrum. The system is believed to host a black hole based upon a variety of spectral and timing characteristics, and displayed an exceptionally bright outburst with plenty of multiwavelength variability in the hard X-ray state (e.g., Baglio et al. 2018; Uttley et al. 2018; Gandhi et al. 2018). Its optical position is reported to be consistent with a star which appeared in *Gaia* DR1

with nominal astrometric uncertainty of a few mas, but its distance had not previously been estimated.

The adopted DR2 distance estimate of  $r_{\text{exp}} = 3.46^{+2.18}_{-1.03}$  kpc corresponds to a height of  $z = 0.61^{+0.38}_{-0.18}$  kpc above the Galactic plane, placing the source within the thin disc population of XRBs, though the posterior probability tail of allowed distances could place the source in the thick disc ( $z > 1$  kpc).

The bolometric (1–100 keV) X-ray flux during this outburst peaked at  $F_{1-100} \approx 1.8 \times 10^{-7}$  erg s $^{-1}$  cm $^{-2}$  in March 2018 (Shidatsu et al. 2018). At the DR2 distance, this yields a luminosity  $L_{1-100} \approx 2.6^{+4.3}_{-1.3} \times 10^{38}$  erg s $^{-1}$ . Assuming a black hole mass  $M_{\text{BH}} = 10 M_{\odot}$  implies a peak hard state luminosity of  $\approx 20^{+34}_{-10}$  % of the Eddington luminosity. Models of hot accretion flows suggest that the hard-to-soft state transition in rising XRB outbursts will occur at Eddington ratios of  $\sim 1$ – $10$  %, though observationally, the range is seen to be broader than this (e.g. Done et al. 2007). During the present outburst, the soft state transition occurred after the source had declined from peak, in July 2018 (Homan et al. 2018). This places MAXI J1820+070 amongst systems capable of maintaining high Eddington ratio hard states. The present uncertainty on the DR2 distance, however, does not allow us to rule out a lower range of Eddington ratios. This source is amongst the brighter of the BHXBs detectable by *Gaia* in quiescence, so future *Gaia* releases are expected to tighten the constraints on distance and improve our understanding of the outburst evolution.

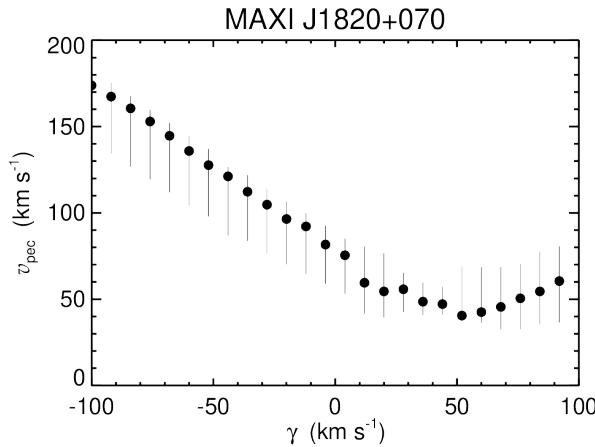
MAXI J1820+070 illustrates the real power of *Gaia*, in that geometric distance estimates are now possible for newly discovered transients (with optical counterparts brighter than  $G \approx 20.5$  or so), giving an incredibly useful first distance estimate in the absence of detailed spectroscopic follow-up.

Finally, we note that its present (projected) velocity is mostly parallel to the Galactic plane (see Fig 3), also suggesting that the source does not venture significantly beyond the disc. DR2 also delivers a  $\approx 30\sigma$  precision on the proper motion of the source. Despite the absence of a published systemic radial velocity ( $\gamma$ ) at present, the proper motion already suggests that the source must have at least a mildly high peculiar velocity at the inferred likely distance. The value of  $v_{\text{pec}} \approx 80$  km s $^{-1}$  for this source (Table 3) is based upon assuming that  $\gamma = 0$  km s $^{-1}$  (cf. other sources in Table 2). Varying  $\gamma$  over the range of  $-100$  to  $+100$  km s $^{-1}$  changes  $v_{\text{pec}}$  between  $\approx 50$ – $140$  km s $^{-1}$ . This trend is shown in Fig. 8.

## 8 SUMMARY

We have investigated *Gaia* DR2 properties for a sample of Galactic X-ray transients that are either dynamically confirmed or likely accreting black hole X-ray binaries. A summary of our findings is as follows:

- (i) DR2 detects the optical counterparts to 18 systems of a sample of 24 BHXBs. The median mag of the detected sample is  $G = 19.35$  mag, i.e. towards the fainter end of the brightness range covered by DR2 (Section 6).
- (ii) Almost all secure DR2 counterparts are within  $1''$  of published source positions. Sources at greater separation are likely to be unassociated neighbours, and we highlight one case of GRS 1009–45 where a widely used public database has made a spurious association.
- (iii) Using best literature distance estimates extracted from a variety of techniques, we derive a characteristic expected exponential scale length of this BHXB population of  $L = 2.17 \pm 0.12$  kpc. This



**Figure 8.** Variation of  $v_{\text{pec}}$  for MAXI J1820+070, for an assumed range of (currently unknown) radial velocities  $\gamma = -100$  to  $+100 \text{ km s}^{-1}$ .

is smaller than the scale length of XRBs in the Galaxy based upon X-ray detections (Section 3), likely a result of the selection effects that go into defining this well-characterised population in the optical. One should keep in mind these systematic selection effects when considering the appropriateness of this scale length for other BHXR studies.

(iv) The DR2 geometric distance estimates are in fair agreement with literature estimates. Of 10 objects with literature estimates, there is agreement at the 90 % confidence level for 8, despite significant parallax measurement uncertainty in most cases (Fig. 2, Section 7.1). This is encouraging, and provides a first geometric test of many literature estimates that are derived from a variety of photometric and spectroscopy techniques.

(v) The two cases of strong disagreement are Cyg X-1 and BW Cir. The former mismatch (Cyg X-1) may be attributable either to intrinsic orbital wobble, or to systematic pipeline measurement uncertainties (Section 7.2). There is no immediate cause for the discrepancy of BW Cir, and we discuss this case at length (Section 7.4) ruling out any obvious source identification issue, pipeline related artefact, and astrometric fit covariances. Other potential issues that we are unable to rule out at present include outburst flux variability, and chance occurrence. Future *Gaia* releases will be crucial for verifying or rejecting the present DR2 solution, and we also highlight the need for complementary multiwavelength follow up in the mean time.

(vi) DR2 presents first proper motions measurements for 7 sources, in addition to 4 sources for which literature estimates exist, and these are used to study the kinematics of this population. Proper motions are typically measured much more precisely than parallaxes (Section 6).

(vii) Combining proper motions with radial velocities, the three-dimensional DR2 peculiar velocities (relative to Galactic rotation) of BHXRbs are estimated, and found to exceed  $v_{\text{pec}} \approx 50 \text{ km s}^{-1}$  for 9 sources (Section 7.3). The median kinetic energy of peculiar motion is  $K_{\text{pec}} \approx 5 \times 10^{47} \text{ erg}$ , or about 0.05 % of typical supernova explosion energies. BW Cir is identified as a potential high velocity system, with  $v_{\text{pec}}$  of about  $\gtrsim 100 \text{ km s}^{-1}$  (Section 7.4).

(viii) Tests for dependence between  $M_{\text{BH}}$  and  $v_{\text{pec}}$  shows a potential anti-correlation, as may be expected in mass-dependent black hole kick formation channels (Section 7.3). However, we do not find a significant dependence of  $K_{\text{pec}}$  against  $M_{\text{BH}}$  in the present

sample, and caution that larger samples will be needed to place the present weak trends on a more robust footing.

Since Galactic X-ray transients typically lie in the faint regime during optical quiescence, the DR2 distance estimates have larger uncertainties as compared to dedicated photometric and spectroscopic studies of individual objects. However, they provide very useful, independent geometric estimates that can be used to validate previous studies, and these estimates will improve with future data releases. But the real advantage of the mission will be in providing immediate distance estimates for new outbursting binaries in years to come, as demonstrated in the case of the recent transient MAXI J1820+070 (Section 7.5).

## 9 ACKNOWLEDGEMENTS

This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. AR acknowledges a Commonwealth Rutherford Fellowship, JP is part supported by funding from a University of Southampton Central VC Scholarship. PG acknowledges support from STFC (ST/R000506/1). This project benefitted from funding by a UGC-UKIERI Phase 3 Thematic Partnership award. We thank J. Casares and P.A. Charles for discussions on the nature of BW Cir, and all coauthors of the BlackCAT catalogue for a useful public resource, in particular J. Corral-Santana. PG also thanks J.C.A. Miller-Jones, R. Plotkin, M. Reynolds, A. Bahramian, J. Steiner, J. Orosz, G. Rate, M. Reid, A.W. Shaw and S. Hodgkin for additional discussions following initial upload of the manuscript to the arXiv server. This research has made use of observatory archival image, including *HST/MAST*, ESO, 2MASS, PanSTARRS1 and NASA/IRSA. The authors acknowledge thorough and useful comments from an anonymous reviewer that helped to improve and broaden the scope of the discussion herein. Finally, we acknowledge prior discussions with Jeffrey McClintock on the nature of Cen X-2 and its potential association with BW Cir.

## REFERENCES

- Allen C. W., 2000, Allen's astrophysical quantities. 4th ed.; New York: AIP Press; Springer, 2000. Edited by Arthur N. Cox. ISBN: 0387987460
- Altmann M., Roeser S., Demleitner M., Bastian U., Schilbach E., 2017, *A&A*, **600**, L4
- Astraatmadja T. L., Bailer-Jones C. A. L., 2016a, *ApJ*, **832**, 137
- Astraatmadja T. L., Bailer-Jones C. A. L., 2016b, *ApJ*, **833**, 119
- Baglio M. C., Russell D. M., Lewis F., 2018, The Astronomer's Telegram, 11418
- Bahcall J. N., Soneira R. M., 1980, *ApJS*, **44**, 73
- Bailer-Jones C. A. L., 2015, *PASP*, **127**, 994
- Bailer-Jones C. A. L., 2017, IA-C8-TN-MPIA-CBJ-081, [https://www.mpia.de/4287840/3D\\_astrometry\\_inference.pdf](https://www.mpia.de/4287840/3D_astrometry_inference.pdf)
- Bailer-Jones C. A. L., Rybizki J., Fouesneau M., Mantelet G., Andrae R., 2018, *AJ*, **156**, 58
- Bailyn C. D., et al., 1995, *Nature*, **374**, 701
- Blaauw A., 1961, Bull. Astron. Inst. Netherlands, **15**, 265
- Brandt N., Podsiadlowski P., 1995, *MNRAS*, **274**, 461
- Brocksopp C., Jonker P. G., Fender R. P., Groot P. J., van der Klis M., Tinney G. S. J., 2001, *MNRAS*, **323**, 517

- Cadolle Bel M., et al., 2007, *ApJ*, **659**, 549  
 Cantrell A. G., et al., 2010, *ApJ*, **710**, 1127  
 Casares J., Charles P. A., 1994, *MNRAS*, **271**, L5  
 Casares J., Jonker P. G., 2014, *Space Sci. Rev.*, **183**, 223  
 Casares J., Zurita C., Shahbaz T., Charles P. A., Fender R. P., 2004, *ApJ*, **613**, L133  
 Casares J., et al., 2009, *ApJS*, **181**, 238  
 Corbel S., Fender R. P., Tzioumis A. K., Nowak M., McIntyre V., Durouchoux P., Sood R., 2000, *A&A*, **359**, 251  
 Corbel S., et al., 2001, *ApJ*, **554**, 43  
 Corral-Santana J. M., Casares J., Shahbaz T., Zurita C., Martínez-Pais I. G., Rodríguez-Gil P., 2011, *MNRAS*, **413**, L15  
 Corral-Santana J. M., Casares J., Muñoz-Darias T., Bauer F. E., Martínez-Pais I. G., Russell D. M., 2016, *A&A*, **587**, A61  
 Dehnen W., Binney J., 1998, *MNRAS*, **294**, 429  
 Denisenko D., 2018, The Astronomer's Telegram, **11400**  
 Dhawan V., Mirabel I. F., Rodríguez L. F., 2000, *ApJ*, **543**, 373  
 Dhawan V., Mirabel I. F., Ribó M., Rodrigues I., 2007, *ApJ*, **668**, 430  
 Done C., Gierliński M., Kubota A., 2007, *A&ARv*, **15**, 1  
 Fender R. P., Hjellming R. M., Tilanus R. P. J., Pooley G. G., Deane J. R., Ogley R. N., Spencer R. E., 2001, *MNRAS*, **322**, L23  
 Fender R., Garrington S., Muxlow T., 2005, The Astronomer's Telegram, **558**  
 Fryer C. L., Kalogera V., 2001, *ApJ*, **554**, 548  
 Gaia Collaboration et al., 2016a, *A&A*, **595**, A1  
 Gaia Collaboration et al., 2016b, *A&A*, **595**, A2  
 Gaia Collaboration et al., 2018, *A&A*, **616**, A1  
 Gallo E., Fender R. P., Miller-Jones J. C. A., Merloni A., Jonker P. G., Heinzel S., Maccarone T. J., van der Klis M., 2006, *MNRAS*, **370**, 1351  
 Gandhi P., et al., 2017, *Nature Astronomy*, **1**, 859  
 Gandhi P., Paice J. A., Littlefair S. P., Dhillon V. S., Chote P., Marsh T. R., 2018, The Astronomer's Telegram, **11437**  
 Garnavich P. M., Stanek K. Z., Berlind P., 1999, IAU Circ., **7276**  
 Gelino D. M., 2002, in American Astronomical Society Meeting Abstracts #200. p. 654  
 Gelino D. M., Harrison T. E., 2003, *ApJ*, **599**, 1254  
 Gelino D. M., Balman Ş., Kızılıoğlu Ü., Yılmaz A., Kalemci E., Tomsick J. A., 2006, *ApJ*, **642**, 438  
 Gies D. R., et al., 2008, The Astrophysical Journal, **678**, 1237  
 González Hernández, J. I. Casares, J., 2010, *A&A*, **516**, A58  
 Grimm H.-J., Gilfanov M., Sunyaev R., 2002, *Astronomy & Astrophysics*, **391**, 923  
 Heida M., Jonker P. G., Torres M. A. P., Chiavassa A., 2017, *ApJ*, **846**, 132  
 Hjellming R. M., Rupen M. P., 1995, *Nature*, **375**, 464  
 Homan J., Wijnands R., Kong A., Miller J. M., Rossi S., Belloni T., Lewin W. H. G., 2006, *MNRAS*, **366**, 235  
 Homan J., et al., 2018, The Astronomer's Telegram, **11820**  
 Jonker P. G., Nelemans G., 2004, *MNRAS*, **354**, 355  
 Kasen D., Woosley S. E., 2009, *ApJ*, **703**, 2205  
 Kawamuro T., et al., 2018, The Astronomer's Telegram, **11399**  
 Kennea J. A., 2018, The Astronomer's Telegram, **11406**  
 Kimura M., et al., 2016, *Nature*, **529**, 54  
 Koljonen K. I. I., Russell D. M., Corral-Santana J. M., Armas Padilla M., Muñoz-Darias T., Lewis F., Coriat M., Bauer F. E., 2016, *MNRAS*, **460**, 942  
 Kuulkers E., et al., 2013, *A&A*, **552**, A32  
 Lindegren L., et al., 2018, *A&A*, **616**, A2  
 Luri X., et al., 2018, *A&A* in press, arXiv:1804.09376  
 MacDonald R. K. D., et al., 2014, *ApJ*, **784**, 2  
 Mata Sánchez D., Muñoz-Darias T., Casares J., Corral-Santana J. M., Shahbaz T., 2015, *MNRAS*, **454**, 2199  
 Mellinger A., 2009, *PASP*, **121**, 1180  
 Miller-Jones J. C. A., 2014, *Publ. Astron. Soc. Australia*, **31**, e016  
 Miller-Jones J. C. A., Jonker P. G., Dhawan V., Brisken W., Rupen M. P., Nelemans G., Gallo E., 2009, *ApJ*, **706**, L230  
 Miller-Jones J. C. A., Jonker P. G., Ratti E. M., Torres M. A. P., Brocksopp C., Yang J., Morrell N. I., 2011, *MNRAS*, **415**, 306  
 Mirabel F., 2017, *New Astron. Rev.*, **78**, 1  
 Mirabel I. F., Rodrigues I., 2003, *Science*, **300**, 1119  
 Mirabel I. F., Dhawan V., Mignani R. P., Rodrigues I., Guglielmetti F., 2001, *Nature*, **413**, 139  
 Mirabel I. F., Mignani R., Rodrigues I., Combi J. A., Rodríguez L. F., Guglielmetti F., 2002, *A&A*, **395**, 595  
 Muñoz-Darias T., et al., 2017, *MNRAS*, **465**, L124  
 Neustroev V. V., Veledina A., Poutanen J., Zharikov S. V., Tsygankov S. S., Sjoberg G., Kajava J. J. E., 2014, *MNRAS*, **445**, 2424  
 Orosz J. A., Jain R. K., Bailyn C. D., McClintock J. E., Remillard R. A., 1998, *ApJ*, **499**, 375  
 Orosz J. A., et al., 2001, *ApJ*, **555**, 489  
 Orosz J. A., Steiner J. F., McClintock J. E., Torres M. A. P., Remillard R. A., Bailyn C. D., Miller J. M., 2011, *ApJ*, **730**, 75  
 Paragi Z., et al., 2013, *MNRAS*, **432**, 1319  
 Rahoui F., Lee J. C., Heinz S., Hines D. C., Pottschmidt K., Wilms J., Grinberg V., 2011, *ApJ*, **736**, 63  
 Ratti E. M., et al., 2012, *MNRAS*, **423**, 2656  
 Reid M. J., et al., 2009, *ApJ*, **700**, 137  
 Reid M. J., McClintock J. E., Narayan R., Gou L., Remillard R. A., Orosz J. A., 2011, *ApJ*, **742**, 83  
 Reid M. J., et al., 2014a, *ApJ*, **783**, 130  
 Reid M. J., McClintock J. E., Steiner J. F., Steeghs D., Remillard R. A., Dhawan V., Narayan R., 2014b, *ApJ*, **796**, 2  
 Repetto S., Igoshev A. P., Nelemans G., 2017, *MNRAS*, **467**, 298  
 Reynolds M. T., Miller J. M., 2011, *ApJ*, **734**, L17  
 Rupen M. P., Dhawan V., Mioduszewski A. J., 2006, The Astronomer's Telegram, **721**  
 Sala G., Greiner J., Ajello M., Bottacini E., Haberl F., 2007, *A&A*, **473**, 561  
 Shidatsu M., et al., 2018, *ApJ* in press, arXiv:1810.11222,  
 Shrader C. R., Wagner R. M., Hjellming R. M., Han X. H., Starrfield S. G., 1994, *ApJ*, **434**, 698  
 Tetarenko B. E., Sivakoff G. R., Heinke C. O., Gladstone J. C., 2016, *ApJS*, **222**, 15  
 Tomsick J. A., Kalemci E., Kaaret P., 2004, *ApJ*, **601**, 439  
 Uttley P., et al., 2018, The Astronomer's Telegram, **11423**  
 Wu J., et al., 2015, *ApJ*, **806**, 92  
 Wu J., Orosz J. A., McClintock J. E., Hasan I., Bailyn C. D., Gou L., Chen Z., 2016, *ApJ*, **825**, 46  
 Yang Y. J., Kong A. K. H., Russell D. M., Lewis F., Wijnands R., 2012, *MNRAS*, **427**, 2876  
 van Paradijs J., White N., 1995, *ApJ*, **447**, L33

## A APPENDIX

### A1 Finding charts for sources with multiple counterparts

Finding charts of the fields of sources with multiple potential counterparts are presented in Figs. A1 and A2. The correct counterpart is identified in the large central green circle. In the latter figure, the correct counterpart to GRS 1009–45 is not detected in DR2.

### A2 Milky Way Prior

Eqs. (4), (5) and (6) from Grimm et al. (2002) are reproduced here for clarity, with differences explained below.

$$\rho_b = \rho_b^0 \cdot \left( \frac{\sqrt{(r_{\text{Gal}}^{\parallel})^2 + \frac{z^2}{q^2}}}{r_0} \right)^{-\gamma} \cdot \exp \left( -\frac{(r_{\text{Gal}}^{\parallel})^2 + \frac{z^2}{q^2}}{r_t^2} \right) \quad (\text{A1})$$

$$\rho_d = \rho_d^0 \cdot \exp \left( -\frac{r_m}{r_d} - \frac{r_{\text{Gal}}^{\parallel}}{r_d} - \frac{|z|}{r_z} \right) \quad (\text{A2})$$

$$\rho_h = \rho_h^0 \cdot \frac{\exp(-b_h \cdot (\frac{r_{\text{Gal}}}{R_e})^{1/4})}{(\frac{r_{\text{Gal}}}{R_e})^{7/8}}, \quad (\text{A3})$$

where  $\rho_b$ ,  $\rho_d$  and  $\rho_h$  represent densities in the bulge, disc, and halo, respectively.  $r_{\text{Gal}}^{\parallel}$  is the distance in the plane from the Galactic centre.  $z$  is the height above the plane (cylindrical coordinates), and  $r_{\text{Gal}}$  is the direct distance from the Galactic centre (spherical coordinates). The constants are listed in Table A1. The normalisations are chosen so that the density of the disc at the position of the Sun satisfies  $\rho_h(r_{\text{Gal}}^{\parallel} = R_{\odot}, z = z_{\odot}) = \frac{1}{500} \cdot \rho_d(r_{\text{Gal}}^{\parallel} = R_{\odot}, z = z_{\odot})$ , and the total mass of the bulge is  $1.3 \times 10^{10} M_{\odot}$ . The integrated disc:bulge:halo ratios are then scaled to be 2.0:1.0:0.8. These reproduce the assumptions adopted in Grimm et al. (2002). All distances are in kpc.

There is, however, one noteworthy difference with respect to Grimm et al. (2002). The first exponential term in Eq. A2 above represents the contribution of the interstellar medium to the disc. While Grimm et al. (2002) list this as decreasing with  $r_{\text{Gal}}^{\parallel}$ , this does not match the final published version of the model in Dehnen & Binney (1998), and appears to be a typographical error. We use the Dehnen & Binney (1998) prescription with the constant  $r_m$  instead.

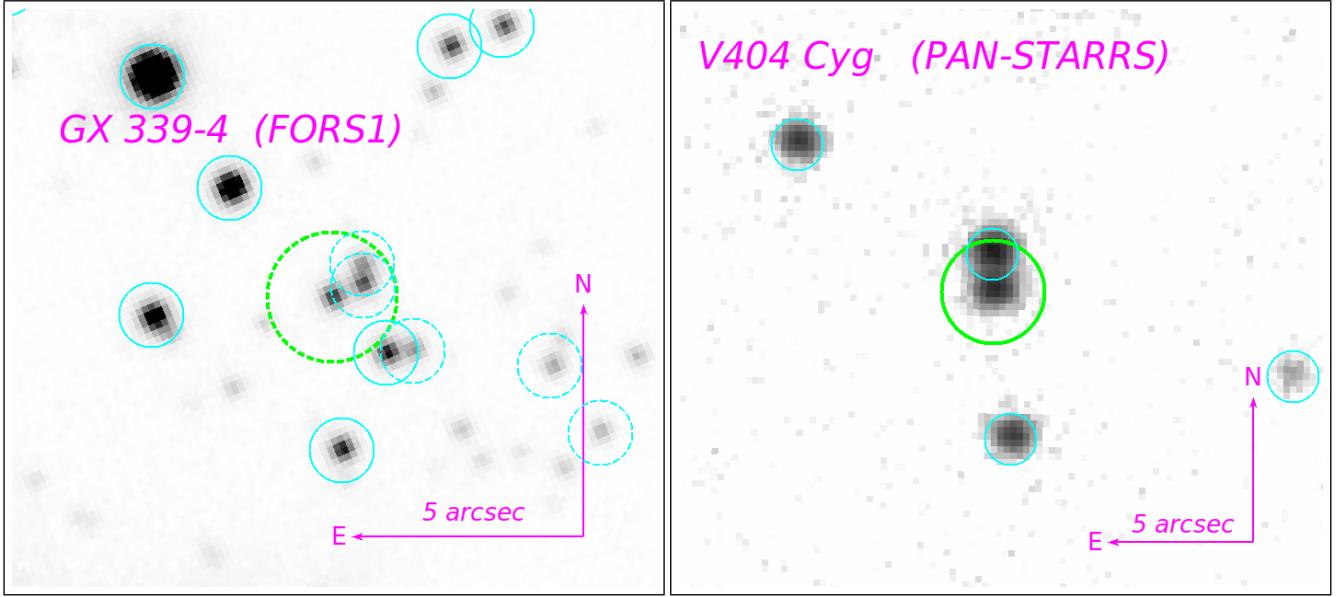
The above model is three-dimensional. The prior of relevance for this work is the one-dimensional  $P(r, l, b)$ , with  $r$  representing distance from the Sun, and is determined by multiplying the densities by the corresponding volume element and subsequently transforming from Galacto-centric to Solar-centric coordinates. Mass density is taken to follow space density. Fig A3 shows a slice through the space density model and examples of the position  $(l, b)$  dependent prior along several representative lines-of-sight.

### A3 Distance estimate comparisons

Fig. A4 displays a comparison between the four distance estimates discussed in the main body of text:  $r_{\text{inv}}$ ,  $r_{\text{exp}}$  and  $r_{\text{MW}}$  are the three DR2 based estimates, and these are compared to the literature estimates  $r_{\text{lit}}$ .

### A4 Distance and peculiar velocity posterior distributions

For each source from Table 2, the figures A5–A15 show the  $r_{\text{exp}}$  distance posterior and the  $v_{\text{pec}}$  posterior, unless otherwise stated.

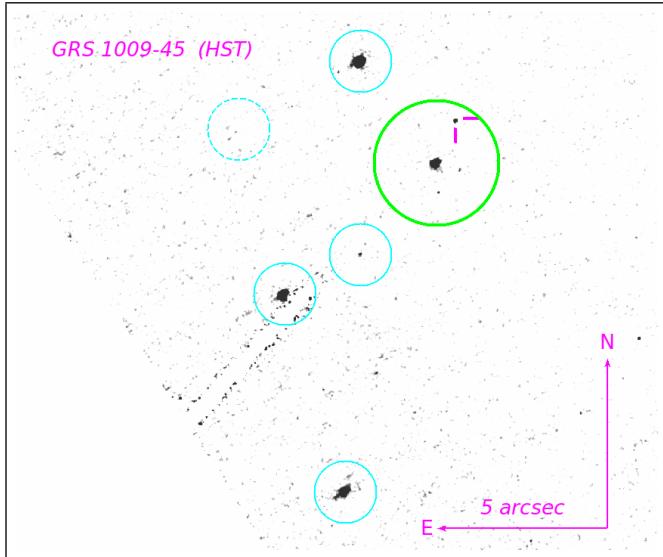


**Figure A1.** Finding charts for the fields of GX 339–4 (*Left*) and V404 Cyg (*Right*), showing the associated DR2 counterpart at the centre of the thick continuous green circle (2'' radius), and neighbouring sources catalogued in DR2 outlined by the smaller cyan circles. The GX 339–4 image is an archival dataset from VLT/FORS1 in the  $r$ (Gunn) filter taken on 2000 June 04, while the V404 Cyg image is from PAN-STARRS in the  $g$  filter taken on 2012 February 09. Sources with a reported parallax in DR2 are outlined by the continuous circles; dotted circles denote sources without a reported parallax.

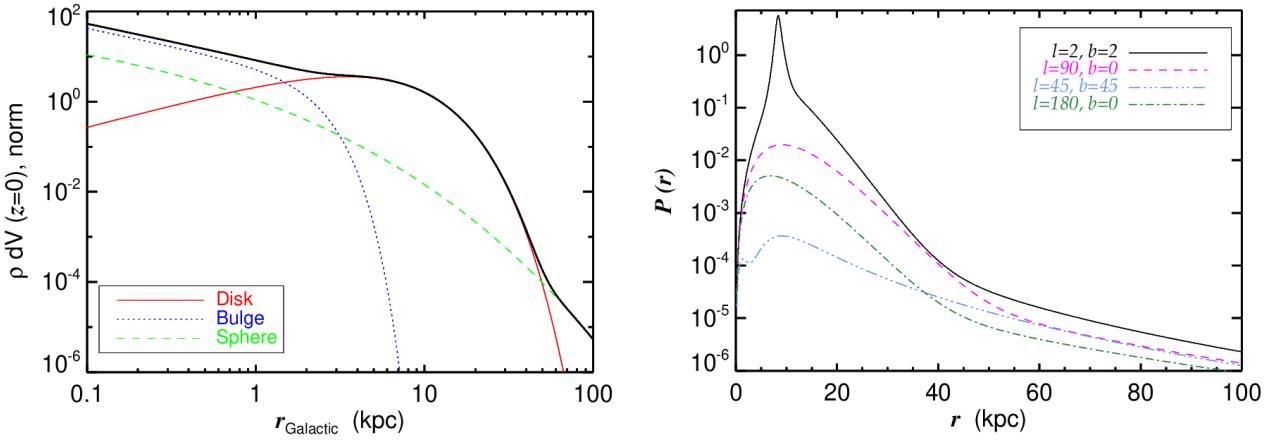
**Table A1.** Table of Galactic and Solar constants

Constant		Value	Units
$R_\odot$	Solar distance from Galactic centre	$8.34 \pm 0.16$	kpc
$z_\odot$	Solar height above Galactic plane	0.024	kpc
$(U_\odot, V_\odot, W_\odot)$	Solar motion	$(10.7 \pm 1.8, 15.6 \pm 6.8, 8.9 \pm 0.9)$	$\text{km s}^{-1}$
$\Theta_\odot$	Galactic rotation speed at $R_\odot$	$240 \pm 8$	$\text{km s}^{-1}$
$q$	oblateness of bulge	0.6	
$\gamma$	–	1.8	
$R_e$	scale length of halo	2.8	kpc
$b_h$	–	7.669	
$r_0$	scale length of bulge	1	kpc
$r_t$	truncation radius of bulge	1.9	kpc
$r_d$	scale length of disc	3.5	kpc
$r_z$	vertical scale of disc	0.41	kpc
$r_m$	inner disc cut-off	6.5	kpc
$\rho_d^0$	disc normalisation factor	2.79	$M_\odot \text{ pc}^{-3}$
$\rho_b^0$	bulge normalisation factor	1.19	$M_\odot \text{ pc}^{-3}$
$\rho_h^0$	halo normalisation factor	22.38	$M_\odot \text{ pc}^{-3}$
$M_b$	Integrated bulge mass	1.3	$10^{10} M_\odot$
	Density ratio Disc:Bulge:Hal	2:1:0.8	

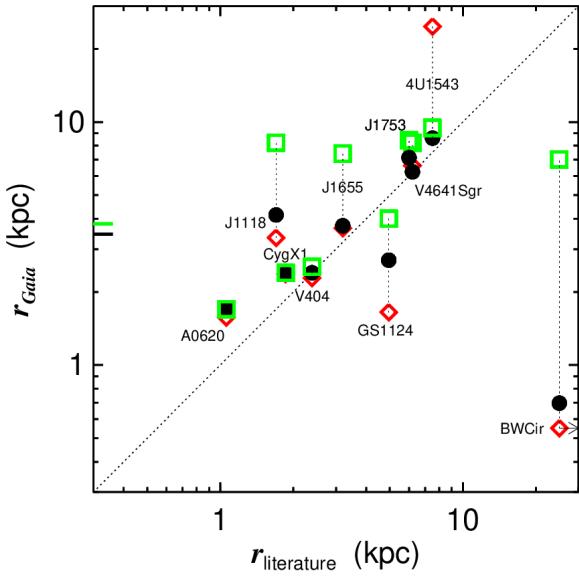
Solar constants taken from [Reid et al. \(2014a\)](#), [Allen \(2000\)](#), and the Milky Way structural constants from [Grimm et al. \(2002\)](#).



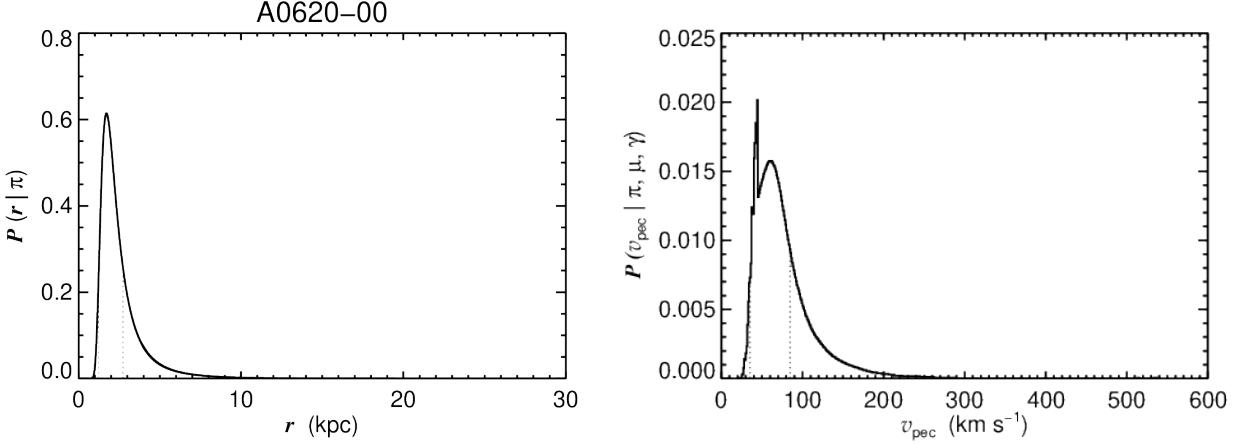
**Figure A2.** Finding chart for GRS 1009–45 from the *Hubble Space Telescope*, showing the correct counterpart in the magenta cross-hair and the nearby confused counterpart in green. Other neighbouring sources from DR2 are highlighted in cyan. Sources with a reported parallax in DR2 are outlined by the continuous circles; dotted circles denote source without a reported parallax.



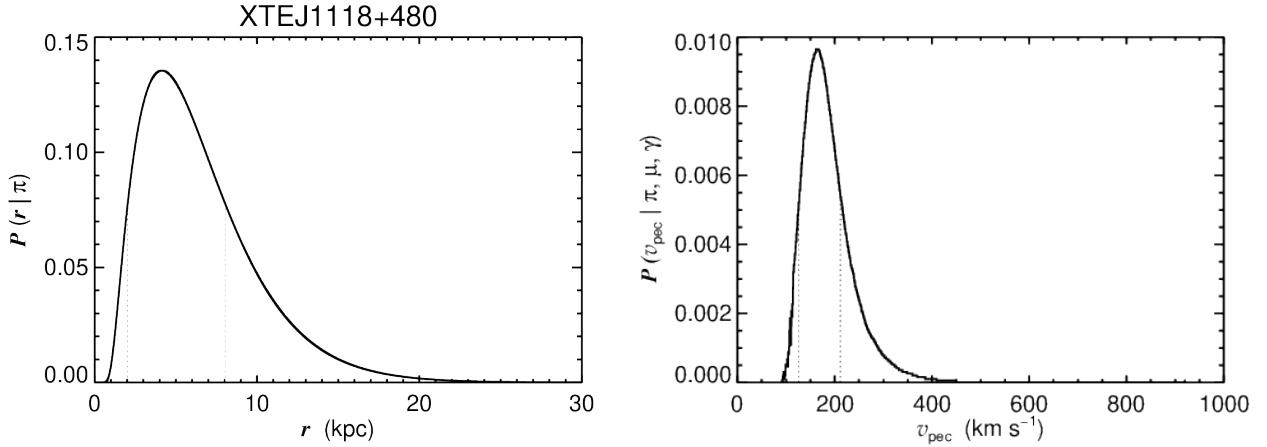
**Figure A3.** (Left:) Galactic model showing the disc, bulge and spheroid (halo) decomposition (space density multiplied by the corresponding volume element), according to Grimm et al. (2002). The model is 3-dimensional, but only a slice at  $z=0$  is shown for clarity. The x-axis represents distance in the plane from the Galactic centre. (Right:) Unnormalised priors of distance  $r$  from Sun resulting from the Galactic model, along several representative lines-of-sight denoted by their Galactic longitude ( $l$ ) and latitude ( $b$ ) expressed in degrees.



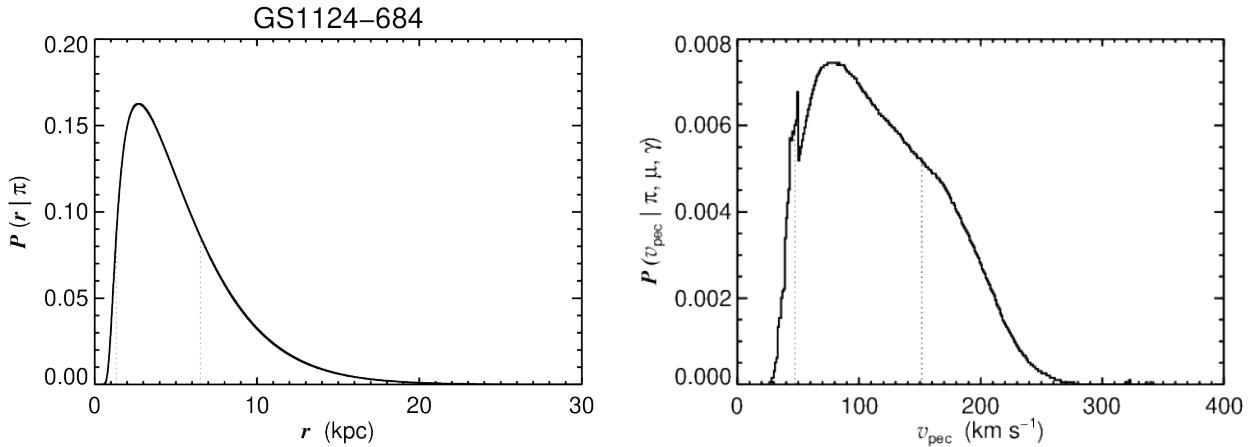
**Figure A4.** Comparison of *Gaia* DR2 distances with previous estimates for our black hole binary targets, where known. The dotted line represents equality between the two axes. The symbols represent:  $r_{\text{exp}}$  as black filled circles (as in Fig. 2),  $r_{\text{inv}}$  as empty red diamonds, and  $r_{\text{MW}}$  as empty green squares. These three estimates are connected with a dotted line for each source. The long dashes on the left-hand vertical axis denote the most likely value of  $r_{\text{exp}}$  (black) and  $r_{\text{MW}}$  (green) for MAXI J1820+070, a new XRB that lacks  $r_{\text{lit}}$  estimates.



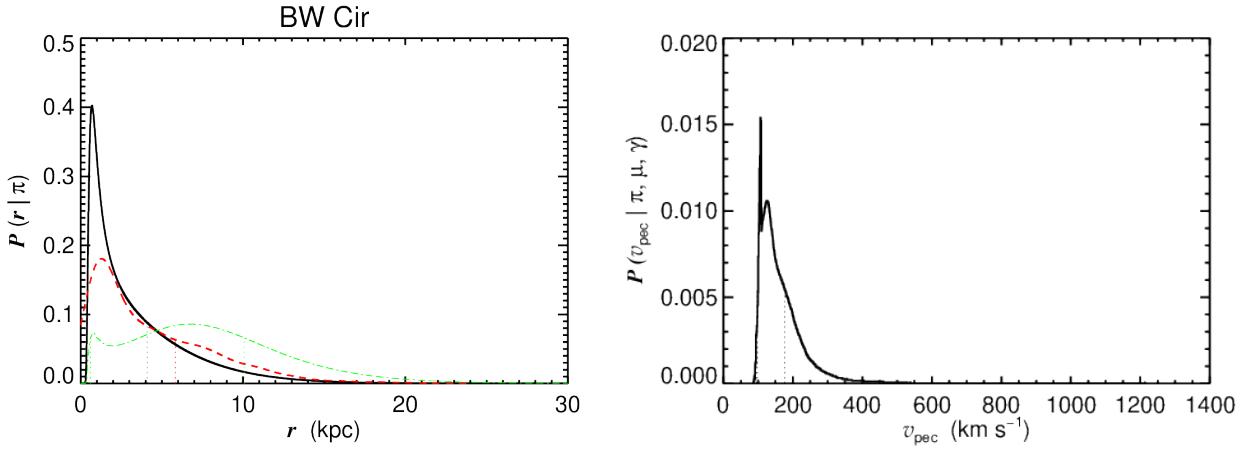
**Figure A5.** A 0620–00. (Left) Distance  $r_{\text{exp}}$  and (Right) peculiar velocity  $v_{\text{pec}}$  posterior distributions. The dotted vertical lines denote the quoted confidence intervals.



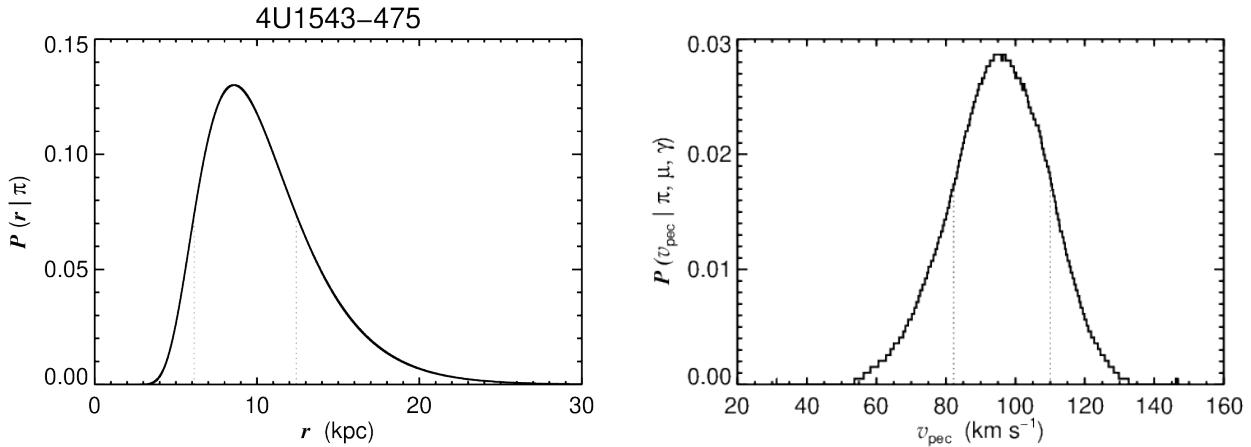
**Figure A6.** XTE J1118+480. (Left) Distance  $r_{\text{exp}}$  and (Right) peculiar velocity  $v_{\text{pec}}$  posterior distributions. The dotted vertical lines denote the quoted confidence intervals. In this case, the fractional DR2 parallax error is particularly poor,  $\frac{\sigma_\pi}{\pi} > 1$ , so  $r_{\text{lit}}$  is combined with the DR2 proper motions and published  $\gamma$  velocity to determine  $v_{\text{pec}}$ .



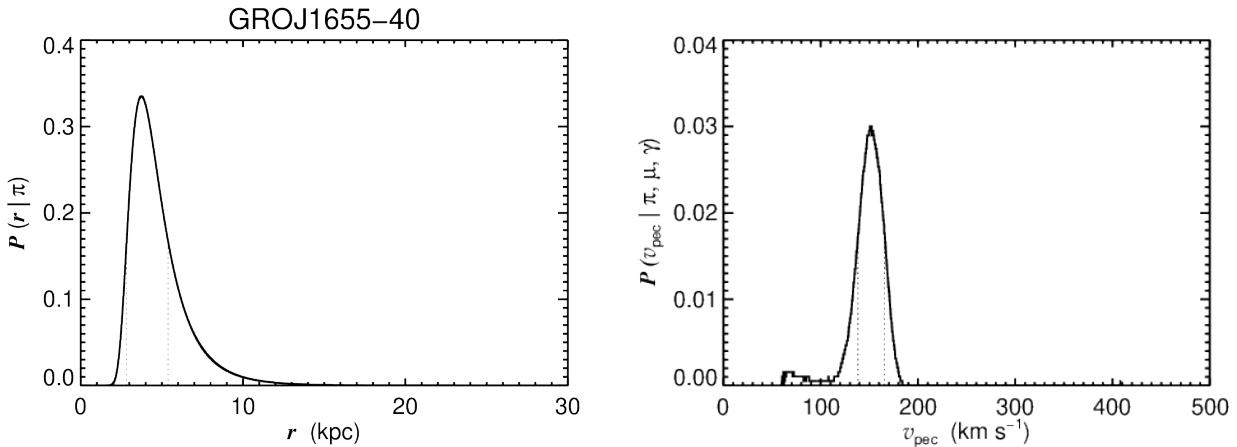
**Figure A7.** GS 1124–684. (Left) Distance  $r_{\text{exp}}$  and (Right) peculiar velocity  $v_{\text{pec}}$  posterior distributions. The dotted vertical lines denote the quoted confidence intervals.



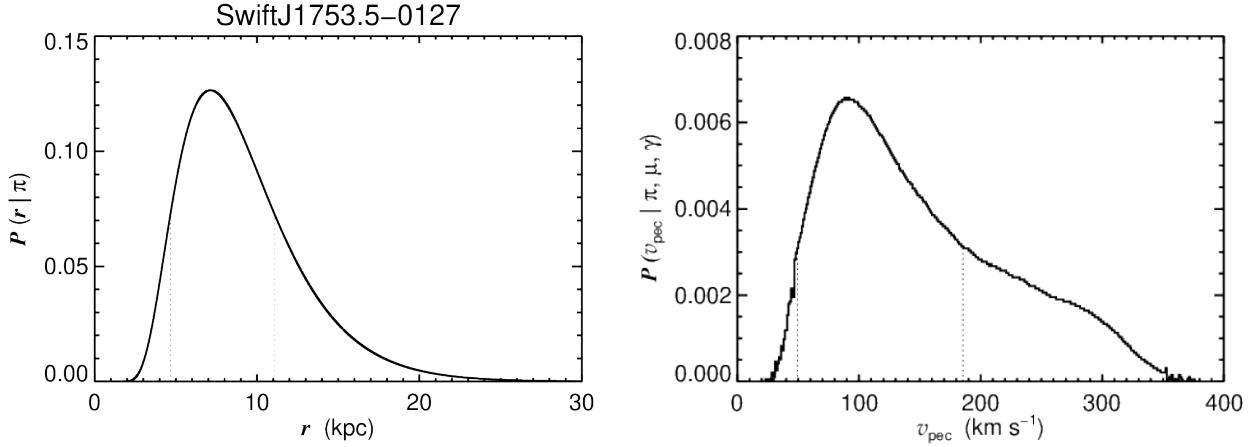
**Figure A8.** BW Cir. (*Left*) Distance  $r_{\text{exp}}$  and (*Right*) peculiar velocity  $v_{\text{pec}}$  posterior distributions. The dotted vertical lines denote the quoted confidence intervals. For this source, the distance plot additionally shows the effect on the posterior when including high potential covariances between the DR2 astrometric fit parameters for this source (red dashed curve), as well as the Milky Way posterior (green dot-dashed curve). Both demonstrate the unlikelihood of the source lying at distances of  $\gtrsim 25$  kpc.



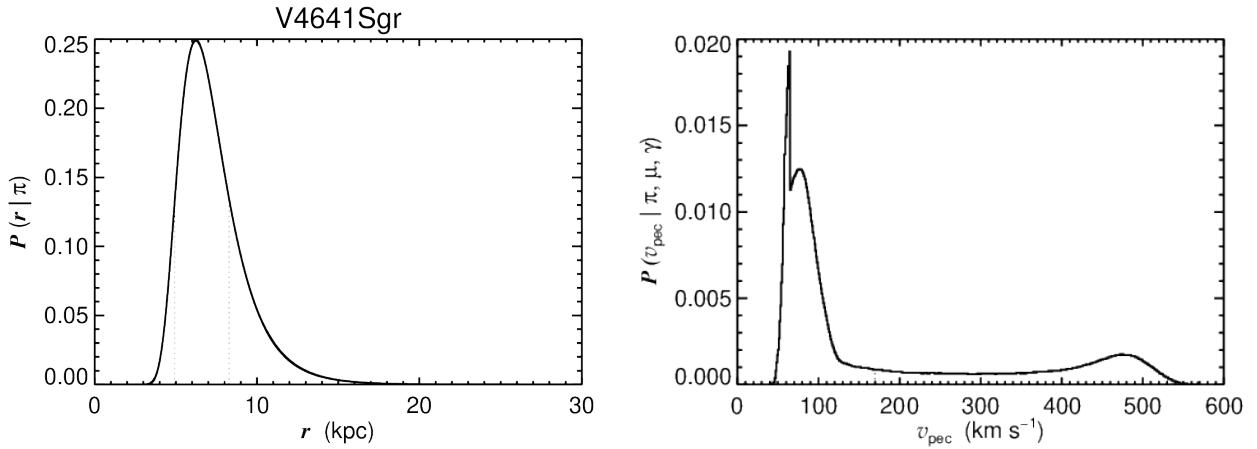
**Figure A9.** 4U 1543–475. (*Left*) Distance  $r_{\text{exp}}$  and (*Right*) peculiar velocity  $v_{\text{pec}}$  posterior distributions. The dotted vertical lines denote the quoted confidence intervals. In this case, the fractional DR2 parallax error is particularly poor,  $\frac{\sigma_\pi}{\pi} > 1$ , so  $r_{\text{lit}}$  is combined with the DR2 proper motions and published  $\gamma$  velocity to determine  $v_{\text{pec}}$ .



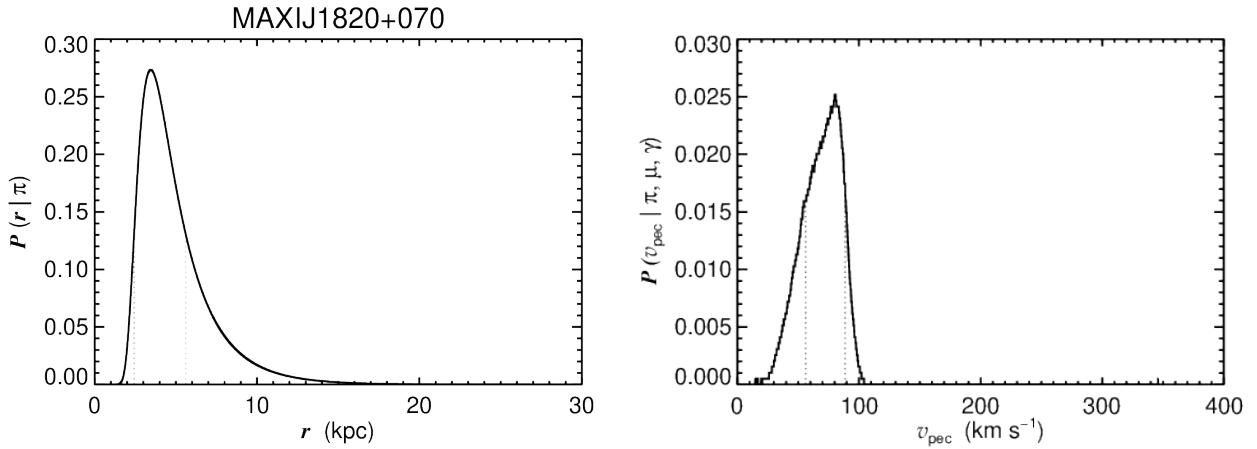
**Figure A10.** GROJ1655-40. (*Left*) Distance  $r_{\text{exp}}$  and (*Right*) peculiar velocity  $v_{\text{pec}}$  posterior distributions. The dotted vertical lines denote the quoted confidence intervals.



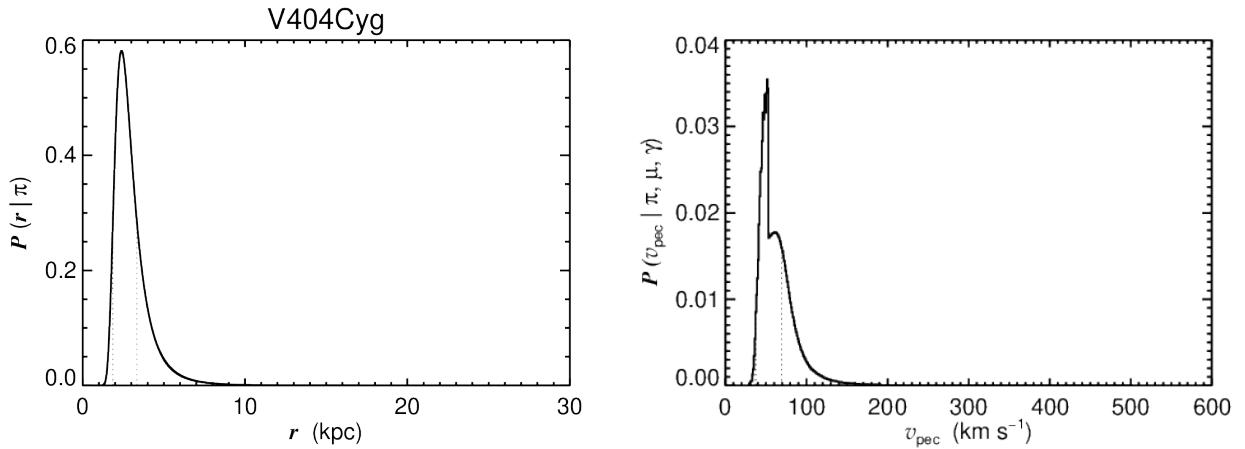
**Figure A11.** Swift J1753.5–0127. (*Left*) Distance  $r_{\text{exp}}$  and (*Right*) peculiar velocity  $v_{\text{pec}}$  posterior distributions. In this case, the fractional DR2 parallax error is particularly poor,  $|\frac{\sigma_\pi}{\pi}| > 1$ , so  $r_{\text{lit}}$  is combined with the DR2 proper motions and published  $\gamma$  velocity to determine  $v_{\text{pec}}$ .



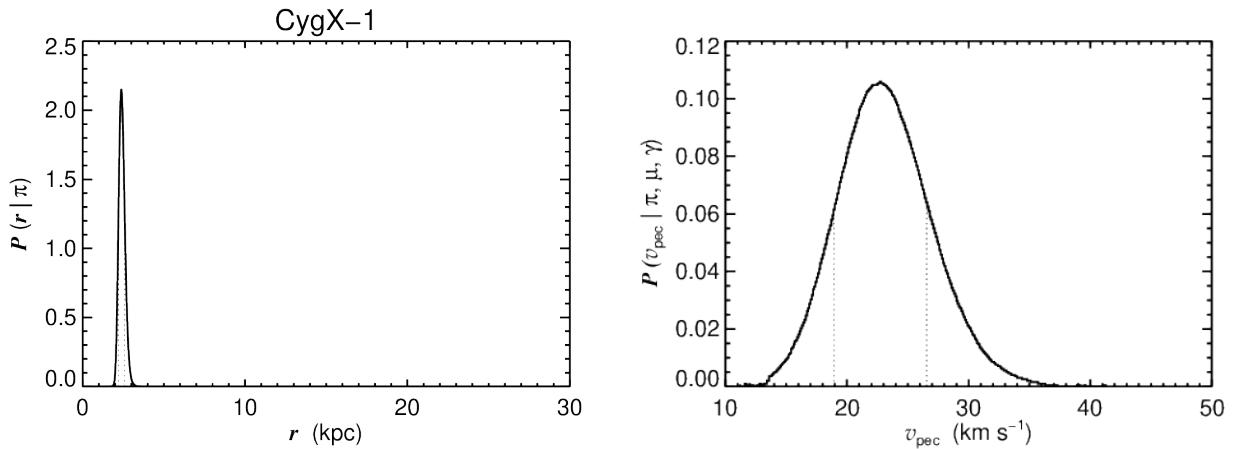
**Figure A12.** V4641 Sgr. (*Left*) Distance  $r_{\text{exp}}$  and (*Right*) peculiar velocity  $v_{\text{pec}}$  posterior distributions. The dotted vertical lines denote the quoted confidence intervals.



**Figure A13.** MAXI J1820+070. (*Left*) Distance  $r_{\text{exp}}$  and (*Right*) peculiar velocity  $v_{\text{pec}}$  posterior distributions. The dotted vertical lines denote the quoted confidence intervals. For this source, the systemic radial velocity is currently unknown, and is assumed to be  $\gamma = 0 \text{ km s}^{-1}$ .



**Figure A14.** V404 Cyg. (Left) Distance  $r_{\text{exp}}$  and (Right) peculiar velocity  $v_{\text{pec}}$  posterior distributions. The dotted vertical lines denote the quoted confidence intervals.



**Figure A15.** Cyg X-1. (Left) Distance  $r_{\text{exp}}$  and (Right) peculiar velocity  $v_{\text{pec}}$  posterior distributions. The dotted vertical lines denote the quoted confidence intervals.