MNRAS **000**, 1–14 (2019) Preprint 4 September 2019 Compiled using MNRAS LATEX style file v3.0

Testing Modified Gravity with Wide Binaries in GAIA DR2

Charalambos Pittordis,[1]⋆ Will Sutherland,[1]†
[1] *School of Physics & Astronomy, Queen Mary University of London, Mile End Road, London E1 4NS, UK.*

Accepted by MNRAS, 05 July 2019; in original form, 21 May 2019

ABSTRACT
Several recent studies have shown that very wide binary stars can potentially provide an interesting test for modified-gravity theories which attempt to emulate dark matter; these systems should be almost Newtonian according to standard dark-matter theories, while the predictions for MOND-like theories are distinctly different, if the various observational issues can be overcome. Here we explore an observational application of the test from the recent GAIA DR2 data release: we select a large sample of $\sim 24,000$ candidate wide binary stars with distance $< 200\,\mathrm{pc}$ and magnitudes $G < 16$ from GAIA DR2, and estimated component masses using a main-sequence mass-luminosity relation. We then compare the frequency distribution of pairwise relative projected velocity (relative to circular-orbit value) as a function of projected separation; these distributions show a clear peak at a value close to Newtonian expectations, along with a long "tail" which extends to much larger velocity ratios; the "tail" is considerably more numerous than in control samples constructed from DR2 with randomised positions, so its origin is unclear. Comparing the velocity histograms with simulated data, we conclude that MOND-like theories *without* an external field effect are strongly inconsistent with the observed data since they predict a peak-shift in clear disagreement with the data; testing MOND-like theories *with* an external field effect is not decisive at present, but has good prospects to become decisive in future with improved modelling or understanding of the high-velocity tail, and additional spectroscopic data.

Key words: gravitation – dark matter – proper motions – binaries:general

1 INTRODUCTION

Einstein's theory of General Relativity (GR) provides the best known description of gravity on all scales. However, much cosmological data (e.g. Ade et al 2016) requires an additional cold, non-baryonic & non-visible dark matter (DM) component to match many observations, in addition to dark energy such as a cosmological constant. At the present time there is no decisive direct detection of DM; this leaves an open window for various modified-gravity theories, which might potentially account for these various observations without the requirement for exotic DM.

The MOdified Newtonian Dynamics (MOND) is a well-known theory that attempts to explain weak-field/non-relativistic gravitational effects without DM. This theory was first proposed by Milgrom (1983) to explain the flat rotation curves observed in most spiral galaxies without requiring DM. The original MOND formulation was non-relativistic and really a fitting function rather than a realistic theory; it has later been incorporated into relativistic theories following from the Tensor-Vector-Scalar (TeVeS) theory proposed by Bekenstein (2004). While the original TeVeS is now excluded by the constraints on time-delay in the neutron-star merger GW 170817 (Boran et al 2017), other versions remain viable; see e.g. Clifton et al. (2012) and Famaey & McGaugh (2012) for reviews of modified gravity, and e.g. Arraut (2014) for a non-local gravity model or Capozziello (2019) for recent cosmological comparisons.

Clearly, a convincing direct detection of dark matter would be the most decisive scenario, but the converse is not true: null results from dark matter experiments can never rule out the paradigm, because the DM interaction cross-section might simply be too small for any practical experiment (or, the cross-section could be weak-like but the DM particle masses could be $\gtrsim 10^9$ GeV, implying a local number density far below the value for conventional TeV-scale WIMPs). Therefore, in the absence of a DM direct detection, new tests which can discriminate between DM and modified-

⋆ E-mail: c.pittordis@qmul.ac.uk
† E-mail: w.j.sutherland@qmul.ac.uk

gravity from direct tests of gravity at the relevant very low accelerations are highly desirable.

Wide-binary stars (separations $\gtrsim 5\,\mathrm{kAU}$) are a promising route to a direct test, since they have low orbital accelerations near or below the typical MOND acceleration scale $a_0 \sim 1.2 \times 10^{-10}\,\mathrm{m\,s^{-2}}$, and are generally presumed to contain negligible dark matter. Studies of wide binaries in general have been explored by e.g. Yoo et al. (2003), Lepine & Bongiorno (2007), Kouwenhoven et al. (2010), Jiang & Tremaine (2010), Dhital et al. (2010), Coronado & Chaname (2015), and others. Previous work concerning tests of MOND-like gravity has been done by Hernandez et al. (2011),Hernandez et al. (2012), Hernandez et al. (2014), Matvienko & Orlov (2015), Scarpa et al. (2017) and Hernandez et al. (2019); these typically give hints of deviations in the direction expected from MOND-like gravity, though due to the limited precision of then-available data, these hints are not yet decisive.

In a recent paper (Pittordis & Sutherland 2018), hereafter PS18, we used simulations to explore the prospects for this test in anticipation of the much improved data expected from the *GAIA* spacecraft (Prusti et al. 2016); PS18 used simulated wide-binary orbits for a variety of acceleration laws, including Newtonian and various MOND models both with and without an external field effect (hereafter ExFE); the general conclusion was that GAIA data provides promising prospects for such a test, since MOND-like models *without* an ExFE should produce large and obvious deviations towards larger velocity differences.

In MOND-like models *with* the ExFE included, as theoretically preferred, the local Galactic acceleration substantially suppresses MOND-like effects, but does not eliminate them. These models give predicted relative velocities much closer to Newtonian, but do still show subtle deviations, most notably a significantly larger fraction of binaries with pairwise velocities in the range $(1.1 - 1.5) \times v_c(r_p)$, where $v_c(r_\mathrm{p})$ is the Newtonian circular velocity at projected separation r_p.

Here we recall two of the main conclusions from PS18: MOND-like theories can allow bound binaries with relative velocities above the Newtonian ceiling, $v_{3D}/v_c(r_p) > \sqrt{2}$, (where v_{3D} is the 3-D pairwise relative velocity); but with the ExFE included the fraction of such systems is predicted to be very small, typically 1 percent or less; so simply counting such systems is unlikely to be a practical test due to possible contamination, observational errors, and small-number statistics. However, the upper percentiles of this velocity ratio, or similarly the fraction of binaries with velocity ratio between $\sim 1.2 - 1.5$ are more promising statistics. In PS18 we noted that no more than 11.1 percent of binaries should have a ratio $v_{3D}/v_c(r) > 1.2$ in Newtonian gravity, for *any* eccentricity distribution, while plausible smooth eccentricity distributions produce a slightly lower percentage; while MOND-like theories can produce a significantly higher percentage. Since the 3D separation r is not a practical observable (since the line-of-sight separation is typically well below the precision of distance measurements) we have to replace it with projected separation r_p as a proxy, which shifts the ratios to lower values depending on viewing angles; but this effect can be readily included in simulations. If future observations were to measure a high-velocity tail of binaries well above the Newtonian prediction (after statistical subtraction of contaminants), this could in principle provide strong evidence in favour of modified gravity.

Note in this paper, since radial velocities are not yet available for the large majority of our binary candidates below, we use 2D sky-projected velocity differences rather than 3D velocities as in PS18; this shifts the most relevant velocity window downward to $\sim 1.1 - 1.5$, and also implies that larger samples will be required to counter the added statistical scatter from random viewing angles. However, this does not substantially change the general principle of the test; see Section 4.1, and also Banik & Zhao (2018) and Banik (2019) for further discussion.

The second main conclusion from PS18 was that, in MOND theories *with* the ExFE, there is an optimal window of projected separation, $5 \lesssim r_\mathrm{p} \lesssim 20\,\mathrm{kAU}$, for practical application of the test. Even wider separations are not favoured in practice because the inclusion of the ExFE causes the MOND-like effects to almost saturate at $r_\mathrm{p} \gtrsim 10\,\mathrm{kAU}$, while several observational issues become proportionally worse at even wider separations.

The plan of this paper is as follows: in Section 2 we describe the selection of candidate wide-binaries from the GAIA DR2 data. In Section 3 we describe some randomised samples used to assess the number of chance projection systems, which turns out to be small. In Section 4 we discuss various simulations of the velocity-ratio distributions for Newtonian and MOND-like binary orbits (with and without the ExFE). In Section 5 we discuss simulations of velocity ratio for co-natal hyperbolic flyby systems. In Section 6 we compare data and models, finding reasonable agreement with a simulated "Newtonian plus flybys" distribution; and we summarise our conclusions in Section 7.

2 GAIA DR2 AND SAMPLE SELECTION

2.1 Preliminary selection

Our starting point is the public GAIA Data Release 2 dataset (hereafter DR2), (GAIA Collaboration 2018), released on 2018 April 25. We initially select all stars with measured parallax $\omega > 5\,\mathrm{mas}$ (i.e. estimated distance $< 200\,\mathrm{pc}$) and GAIA broadband magnitude $G < 16$, yielding a sample of 970,760 stars. (Data quality cuts are applied later on, in order that these may be adjusted post-selection). The parallax and magnitude cuts above are chosen to provide a large enough volume to contain a usefully large statistical sample of wide binaries; while the moderate distance limit and relatively bright magnitude limit ensures that GAIA provides high precision on distances and transverse velocities. Finally, the $G < 16$ cut ensures good feasibility for future follow-up high-resolution spectroscopy on selected subsamples. The sky distribution of these stars shows a fairly uniform distribution, with some enhancements near the Galactic Plane and some well-known open clusters.

We then search this nearby-star sample for pairs of stars with projected separation $\leq 40\,\mathrm{kAU}$ (calculated at the mean distance of each candidate pair), parallaxes of both stars consistent with each other within $4\times$ the combined uncertainty, and projected velocity difference $\leq 3\,\mathrm{km\,s^{-1}}$ as inferred from the difference in proper motions; here, the projected velocity difference is computed assuming *both* stars

in each candidate pair are actually at the mean of the two estimated distances.

We note here that this common-distance assumption is important: if the relative velocities are calculated using individual parallax distances, then an example random 1 percent difference in parallax for a system with transverse velocity $40\,\mathrm{km\,s^{-1}}$ scales to a $0.4\,\mathrm{km\,s^{-1}}$ transverse velocity difference, which is similar to or larger than the orbital velocities of interest below. However, since we are almost entirely interested in the velocity *difference* within a binary, the common-distance assumption leads to an error in estimated relative velocity proportional to the unknown *true* fractional distance difference, $(d_1 - d_2)/d$ (see also Shaya & Olling (2011), Section 2.4 of PS18 and El-Badry (2019) for related effects). For true binaries with random orientation we expect $|d_1 - d_2| \leq r_\mathrm{p}$ for 71 percent of systems, and $\leq 2r_\mathrm{p}$ for 90 percent. Then for a typical binary (see below) with $r_\mathrm{p} \sim 10\,\mathrm{kAU}$ and $d \sim 130\,\mathrm{pc}$ we have $r_\mathrm{p}/d \sim 3.7 \times 10^{-4}$; this is much smaller than the fractional uncertainty of the parallaxes, so choosing the common-distance assumption yields a much more precise estimate of the relative velocity for *genuine* binaries.

This search results in a first-cut sample of 50,003 candidate binaries, which is then pruned with additional cuts as described in the following subsections.

2.2 Sky cuts

Inspection of the initial sample showed a roughly uniform distribution across the sky, with some enhancement near the galactic plane and around some well-known open clusters, i.e. the Hyades, Praesepe and Upper Sco.

We therefore applied sky cuts to eliminate galactic latitudes $|b| \leq 15\,\mathrm{deg}$, and regions around the above clusters. Our galactic-latitude cut removes about 1/4 of the sky, which is a moderate reduction in sample size but should significantly reduce confusion or contamination issues at high source densities. These sky cuts reduce the sample to 33,667 candidate binaries.

2.3 Triple and higher systems

To reject the majority of "moving groups" or similar, we searched our binary sample for any star in common between two or more candidate binaries; if so, both or all those binaries were rejected, leaving a sample of 30,550 candidate binaries with no star common to more than one candidate binary.

We also searched for additional co-moving companion stars to a fainter limit: we selected a "faint star" sample of GAIA stars with $G \leq 20$ and measured parallax $\omega \geq 4.2\,\mathrm{mas}$; for each star in a candidate binary, we then searched for faint-star companions with the following criteria:

(i) Parallax consistent with the main star at 4σ.

(ii) Angular separation less than 2/3 of the main-binary separation (since hierarchical triples are expected to be unstable for inner-orbit separation above $\sim 0.4\times$ the outer separation); and angular separation above 0.5 arcsec to avoid barely-resolved companions.

(iii) Measured projected velocity difference from the main star $\leq 5\,\mathrm{km\,s^{-1}}$.

If any such "third star" was found, (in 375 cases), we rejected the candidate binary since a hierarchical triple will generally boost the projected velocity difference of the wide pair; this left a de-tripled sample of 30,175 candidate binaries.

Clearly, the third-star search above will not reject third stars which are either very faint or unresolved from one of our binary members; this will need to be considered for possible followup observations later, but is unlikely to be the dominant source of contamination as we see below.

2.4 Data quality cuts

We next applied data-quality cuts based on the GAIA parameters, as Arenou et al (2018) Equation 1 as follows:

$$\begin{aligned} \chi^2 &\equiv \texttt{astrometric\_chi2\_al} \\ \nu &\equiv \texttt{astrometric\_n\_good\_obs\_al} - 5 \\ u &\equiv \sqrt{\chi^2/\nu} \\ u &\leq 1.2 \times \max(1, \exp\left[-0.2(G-19.5)\right]) \end{aligned} \quad (1)$$

We rejected binaries where either star did not satisfy Eq. 1; this rejected another 5,270 candidate binaries, leaving a final cleaned sample of 24,282 candidate binaries which we use for the main analysis below.

2.5 Results and scaled velocities

For the surviving 24,282 candidate binaries, we show a plot of projected velocity difference vs projected separation in Figure 4; this shows a clear excess approximately as expected for bound binaries, with an overdense cloud following a locus $v_\mathrm{p} \sim 1\,\mathrm{km\,s^{-1}}(r_\mathrm{p}/1\,\mathrm{kAU})^{-0.5}$. We note that our sample starts to miss true binaries at projected separations below $r_\mathrm{p} \lesssim 0.6\,\mathrm{kAU}$, due to the $3\,\mathrm{km\,s^{-1}}$ velocity threshold, but this r_p is much smaller than the separations of interest below. At $r_\mathrm{p} > 5\,\mathrm{kAU}$ the threshold includes pairs with velocity difference far above the bound limit, which are interesting for assessing sample contamination as seen below.

It is more informative to rescale to the typical Newtonian orbit velocity, so we next estimate masses for each binary using an estimated mass/luminosity relation: here, we adopt the main-sequence $M_I(mass)$ relation from Pecaut & Mamajek (2013), and the $V-I, M_I$ colour relation from the same, where M_I denotes absolute magnitude. From those we apply the colour relation given in Table A2 of Evans et al. (2018) to predict G magnitude from V and I magnitudes as

$$\begin{aligned} G \simeq\ & V - 0.01746 + 0.008092(V-I) - 0.2810(V-I)^2 \\ & +0.03655(V-I)^3 \end{aligned} \quad (2)$$

to obtain a predicted relationship between absolute GAIA magnitude M_G vs mass; we then fit to this to obtain an approximate mass/M_G relation

$$\frac{M}{M_\odot} = 10^{0.0725(4.76-M_G)} \quad (3)$$

Then, for each star we have M_G directly from G and parallax distance, hence an estimated mass follows. Since the luminosity(mass) relation is rather steep, small errors in G or distance have relatively little effect on mass estimates below. A histogram of estimated distances for our binary sample is shown in Figure 2, and a histogram of estimated masses is shown in Figure 3.

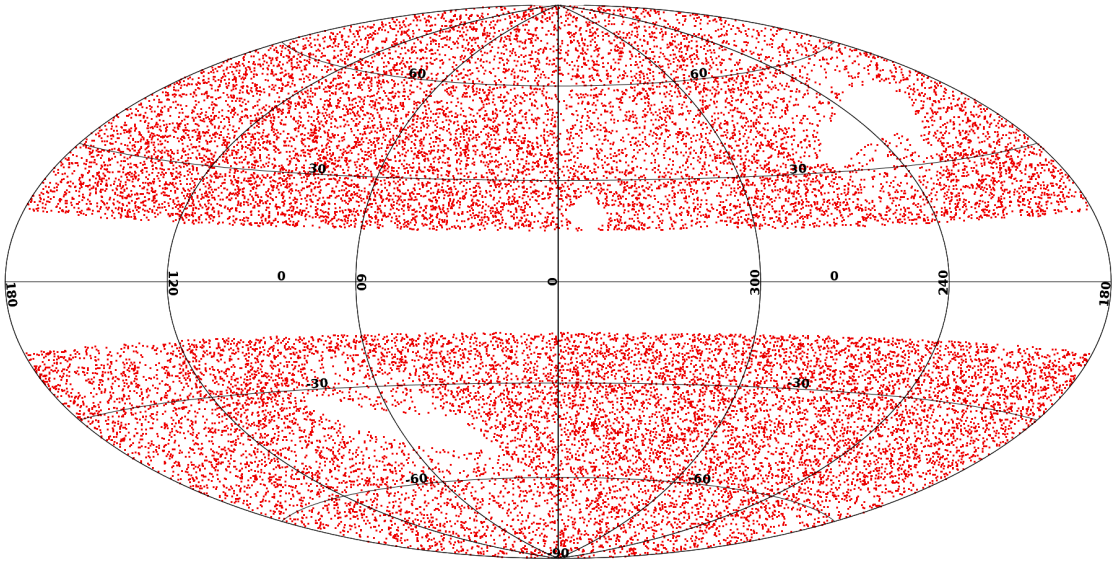

Figure 1. The sky distribution in Galactic coordinates for candidate binaries surviving all cuts in Section 2. The two larger holes are due to regions with fewer GAIA scans.

For each candidate binary we then define

$$v_c(r_\mathrm{p}) \equiv \sqrt{GM_{tot}/r_p} \tag{4}$$

as the estimated circular-orbit velocity at the current *projected* separation; for each candidate binary, we then divide the measured projected velocity difference by the above to obtain a dimensionless ratio $v_p/v_c(r_\mathrm{p})$; a scatter plot is shown in Figure 5, and various histograms of this ratio are compared with models below.

2.6 Transverse velocity errors

We have estimated relative-velocity errors assuming uncorrelated errors between the two components of the binary, simply from the root-sum-square of the quoted rms errors in μ_α and μ_δ for each of the two stars in each binary, and multiplying by distance to obtain the transverse-velocity error. (This should be reasonable as long-range correlated errors should mostly cancel between the two stars). The median of this for the binary sample above is $\sigma(v_\mathrm{p}) \approx 0.09\,\mathrm{km\,s^{-1}}$, which is already impressively small. A scatter plot of $\sigma(v_\mathrm{p})$ versus distance is shown in Figure 6; the trend with distance is clear, but most systems have $\sigma(v_\mathrm{p}) \lesssim 0.15\,\mathrm{km\,s^{-1}}$ even near our 200 pc limit.

Converting to the ratio to circular-orbit velocity, $\sigma(v_\mathrm{p})/v_c(r_\mathrm{p})$, the median for the full candidate sample is 0.08 and the 80th percentile is 0.14; for the "wide" subsample with $5 < r_\mathrm{p} < 20\,\mathrm{kAU}$, the median is 0.23 and the 80th percentile is 0.39.

The latter values are significantly smaller than 1, but not very small, so the effect of random proper motion errors will affect the detailed shape of the distributions below. However, in future GAIA data releases these values are expected to reduce by factors of at least 2–4 as proper motion precision scales as $\propto t^{-3/2}$, so the random errors in proper motions are likely to become relatively unimportant in the medium-term future.

We note that for a "typical" binary below at $r_\mathrm{p} \sim 10\,\mathrm{kAU}$ and $d \sim 130\,\mathrm{pc}$, the angular separation is 0.37 mrad or 77 arcsec, so these are very well resolved and the uncertainty on r_p is essentially the same as the error on the mean distance, typically well below 1 percent and almost negligible. The error on v_p is dominated by random errors on the proper motions, assuming that correlated systematic errors mostly cancel between the two components of the binary. Since we are mostly interested in statistical distributions, the effect of random errors is modest as long as these are not larger than ~ 0.25 in $v_p/v_c(r_\mathrm{p})$. Note that for systems with small observed ratios $v_\mathrm{p}/v_c(r_\mathrm{p}) \sim 0.5$, the *fractional* uncertainty in this ratio is rather large; however such systems still have a high probability of the true ratio being $\lesssim 0.8$, so this scatter is relatively unimportant. For systems with $v_\mathrm{p}/v_c(r_p) \gtrsim 1$, the fractional uncertainty is relatively modest; though possible non-Gaussian errors in the GAIA data remain a concern, this should improve in future GAIA releases as more observing epochs become available to reject outliers.

2.7 Comparison with El-Badry & Rix

Here we note that a catalogue of candidate binaries in GAIA DR2 has been published by El-Badry & Rix (2018) (hereafter ER18). Much of our sample selection was completed independently prior to the appearance of ER18, but here we give a brief comparison.

ER18 chose (coincidentally) the same 200 pc limit as here, but there are two main differences: firstly, our selection adopts a fixed threshold $v_p \leq 3\,\mathrm{km\,s^{-1}}$, whereas ER18 use a separation-dependent threshold which translates to $v_p \leq 2.1\,\mathrm{km\,s^{-1}}(r_\mathrm{p}/1\,\mathrm{kAU})^{-0.5}$ (equivalent to the Newtonian bound limit $\sqrt{2}v_\mathrm{C}$ for a $2.5\,M_\odot$ system, or just above at $1.83v_\mathrm{C}$ for a more typical system with mass $1.5\,M_\odot$). This means that our sample extends to substantially higher (unbound) velocity ratios at $r_\mathrm{p} \gtrsim 3\,\mathrm{kAU}$, which turns out to be useful below for investigating the high-velocity tail. Secondly, our sample uses a magnitude limit $G < 16$, while

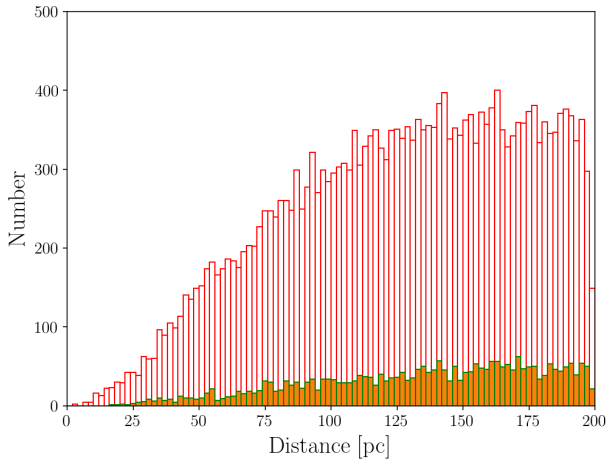

Figure 2. Histogram of average distance for the candidate binaries. The open histogram shows all 24,282 binaries passing the cuts, while the filled histogram shows the subset with $5 < r_{\rm p} < 20\,{\rm kAU}$.

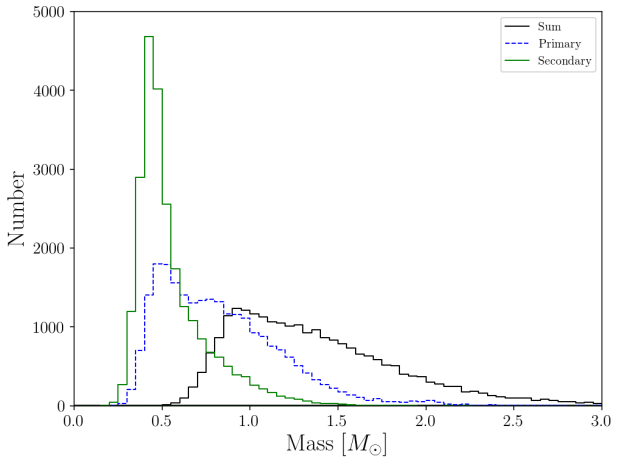

Figure 3. Histogram of estimated masses for the candidate binaries. The black solid line shows combined system mass; dashed blue line shows the primary (more massive) star, and green line shows the secondary star.

ER18 have a $G < 20$ limit (but also other cuts on relative errors, which do introduce some implicit magnitude-dependence).

There are additional differences in how we cut for clusters, triples, etc, but these turn out to be relatively less important.

We have done a cross-match of our sample to ER18 as follows: selecting the subset of the ER18 sample where both stars have $G < 16$, our galactic cut $|b| > 15\,{\rm deg}$ and $r_{\rm p} < 40\,{\rm kAU}$ gives a subsample of 18,513 candidate binaries from ER18 which could in principle pass our other cuts. Cross-matching those against our cleaned sample of 24,282 candidates, we find that 15,652 are in common. Most of the additional binaries in our sample are either at larger velocity differences (above the ER18 limit and below our $3\,{\rm km\,s^{-1}}$ limit), or at smaller separations $r_{\rm p} < 0.5\,{\rm kAU}$ which we do not consider below.

Considering a subsample of our candidates with $3 \leq r_{\rm p} \leq 20\,{\rm kAU}$ and velocity ratio ≤ 2, we find 3380 candidates; of those, 3138 are common to the ER18 sample, while only 242 are not in ER18. (There are an additional 606 candidates in ER18 satisfying the above criteria, but not in our cut sample above). This indicates that the different selection criteria have not had a major influence in this region.

3 RANDOM SAMPLES

3.1 Construction of random samples

It is clearly important to estimate the level of contamination of our sample by random chance projections of unrelated stars which just happen to have chance small velocity differences. To do this, we have constructed several randomised samples by first removing one star from each binary, randomising the true RA/Dec values by a few degrees (see below) in each coordinate, then re-running the binary search on the position-randomised sample.

Here, the removal of one star from each binary was chosen since otherwise close binaries can be "scattered" and re-selected as wide binaries in the randomised list. Also, the choice of few degree position shifts is small enough to preserve the global distributions with respect to galactic coordinates, but is large enough to eliminate most truly associated stars.

We show an example histogram for a set of randomised samples compared to the data in Figure 7. Since there the random counts are much smaller than the data and hard to see, the following Figure 8 shows the same with the random counts multiplied by 10× for visibility.

3.2 Comparison of data and randomised sample

The main result from our randomised samples is that they contain far fewer candidate "binaries" than the actual GDR2 data: the data samples outnumber the mean randoms by a factor of 124 in the separation bin 5 to 7.1 kAU, falling to 11.6 in our widest bin 14 to 20 kAU. As expected the randomised samples show no peak at small velocity ratios, but a fairly smooth distribution with a gradual rise towards larger velocity ratios.

We note that this large excess number in the real sample is present especially at "bound" velocity ratios $< \sqrt{2}$, but also persists to large velocity ratios $\gtrsim 2$.

It seems reasonable to assume that the clear peak in the data at small velocity ratio ~ 0.6 is dominated by genuine bound binaries, while the "tail" at larger velocity ratios is rather unexpected: the tail is much more populous than our randomised samples, so it is clearly not due to chance projections of unrelated stars. But the tail extends smoothly to velocity ratios much larger than attainable even in reasonable modified-gravity models; thus the tail may be dominated by stars with some common origin, e.g. co-natal pairs of stars born in the same open cluster which has subsequently dissolved, while the two star velocities retain a memory of their common origin, leading to unbound but correlated pairs in

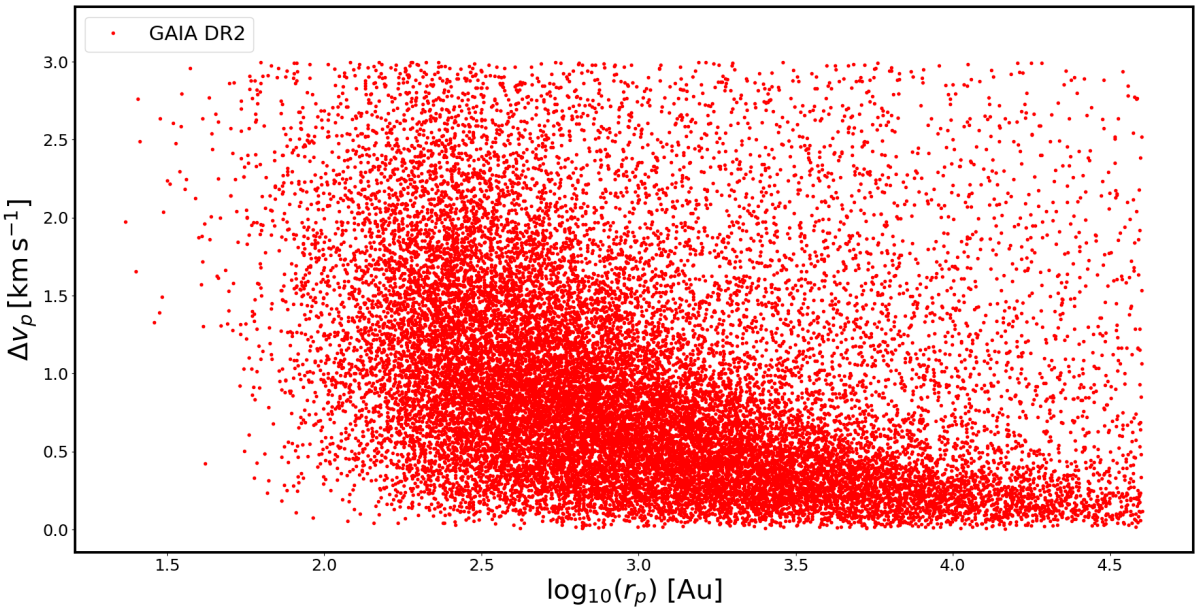

Figure 4. Scatter plot of projected velocity $v_{\rm p}$ (y-axis) vs projected separation (log scale, x-axis) for the cleaned binary sample. The main selection cuts are visible at top and right.

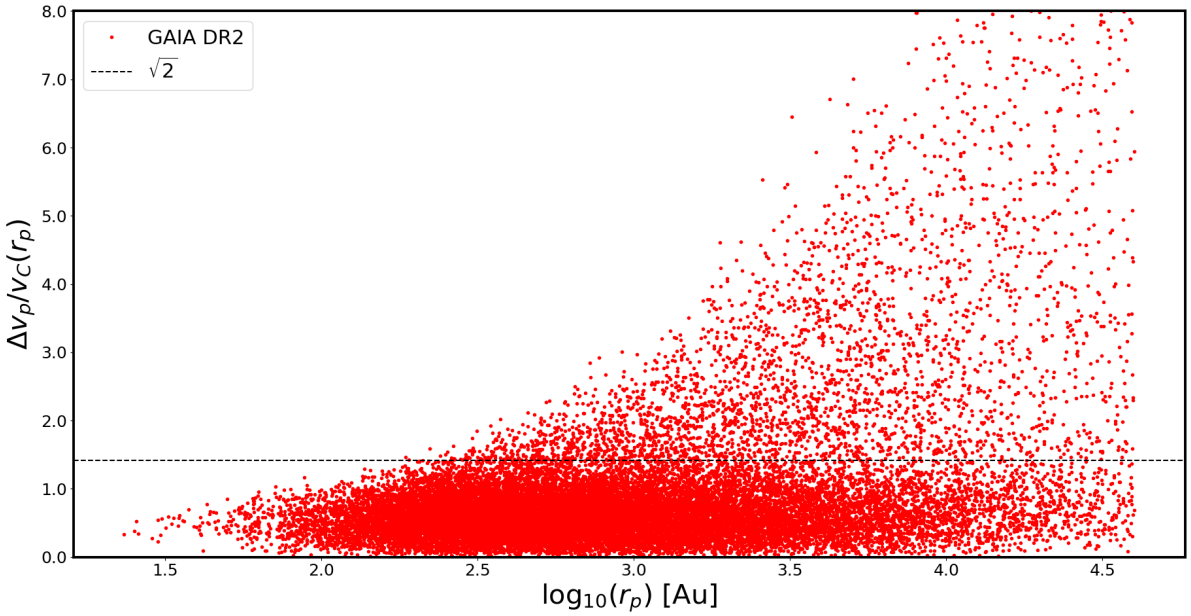

Figure 5. Scatter plot of projected velocity relative to Newtonian, $v_{\rm p}/v_c(r_{\rm p})$, vs projected separation. The dashed line at $\sqrt{2}$ indicates the Newtonian limit. The upper cutoff is now slightly fuzzy due to the additional dependence on mass.

phase-space. It seems clear that understanding and modelling the origin of this tail will be crucial to use the binaries for a gravity test: this is modelled later in Section 5. Before investigating the tail, we next discuss simulations of velocity ratios for genuine bound binaries.

4 ORBIT SIMULATIONS

4.1 Velocity ratios

We here recall that, unlike PS18 where we used 3D relative velocities and 2D projected separations, in this paper we are using 2D sky-projected velocities as well as 2D projected separations; this is simply because radial velocities are not yet available for the large majority of our candidate WBs, though they should become available in future with large spectroscopic surveys such as 4MOST, PFS, WEAVE and MSE, and targeted followup.

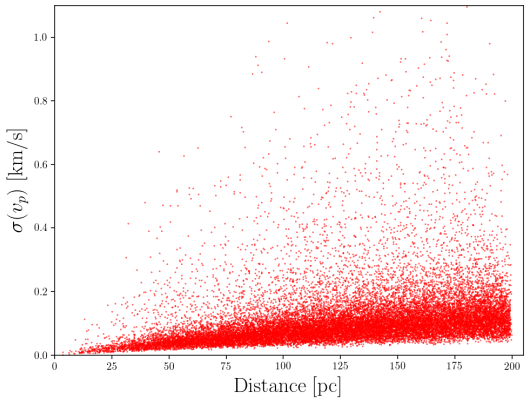

Figure 6. Scatter plot of rms velocity uncertainty $\sigma(v_{\rm p})$ versus mean distance for the 24,282 binaries surviving all cuts.

For the 2D velocity ratios as above, note that the true value (in the absence of perspective rotation effects and/or observational errors), is always smaller than the corresponding (unknown) value in 3D, since $v_{\rm p} \leq v_{3D}$, and $r_p \leq r$ so $v_c(r_p) \geq v_C(r)$, by a factor which depends on the unknown alignment and orientation of the orbit.

We note here that for circular orbits the 3D ratio is 1 by definition, and the above 2D ratio is readily derived as

$$v_{\rm p}/v_c(r_{\rm p}) = \sqrt{1-\sin^2 i \sin^2\phi}\left[1-\sin^2 i\cos^2\phi\right]^{1/4}$$

where i is inclination and ϕ is orbit angle from conjunction; here the first factor is due to the deletion of the line-of-sight velocity component, while the second factor is $\sqrt{r_p/r}$ due to projected separation. A histogram of this ratio for random angles is shown in Figure 9; this is explained because nearly face-on orbits $\sin^2 i \lesssim 0.36$ produce a ratio close to 1 at all orbit phases, while nearly edge-on orbits $\sin^2 i \gtrsim 0.8$ produce a maximum value near 0.62 at intermediate angles $\phi \sim 0 \pm 0.62$ or $\pi \pm 0.62$ rad from conjunction, dropping to small values near conjunction or greatest elongation; thus high-i orbits create the caustic spike and the tail to low ratios, while low-i orbits fill the plateau from 0.7 to 1.

The above shows a simple but unrealistic case of pure circular orbits: for realistic distributions of non-zero orbit eccentricity, we calculate the resulting distribution via numerical simulations, similar to PS18 but using 2D projected velocity differences instead of 3D, i.e. suppressing the radial-velocity information. Some example results for Newtonian gravity and three selected eccentricity distributions are shown in Figure 10; the eccentricity distributions are flat ($f(e) = 1$), the observational fit of Tokovinin & Kiyaeva (2016) which is $f(e) = 0.4 + 1.2e$; and finally $f(e) = 2e$. As expected the distributions become broader and extend to values > 1, while the circular-orbit spike near 0.62 is smeared out, resulting in a broad peak at a ratio around 0.6 and a steeply-declining tail at values above 1.

A notable point here is that the tail is relatively small, since a ratio $\gtrsim 1.1$ requires a simultaneous combination of three factors: a fairly eccentric orbit, a viewing angle roughly face-on, and a epoch fairly close to orbit pericenter today. Each of these singly is only mildly improbable, but they are expected to be almost independent of each other so the combination of all three has a rather small probability. Also, Figure 10 shows that the tail at velocity ratio ≥ 1.1 is relatively insensitive to the details of the eccentricity distribution, similar to the PS18 result for 3D velocities (see also Fig. 3 of BZ18 for a similar conclusion); this is because extreme-e orbits $e > 0.9$ achieve higher maximum ratios, but spend less time near pericenter, so much of the tail near velocity ratio ~ 1.1 is contributed by moderate-e orbits.

In particular, it is possible to get *fewer* binaries at high ratios if orbits are preferentially near-circular, but the converse is not true i.e. the Tokovinin or $f(e) = 2e$ distributions are almost a ceiling on the fraction of bound binaries in the high-velocity-ratio tail [1]. Also, the predicted distribution is steeply declining at ratios $\sim 1.0 - 1.3$; this feature is useful later, since a rather moderate shift in velocities from MOND-like gravity with the ExFE (below, we find typically $\sim 15\%$ velocity boost above Newtonian) translates into a rather large multiplicative enhancement (typically $2\times$ to $2.5\times$) in the predicted fraction of true binaries with velocity ratio $\gtrsim 1.1$; this large enhancement cannot be mimicked in standard gravity by varying the eccentricity distribution, in agreement with BZ18.

4.2 Modified gravity orbits

Here we simulate a large sample of $\sim 5\times10^6$ orbits with random values of a, e in several modified-gravity theories then study the joint distribution of observables, in particular projected separation r_p and relative velocity ratio $v_{\rm p}/v_c(r_p)$, as defined above. Here we use the Tokovinin & Kiyaeva (2016) eccentricity distribution, as the intermediate of the 3 example cases above.

In the case of modified gravity models, the orbits are generally not closed ellipses, so they are not strictly defined by the standard Keplerian parameters a, e, but we still need to simulate a distribution in size and shape of orbits. To deal with this, as in PS18, for a modified-gravity orbit we define an "effective" orbit size $\hat{a}$ and quasi-eccentricity $\hat{e}$ as follows: we define $\hat{a}$ to be the separation at which the simulated relative velocity is equal to the circular-orbit velocity (in the current modified-gravity model), then we define $\theta_{\rm circ}$ to be the angle between the relative velocity vector and the tangential direction when the orbital separation crosses $\hat{a}$, and then $\hat{e} \equiv \sin\theta_{\rm circ}$; these definitions coincide with the usual Keplerian a, e in the case of standard gravity.

After integrating these orbits using one of a selected set of gravity laws (Newton/GR, MoND with/without ExFE) and a chosen value for external field g_e, we "observe" the resulting binaries at many random times and random inclinations to the line-of-sight.

For each simulated orbit/epoch snapshot, we produce simulated observables including the projected separation r_p, projected relative velocity $v_{\rm p}$, and also $v_{\rm p}/v_C(r_p)$ corresponding to our observable from GAIA.

[1] Note that it may be possible to get a "false negative" conclusion if a MOND-like gravity modification is correct, but there is also a strong bias against eccentric orbits at larger separations which cuts off the high-velocity tail. This seems rather contrived, and may be testable by looking at detailed shapes of the distributions.

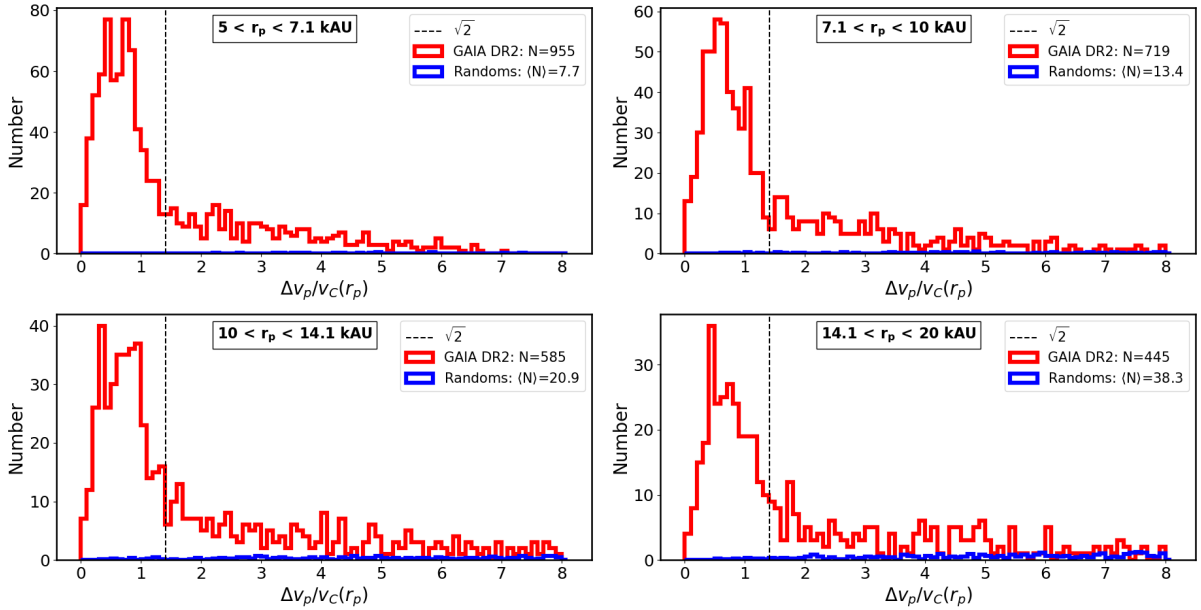

Figure 7. Histograms of velocity ratio $v_{\rm p}/v_{\rm c}(r_{\rm p})$ for binaries in GAIA data (red), and the mean of 9 randomised samples (blue). The four panels show ranges of projected separation, $5-7.1\,$kAU up to $14.1-20\,$kAU as labelled. Sample numbers are shown in the legend.

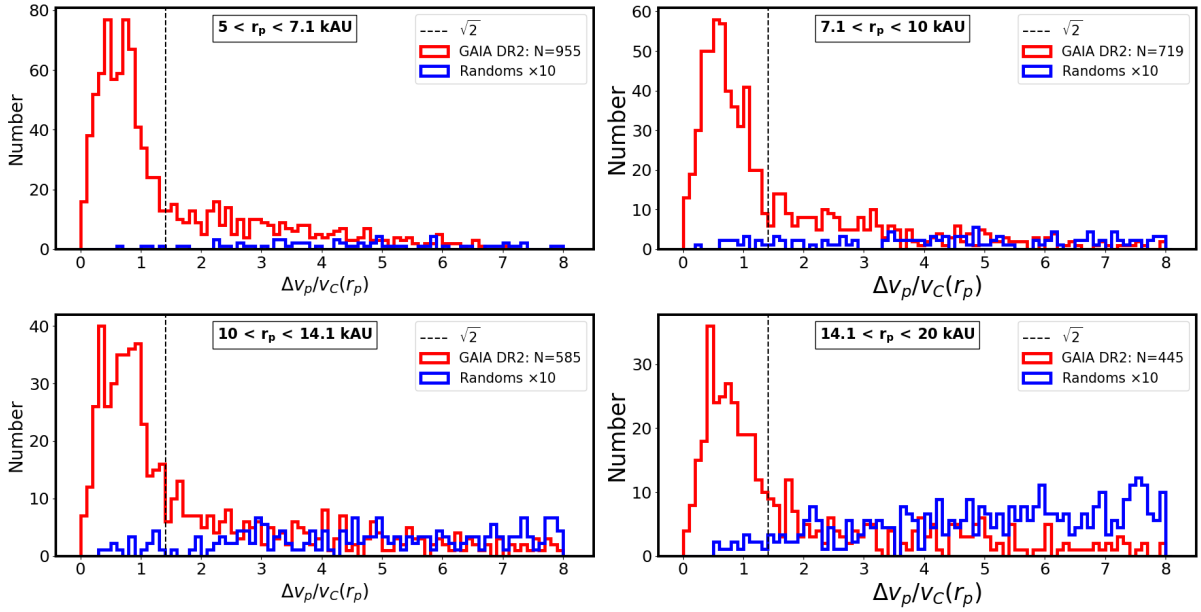

Figure 8. Same as Figure 7, but with the mean random samples artificially scaled-up by 10× to enhance visibility.

The radial acceleration law is chosen according to the selected gravity theory under consideration, and also with the external field effect turned off or on (see below). For the Newtonian/GR case, we have the standard

$$g_N = \frac{G(M_1+M_2)}{r^2} \tag{5}$$

For the MOND-like case (both with and without the ExFE), we use the fitting function of McGaugh et al (2016) (hereafter MLS), sometimes known as the "radial acceleration relation", given by

$$\begin{aligned} g_{MLS} &= g_N\,\nu(g_N/a_0)\ ; \\ \nu(y) &= \frac{1}{1-\exp(-\sqrt{y})}\ ; \end{aligned} \tag{6}$$

we refer to this ν function as the MLS interpolating function below. This function is shown by MLS to produce a good fit to rotation curves for a large sample of disc galaxies spanning a range of masses; it also has the desirable feature that the function $\nu(y)$ converges very rapidly to 1 when $y \gtrsim 20$, so deviations on Solar System scales are predicted to be vanishingly small, consistent with observational limits.

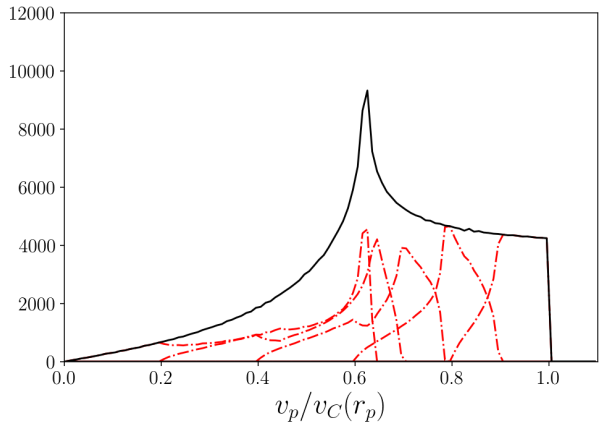

Figure 9. Distribution of 2D velocity ratio $v_{\rm p}/v_{\rm c}(r_{\rm p})$ for randomly-viewed circular orbits. The black solid curve shows the full distribution, while red dash-dot curves show contributions from 5 equal bins of $|\cos i|$, from $0 < |\cos i| < 0.2$ (left) to $0.8 < |\cos i| \leq 1$ (right).

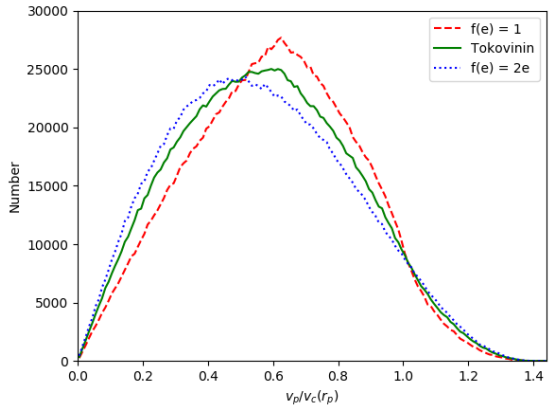

Figure 10. Distribution of 2D velocity ratio $v_{\rm p}/v_{\rm c}(r_{\rm p})$ for randomly-viewed elliptical orbits, with 3 example eccentricity distributions: dashed red curve shows flat distribution $f(e) = 1$; solid green curve shows Tokovinin distribution $f(e) = 0.4{+}1.2e$; dotted blue curve shows $f(e) = 2e$.

For the case without the ExFE, we apply Eq. 6 directly, with the conventional value $a_0 = 1.2 \times 10^{-10}\,{\rm m\,s^{-2}}$.

To apply the External Field Effect (ExFE), we use the approximation of Banik & Zhao (2018); this is given by

$$g_{N,int} = G(M_1 + M_2)/r^2 \quad (7)$$

$$g_{N,gal} = 1.2\,a_0 \quad (8)$$

$$g_{N,tot} = \left(g_{N,int}^2 + g_{N,gal}^2\right)^{1/2} \quad (9)$$

$$g_{i,EFE} = g_{N,int}\nu(g_{N,tot}/a_0)\left(1 + \frac{\kappa(g_{N,tot})}{3}\right) \quad (10)$$

$$\kappa \equiv \frac{\partial \ln \nu}{\partial \ln g_N} \quad (11)$$

where $g_{N,int}$ is the internal Newtonian acceleration of the binary; $g_{N,gal}$ is the external (Galactic) Newtonian acceleration, $g_{N,tot}$ is the quadrature sum of these, ν is the MLS function from Equation (6) and $g_{i,EFE}$ is our model MOND-ian internal acceleration, approximating the application of the external field effect. (This is not quite an exact solution of the MOND-like equations, but is shown by BZ18 to be a good approximation to the full numerical solution).

Above, the observed Galactic rotation values $v_{LSR} \simeq 232\,{\rm km\,s^{-1}}$ and $R_0 \simeq 8.1\,{\rm kpc}$ imply a total Galactic acceleration close to $1.75\,a_0$, hence we require $g_{N,gal}\nu(g_{N,gal}/a_0) \approx 1.75\,a_0$. Solving this leads to $g_{N,gal} \approx 1.16\,a_0$ as above and $\nu \approx 1.51$, in reasonably good agreement with the estimated baryonic contribution to the Galactic rotation (as expected, since the MLS fitting function was derived by fitting to a sample of external spiral galaxies with well-observed rotation curves, so this is consistent with our Galaxy being typical).

For the MLS ν function and ExFE approximation as above, some example values for our wide-binaries are $g_{N,int} \sim 0.8a_0$, $g_{N,gal} \sim 1.2\,a_0$, so the internal and galactic accelerations are both comparable to a_0. This leads to numerical values $\nu(g_{N,tot} \sim 1.41\,a_0) \sim 1.44$ and $\kappa \approx -0.26$; this example leads to $g_{i,EFE} \simeq 1.35\,g_{N,int}$, hence Eq. 11 predicts that the wide binaries in ExFE will obtain a $\sim 35\%$ boost above Newtonian acceleration, equivalent to a $\sim 16\%$ boost in typical binary orbital velocities. This boost is significantly larger than we found in PS18 using the previous less accurate ExFE approximation (Eq. 45 of PS18); hence this improves the prospects for a decisive test (as was also briefly noted in PS18 Section 4.3).

In Figure 11, we show histograms of the ratio $v_{\rm p}/v_{\rm C}(r_{\rm p})$ for four bins of projected separation, and three gravity acceleration models: standard Newtonian, and MLS both without and with the ExFE, compared with the GAIA DR2 data. Model histograms are normalised to match the total of the data at velocity ratio $\leq \sqrt{2}$.

On inspection of Figure 11, several features are immediately clear:

(i) The simulated histograms for MOND *without* ExFE show a large and obvious shift of the peak to larger velocity ratios, especially in the wider separation bins where the Newtonian acceleration is well below a_0; this shifted peak appears clearly inconsistent with the data, and plausible observational errors or sample contamination appear unlikely to remove this inconsistency.

(ii) The simulated histogram for MOND *with* ExFE is much closer to the Newtonian case, but shows a distinctly larger fraction of binaries at velocity ratio $1.1 - 1.5$; the excess is increasing only moderately with projected separation, and saturates at still larger separations (beyond 20 kAU).

(iii) The actual data show a clear peak at ~ 0.6 as above, then a fall towards $\sqrt{2}$; followed by a prominent tail which slowly declines to much larger ratios $\gtrsim 5$. Clearly, the presence of this tail makes it hard to decide a preference between Newtonian or MOND-with-ExFE: a smooth downward extrapolation of the tail below $\sqrt{2}$ can account for at least 50 percent, possibly as many as 70 percent, of the observed number of systems with velocity ratio $1.1 < v_{\rm p}/v_{\rm C}(r_{\rm p}) < 1.4$, so the tail is not well understood but has a major impact on the statistics. This tail is discussed and modelled in the next Section.

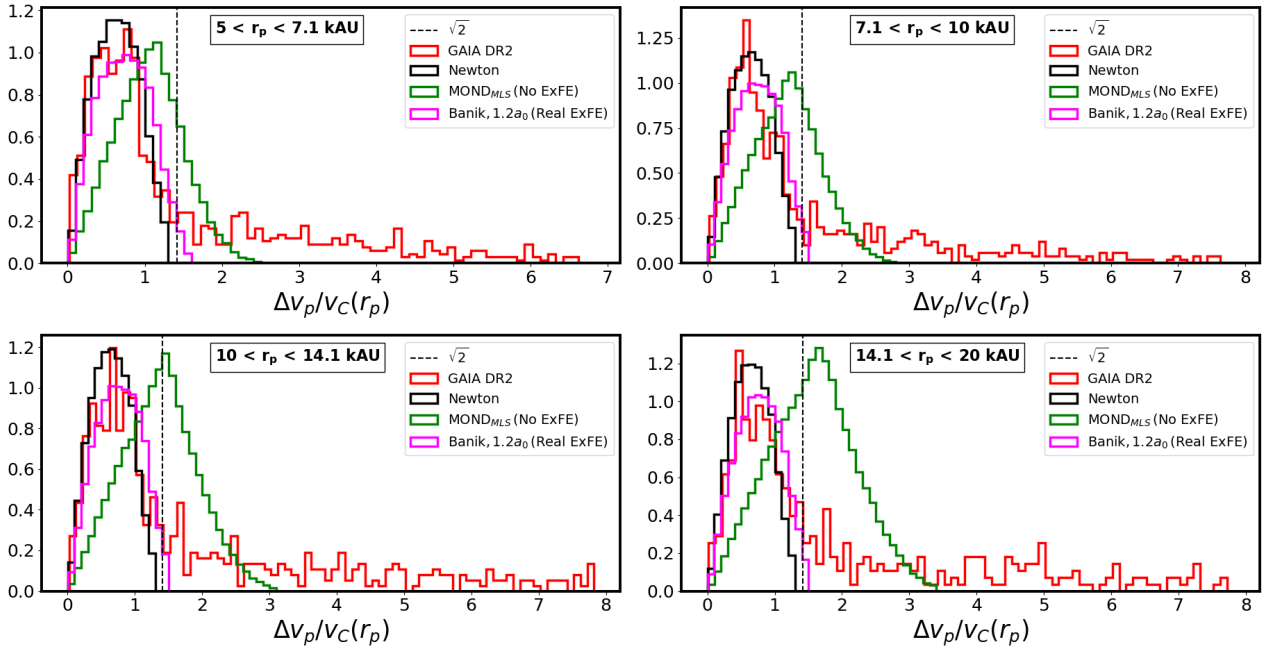

Figure 11. Histograms of velocity ratio, $v_{\rm p}/v_{\rm C}(r_{\rm p})$, for observed and simulated binaries: the four panels show four bins of projected separation as labelled in the legend. The histograms show our observed GAIA sample (red); Newtonian simulated orbits (black); MOND *with* ExFE (magenta) and MOND *without* ExFE (green). The simulated histograms are normalised to the number of observed systems at ratio $\leq \sqrt{2}$.

5 FLYBY SIMULATIONS

We saw above that the observed histograms (Figure 11, red lines) of velocity ratios in GAIA DR2 can be qualitatively described by a two-component "hump + tail" structure, with the hump peaking near 0.6 in approximate agreement with Newtonian orbit expectations, while the smooth "tail" extends to much larger velocity ratios. Our randomised-position samples show that the tail is much too populous to be explained by random chance projections of unrelated stars; one plausible explanation appears to be pairs of co-natal stars born in the same open cluster, which therefore have similar velocities and are currently undergoing a chance close flyby. Evidence for a population of such "cold streams" has been given by Oh et al (2017).

In this case, we would expect two of the three velocity components (perpendicular to the escape direction) must be similar in order to get a close flyby, while the velocity difference in the escape direction should approximately reflect the distribution in ejection velocities from the cluster and the time difference between the two ejections.

During a flyby, the relative velocity will speed up according to a hyperbolic flyby orbit. To simulate this, we generated random flyby encounters as follows:

(i) We chose a distribution of impact parameter b with $dn/db \propto b$ up to a maximum value of 300 kAU or 1.45 pc.

(ii) We chose various distributions of asymptotic velocity difference v_∞, including a uniform distribution up to $2\,{\rm km\,s^{-1}}$ (hereafter the Flat distribution), two Maxwell-Boltzmann distributions with $\sigma = 0.75, 1.5\,{\rm km\,s^{-1}}$, and an exponential distribution $\propto \exp(-v_\infty/1\,{\rm km\,s^{-1}})$.

(iii) Given values of b, v_∞ generated with distributions as above, each such pair produces one flyby encounter. We sampled each hyperbolic flyby at random times while the 3D separation was ≤ 300 kAU, and computed the 3D relative separation and relative velocity vectors. We "observed" these from random viewing angles to produce $v_{\rm p}, r_{\rm p}$.

(iv) We truncated the sample as for the DR2 data, at $r_{\rm p} < 40$ kAU and $v_{\rm p} \leq 3\,{\rm km\,s^{-1}}$, and computed the resulting velocity ratio $v_{\rm p}/v_c(r_{\rm p})$.

Clearly for hyperbolic flyby encounters the 3D velocity ratio $v_{3D}/v_c(r)$ is always $\geq \sqrt{2}$, with simulations showing a modest pile-up in the distribution just above this value; this pile-up arises from flybys with eccentricity not much larger than 1, which speed-up substantially but have velocity ratios decreasing towards $\sqrt{2}$ as they approach pericenter. The projection to 2D smears this distribution to lower ratios, and thus fills-in the gap below $\sqrt{2}$; the result is that simulated fly-bys with a flat distribution of v_∞ produce a smooth maximum in the distribution at a ratio below 1.0, with a gently declining tail at larger ratios.

6 DATA VS MODEL COMPARISONS

We now turn to a comparison between data and models with a Binaries + Flybys population, for the histograms of observed velocity ratio $v_{\rm p}/v_c(r_{\rm p})$. We have produced histograms sliced in $\sqrt{2}$ bins of $r_{\rm p}$, from 5 kAU to 20 kAU; hence four bins respectively $5-7.1; 7.1-10; 10-14.1; 14.1-20$ kAU. These are chosen spanning the most favourable range for the gravity test, since in PS18 it was found that modified-gravity effects with the ExFE included should largely saturate above 10 kAU; while the statistics become rather poor

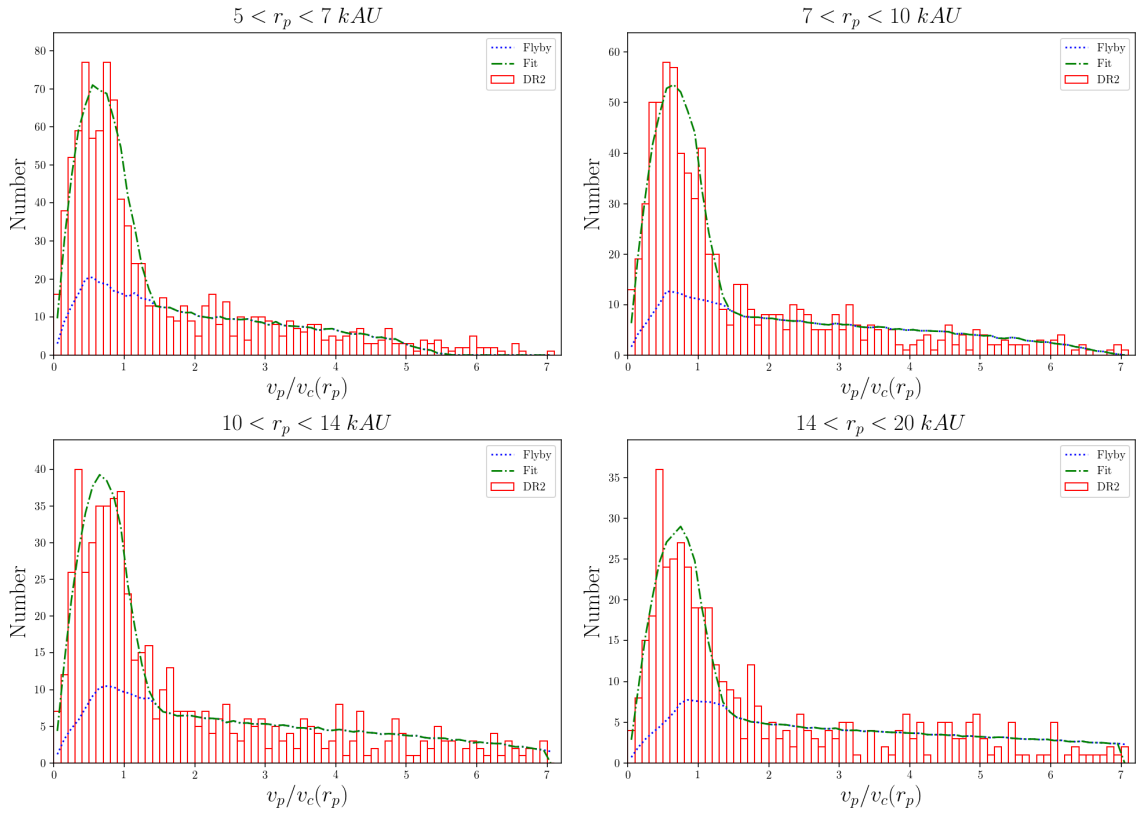

Figure 12. Fits of velocity ratios using a "Binaries + Flybys" model; four panels show four bins of projected separation as labelled. Here the Flyby model adopts a flat distribution of v_∞ up to $2\,\mathrm{km\,s^{-1}}$; fits adopt arbitrary normalisation for both binaries and flybys. The red histogram shows the data. The blue-dotted line is the fitted flyby population; green dashed line is the total, so the difference is fitted true binaries.

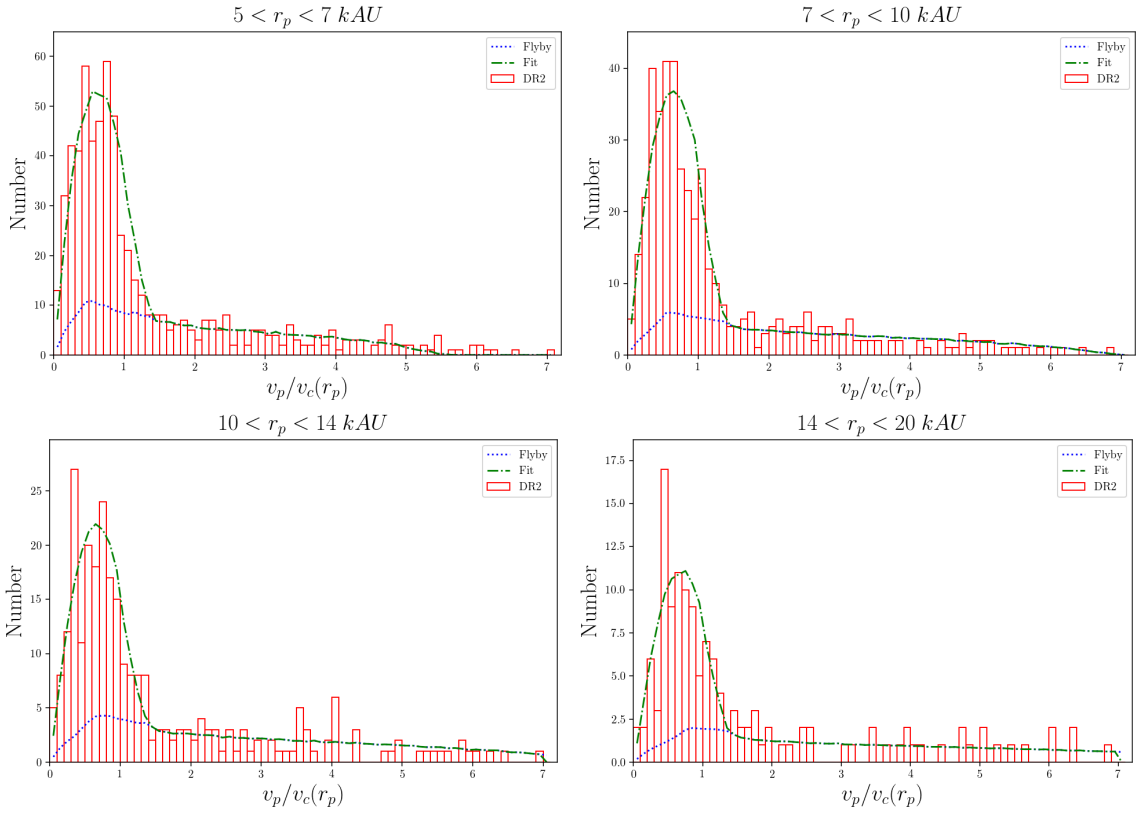

Figure 13. Same as Figure 12, for the subsample of binaries with $\sigma(v_\mathrm{p})/v_\mathrm{C}(r_\mathrm{p}) < 0.25$.

and contamination worsens beyond $r_{\rm p} > 20\,$kAU. More recent simulations from Banik & Zhao (2018) (which used a more realistic formulation of the ExFE compared to PS18) led to a similar conclusion on this optimal separation range.

For fitting, we take the simulated Newtonian binaries with the Tokovinin eccentricity distribution, and the simulated hyperbolic flybys with the Flat v_∞ distribution. We produce a 2-parameter fit by keeping the shapes of the Binary and Flyby populations fixed, and simply adjusting the relative normalisations of each of the Binary and Flyby populations, fitting the sum of these to the observed histogram assuming Poisson errors to provide a maximum-likelihood fit. Results of this fit are shown in Figure 12, and it is seen that the fits match the data quite well given the statistical noise; this is confirmed by the χ^2 values which are acceptable. Counts of the observed and fitted histograms in some selected ranges of velocity ratio are given in Table 1.

The fits indicate that over the interval $0 \leq v_{\rm p}/v_c(r_p) \leq \sqrt{2}$, the Binary population dominates over the model Flybys by a factor of 2 - 4, i.e. most but not all such systems are genuine binaries.

However, it is also seen from the fits that the situation is reversed in the most interesting range $1.1 - 1.4$; here the model Flyby population dominates over the model Binary population by a factor of $\sim 1.5 - 2.5$. While this is not necessarily fatal since the flyby population can be statistically subtracted, it does imply that a better physical understanding or modelling of the flyby population will be required to derive conclusions about MOND with ExFE. We postpone this modelling to a future paper.

Concerning the effect of proper-motion errors on the velocity histograms, we have repeated the above procedure for the subsample of binary candidates with an additional cut $\sigma(v_{\rm p})/v_{\rm C}(r_{\rm p}) < 0.25$; the corresponding results are shown in Figure 13 and Table 2. This does reduce the sample size quite substantially, by a factor $\sim 2/3$ in the first bin down to $\sim 1/3$ in the widest bin; however, the tail is still present and the general appearance of the histograms is approximately unchanged. The fitted ratios of Flyby vs Newtonian binaries in the range of velocity ratio 1.1 – 1.4 do slightly reduce, but this is countered by the small-number statistics so again no firm conclusion concerning MOND with ExFE can be made at present. However, with the substantial reduction of proper motion errors expected in the future GAIA DR3, a large majority of candidate binaries will be expected to survive the above cut.

We also note that while the flyby model does provide an acceptable fit to each of the individual histograms, there is a potential inconsistency in that the fitted number of "flyby" events is slightly decreasing with $r_{\rm p}$, while the simulations predict a rising distribution. This suggests that there may be an additional contribution to the tail from e.g. undetected hierarchical triples or non-Gaussian errors in the GAIA proper motions. Further work e.g. with radial velocities and future GAIA data releases is probably required to understand the origin of the tail.

6.1 Discussion and future prospects

We have seen that the observed distributions of velocity ratio for our candidate binaries appears to be fairly strongly inconsistent with MOND *without* ExFE, since the observed peak stays close to the Newtonian prediction ~ 0.6 independent of $r_{\rm p}$, and there are many more observed systems with $v_{\rm p}/v_{\rm C}(r_{\rm p}) < 1$ compared to $1 < v_{\rm p}/v_{\rm C}(r_{\rm p}) < 2$, contrary to the model values of MOND without ExFE. Realistic contamination or observational errors is expected to produce more contaminants at $1 < v_{\rm p}/v_{\rm C}(r_{\rm p}) < 2$ than $v_{\rm p}/v_{\rm C}(r_{\rm p}) < 1$, hence is unlikely to erase this discrepancy.

We also find that the number of binary candidates is sufficient *in principle* for the more challenging case of testing MOND *with* ExFE, e.g. we have a total of 1724 binary candidates with $5 < r_{\rm p} < 20\,$kAU and $v_{\rm p}/v_{\rm C}(r_{\rm p}) < \sqrt{2}$; given this, Newtonian models predict $\lesssim 95$ above ratio 1.1 while MOND-ExFE predicts $\gtrsim 200$, values which would be easily separable at high significance if simple Poisson statistics applied.

But, the presence of the high-velocity tail is currently poorly understood; the tail is much too populous to be explained by chance coincidences, but extends to high velocity ratios. We note that the presence of this tail will severely contaminate any statistic based mainly on *rms* velocity differences; thus, previous hints of excess binary *rms* velocities in the literature (e.g. Hernandez et al. 2019) may well be caused by this tail, rather than an actual modification of gravity. It therefore appears that improved modelling or understanding of the origins of this tail will be crucial in future when trying to test Newtonian gravity against MOND-ExFE type modifications; in particular, it will be necessary to quantify the contribution of non-bound flyby systems or hierarchical triples to the important velocity window $v_{\rm p}/v_{\rm C}(r_{\rm p}) \sim 1.1 - 1.4$, which may be challenging.

Hierarchical triples with an undetected third object are another possible contributor to the tail, though the tail seems too populous for hierarchical triples to contribute a majority of the tail. Most such systems can be detected in principle by either radial velocity variation over the years (for close-in third stars), or detection in high-resolution imaging (for wider third stars).

The prospects for improved data in the future are good: the anticipated GAIA DR3 should provide a factor-2 improvement in proper-motion precision, along with many more epochs to weed out anomalous non-Gaussian errors; the extended mission to 2022+ should provide another factor-2. This will also allow an expanded sample, e.g. pushing the distance limit moderately outward to $\sim 250 - 300\,$pc. Spectroscopic observations can deliver the missing radial velocity information; this improves the statistics, by increasing the predicted fraction of bound binaries in the key range $v_{\rm p}/v_{\rm C}(r_{\rm p}) \sim 1.1 - 1.4$; while most (but not all) unbound flyby systems should move to ratios above 1.5 with radial velocities included, so this will help to sharpen the discrimination.

Also, modelling of "cold streams" may be helpful: it is currently difficult to detect poor cold streams due to angle-dependent projection effects in 2D data, but if modest-precision radial velocities ($\sigma \sim 2\,{\rm km\,s^{-1}}$) become available for a good fraction of stars with $G \lesssim 16$, this would allow much cleaner selection of cold streams via matching 3D velocities; we could then reject "binaries" consistent with membership in streams. (We note that much higher RV precision $\sim 0.1\,{\rm km\,s^{-1}}$ is required to get 3D velocities for surviving wide binary candidates, but rejection of most cold streams should be possible with only $2\,{\rm km\,s^{-1}}$ RV precision).

Table 1. Number of candidate binaries in selected ranges of projected separation and velocity ratio: data, and model fits for combined Newtonian (N) and flyby (F) populations. Newtonian+flyby fits. Rows are range of projected separation, as in Column 1. Columns 2-4 are for all velocity ratios, columns 5-7 for ratios < 1.4, and columns 8-10 for ratios between 1.1 and 1.4.

	$v_{\rm p}/v_{\rm C}(r_{\rm p}) < 7$			$v_{\rm p}/v_{\rm C}(r_{\rm p}) < 1.4$			$1.1 < v_{\rm p}/v_{\rm C}(r_{\rm p}) < 1.4$		
$r_{\rm p}$ range	Data	Fit(N)	Fit(F)	Data	Fit(N)	Fit(F)	Data	Fit(N)	Fit(F)
$5-7.1\,$kAU	955	440.5	513.4	638	440.5	213.5	61	28.3	45.8
$7.1-10\,$kAU	711	352.3	390.4	474	352.3	132.5	49	22.4	30.9
$10-14.1\,$kAU	569	244.6	353.4	352	244.6	108.8	45	15.0	26.8
$14.1-20\,$kAU	434	183.6	282.5	260	183.6	77.2	41	11.0	21.8

Table 2. Same as Table 1, but for the subsample of binaries with relative velocity error ratio $\sigma(v_{\rm p})/v_{\rm C}(r_{\rm p}) < 0.25$.

	$v_{\rm p}/v_{\rm C}(r_{\rm p}) < 7$			$v_{\rm p}/v_{\rm C}(r_{\rm p}) < 1.4$			$1.1 < v_{\rm p}/v_{\rm C}(r_{\rm p}) < 1.4$		
$r_{\rm p}$ range	Data	Fit(N)	Fit(F)	Data	Fit(N)	Fit(F)	Data	Fit(N)	Fit(F)
$5-7.1\,$kAU	629	367.0	272.5	463	367.0	113.3	35	23.5	24.3
$7.1-10\,$kAU	428	265.7	183.5	320	265.7	62.2	29	16.9	14.5
$10-14.1\,$kAU	270	149.7	144.8	190	149.7	44.5	24	9.2	11.0
$14.1-20\,$kAU	134	78.2	72.2	91	78.2	19.7	10	4.7	5.6

Selecting stars by age, if possible, should also be helpful: most binaries at 10 kAU should survive for the age of the Galaxy, while common-origin unbound pairs or cold streams should disperse over Gyr timescales and thus show a strong bias to young ages. Thus, if it were possible to add an additional selection cut on age $\gtrsim 3$ Gyr, we could substantially reduce contamination from unbound co-natal flyby systems.

To summarise, we have seen that the observed population of wide binaries in GAIA DR2 already disfavours MOND *without* ExFE. For the more interesting case of MOND *with* ExFE, there are likely an adequate number of probable bound binaries to carry out the test as outlined in PS18; the current velocity precision in GDR2 is slightly marginal, but the future GAIA DR3/DR4 should be readily precise enough, that observational followup data on candidate binaries obtainable within a 5-7 year timescale has good prospects for actually supporting or refuting acceleration-based models of modified gravity similar to MOND-with-ExFE, subject to a better understanding of the origins of the high-velocity tail.

7 CONCLUSIONS

We have used the recent GAIA Data Release 2 to select a large sample of candidate wide binary stars at $d \lesssim 200\,$pc and magnitude $G < 16$, which are suitable for testing modifications of gravity at $g \lesssim a_0$. We applied various cuts to minimise contamination, leading to a cleaned sample of 24,282 candidate binaries, with 2,749 in the wide-separation range $5-20\,$kAU. After estimating masses from a main-sequence mass/luminosity relation, we derived for each candidate binary the ratio $v_{\rm p}/v_C(r_{\rm p})$. We compared this sample to various control samples with randomised positions; we find that the real sample is much more numerous, hence chance projection systems are relatively negligible. We then cut the samples into $\sqrt{2}$ bins of $r_{\rm p}$, and explored the histograms of velocity ratio for each separation bin, and made various comparisons with models for Newtonian and MOND-like gravity models, and contamination from flyby populations.

Our main conclusions are as follows:

(i) The number of candidate binaries in the real data is much larger than in the randomised control samples, by a factor ~ 100 at $5 \le r_{\rm p} \le 7.1\,$kAU, down to ~ 11 in the bin from $14-20\,$kAU. Thus, the large majority of our candidate binaries are clearly physically associated in some way, in agreement with Andrews et al (2017) and Andrews et al (2018). The frequency of "probable bound" binaries is declining beyond 5 kAU, approximately in agreement with Andrews et al (2018).

(ii) The histograms of relative velocity are well described by a quasi-Newtonian "peak" (presumably bound binaries), plus a gently declining "tail" which extends to substantially larger velocity ratios $\gtrsim 5$. The tail extends to velocity ratios larger than any reasonable modified-gravity model; we speculate that the tail is likely to result from co-natal star pairs originating in the same open cluster, which are unbound but currently undergoing a close flyby at a relative velocity $\sim 1-3\,{\rm km\,s^{-1}}$.

(iii) The existence of the high-velocity tail implies that previous hints of anomalous velocities based on *rms* statistics are likely contaminated.

(iv) MOND-like theories *without* an ExFE predict a substantial shift of the peak at large separations $\sim 10-20\,$kAU, which is not observed; these models appear to be strongly disfavoured by the current data.

(v) Precision tests of MOND-like theories *with* the ExFE are not quite practical giving the current data, but should be possible in the near future, given improved modelling and understanding of the high-velocity tail. Further data such as high-precision radial velocities, removal of cold streams and age estimates should be helpful in this direction.

(vi) The available data should improve quite rapidly in the near future: in particular, the upcoming GAIA DR3 will provide a substantial improvement in proper motion precision, allowing a larger usable distance limit and a larger sample. Future large-multiplex $R \sim 20,000$ spectrographs such as 4MOST, WEAVE and MSE should allow high-quality spectra to be obtained for most of these candidate binaries. Thus with anticipated data over the next ~ 5 years, it looks promising that wide binaries can provide an interesting and

direct observational test for possible modifications of gravity at low accelerations $\sim 10^{-10}\,\mathrm{m\,s^{-2}}$.

ACKNOWLEDGEMENTS

We thank Indranil Banik and Tim Clifton for helpful discussions. CP has been supported by an STFC studentship.

This is an author-produced, non-copy-edited version of the paper as accepted by MNRAS. The version of record is available at DOI:10.1093/mnras/stz1898 .

REFERENCES

Ade P. et al (Planck Collaboration), 2016, A&A, 594, A13. (arXiv:1502.01589)
Andrews J.J., Chanamé J., Agüeros M.A., 2017, MNRAS, 472, 675. (arXiv:1704.07829)
Andrews J.J., Chanamé J., Agüeros M.A., 2018, Res. Notes AAS, 2, 29.
Arenou F., Luri X., Babusiaux C. et al, 2018, A&A, 616, A17.
Arraut I., 2014, IJMPD, 23, 1450008.
Banik I., preprint, arXiv.org:1902.01857
Banik I., Zhao H.-S., MNRAS, 480, 2660. (BZ18).
Bekenstein J.D., 2004, Phys. Rev. D, 70, 083509. (arxiv:astro-ph/0403694v6)
Bekenstein J.D., Milgrom M., 1984, ApJ, 286, 7.
Boran S., Desai S., Kahya E.O., Woodard R.P., 2018, Phys.Rev. D, 97, 041501 (arXiv:1710.06168)
Capozziello S., D'Agostino R., Luongo O., 2019, preprint, arXiv.org:1904.01427
Coronado J., Chanamé J., 2015. VI Reunion de Astronomia Dinamica en Latinoamerica (ADeLA 2014) (Eds. K. Vieira, W. van Altena & R.A. Mendez), Revista Mexicana de Astronomia y Astrofisica (Serie de Conferencias), 46, 61.
Clifton, T., Ferreira, P. G., Padilla, A., Skordis, C., 2012, Phys. Rep., 513, 1. (arXiv:1106.2476)
Dhital S., West A.A., Stassun K.G., Bochanski J., 2010, AJ, 139, 2566.
El-Badry K., 2019, MNRAS, 482, 5018.
El-Badry K., Rix H.-W., 2018, MNRAS, 480, 4884. (ER18)
Evans D.W., Riello M., De Angeli F. et al, 2018, A&A, 616, A4.
Famaey B., McGaugh S.S., 2012, Living Rev. Relativity 15, 10. (doi:10.12942/lrr-2012-10) (arXiv:1112.3960)
GAIA Collaboration (Brown et al), 2018, A&A, 616, A1.
Hernandez X., Jimenez A., Allen C., 2011, European Phys. J. C, 72, 1884. (arXiv:1105.1873)
Hernandez X., Jimenez A., Allen C., 2012, Journal of Physics: Conf. Series, 405, 012018. (arXiv:1205.5767)
Hernandez X., Jimenez A., Allen C., 2014, in "Accelerated Cosmic Expansion", Astrophys. and Space Sci. Procs., 38, 43 (Springer). (arXiv:1401.7063)
Hernandez X., Cortes R.A.M., Allen C., Scarpa R., 2019, IJMPD, in press. (arXiv:1810.08696)
Jiang Y.-F., Tremaine S., 2010, MNRAS, 401, 977.
Kouwenhoven M.B.N., Goodwin S.P., Parker R.J., Davies M.B., Malmberg D., Kroupa P., 2010, MNRAS, 404, 1835.
Lepine S., Bongiorno B., 2007, AJ, 133, 889.
McGaugh S.S., Lelli F., Schombert J.M., 2016, Phys. Rev. Lett.117, 201101. (arXiv:1609.05917) (MLS)
Milgrom M., 1983, ApJ, 270, 365.
Matvienko A.S., Orlov V.V., 2015, Astronomy Letters, 41, 824.
Oh S., Price-Whelan A.M., Hogg D.W., Morton T.D., Spergel D., 2017, AJ, 153, 257.
Pecaut M.J., Mamajek E.E., 2013, ApJS, 208, 9.
Pittordis C., Sutherland W., 2018, MNRAS, 480, 1778.
Prusti T., de Bruijne J.H.J. et al (GAIA collaboration), 2016, A&A, 595, A1.
Quinn D.P., Wilkinson M.I., Irwin M.J. et al, 2009, MNRAS, 396, L11. (arXiv:0903.1644)
Scarpa R., Ottolino R., Falomo R., Treves A., 2017, Int. J. Mod. Phys. D, 26, 1750067 (arXiv:1611.08635)
Shaya E.J., Olling R.P., 2011, ApJS, 192, 2. (arXiv:1007.0425)
Tokovinin A., Kiyaeva O., 2016, MNRAS, 456, 2070. (arXiv:1512.00278)
Yoo J., Chaname J., Gould A., 2003. ApJ, 601, 311.