Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **490,** 157–171 (2019)
Advance Access publication 2018 September 20

doi:10.1093/mnras/sty2592

Gaia DR2 in 6D: searching for the fastest stars in the Galaxy

T. Marchetti,* E. M. Rossi and A. G. A. Brown
Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands

Accepted 2018 September 17. Received 2018 September 12; in original form 2018 April 27

ABSTRACT

We search for the fastest stars in the subset of stars with radial velocity measurements of the second data release (DR2) of the European Space Agency mission *Gaia*. Starting from the observed positions, parallaxes, proper motions, and radial velocities, we construct the distance and total velocity distribution of more than 7 million stars in our Milky Way, deriving the full 6D phase space information in Galactocentric coordinates. These information are shared in a catalogue, publicly available at http://home.strw.leidenuniv.nl/~marchetti/research.html. To search for unbound stars, we then focus on stars with a probability greater than 50 per cent of being unbound from the Milky Way. This cut results in a clean sample of 125 sources with reliable astrometric parameters and radial velocities. Of these, 20 stars have probabilities greater than 80 per cent of being unbound from the Galaxy. On this latter subsample, we perform orbit integration to characterize the stars' orbital parameter distributions. As expected given the relatively small sample size of bright stars, we find no hypervelocity star candidates, stars that are moving on orbits consistent with coming from the Galactic Centre. Instead, we find seven hyperrunaway star candidates, coming from the Galactic disc. Surprisingly, the remaining 13 unbound stars cannot be traced back to the Galaxy, including two of the fastest stars (around 700 km s^{-1}). If conformed, these may constitute the tip of the iceberg of a large extragalactic population or the extreme velocity tail of stellar streams.

Key words: Galaxy: centre – Galaxy: kinematics and dynamics – Galaxy: stellar content.

1 INTRODUCTION

Stars with extremely high velocities have been long studied to probe our Galaxy. The interest in the high-velocity tail of the total velocity distribution of stars in our Milky Way (MW) is twofold. First, it flags the presence of extreme dynamical and astrophysical processes, especially when the velocity of a star is so high that it approaches (or even exceeds) the escape speed from the Galaxy at its position. Secondly, high-velocity stars, spanning a large range of distances, can be used as dynamical tracers of integral properties of the Galaxy. The stellar high-velocity distribution has for example been used to trace the local Galactic escape speed and the mass of the MW (e.g. Smith et al. 2007; Gnedin et al. 2010; Piffl et al. 2014). To put the concept of *high velocity* in context, the value of the escape speed is found to be ∼530 km s^{-1} at the Sun position, it increases up to ∼600 km s^{-1} in the central regions of the Galaxy, and then falls down to $\lesssim$400 km s^{-1} at Galactocentric distances ∼50 kpc (Williams et al. 2017).

A first class of objects that can be found in the high tail of the total velocity distribution is fast halo stars. Their measured dispersion velocity is around 150 km s^{-1} (Smith et al. 2009; Evans et al. 2016), therefore 3 σ outliers can exceed 450 km s^{-1}, while remaining bound. Halo stars could also reach unbound velocities, when they are part of the debris of tidally disrupted satellite galaxies, like the Sagittarius Dwarf galaxy, that has not yet virialized (e.g. Abadi, Navarro & Steinmetz 2009). Velocities outliers in the bulge and disc velocity distribution may also exist and become apparent in a large data set.

'Runaway stars' (RSs) form an another class of high-velocity stars. They were originally introduced as O and B type stars ejected from the Galactic disc with velocities higher than 40 km s^{-1} (Blaauw 1961). Theoretically, there are two main formation channels: (i) dynamical encounters between stars in dense stellar systems such as young star clusters (e.g. Poveda, Ruiz & Allen 1967; Leonard & Duncan 1990; Gvaramadze, Gualandris & Portegies Zwart 2009), and (ii) supernova explosions in stellar binary systems (e.g. Blaauw 1961; Portegies Zwart 2000). Both mechanisms have been shown to occur in our Galaxy (Hoogerwerf, de Bruijne & de Zeeuw 2001). Typical velocities attained by the two formation channels are of the order of a few tens of km s^{-1}, and even if several hundreds of km s^{-1} can be attained for the most extreme systems (Portegies Zwart 2000; Przybilla et al. 2008; Gvaramadze et al. 2009; Gvaramadze & Gualandris 2011; Silva & Napiwotzki 2011), simulations indicate that the majority of RSs from dynamical encounters have ejection velocities $\lesssim$200 km s^{-1} (Perets & Šubr 2012). Recent results show that it is possible to achieve ejection velocities up to ∼1300 km s^{-1} for low-mass G/K type stars in very compact binaries (Tauris 2015).

---

* E-mail: marchetti@mail.strw.leidenuniv.nl

Published by Oxford University Press on behalf of the Royal Astronomical Society

Nevertheless, the rate of production of unbound RSs, referred to as *hyper runaway stars* (HRSs), is estimated to be as low as 8×10^{-7} yr^{-1} (Perets & Šubr 2012; Brown 2015).

As a class, the fastest stars in our Galaxy are expected to be hypervelocity stars (HVSs). These were first theoretically predicted by Hills (1988) as the result of a three-body interaction between a binary star and the massive black hole in the Galactic Centre (GC), Sagittarius A∗. Following this close encounter, a star can be ejected with a velocity $\sim$1000 km s^{-1}, sufficiently high to escape from the gravitational field of the MW (Kenyon et al. 2008; Brown 2015). The first HVS candidate was discovered by Brown et al. (2005); a B-type star with a velocity more than twice the Galactic escape speed at its position. Currently about $\sim$20 unbound HVSs with velocities $\sim$300–700 km s^{-1} have been discovered by targeting young stars in the outer halo of the MW (Brown, Geller & Kenyon 2014). In addition, tens of mostly bound candidates have been found at smaller distances but uncertainties prevent the precise identification of the GC as their ejection location (e.g. Hawkins et al. 2015; Vickers, Smith & Grebel 2015; Zhang, Smith & Carlin 2016; Marchetti et al. 2017; Ziegerer et al. 2017). HVSs are predicted to be ejected from the GC with an uncertain rate around 10^{-4} yr^{-1} (Yu & Tremaine 2003; Zhang, Lu & Yu 2013), two orders of magnitude larger than the rate of ejection of RSs with comparable velocities from the stellar disc (Brown 2015). Because of their extremely high velocities, HVS trajectories span a large range of distances, from the GC to the outer halo. Thus, HVSs have been proposed as tools to study the matter distribution in our Galaxy (e.g. Gnedin et al. 2005; Sesana, Haardt & Madau 2007; Kenyon et al. 2014; Fragione & Loeb 2017; Rossi et al. 2017; Contigiani, Rossi & Marchetti 2018) and the GC environment (e.g. Zhang et al. 2013; Madigan et al. 2014), but a larger and less observationally biased sample is needed in order to break degeneracies between the GC binary content and the Galactic potential parameters (Rossi et al. 2017). Using the fact that their angular momentum should be very close to zero, HVSs have also been proposed as tools to constrain the solar position and velocity (Hattori, Valluri & Castro 2018a). Other possible alternative mechanisms leading to the acceleration of HVSs are the encounter between a single star and a massive black hole binary in the GC (e.g. Yu & Tremaine 2003; Sesana, Haardt & Madau 2006, 2008), the interaction between a globular cluster with a single or a binary massive black hole in the GC (Capuzzo-Dolcetta & Fragione 2015; Fragione & Capuzzo-Dolcetta 2016), and the tidal interaction of a dwarf galaxy near the centre of the Galaxy (Abadi et al. 2009). Another possible ejection origin for HVSs and high-velocity stars in our Galaxy is the Large Magellanic Cloud (LMC; Boubert & Evans 2016; Boubert et al. 2017; Erkal et al. 2018), orbiting the MW with a velocity $\sim$380 km s^{-1} (van der Marel & Kallivayalil 2014).

In addition to the unbound population of HVSs, all the ejection mechanisms mentioned above predict also a population of *bound* HVSs (BHVSs): stars sharing the same formation scenario as HVSs, but with an ejection velocity that is not sufficiently high to escape from the whole MW (e.g. Bromley et al. 2006). Most of the deceleration occurs in the inner few kpc due to the bulge potential (Kenyon et al. 2008), and the minimum velocity necessary at ejection to be unbound is of the order of $\sim$800 km s^{-1} (a precise value depends on the choice of the Galactic potential; Brown 2015; Rossi et al. 2017). If we consider the Hills mechanism, this population of bound stars is expected to be dominant over the sample of HVSs (Rossi, Kobayashi & Sari 2014; Marchetti et al. 2018).

At the moment, the fastest star discovered in our Galaxy is US 708, travelling away from the MW with a total velocity $\sim$1200 km s^{-1} (Hirsch et al. 2005). Its orbit is not consistent with coming from the GC (Brown et al. 2015), and the most likely mechanism responsible for its acceleration is the explosion of a thermonuclear supernova in an ultracompact binary in the Galactic disc (Geier et al. 2015).

The second data release (DR2) of the European Space Agency satellite *Gaia* (Gaia Collaboration et al. 2016, 2018a) gives us the first opportunity to look for extremely high-velocity stars in our MW, using an unprecedented sample of precisely and accurately measured sources. On 2018 April 25, *Gaia* provided positions (α, δ), parallaxes ϖ and proper motions ($\mu_{\alpha*}$, μ_δ) for more than 1.3 billion of stars, and, notably, radial velocities $v_{\rm rad}$ for a subset of 7224 631 stars brighter than the 12th magnitude in the *Gaia* Radial Velocity Spectrograph (RVS) passband (Cropper et al. 2018; Katz et al. 2018). Radial velocities are included in the *Gaia* catalogue for stars with an effective temperature $T_{\rm eff}$ from 3550 to 6990 K, and have typical uncertainties of the order of few hundreds of m s^{-1} at the bright end of the magnitude distribution (*Gaia* *G*-band magnitude $\approx$4), and of a few km s^{-1} at the faint end ($G \approx 13$).

Using *Gaia* DR2 data, Boubert et al. (2018) show that almost all the previously discovered late-type HVS candidates are most likely bound to the Galaxy, and their total velocity was previously overestimated because of inaccurate parallaxes and/or proper motions. Only one late-type star, LAMOST J115209.12+120258.0 (Li et al. 2015), is most likely unbound, but the Hills mechanisms is ruled out as a possible explanation of its extremely high velocity. The majority of B-type HVSs from Brown et al. (2014, 2015) are still found to be consistent with coming from the GC when using *Gaia* DR2 proper motions (Erkal et al. 2018).

In this paper, we search for the fastest stars in the MW, within the sample of $\sim$7 million stars with a six-dimensional phase space measurement in *Gaia* DR2. Since the origin of high-velocity stars in our Galaxy is still a puzzling open question, we simply construct the total velocity distribution in the Galactic rest frame in order to identify and characterize the high-velocity tail. In doing so, we do not bias our search towards any specific class of high-velocity stars.

This manuscript is organized as follows. In Section 2, we explain how we determine distances and total velocities in the Galactic rest frame for the whole sample of stars. We present results in terms of stellar total velocity in Section 3. In Section 4, we focus on the high-velocity stars in the sample, and then in Section 5 we concentrate on the stars with a probability greater than 80 per cent of being unbound from the Galaxy, discussing individually the most interesting candidates. Finally, we conclude and discuss our results and findings in Section 6.

2 DISTANCE AND TOTAL VELOCITY DETERMINATION

The *Gaia* catalogue provides parallaxes, and thus a conversion to a distance is required to convert the apparent motion of an object on the celestial sphere to a physical motion in space, which is needed to determine the total velocity of a star. Bailer-Jones (2015) discusses in details how this operation is not trivial when the relative error in parallax, $f \equiv \sigma_\varpi/\varpi$, is either above 20 per cent or it is negative. We choose to separate the discussion on how we determine distances and total velocities of stars with $0 < f \leq 0.1$ (the '*low-f* sample') and of those with either $f > 0.1$ or $f < 0$ (the '*high-f* sample'). There are 7183 262 stars with both radial velocity and the astrometric parameters (parallax and proper motions) in *Gaia* DR2, therefore in the following we will focus on this subsample of stars.

2.1 The '*low-f* sample'

5 393 495 out of 7 183 262 stars (∼75 per cent) with radial velocity measurement in *Gaia* DR2 have a relative error in parallax $0 < f \leq 0.1$. For this majority of stars, we can get an accurate determination of their distance just by inverting the parallax: $d = 1/\varpi$ (Bailer-Jones 2015). We then model the proper motions and parallax distribution as a multivariate Gaussian with mean vector:

$$\boldsymbol{m} = [\mu_{\alpha *}, \mu_\delta, \varpi] \tag{1}$$

and with covariance matrix:

$$\Sigma = \begin{pmatrix} \sigma^2_{\mu_{\alpha *}} & \sigma_{\mu_{\alpha *}}\sigma_{\mu_\delta}\rho(\mu_{\alpha *}, \mu_\delta) & \sigma_{\mu_{\alpha *}}\sigma_\varpi \rho(\mu_{\alpha *}, \varpi) \\ \sigma_{\mu_{\alpha *}}\sigma_{\mu_\delta}\rho(\mu_{\alpha *}, \mu_\delta) & \sigma^2_{\mu_\delta} & \sigma_{\mu_\delta}\sigma_\varpi \rho(\mu_\delta, \varpi) \\ \sigma_{\mu_{\alpha *}}\sigma_\varpi \rho(\mu_{\alpha *}, \varpi) & \sigma_{\mu_\delta}\sigma_\varpi \rho(\mu_\delta, \mu_\varpi) & \sigma^2_\varpi \end{pmatrix}, \tag{2}$$

where $\rho(i, j)$ denotes the correlation coefficient between the astrometric parameters i and j, and it is provided in the *Gaia* DR2 catalogue. Radial velocities are uncorrelated to the astrometric parameters, and we assume them to follow a Gaussian distribution centered on $v_{\rm rad}$, and with standard deviation $\sigma_{v_{\rm rad}}$. We then draw 1000 Monte Carlo (MC) realizations of each star's observed astrometric parameters, and we simply compute distances by inverting parallaxes.

Total velocities in the Galactic rest frame are computed correcting radial velocities and proper motions for the solar and the local standard of rest (LSR) motion (Schönrich 2012). In doing so, we assume that the distance between the Sun and the GC is $d_\odot = 8.2$ kpc, and that the Sun has an height above the stellar disc of $z_\odot = 25$ pc (Bland-Hawthorn & Gerhard 2016). We assume a rotation velocity at the Sun position $v_{\rm LSR} = 238\ {\rm km\,s^{-1}}$ and a Sun's peculiar velocity vector $\boldsymbol{v}_\odot = [U_\odot, V_\odot, W_\odot] = [14.0, 12.24, 7.25]$ ${\rm km\,s^{-1}}$(Schönrich, Binney & Dehnen 2010; Schönrich 2012; Bland-Hawthorn & Gerhard 2016). To save computational time, we do not sample within the uncertainties of the solar position and motion. We verify that this does not considerably affect our results. We then derive Galactic rectangular velocities (U, V, W) adopting the following convention: U is positive when pointing in the direction of the GC, V is positive along the direction of the Sun rotation around the Galaxy, and W is positive when pointing towards the North Galactic Pole (Johnson & Soderblom 1987). Starting from the MC samples on proper motions, distances, and radial velocities, we then compute total velocities in the Galactic rest frame $v_{\rm GC} = v_{\rm GC}(\alpha, \delta, \mu_{\alpha *}, \mu_\delta, d, v_{\rm rad})$ summing in quadrature the three velocity components (U, V, W).

Finally, for each star we estimate the probability $P_{\rm ub}$ of being unbound from the Galaxy as the fraction of MC realizations that result in a total velocity $v_{\rm GC}$ greater than the escape speed from the MW at that given position. We compute the escape velocity from the Galaxy at each position using the Galactic potential model introduced and discussed in Section 4.1.

2.2 The '*high-f* sample'

A more careful analysis is required for 1789 767 stars (∼25 per cent) with either $f > 0.1$ or with a negative measured parallax. For these stars, we follow the approach outlined in Bailer-Jones (2015), Astraatmadja & Bailer-Jones (2016a,b), Luri et al. (2018), and Bailer-Jones et al. (2018). We use a full Bayesian analysis to determine the posterior probability $P(d|\varpi, \sigma_\varpi)$ of observing a star at a distance d, given the measured parallax ϖ and its Gaussian uncertainty σ_ϖ. The authors show how the choice of the prior probability on distance $P(d)$ can seriously affect the shape of the posterior distribution, and therefore lead to significantly different values for the total velocity of a star. We decide to adopt an *exponentially decreasing prior*:

$$P(d) \propto d^2 \exp\left(-\frac{d}{L}\right), \tag{3}$$

which has been shown to perform best for stars further out than ∼2 kpc (Astraatmadja & Bailer-Jones 2016b), which is the expected distance of stars with a large relative error on parallax (see Appendix A). The value of the scale length parameter L is fixed to 2600 pc, and we refer the reader to the discussion in Appendix A for the reasons behind our choice of this particular value. By means of Bayes' theorem, we can then express the posterior distribution on distances as

$$P(d|\varpi, \sigma_\varpi) \propto P(\varpi|d, \sigma_\varpi)P(d), \tag{4}$$

where the likelihood probability $P(\varpi|d, \sigma_\varpi)$ is a Gaussian distribution centered on $1/d$:

$$P(\varpi|d, \sigma_\varpi) \propto \exp\left[-\frac{1}{2\sigma^2_\varpi}\left(\varpi - \frac{1}{d}\right)\right]. \tag{5}$$

In our case, we decide to fully include the covariance matrix between the astrometric properties, following the approach introduced in Marchetti et al. (2017). In this case, for each star the likelihood probability is a three-dimensional multivariate Gaussian distribution with mean vector:

$$\boldsymbol{m} = [\mu_{\alpha *}, \mu_\delta, 1/d] \tag{6}$$

and covariance matrix given by equation (2). The prior distribution on distance is given by equation (3), and we assume uniform priors on proper motions. We then draw proper motions and distances from the resulting posterior distribution using the affine invariant ensemble Markov chain Monte Carlo (MCMC) sampler EMCEE (Goodman & Weare 2010; Foreman-Mackey et al. 2013). We run each chain using 32 walkers and 100 steps, for a total of 3200 random samples drawn from the posterior distribution. We initialize the walkers to random positions around the mean value of the proper motions and of the inverse of the mode of the posterior distribution in distance, equation (4), to achieve a fast convergence of the chain. We run 500 burn-in steps to let the walkers explore the parameter space, and then we use the final positions as initial conditions for the proper MC chain. We then directly use this MC sampling to derive a distribution for the total velocity in the Galactic rest frame of each star, assuming the same parameters for the Sun presented in Section 2.1. We check that the mean acceptance fraction (i.e. the fraction of steps accepted for each walker) is between 0.25 and 0.5 as a test for the convergence of each MC chain (Foreman-Mackey et al. 2013).

3 THE TOTAL VELOCITY DISTRIBUTION OF STARS IN *GAIA* DR2

Using the approach discussed in Section 2, we publish a catalogue with distances and velocities in the Galactocentric frame for all the 7183 262 stars analyzed in this paper. This is publicly available at http://home.strw.leidenuniv.nl/~marchetti/research.html. A full description of the catalogue content can be found in Appendix B.

In order to filter out the more uncertain candidates, for which it would be difficult to constrain the origin, we will now only discuss and plot results for stars with a relative error on total velocity $\sigma_{v_{\rm GC}}/v_{\rm GC} < 0.3$, where $\sigma_{v_{\rm GC}}$ is estimated summing in quadrature the lower and upper uncertainty on $v_{\rm GC}$. This cut results into a total of 6884 304 stars, ∼96 per cent of the original sample of stars.

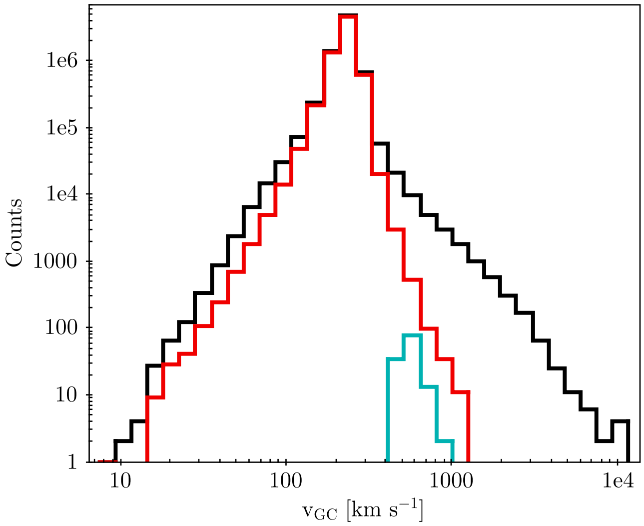

Figure 1. Histogram of median total velocities in the Galactic rest frame for all the ∼7 million stars with three-dimensional velocity by *Gaia* DR2 (black). The red line corresponds to those stars with a relative error on total velocity in the Galactic rest frame below 30 per cent, while the cyan line refers to our 'clean' sample of high-velocity stars (see discussion in Section 4).

Fig. 1 shows the total velocity distribution of the median Galactic rest-frame total velocity $v_{\rm GC}$ for the original sample of 7183 262 stars (black line) and for the stars with a relative error on total velocity below 30 per cent (red line). We can see how this cut filters out most of the stars with extremely high velocities, which are likely to be outliers with relatively more uncertain measurements by *Gaia*. Nevertheless, we note the presence of a high-velocity tail extending up to and above ∼1000 km s^{-1} surviving the cut. We will now focus only on stars with $\sigma_{v_{\rm GC}}/v_{\rm GC} < 0.3$.

To highlight visually possibly unbound objects, we plot in Fig. 2 the total velocity for all stars as a function of the Galactocentric distance $r_{\rm GC}$, and we overplot the median escape speed from the Galaxy with a green solid line, computed using the Galactic potential model introduced in Section 4.1. Datapoints correspond to the medians of the distributions, with lower and upper uncertainties derived, respectively, from the 16th and 84th percentiles. Most of the stars are located in the solar neighbourhood, and have typical velocities of the order of the LSR velocity. We find 510 stars to have probabilities greater than 50 per cent of being unbound from the Galaxy (but note the large error bars). In particular, 212 (103) stars are more than 1 σ (3 σ) away from the Galactic escape speed.

Fig. 3 shows the Toomre diagram for all the ∼7 million stars, a plot that is useful to distinguish stellar populations based on their kinematics. On the x-axis, we plot the component V of the Galactocentric Cartesian velocity, and on the y-axis the component orthogonal to it, $\sqrt{U^2+W^2}$. Not surprisingly, most of the stars behave kinematically as disc stars on rotation-supported orbits, with V values around the Sun's orbital velocity (see Gaia Collaboration et al. 2018b). A subdominant, more diffuse, population of stars with halo-like kinematics is also present, centered around $V = 0$ and with a larger spread in total velocity.

4 HIGH-VELOCITY STARS IN *Gaia* DR2

We now focus our interest towards high-velocity stars, which we define as stars with a probability $P_{\rm ub} > 0.5$. Since we are interested in kinematic outliers, we have to pay particular attention not to be contaminated by data processing artefacts and/or spurious measurements. We therefore choose to adopt the following conservative cuts on the columns of the *Gaia* DR2 GAIA_SOURCE catalogue (in addition to the selection $\sigma_{v_{\rm GC}}/v_{\rm GC} < 0.3$ introduced in Section 3):

(i) ASTROMETRIC_GOF_AL < 3;
(ii) ASTROMETRIC_EXCESS_NOISE_SIG ≤ 2;
(iii) $-0.23 \leq$ MEAN_VARPI_FACTOR_AL ≤ 0.32;
(iv) VISIBILITY_PERIODS_USED > 8;
(v) RV_NB_TRANSITS > 5.

The first cut ensures that statistic astrometric model resulted in a good fit to the data, while the second cut selects only astrometrically well-behaved sources (refer to Lindegren et al. 2012 for a detailed explanation of the excess noise and its significance). The third and the fourth cuts are useful to exclude stars with parallaxes more vulnerable to errors. Finally, the final selection ensures that each source was observed a reasonable number of times (5) by *Gaia* to determine its radial velocity. Further details on the parameters used to filter out possible contaminants and the reasons behind the adopted threshold values can be found in the *Gaia* data model.[1] Applying these cuts and with the further constrain on the unbound probability $P_{\rm ub} > 0.5$, we are left with a clean final sample of 125 high-velocity stars. We also verify that the quality cuts C.1 and C.2 introduced in appendix C of Lindegren et al. (2018a), designed to select astrometrically clean subsets of objects, are already verified by our sample of high-velocity stars. In addition, selection N in Appendix C of Lindegren et al. (2018a) does not select any of our candidates. Looking at Fig. 2, where this clean sample of 125 stars is highlighted with blue squares, we can see how these cuts filter out most of the stars with exceptionally high velocities, which are therefore likely to be instrumental artefacts. This is also evident in Fig. 1, where the Galactic rest-frame total velocity distribution of the 125 high-velocity stars is shown with a cyan line.

We present distances, total velocities, and probability of being unbound for all the 105 stars with $0.5 < P_{\rm ub} \leq 0.8$ in Appendix C, Table C1. Stars with $P_{\rm ub} > 0.8$ are presented and discussed in detail in Section 5.

The spatial distribution of these 125 high-velocity stars in our Galaxy is shown in Fig. 4, where we overplot the position on the Galactic plane of this subset of stars with blue markers above the underlying distribution of the ∼7 million stars used in this paper. We can see how the majority of high-velocity stars lies in the inner region of the Galaxy, with typical distances $\lesssim$15 kpc from the GC. Most of these stars are on the faint end of the magnitude distribution because of extinction due to dust in the direction of the GC, and thus they have large relative errors on parallax. This in turn translates into larger uncertainties on total velocity, which may cause the stars to be included into our high-velocity cut. Another small overdensity corresponds to the Sun's position, correlating with the underlying distribution of all the stars. In Fig. 5, we plot the same but in the $(x_{\rm GC}, z_{\rm GC})$ plane. Most of our high-velocity stars lie away from the stellar disc.

Fig. 6 shows the Hertzsprung–Russell (HR) diagram for all the sources with a radial velocity measurement, with the high-velocity star sample overplotted in blue. On the x-axis, we plot the colour index in the *Gaia* Blue Pass (BP) and Red Pass (RP) bands $G_{\rm BP}-G_{\rm RP}$, while on the y-axis we plot the absolute magnitude in the *Gaia* G band M_G, computed assuming the median of the posterior distance distribution. Note that we did not consider extinction to

---

[1] https://gea.esac.esa.int/archive/documentation/GDR2/Gaia_archive/chap_datamodel/

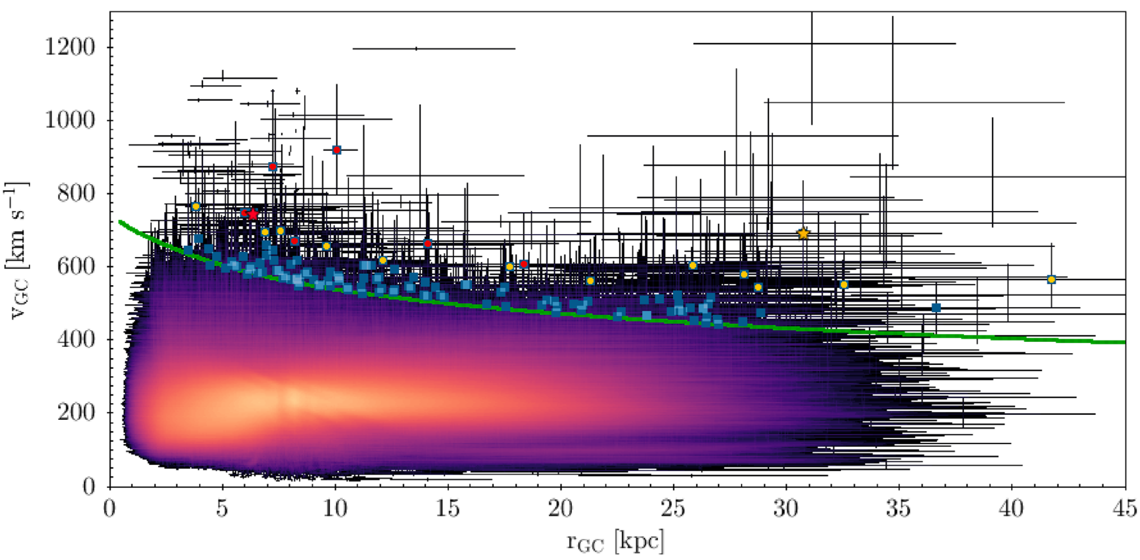

Figure 2. Total velocity in the Galactic rest-frame $v_{\rm GC}$ as a function of Galactocentric distance $r_{\rm GC}$ for all the 6884 304 stars in *Gaia* DR2 with relative error on total velocity < 0.3. Colour is proportional to the logarithmic number density of stars. The green solid line is the median posterior escape speed from the adopted Galactic potential (Section 4.1). We overplot in blue the 'clean' high-velocity star sample introduced in Section 4. Red and yellow points correspond, respectively, to the Galactic and extragalactic candidates discussed in Section 5. *Gaia* DR2 5932173855446728064 (*Gaia* DR2 1396963577886583296) is marked with a red (yellow) star.

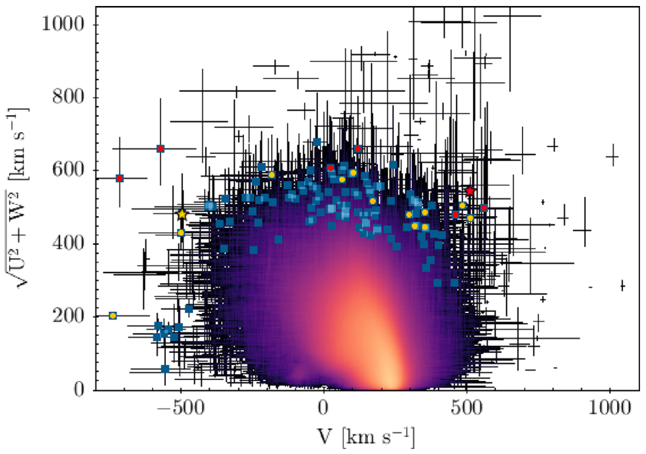

Figure 3. Toomre diagram for the same stars plotted in Fig. 2.

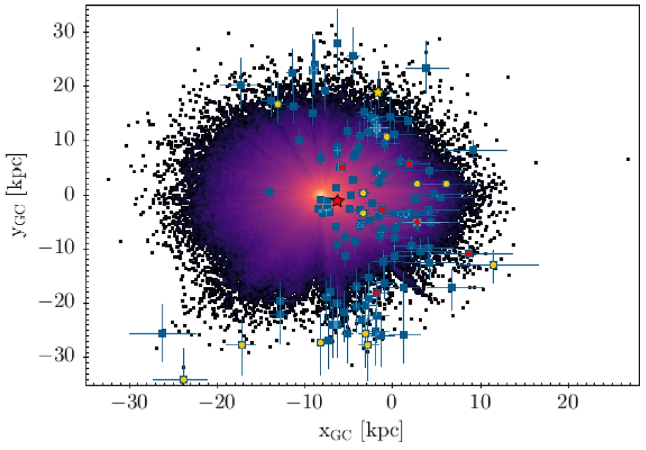

Figure 4. Distribution of the ~7 million stars on the Galactic plane. The Sun is located at ($x_{\rm GC}$, $y_{\rm GC}$) = (−8.2, 0) kpc. Colours are the same as in Fig. 2.

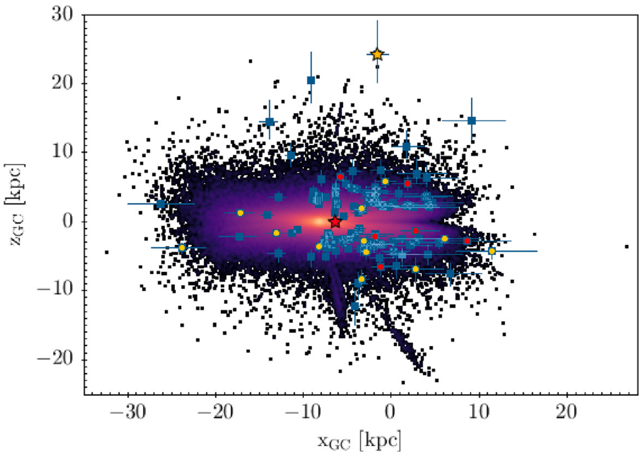

Figure 5. Same as Fig. 4, but showing the distribution of the stars in the ($x_{\rm GC}, z_{\rm GC}$) plane. The Sun is located at ($x_{\rm GC}, z_{\rm GC}$) = (−8200, 25) pc. Colours are the same as in Fig. 2.

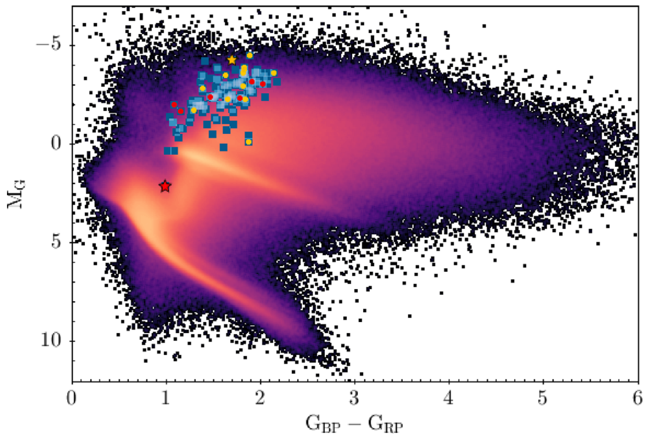

Figure 6. HR diagram for all the ~7 million stars in *Gaia* DR2 with a radial velocity measurement. Colours are the same as in Fig. 2.

Table 1. Parameters for the GALA potential *MilkyWayPotential*.

Component	Parameters
Bulge	$M_\mathrm{b} = 5.00 \times 10^9\ \mathrm{M}_\odot$ $r_\mathrm{b} = 1.00$ kpc
Nucleus	$M_\mathrm{n} = 1.71 \times 10^9\ \mathrm{M}_\odot$ $r_\mathrm{n} = 0.07$ kpc
Disc	$M_\mathrm{d} = 6.80 \times 10^{10}\ \mathrm{M}_\odot$ $a_\mathrm{d} = 3.00$ kpc $b_\mathrm{d} = 0.28$ kpc
Halo	$M_\mathrm{h} = 5.40 \times 10^{11}\ \mathrm{M}_\odot$ $r_\mathrm{s} = 15.62$ kpc

construct the HR diagram, because of the caveats with using the line-of-sight extinction in the G band A_G for individual sources (Andrae et al. 2018). We can see that the great majority of our stars are giants stars. This is consistent with recent findings of Hattori et al. (2018b) and Hawkins & Wyse (2018), which confirm some of these candidates as being old (>1 Gyr), metal-poor giants ($2 \leq$ [Fe/H] ≤ 1).

4.1 Orbital integration

In order to get hints on the ejection location of our sample of high-velocity stars, we perform numerical orbit integration of their trajectories back in time using the PYTHON package GALA (Price-Whelan 2017). For each star, we use 1000 random samples from the proper motions, distance, and radial velocity MC sampling discussed in Section 2. We integrate each orbit back in time for a total time of 1 Gyr, with a fixed time-step of 0.1 Myr, using the GALA potential *MilkyWayPotential*. This is a four components Galactic potential model consisting of a Hernquist bulge and nucleus (Hernquist 1990):

$$\phi_b(r_\mathrm{GC}) = -\frac{GM_i}{r_\mathrm{GC} + r_i}, \tag{7}$$

where $i = b$, n for the bulge and the nucleus, respectively, a Miyamoto–Nagai disc (Miyamoto & Nagai 1975):

$$\phi_\mathrm{d}(R_\mathrm{GC}, z_\mathrm{GC}) = -\frac{GM_\mathrm{d}}{\sqrt{R^2_\mathrm{GC} + \left(a_\mathrm{d} + \sqrt{z^2_\mathrm{GC} + b^2_\mathrm{d}}\right)^2}}, \tag{8}$$

and a Navarro–Frenk–White halo (Navarro, Frenk & White 1996):

$$\phi_\mathrm{h}(r_\mathrm{GC}) = -\frac{GM_\mathrm{h}}{r_\mathrm{GC}} \ln\left(1 + \frac{r_\mathrm{GC}}{r_\mathrm{s}}\right). \tag{9}$$

The parameters are chosen to fit the enclosed mass profile of the MW (Bovy 2015) and are summarized in Table 1. We then derive the pericentre distance and, for bound MC realizations, the apocentre distance and the eccentricity of the orbit. We also record the energy and the angular momentum of each MC orbit. We check for energy conservation as a test of the accuracy of the numerical integration.

In Fig. 7, we plot the maximum height above the Galactic disc Z_max as a function of the eccentricity of the orbit for our sample of high-velocity stars. This plot is useful to identify similar stars based on their orbits (e.g. Boeche et al. 2013; Hawkins et al. 2015). The dashed red line at $Z_\mathrm{max} = 3$ kpc denotes the typical scale height of the thick disc (Carollo et al. 2010). Not surprisingly, high-velocity stars are on highly eccentric orbits, with a mean eccentricity of the sample ∼0.8. Most of these stars span a large range of Z_max, with values up to hundreds of kpc, reflecting the large amplitude of the vertical oscillations.

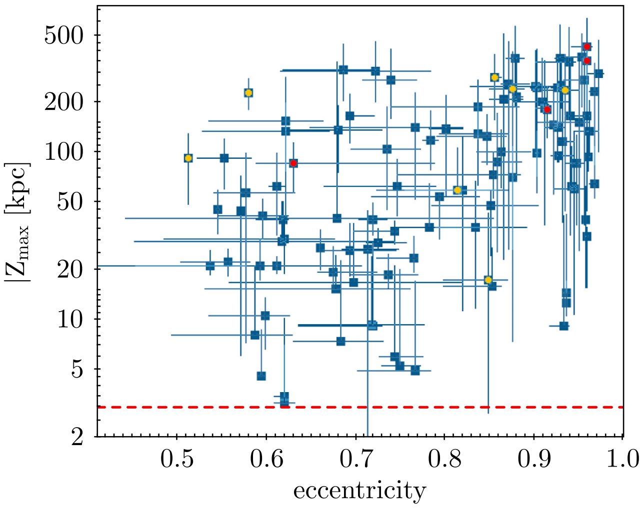

Figure 7. Absolute value of the maximum height above the Galactic plane $|Z_\mathrm{max}|$ as a function of eccentricity for the high-velocity sample of stars. The yellow horizontal dashed line corresponds to $Z_\mathrm{max} = 3$ kpc, the edge of the thick disc (Carollo et al. 2010). Colours are the same as in Fig. 2.

In our search for HVSs, we keep track of each disc crossing (Cartesian Galactocentric coordinate $z_\mathrm{GC} = 0$) in the orbital trace-back of our high-velocity star sample. For each MC realization, we then define the crossing radius r_c as

$$r_\mathrm{c} = \sqrt{x_\mathrm{c}^2 + y_\mathrm{c}^2}, \tag{10}$$

where x_c and y_c are the Galactocentric coordinates of the orbit (x_GC, y_GC) at the instant when $z_\mathrm{GC} = 0$. In the case of multiple disc crossings during the orbital trace-back, we define r_min as the minimum crossing radius attained in that particular MC realization of the star's orbit. This approach allows us to check for the consistency of the GC origin hypothesis for our sample of high-velocity stars. We also record the ejection velocity v_ej: the velocity of the star at the minimum crossing radius, and the flight time t_f: the time needed to travel from the observed position to the disc crossing happening closest to the GC.

In Fig. 8, we plot r_min as a function of the orbital energy E. The red dashed line coincides with the separation region between bound and unbound orbits. The majority of candidates are travelling on unbound orbits ($E > 0$), and we can see a few stars with remarkably high values of the energy: 25 stars are unbound at more than $1\,\sigma$ significance, and 1 star (*Gaia* DR2 5932173855446728064) is unbound at more than $3\,\sigma$ significance.

5 UNBOUND STARS: HVS AND HRS CANDIDATES

We now focus our search on possible unbound stars, defined as the subsample of clean high-velocity stars with $P_\mathrm{ub} > 80$ per cent. This amounts to a total of 20 objects. Observed properties from *Gaia* DR2, distances, and total velocities for these stars are summarized in Table 2. Fig. 9 shows the position in Galactocentric cylindrical coordinates of these high-velocity star candidates. The length of the arrows is proportional to the total velocity of each star in the Galactic rest frame. We note that for most of our candidates (18 out of 20 stars), the parallax uncertainty is smaller than the quoted parallax zero-point of −0.029 mas, as estimated by *Gaia* 's observations of quasars (Lindegren et al. 2018a). We discuss the

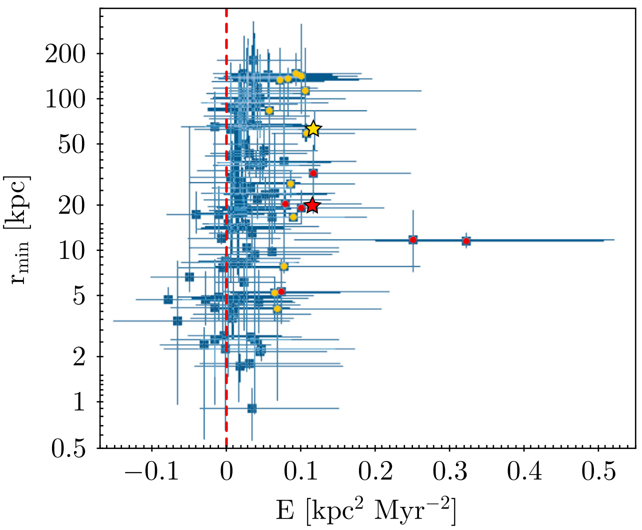

Figure 8. Minimum crossing radius $r_{\rm min}$ versus energy E for the 125 high-velocity stars. The vertical dashed line separates unbound ($E > 0$) from bound ($E < 0$) orbits. Colours are the same as in Fig. 2.

impact of considering this negative offset in the analysis of our stars in Appendix D. We further discuss the impact of systematic errors for our sample of 20 unbound candidates in Appendix E.

If a star on an unbound orbit was ejected either from the stellar disc (HRS) or from the GC (HVS), then its distribution of minimum crossing radii $r_{\rm min}$ should fall within the edge of the MW disc. To maximize the probability of a disc crossing during the orbital traceback, we integrate the orbits of these stars for a maximum time of 5 Gyr. We then define the probability $P_{\rm MW}$ for a star to come from the MW as the fraction of MC realizations resulting in a minimum crossing radius within the edge of the stellar disc: $r_{\rm min} < r_{\rm disc}$, where $r_{\rm disc} = 25$ kpc (Xu et al. 2015). This probability is useful to flag candidates of possible extragalactic origin, which we define as those stars with $P_{\rm MW} < 0.5$. This subset of 13 stars, if their high velocity is confirmed, could either originate as RS/HRS/HVS from the LMC (Boubert & Evans 2016; Boubert et al. 2017; Erkal et al. 2018), or could be the result of the tidal disruption of a dwarf galaxy interacting with the MW (Abadi et al. 2009). Stars with a Galactic and extragalactic origin are marked in Fig. 9 with red and yellow points, respectively. Stars with a Galactic origin have trajectories pointing away from the stellar disc. On the other hand, extragalactic stars are pointing either towards the disc, or are consistent with coming from regions of no current active star formation (i.e. the outer halo).

5.1 Galactic stars

7 of the 20 possible unbound stars have $P_{\rm MW} > 0.5$, and therefore are consistent with being ejected from the stellar disc of the MW. These stars, given their extremely high velocities, could be either HVS or HRS candidates.

We then classify a star as an HVS (HRS) candidate if we cannot (can) exclude the hypothesis of GC origin, which we define by the condition $r_{\rm min} - \sigma_{r_{\rm min,l}} < 1$ kpc ($r_{\rm min} - \sigma_{r_{\rm min,l}} > 1$ kpc), where $r_{\rm min}$ denotes the median of the distribution, and $\sigma_{r_{\rm min,l}}$ is the lower uncertainty on the minimum crossing radius. In this way, we are testing whether, within its error bars, a star is consistent with coming from the central region of the Galaxy. Fig. 10 shows the histogram of the median minimum disc crossing $r_{\rm min}$ minus the lower uncertainty $\sigma_{r_{\rm min,l}}$ for all the 20 stars with $P_{\rm ub} > 0.8$. A vertical red dashed line corresponds to the value 1 kpc, which we use to define HVS candidates.

We find that all of these seven stars have orbits that, when integrated back in time, are not consistent with coming from the GC. Therefore, according to our classification criterion, there are no stars classified as HVS candidates. The absence of HVS candidates in the subset of *Gaia* DR2 with radial velocities was anticipated by predictions by Marchetti et al. (2018), analyzing the Hills mock catalogue of HVSs. This is due to the fact that the expected number density of HVSs generated via the Hills' mechanism is expected to increase linearly with increasing Galactocentric distance (Brown 2015), and the majority of HVSs in the MW are too faint to have a radial velocity measurement from *Gaia* DR2. We cannot exclude the presence of *bound* HVSs in the subset of ∼7 million stars considered in this work, but their identification is not trivial because of their complex orbits and lower velocities. About 20 BHVSs are expected to have radial velocities from *Gaia* DR2 (Marchetti et al. 2018), but their identification is beyond the scope of this manuscript.

All the seven Galactic stars are therefore HRS candidates (red circles in Fig. 2 and following plots). One particular HRS candidate that is worth mentioning is *Gaia* DR2 5932173855446728064 (marked with a red star in Fig. 2 and following). This star has an exceptionally well-constrained total velocity,[2] $v_{\rm GC} = 747^{+2}_{-3}\,{\rm km\,s^{-1}}$, which results in a probability of being unbound ≈1. This star most likely was ejected in the thin disc of the MW.

We note that five of the seven HRS candidates with a Galactic origin have $P_{\rm ub} > 90$ per cent. Such exceptionally high velocities are thought to be very uncommon in our Galaxy for HRSs, which are predicted to be much rarer than HVSs (Brown 2015). This is correct in the context of the MW as a whole. In this study, we only focus on bright sources ($G_{\rm RVS} < 12$), therefore we maximize the probability of observing stars ejected from the stellar disc. The HVS population is instead expected to be much fainter than this magnitude cut (Marchetti et al. 2018). Since estimates on the expected HRS population in *Gaia* are currently missing, at the moment it is not clear whether this tension is real, and/or if other ejection mechanisms are needed (e.g. Irrgang, Kreuzer & Heber 2018).

5.2 Extragalactic stars

13 of the 20 $P_{\rm ub} > 80$ per cent stars have probabilities < 50 per cent of intersecting the MW stellar disc when traced back in time, therefore an extragalactic origin is preferred. A possible ejection location could be the LMC, or otherwise spatial correlations with the density of surrounding stars could help identifying them as the high-velocity tail of a stellar stream produced by the effect of the gravitational field of the MW on a dwarf satellite galaxy (Abadi et al. 2009).

The extragalactic star with a highest probability of being unbound from our Galaxy is Gaia DR2 1396963577886583296, with a total velocity $\sim 700\,{\rm km\,s^{-1}}$, resulting in a probability $P_{\rm ub} = 0.98$. We mark this source with a yellow star in Fig. 2 and following. This star is at ∼30 kpc from the GC, with an elevation of ∼25 kpc above the Galactic plane.

---

[2] Because of the small uncertainties, we repeat the total velocity determination for *Gaia* DR2 5932173855446728064 sampling within the uncertainties of the Sun position and motion (see discussion in Section 2.1). The result is $v_{\rm GC} = (747 \pm 7)\,{\rm km\,s^{-1}}$, in agreement with the previous estimate.

Table 2. Observed and derived properties for the 20 ‘clean’ high-velocity star candidates with a probability >80% of being unbound from the Galaxy. Stars are sorted by decreasing P_{ub}.

Gaia DR2 ID	(RA, Dec.) (°)	ϖ (mas)	$\mu_{\alpha *}$ (mas yr^{-1})	μ_{δ} (mas yr^{-1})	v_{rad} (km s^{-1})	G (mag)	d (pc)	r_{GC} (pc)	v_{GC} (km s^{-1})	P_{MW}	P_{ub}
Galactic											
5932173855446728064	(244.1181, −54.44045)	0.454 ± 0.029	− 2.676 ± 0.043	− 4.991 ± 0.034	− 614.286 ± 2.492	13.81	2197^{+162}_{-120}	6397^{+92}_{-123}	747^{+2}_{-3}	1.00	1.00
1383279090527227264	(240.33735, 41.16677)	0.118 ± 0.016	− 25.759 ± 0.025	− 9.745 ± 0.04	− 180.902 ± 2.421	13.01	8491^{+1376}_{-951}	10064^{+908}_{-561}	921^{+179}_{-124}	1.00	1.00
6456587609813249536	(317.36089, −57.9124)	0.099 ± 0.019	13.002 ± 0.029	− 18.263 ± 0.03	− 15.851 ± 2.833	13.01	10021^{+2023}_{-1480}	7222^{+1350}_{-761}	875^{+212}_{-155}	0.98	0.99
5935868592404029184	(253.90291, −53.29868)	0.074 ± 0.021	5.47 ± 0.032	6.358 ± 0.026	308.412 ± 1.212	13.08	12150^{+2919}_{-1909}	5985^{+2516}_{-1380}	747^{+110}_{-73}	0.83	0.98
5831614858352694400	(247.45238, −59.96738)	− 0.008 ± 0.025	4.405 ± 0.032	1.532 ± 0.03	258.295 ± 1.245	13.37	20196^{+6006}_{-4394}	14113^{+5781}_{-4061}	664^{+130}_{-93}	0.94	0.92
5239334504523094784	(158.89457, −65.46548)	0.038 ± 0.013	− 6.77 ± 0.025	2.544 ± 0.022	22.464 ± 1.891	13.39	19353^{+4247}_{-2940}	18351^{+3923}_{-2617}	609^{+140}_{-94}	0.77	0.88
4395399303719163904	(258.75009, 8.73145)	0.073 ± 0.019	− 9.911 ± 0.029	4.848 ± 0.029	24.364 ± 1.484	13.19	12848^{+2766}_{-2262}	8194^{+2309}_{-1620}	671^{+136}_{-106}	1.00	0.84
Extragalactic											
1396963577886583296	(237.73164, 44.4357)	− 0.017 ± 0.014	− 1.649 ± 0.023	− 4.966 ± 0.029	− 412.464 ± 1.002	13.24	31374^{+6332}_{-5185}	30720^{+6150}_{-4970}	693^{+145}_{-113}	0.00	0.98
5593107043671135744	(113.26944, −31.3792)	− 0.1 ± 0.017	− 1.582 ± 0.03	2.113 ± 0.028	104.437 ± 1.511	13.39	37681^{+8295}_{-6444}	41753^{+8183}_{-6322}	567^{+100}_{-76}	0.00	0.97
5546986344820400512	(125.63998, −32.62)	− 0.08 ± 0.022	− 1.986 ± 0.028	2.747 ± 0.035	79.255 ± 1.273	13.82	29062^{+5928}_{-4950}	32552^{+5782}_{-4781}	551^{+90}_{-75}	0.00	0.93
5257182876777912448	(144.73682, −60.53137)	− 0.012 ± 0.017	− 3.736 ± 0.029	3.444 ± 0.027	22.64 ± 1.723	13.49	26140^{+6400}_{-4240}	25824^{+6144}_{-3989}	605^{+148}_{-93}	0.03	0.92
4326973843264734208	(248.8923, −14.51844)	0.199 ± 0.031	− 20.546 ± 0.05	− 33.974 ± 0.033	− 220.392 ± 2.052	13.5	5257^{+881}_{-677}	3842^{+450}_{-465}	766^{+163}_{-122}	0.04	0.91
5298599521278293504	(140.14259, −62.46243)	− 0.053 ± 0.02	− 2.373 ± 0.071	3.883 ± 0.055	54.363 ± 1.17	13.39	28525^{+6774}_{-5110}	28145^{+6545}_{-4850}	579^{+139}_{-104}	0.03	0.88
6700075834174889472	(304.32289, −32.41577)	0.054 ± 0.037	− 7.243 ± 0.065	4.955 ± 0.047	22.491 ± 2.057	12.76	13068^{+3816}_{-3123}	7584^{+3330}_{-2219}	698^{+152}_{-120}	0.10	0.84
4073247619504712192	(280.26863, −26.28806)	0.05 ± 0.024	− 3.596 ± 0.046	6.231 ± 0.039	− 191.767 ± 2.735	13.58	14653^{+4331}_{-2807}	6884^{+4240}_{-2648}	695^{+139}_{-88}	0.11	0.84
6492391900301222656	(348.64665, −58.42957)	0.095 ± 0.018	7.502 ± 0.027	− 15.822 ± 0.026	− 149.856 ± 1.163	13.36	10276^{+1878}_{-1541}	9641^{+1335}_{-944}	658^{+149}_{-117}	0.06	0.84
4596514892566325504	(268.57736, 29.12348)	0.064 ± 0.013	− 1.086 ± 0.019	− 10.512 ± 0.023	− 112.792 ± 1.093	13.49	14255^{+2485}_{-1839}	12120^{+2106}_{-1453}	617^{+121}_{-90}	0.07	0.84
5830109386395388544	(249.9792, −61.90285)	− 0.006 ± 0.019	− 1.072 ± 0.027	3.932 ± 0.029	143.395 ± 0.633	13.14	23852^{+6287}_{-4917}	17735^{+6123}_{-4680}	600^{+118}_{-88}	0.08	0.84
1990547230937629696	(344.00637, 53.61551)	0.043 ± 0.017	− 4.769 ± 0.028	− 2.83 ± 0.027	− 83.38 ± 1.158	13.31	17543^{+4372}_{-3415}	21331^{+4114}_{-3130}	563^{+112}_{-84}	0.05	0.83
5321157479786017280	(128.82063, −53.20458)	− 0.023 ± 0.018	− 2.518 ± 0.032	3.224 ± 0.034	81.295 ± 0.668	13.59	27523^{+6086}_{-5176}	28715^{+5877}_{-4914}	545^{+110}_{-95}	0.08	0.83

Note. Distances and total velocities are quoted in terms of the median of the distribution, with uncertainties derived from the 16th and 84th percentiles.

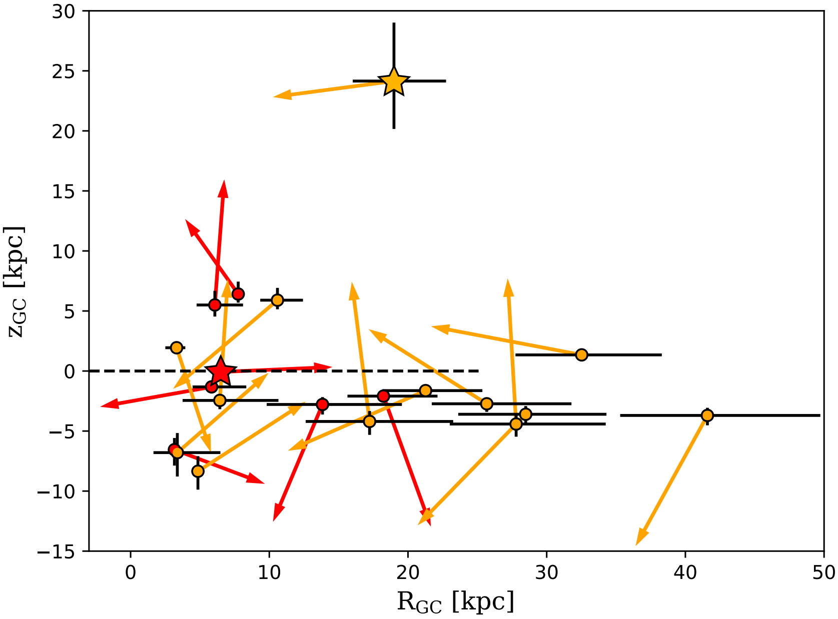

Figure 9. Position of the 20 high-velocity stars with $P_{\rm ub} > 80$ per cent in Galactocentric cylindrical coordinates ($R_{\rm GC}$, $z_{\rm GC}$). Arrows point to the direction of the velocity vector of the stars in this coordinate system, and the arrow's length is proportional to the total velocity of the star in the Galactic rest frame. Red (yellow) points and arrows mark the 7 (13) Galactic (extragalactic) candidates with $P_{\rm MW} > 0.5$ ($P_{\rm MW} < 0.5$). *Gaia* DR2 5932173855446728064 (*Gaia* DR2 1396963577886583296) is marked with a red (yellow) star. The Sun is located at ($R_{\rm GC}$, $z_{\rm GC}$) = (8200, 25) pc. The horizontal dashed line denotes the position of the Galactic plane, and extends up to the edge of the stellar disc, which we take to be at 25 kpc (Xu et al. 2015).

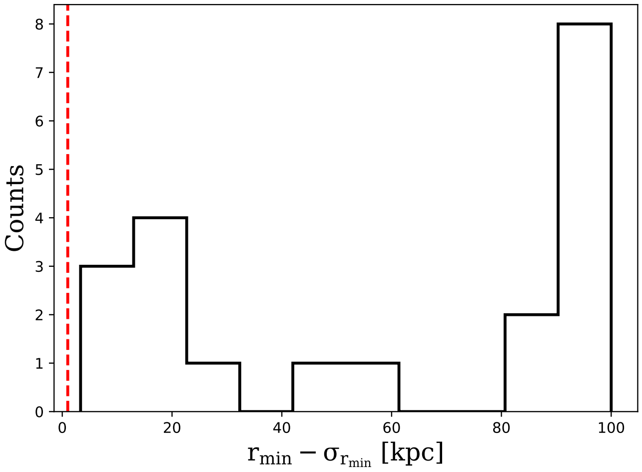

Figure 10. Histogram of the median minimum crossing radius $r_{\rm min}$ minus the correspondent lower uncertainty $\sigma_{r_{\rm min,l}}$ for the sample of 20 high-velocity stars with $P_{\rm ub} > 0.8$. The vertical dashed line corresponds to $(r_{\rm min} - \sigma_{r_{\rm min,l}}) = 1$ kpc, our boundary condition for not rejecting the GC origin hypothesis for the HVS candidates (see discussion in Section 5). $(r_{\rm min} - \sigma_{r_{\rm min,l}}) > 1$ kpc for all the 20 stars, therefore there are no HVS candidates.

6 CONCLUSIONS

We derived distance and total velocities for all the 7183 262 stars with a full phase space measurement in the *Gaia* DR2 catalogue, in order to find unbound objects and velocity outliers. We defined our sample of high-velocity stars as those stars with an estimated probability of being unbound from the MW $P_{\rm ub} > 50$ per cent, resulting in a total of 125 stars with reliable astrometric parameters and radial velocities. We traced back the high-velocity stars in the Galactic potential to derive orbital parameters. Out of these 125 stars, we found the following.

(i) 20 stars have predicted probabilities $P_{\rm ub} > 80$ per cent. The observed and derived kinematic properties of these stars are summarized in Table 2, and are discussed in Section 5.

(ii) None of these 20 stars is consistent with coming from the inner 1 kpc, so there are no HVS candidates. This is consistent with estimates presented in Marchetti et al. (2018).

(iii) 7 out of the 20 stars with $P_{\rm ub} > 0.8$, when traced back in time in the Galactic potential, originate from the stellar disc of the MW. These stars are HRS candidates.

(iv) 13 out of the 20 unbound candidates have probabilities < 50 per cent to originate from the stellar disc of the Galaxy. This surprising and unexpected population of stars could be either produced as RSs/HRSs/HVSs from the LMC, thanks to its high orbital velocity around the MW, or could be members of dwarf galaxies tidally disrupted by the gravitational interaction with the Galaxy. Further analyses are required in order to identify their origin.

Another possibility that we cannot rule out is that a subset of these 20 stars is actually gravitationally *bound* to the MW. Recent high-resolution spectroscopic follow-ups showed that some of these stars are actually indistinguishable from halo stars from a chemical point of view (Hawkins & Wyse 2018), therefore if they are actually bound, this would in turn imply a more massive MW (Hattori et al. 2018b; Monari et al. 2018), a possibility that cannot be ruled out (e.g. Wang et al. 2015). Otherwise, a confirmation of the global parallax zero-point measured with quasars could lower down their total velocities, resulting in the same effect as discussed in Appendix D, including this parallax offset results in 14 (4) stars with an updated $P_{\rm ub} > 50$ per cent ($P_{\rm ub} > 80$ per cent). The choice of not considering the parallax zero-point in the main text is therefore a conservative choice, which ensures us that all the high-velocity stars in the subset of *Gaia* DR2 with radial velocities are actually included in this work. In Appendix E, we show how including systematic errors in parallax can significantly lower the distances and total velocities for our candidates, but we want to stress that the adopted parameters might be too pessimistic for the stars considered in this paper (Lindegren et al. 2018b). Follow-up observations with ground-based facilities and/or future data releases of the *Gaia* satellite will help us confirming or rejecting their interpretation as kinematic outliers.

This paper is just a first proof of the exciting discoveries that can be made mining the *Gaia* DR2 catalogue. We only limited our search to the ∼7 million stars with a full phase space information, a small catalogue compared to the full 1.3 billion sources with proper motions and parallaxes. Synergies with existing and upcoming ground-based spectroscopic surveys will be essential to obtain radial velocities and stellar spectra for subsets of these stars (e.g. Dalton 2016; de Jong et al. 2016; Kunder et al. 2017; Martell et al. 2017). For what concerns HVSs, Marchetti et al. (2018) show how the majority of HVSs expected to be found in the *Gaia* catalogue are actually fainter than the limiting magnitude for radial velocities in DR2. We therefore did not expect to discover the bulk of the HVS population with the method outlined in this paper, but other data mining techniques need to be implemented in order to identify them among the dominant background of bound, low-velocity stars (see for example Marchetti et al. 2017). We also show how particular attention needs to be paid to efficiently filter out contaminants and instrumental artefacts, which might mimic high-velocity stars at a first inspection.

ACKNOWLEDGEMENTS

We thank the anonymous referee for his/her comments, which greatly improved the quality of this manuscript. We also thank E. Zari and the *Gaia* group meeting at Leiden Observatory for useful comments, suggestions, and discussions during the preparation of this paper. TM and EMR acknowledge support from NWO TOP grant Module 2, project number 614.001.401. This project was developed in part at the 2017 Heidelberg *Gaia* Sprint, hosted by the Max-Planck-Institut für Astronomie, Heidelberg. 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. This research made use of ASTROPY, a community-developed core PYTHON package for Astronomy (Astropy Collaboration et al. 2013). All figures in the paper were produced using MATPLOTLIB (Hunter 2007) and TOPCAT (Taylor 2005). This work would not have been possible without the countless hours put in by members of the open-source community all around the world.

REFERENCES

Abadi M. G., Navarro J. F., Steinmetz M., 2009, ApJ, 691, L63
Andrae R. et al., 2018, A&A, 616, A8
Astraatmadja T. L., Bailer-Jones C. A. L., 2016a, ApJ, 832, 137
Astraatmadja T. L., Bailer-Jones C. A. L., 2016b, ApJ, 833, 119
Astropy Collaboration et al., 2013, A&A, 558, A33
Bailer-Jones C. A. L., 2015, PASP, 127, 994
Bailer-Jones C. A. L., Rybizki J., Fouesneau M., Mantelet G., Andrae R., 2018, ApJ, 156, 58
Blaauw A., 1961, Bull. Astron. Inst. Neth., 15, 265
Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529
Boeche C. et al., 2013, A&A, 553, A19
Boubert D., Evans N. W., 2016, ApJ, 825, L6
Boubert D., Erkal D., Evans N. W., Izzard R. G., 2017, MNRAS, 469, 2151
Boubert D., Guillochon J., Hawkins K., Ginsburg I., Evans N. W., 2018, MNRAS, 479, 2789
Bovy J., 2015, ApJS, 216, 29
Bromley B. C., Kenyon S. J., Geller M. J., Barcikowski E., Brown W. R., Kurtz M. J., 2006, ApJ, 653, 1194
Brown W. R., 2015, ARA&A, 53, 15
Brown W. R., Geller M. J., Kenyon S. J., Kurtz M. J., 2005, ApJ, 622, L33
Brown W. R., Geller M. J., Kenyon S. J., 2014, ApJ, 787, 89
Brown W. R., Anderson J., Gnedin O. Y., Bond H. E., Geller M. J., Kenyon S. J., 2015, ApJ, 804, 49
Capuzzo-Dolcetta R., Fragione G., 2015, MNRAS, 454, 2677
Carollo D. et al., 2010, ApJ, 712, 692
Contigiani O., Rossi E. M., Marchetti T., 2018, preprint (arXiv:1807.04468)
Cropper M. et al., 2018, AAP, 616, A5
Dalton G., 2016, in Skillen I., Balcells M., Trager S., eds, ASP Conf. Ser. Vol. 507, Multi-Object Spectroscopy in the Next Decade: Big Questions, Large Surveys, and Wide Fields. Astron. Soc. Pac., San Francisco, p. 97
de Jong R. S. et al., 2016, in Evans C. J., Simard L., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI. SPIE, Bellingham, p. 99081O
Erkal D., Boubert D., Gualandris A., Evans N. W., Antonini F., 2018, MNRAS, preprint (arXiv:1804.10197)
Evans N. W., Sanders J. L., Williams A. A., An J., Lynden-Bell D., Dehnen W., 2016, MNRAS, 456, 4506
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Fragione G., Capuzzo-Dolcetta R., 2016, MNRAS, 458, 2596
Fragione G., Loeb A., 2017, New Astron., 55, 32
Gaia Collaboration et al., 2016, A&A, 595, A2
Gaia Collaboration et al., 2018a, A&A, 616, A1
Gaia Collaboration et al., 2018b, A&A, 616, A11
Geier S. et al., 2015, Science, 347, 1126
Gnedin O. Y., Gould A., Miralda-Escudé J., Zentner A. R., 2005, ApJ, 634, 344
Gnedin O. Y., Brown W. R., Geller M. J., Kenyon S. J., 2010, ApJ, 720, L108
Goodman J., Weare J., 2010, Comm. Appl. Math. Comput. Sci., 5, 65
Gvaramadze V. V., Gualandris A., 2011, MNRAS, 410, 304
Gvaramadze V. V., Gualandris A., Portegies Zwart S., 2009, MNRAS, 396, 570
Hattori K., Valluri M., Castro N., 2018a, preprint (arXiv:1804.08590)
Hattori K., Valluri M., Bell E. F., Roederer I. U., 2018b, preprint (arXiv:1805.03194)
Hawkins K., Wyse R. F. G., 2018, MNRAS, 481, 1028
Hawkins K. et al., 2015, MNRAS, 447, 2046
Hernquist L., 1990, ApJ, 356, 359
Hills J. G., 1988, Nature, 331, 687
Hirsch H. A., Heber U., O'Toole S. J., Bresolin F., 2005, A&A, 444, L61
Hoogerwerf R., de Bruijne J. H. J., de Zeeuw P. T., 2001, A&A, 365, 49
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Irrgang A., Kreuzer S., Heber U., 2018, preprint (arXiv:1807.05909)
Johnson D. R. H., Soderblom D. R., 1987, AJ, 93, 864
Katz D. et al., 2018, preprint (arXiv:1804.09372)
Kenyon S. J., Bromley B. C., Geller M. J., Brown W. R., 2008, ApJ, 680, 312
Kenyon S. J., Bromley B. C., Brown W. R., Geller M. J., 2014, ApJ, 793, 122
Kunder A. et al., 2017, AJ, 153, 75
Leonard P. J. T., Duncan M. J., 1990, AJ, 99, 608
Li Y.-B. et al., 2015, Res. Astron. Astrophys., 15, 1364
Lindegren L., Lammers U., Hobbs D., O'Mullane W., Bastian U., Hernández J., 2012, A&A, 538, A78
Lindegren L. et al., 2016, A&A, 595, A4

Lindegren L. et al., 2018a, A&A, 616, A2
Lindegren L. et al., 2018b
Luri X. et al., 2014, A&A, 566, A119
Luri X. et al., 2018, A&A, 616, A9
Madigan A.-M., Pfuhl O., Levin Y., Gillessen S., Genzel R., Perets H. B., 2014, ApJ, 784, 23
Marchetti T., Rossi E. M., Kordopatis G., Brown A. G. A., Rimoldi A., Starkenburg E., Youakim K., Ashley R., 2017, MNRAS, 470, 1388
Marchetti T., Contigiani O., Rossi E. M., Albert J. G., Brown A. G. A., Sesana A., 2018, MNRAS, 476, 4697
Martell S. L. et al., 2017, MNRAS, 465, 3203
Miyamoto M., Nagai R., 1975, PASJ, 27, 533
Monari G. et al., 2018, A&A, 616, L9
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Perets H. B., Šubr L., 2012, ApJ, 751, 133
Piffl T. et al., 2014, A&A, 562, A91
Portegies Zwart S. F., 2000, ApJ, 544, 437
Poveda A., Ruiz J., Allen C., 1967, Bol. Obs. Tonantzintla Tacubaya, 4, 86
Price-Whelan A. M., 2017, The Journal of Open Source Software, 2, 388
Przybilla N., Fernanda Nieva M., Heber U., Butler K., 2008, ApJ, 684, L103
Robin A. C. et al., 2012, A&A, 543, A100
Rossi E. M., Kobayashi S., Sari R., 2014, ApJ, 795, 125
Rossi E. M., Marchetti T., Cacciato M., Kuiack M., Sari R., 2017, MNRAS, 467, 1844
Schönrich R., 2012, MNRAS, 427, 274
Schönrich R., Binney J., Dehnen W., 2010, MNRAS, 403, 1829
Sesana A., Haardt F., Madau P., 2006, ApJ, 651, 392
Sesana A., Haardt F., Madau P., 2007, MNRAS, 379, L45
Sesana A., Haardt F., Madau P., 2008, ApJ, 686, 432
Silva M. D. V., Napiwotzki R., 2011, MNRAS, 411, 2596
Smith M. C. et al., 2007, MNRAS, 379, 755
Smith M. C. et al., 2009, MNRAS, 399, 1223
Tauris T. M., 2015, MNRAS, 448, L6
Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, ASP Conf. Ser. Vol. 347, Astronomical Data Analysis Software and Systems XIV. Astron. Soc. Pac., San Francisco, p. 29
van der Marel R. P., Kallivayalil N., 2014, ApJ, 781, 121
Vickers J. J., Smith M. C., Grebel E. K., 2015, AJ, 150, 77
Wang W., Han J., Cooper A. P., Cole S., Frenk C., Lowing B., 2015, MNRAS, 453, 377
Williams A. A., Belokurov V., Casey A. R., Evans N. W., 2017, MNRAS, 468, 2359
Xu Y., Newberg H. J., Carlin J. L., Liu C., Deng L., Li J., Schönrich R., Yanny B., 2015, ApJ, 801, 105
Yu Q., Tremaine S., 2003, ApJ, 599, 1129
Zhang F., Lu Y., Yu Q., 2013, ApJ, 768, 153
Zhang Y., Smith M. C., Carlin J. L., 2016, ApJ, 832, 10
Ziegerer E., Heber U., Geier S., Irrgang A., Kupfer T., Fürst F., Schaffenroth J., 2017, A&A, 601, A58

APPENDIX A: CHOICE OF THE PRIOR PROBABILITY ON DISTANCES

In this appendix, we discuss the choice of the prior probability on distances $P(d)$ that gives the most accurate results on the subsample of bright stars in *Gaia* DR2 with a large relative error on parallax (the *high-f* sample introduced in Section 2). We cross-match the *Gaia* Universe Model Snapshot (GUMS; Robin et al. 2012) and the *Gaia* Object Generator (GOG; Luri et al. 2014) catalogues based on the value of the source identifier, to get a resulting sample of 7×10^6 stars with $G_{\rm RVS} < 12.2$. We use the latest versions of these mock catalogues, GUMS-18 and GOG-18.[3] The resulting combined catalogue contains positions, parallaxes, proper motions, radial velocities, and distances for all stars, with corresponding uncertainties. We extend the limiting magnitude to $G_{\rm RVS} = 12.2$ to take into account the fact that *Gaia* does take spectra of some stars that are fainter than the limiting magnitude. In particular, these faint stars are the one with the largest error on parallax, so we want to be sure to include them, in order to derive accurate distances for the stars in *Gaia* DR2. We multiply the uncertainties on parallax and radial velocity by a factor $(60/22)^{0.5}$, and the ones on both proper motions by a factor $(60/22)^{1.5}$, to simulate the reduced performance of the *Gaia* satellite on 22 months of collected data.

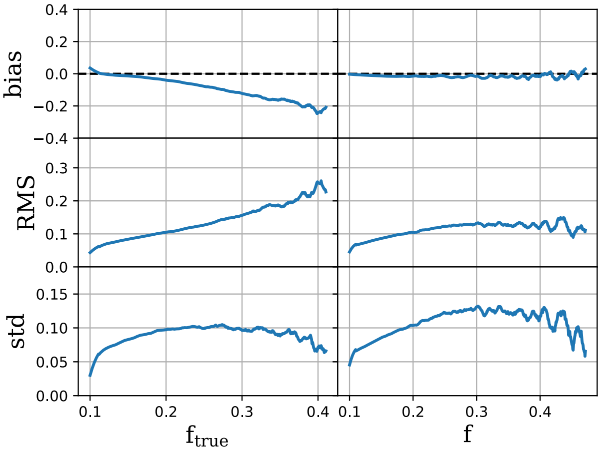

Figure A1. Bias, rms, and standard deviation of the estimator x_0 as a function of $f_{\rm true} = \sigma_\varpi d_{\rm true}$ (left-hand panel) and $f = \sigma_\varpi/\varpi$ (right-hand panel). The modes of the posterior distributions are estimated using the exponentially decreasing prior with a characteristic scale length $L = 2600$ pc.

We find 352 010 of the 7 million stars to have $f = \sigma_\varpi/\varpi > 0.1$. We can see that this value is about five times smaller than the one found in *Gaia* DR2 (see Section 2.2). All these stars are found at distance larger than ∼4.5 kpc from the Sun, and therefore we choose to adopt the exponentially decreasing prior to derive their distances (Astraatmadja & Bailer-Jones 2016b), see equation (3). The mode of the posterior distribution in equation (4) can be determined by numerically finding the roots of the implicit equation (Bailer-Jones 2015):

$$\frac{d^3}{L} - 2d^2 + \frac{\varpi}{\sigma_\varpi^2} d - \frac{1}{\sigma_\varpi^2} = 0. \tag{A1}$$

We compute the mode $d_{\rm Mo,\,i}$ for each star i in the simulated catalogue for different values of the scaling length L. We then determine the best-fitting value of the parameter L as the one minimizing the quantity $\sum_i x_i^2$, where the scaled residual x_i is computed as (Astraatmadja & Bailer-Jones 2016a)

$$x_i = \frac{d_{\rm Mo,i} - d_{\rm true,i}}{d_{\rm true,i}}, \tag{A2}$$

where $d_{\rm true,i}$ denotes the *true* simulated distance of the ith star. We find the value for the scale length $L = 2600$ pc to work best on this sample of ∼352 000 simulated stars. In Fig. A1, we plot the mean value of the bias $\bar{x}$, the root mean squared (rms) $\overline{x^2}^{1/2}$, and the standard deviation of the residual x for each bin of $f_{\rm true} = \sigma_\varpi d_{\rm true}$ (left-hand panel) and f (right-hand panel). We can see that, with this choice of prior, the mode of the posterior distribution on distances is an unbiased estimator for all the range of *observed* relative errors in parallax f, even if it shows a negative bias of ∼20 per cent for stars with large values of the *true* relative error $f_{\rm true}$.

---

[3] https://wwwhip.obspm.fr/gaiasimu/

Table B1. Catalogue description. Derived distances and velocities correspond to the median of the distribution, and lower and upper uncertainties are derived, respectively, from the 16th and 84thpercentiles of the distribution function. Entries labelled[1] are derived in this paper, while entries labelled[2] are taken from the *Gaia* DR2 catalogue (Gaia Collaboration et al. 2018a).

Column	Units	Name	Description
1	–	source_id	*Gaia* DR2 identifier[2]
2	deg	RA	Right ascension[2]
3	deg	Dec.	Declination[2]
4	mas	parallax	Parallax[2]
5	mas	e_parallax	Standard uncertainty in parallax[2]
6	mas yr^{-1}	pmra	Proper motion in right ascension[2]
7	mas yr^{-1}	e_pmra	Standard uncertainty in proper motion in right ascension[2]
8	mas yr^{-1}	pmdec	Proper motion in declination[2]
9	mas yr^{-1}	e_pmdec	Standard uncertainty in proper motion declination[2]
10	km s^{-1}	vrad	Radial velocity[2]
11	km s^{-1}	e_vrad	Radial velocity error[2]
12	mag	GMag	G-band mean magnitude[2]
13	pc	dist	Distance estimate[1]
14	pc	el_dist	Lower uncertainty on distance[1]
15	pc	eu_dist	Upper uncertainty on distance[1]
16	pc	rGC	Spherical Galactocentric radius[1]
17	pc	el_rGC	Lower uncertainty on spherical Galactocentric radius[1]
18	pc	eu_rGC	Upper uncertainty on spherical Galactocentric radius[1]
19	pc	RGC	Cylindrical Galactocentric radius[1]
20	pc	el_RGC	Lower uncertainty on cylindrical Galactocentric radius[1]
21	pc	eu_RGC	Upper uncertainty on cylindrical Galactocentric radius[1]
22	pc	xGC	Cartesian Galactocentric x-coordinate[1]
23	pc	el_xGC	Lower uncertainty on Cartesian Galactocentric x-coordinate[1]
24	pc	eu_xGC	Upper uncertainty on Cartesian Galactocentric x-coordinate[1]
25	pc	yGC	Cartesian Galactocentric y-coordinate[1]
26	pc	el_yGC	Lower uncertainty on Cartesian Galactocentric y-coordinate[1]
27	pc	eu_yGC	Upper uncertainty on Cartesian Galactocentric y-coordinate[1]
28	pc	zGC	Cartesian Galactocentric z-coordinate[1]
29	pc	el_zGC	Lower uncertainty on Cartesian Galactocentric z-coordinate[1]
30	pc	eu_zGC	Upper uncertainty on Cartesian Galactocentric z-coordinate[1]
31	km s^{-1}	U	Cartesian Galactocentric x-velocity[1]
32	km s^{-1}	el_U	Lower uncertainty on Cartesian Galactocentric x-velocity[1]
33	km s^{-1}	eu_U	Upper uncertainty on Cartesian Galactocentric x-velocity[1]
34	km s^{-1}	V	Cartesian Galactocentric y-velocity[1]
35	km s^{-1}	el_V	Lower uncertainty on Cartesian Galactocentric y-velocity[1]
36	km s^{-1}	eu_V	Upper uncertainty on Cartesian Galactocentric y-velocity[1]
37	km s^{-1}	W	Cartesian Galactocentric z-velocity[1]
38	km s^{-1}	el_W	Lower uncertainty on Cartesian Galactocentric z-velocity[1]
39	km s^{-1}	eu_W	Upper uncertainty on Cartesian Galactocentric z-velocity[1]
40	km s^{-1}	UW	Cartesian Galactocentric xz-velocity[1]
41	km s^{-1}	el_UW	Lower uncertainty on Cartesian Galactocentric xz-velocity[1]
42	km s^{-1}	eu_UW	Upper uncertainty on Cartesian Galactocentric xz-velocity[1]
43	km s^{-1}	vR	Cylindrical Galactocentric R-velocity[1]
44	km s^{-1}	el_vR	Lower uncertainty on cylindrical Galactocentric R-velocity[1]
45	km s^{-1}	eu_vR	Upper uncertainty on cylindrical Galactocentric R-velocity[1]
46	km s^{-1}	vtot	Total velocity in the Galactic rest frame[1]
47	km s^{-1}	el_vtot	Lower uncertainty on total velocity in the Galactic rest frame[1]
48	km s^{-1}	eu_vtot	Upper uncertainty on total velocity in the Galactic rest frame[1]
49	–	P_ub	Probability of being unbound from the Galaxy[1]

The reason why we choose not to use distances from Bailer-Jones et al. (2018) is that the authors fit the values of the scale length L to a full three-dimensional model of the Galaxy.[4] Their values are therefore driven by nearby, bright disc stars, with $f \ll 1$. Such an approach would underestimate distances (and therefore total velocities) to faint distant stars, the ones we are more interested in.

APPENDIX B: CONTENT OF THE DISTANCE AND VELOCITY CATALOGUE

Table B1 provides an explanation of the content of the catalogue containing distances and velocities for the 7183 262 stars with a radial velocity measurement in *Gaia* DR2. The catalogue is publicly available at http://home.strw.leidenuniv.nl/ marchetti/research.html.

---

[4] Note that Bailer-Jones et al. (2018) adopt a scale length that varies smoothly with Galactic longitude and latitude.

APPENDIX C: LIST OF HIGH-VELOCITY STARS WITH $0.5 < P_{\rm UB} \le 0.8$.

In Table C1, we present *Gaia* identifiers, distances, and total velocities for the 105 high-velocity stars discussed in Section 4, with $0.5 < P_{\rm ub} \le 0.8$.

Table C1. Distances and total velocities in the Galactic rest frame for the 105 'clean' high-velocity star candidates with $0.5 < P_{\rm ub} \le 0.8$. Sources are sorted by decreasing $P_{\rm ub}$.

Gaia DR2 ID	d (pc)	$v_{\rm GC}$ (km s^{-1})	$P_{\rm ub}$
5718618735518384768	$31\,308^{+6464}_{-6622}$	488^{+69}_{-73}	0.79
4532372476587492608	$14\,132^{+3798}_{-2288}$	606^{+142}_{-83}	0.78
4366218814874247424	7506^{+1521}_{-955}	678^{+137}_{-86}	0.78
5244448023850619648	$16\,553^{+3638}_{-2756}$	552^{+91}_{-66}	0.77
1994938164981988864	$22\,185^{+5526}_{-4751}$	516^{+85}_{-70}	0.77
2159020415489897088	7686^{+1651}_{-1293}	603^{+123}_{-97}	0.77
2112308930997657728	6114^{+999}_{-712}	619^{+119}_{-84}	0.77
5802638672467252736	9985^{+1804}_{-1322}	647^{+150}_{-108}	0.76
5996908319666721792	$13\,616^{+3593}_{-2595}$	662^{+151}_{-108}	0.75
5316722526615701504	$24\,242^{+6103}_{-4691}$	525^{+123}_{-89}	0.74
2095259117723646208	$13\,359^{+2970}_{-2614}$	594^{+134}_{-112}	0.73
5839686407534279808	7346^{+1033}_{-839}	633^{+112}_{-92}	0.72
1333199496978208128	$20\,038^{+4062}_{-3076}$	543^{+120}_{-86}	0.72
2089995308886282880	$13\,397^{+2700}_{-1874}$	573^{+121}_{-81}	0.71
2045752026157687040	$11\,799^{+2705}_{-2004}$	604^{+144}_{-106}	0.71
6431596947468407552	$11\,356^{+2099}_{-1531}$	590^{+66}_{-47}	0.71
5247579810921207680	$27\,357^{+5878}_{-4547}$	499^{+115}_{-85}	0.7
5298494930231856512	$23\,913^{+5493}_{-4057}$	510^{+119}_{-85}	0.7
2095397827987170816	14751^{+2839}_{-2301}	574^{+122}_{-98}	0.7
4656931544705794816	$24\,368^{+5597}_{-4637}$	514^{+118}_{-95}	0.7
6642234513167197824	6836^{+1252}_{-1037}	649^{+117}_{-91}	0.69
5399966178291369728	$10\,155^{+2090}_{-1430}$	566^{+121}_{-81}	0.69
5374177064347894272	6225^{+1109}_{-879}	587^{+97}_{-76}	0.68
2072048770884296704	$16\,139^{+3291}_{-2678}$	552^{+118}_{-94}	0.68
6116555426949827200	7741^{+1164}_{-1011}	628^{+118}_{-102}	0.67
6500989806352727936	$10\,407^{+2456}_{-1809}$	577^{+128}_{-90}	0.67
5217818333256869376	8642^{+1631}_{-1139}	585^{+118}_{-81}	0.67
2106519830479009920	8213^{+1326}_{-1065}	570^{+85}_{-67}	0.67
6397497209236655872	5802^{+643}_{-487}	587^{+54}_{-41}	0.66
2044224735768501760	$15\,167^{+3227}_{-2538}$	560^{+124}_{-96}	0.66
5303927273594669056	$20\,331^{+5200}_{-3372}$	508^{+118}_{-73}	0.66
1966103266381646720	$28\,232^{+6210}_{-5780}$	474^{+88}_{-76}	0.65
6241406793347941504	$14\,098^{+4035}_{-3000}$	609^{+139}_{-98}	0.65
5627896072604568960	$22\,754^{+5478}_{-4591}$	490^{+101}_{-83}	0.65
5415267600583814912	$24\,505^{+6046}_{-4520}$	498^{+115}_{-87}	0.65
5856098302217892352	$19\,735^{+4562}_{-3404}$	529^{+127}_{-93}	0.65
6444276683058885248	$11\,413^{+3064}_{-2202}$	617^{+147}_{-103}	0.65
2094386346009409280	$14\,643^{+2968}_{-2007}$	549^{+125}_{-82}	0.64
5309766504975294592	$25\,956^{+5528}_{-5114}$	490^{+106}_{-96}	0.64
3905884598043829504	2709^{+385}_{-289}	580^{+115}_{-86}	0.63
2038012426369296128	$16\,453^{+4086}_{-3062}$	543^{+127}_{-88}	0.63
5317203154946837760	$18\,068^{+3537}_{-3079}$	510^{+94}_{-80}	0.63
5897201311028035456	$17\,717^{+4423}_{-4116}$	543^{+83}_{-70}	0.62
5823425661366917376	$15\,652^{+4759}_{-3695}$	568^{+127}_{-97}	0.62
5807202126764572288	$14\,365^{+3602}_{-2776}$	563^{+97}_{-74}	0.62
3705761936916676864	3756^{+371}_{-300}	566^{+59}_{-46}	0.62
2183775885439262592	$23\,213^{+5580}_{-4338}$	480^{+102}_{-78}	0.62
5317776481532378240	$19\,139^{+4400}_{-3115}$	500^{+112}_{-79}	0.62
6077622510498751616	$14\,503^{+3852}_{-2502}$	538^{+84}_{-46}	0.62
4531575708618805376	$12\,030^{+2748}_{-1974}$	562^{+80}_{-56}	0.62
1956680279930601344	$23\,550^{+6723}_{-4451}$	480^{+113}_{-75}	0.62
6010197124582216832	$10\,863^{+3441}_{-1945}$	629^{+118}_{-65}	0.62
5232568213032618496	$27\,921^{+5690}_{-4842}$	487^{+111}_{-92}	0.61
5249820306388948992	$26\,092^{+6478}_{-4213}$	478^{+117}_{-78}	0.61
5779439836114210304	$23\,901^{+5743}_{-4509}$	492^{+69}_{-53}	0.6
5247264629041172608	$20\,274^{+3940}_{-3336}$	507^{+100}_{-80}	0.6
5912922197004254848	$12\,401^{+3128}_{-2696}$	610^{+122}_{-99}	0.6
5247811567357582336	$21\,321^{+4641}_{-3453}$	497^{+114}_{-86}	0.59
4489509905555953408	$11\,610^{+2734}_{-2257}$	590^{+117}_{-91}	0.59
2121857472227927168	$13\,251^{+2401}_{-1679}$	522^{+92}_{-63}	0.59
1989862986804105344	$10\,429^{+2057}_{-1607}$	523^{+107}_{-82}	0.58
6677910160794903296	4345^{+554}_{-396}	604^{+106}_{-76}	0.58
6229070238523155328	$13\,987^{+4361}_{-2810}$	567^{+142}_{-87}	0.58
4452929978332889216	$24\,168^{+5324}_{-4537}$	496^{+108}_{-88}	0.58
5785402796909679744	$14\,723^{+3134}_{-2187}$	543^{+132}_{-89}	0.58
5362114562797004544	$23\,461^{+5342}_{-4015}$	479^{+113}_{-80}	0.57
1331585993728475264	$10\,902^{+2413}_{-1920}$	544^{+115}_{-87}	0.57
6733156428223193856	$13\,978^{+3684}_{-2829}$	601^{+122}_{-92}	0.57
6221350429945324032	8878^{+2117}_{-1582}	593^{+141}_{-104}	0.57
3454083549225619712	5943^{+794}_{-627}	522^{+100}_{-77}	0.57
6868478546915992320	$14\,043^{+4460}_{-3582}$	576^{+130}_{-101}	0.57
4127621699294858368	$13\,174^{+3602}_{-2904}$	615^{+128}_{-98}	0.56
1364548016594914560	$10\,327^{+1989}_{-1642}$	531^{+66}_{-50}	0.56
4609875745549298688	$10\,640^{+1380}_{-1204}$	544^{+76}_{-66}	0.56
5212817273334550016	3811^{+330}_{-283}	565^{+59}_{-51}	0.56
1268023196461923712	4586^{+500}_{-390}	568^{+79}_{-61}	0.56
1696697285206197248	$23\,235^{+5014}_{-3909}$	464^{+111}_{-81}	0.56
6034352158118691072	$11\,013^{+2964}_{-2267}$	646^{+157}_{-104}	0.56
2098831980759357696	$15\,685^{+3439}_{-2694}$	518^{+119}_{-92}	0.56
5354094037807264384	$11\,683^{+2120}_{-1758}$	533^{+111}_{-90}	0.56
4220617568115374848	4978^{+814}_{-677}	603^{+114}_{-92}	0.56
5779919841659989120	$10\,641^{+2101}_{-1505}$	568^{+135}_{-95}	0.55
5317675979297751040	$27\,098^{+5311}_{-4561}$	451^{+81}_{-70}	0.55
3891412241883772928	7004^{+1531}_{-1150}	539^{+88}_{-65}	0.55
4916199478888664320	5579^{+725}_{-629}	549^{+66}_{-56}	0.55
2255126837089768192	$24\,623^{+4714}_{-4286}$	456^{+85}_{-74}	0.55

Table C1 – *continued*

Gaia DR2 ID	d (pc)	$v_{\rm GC}$ (km s^{-1})	$P_{\rm ub}$
5511130239834500864	$20\,579^{+5603}_{-3668}$	467^{+100}_{-68}	0.55
3784964943489710592	4031^{+733}_{-505}	552^{+92}_{-61}	0.55
2038818952503671424	$26\,358^{+5535}_{-5090}$	469^{+105}_{-92}	0.55
1954400884950622464	$19\,455^{+4960}_{-3498}$	482^{+108}_{-75}	0.54
5846560382443820032	7054^{+936}_{-629}	585^{+96}_{-64}	0.54
6130863887159694848	9639^{+2070}_{-1335}	550^{+133}_{-85}	0.54
5231000034569444992	$18\,206^{+3340}_{-3437}$	501^{+101}_{-105}	0.53
2186887606421426816	$24\,376^{+4607}_{-4292}$	454^{+74}_{-67}	0.53
5818738237122521344	$11\,884^{+3059}_{-2216}$	559^{+136}_{-89}	0.53
5249917441371959040	$17\,540^{+4063}_{-3149}$	494^{+116}_{-85}	0.53
6639557580310606976	$11\,135^{+3975}_{-2226}$	579^{+108}_{-55}	0.53
4210389120686616832	7886^{+2550}_{-1822}	599^{+143}_{-88}	0.52
1191989287342960640	$10\,798^{+2233}_{-1691}$	549^{+131}_{-96}	0.52
6098331056080412416	$16\,089^{+3894}_{-3358}$	528^{+89}_{-72}	0.52
2086507417487662976	$26\,304^{+5278}_{-4208}$	448^{+90}_{-72}	0.51
5303240216263896192	$21\,972^{+5482}_{-3995}$	464^{+111}_{-79}	0.51
2000253135474943616	$16\,537^{+3984}_{-3129}$	475^{+89}_{-69}	0.51
6035120957243593600	$10\,873^{+3525}_{-2307}$	603^{+124}_{-76}	0.51
1612628419987892096	$25\,402^{+5063}_{-3992}$	442^{+104}_{-79}	0.5

Table D1. Distances and total velocities in the Galactic rest frame for the 20 'clean' high-velocity star candidates with $P_{\rm ub} > 0.8$ presented in Table 2, including the -0.029 mas global parallax offset. For comparison, stars are sorted as in Table 2.

Gaia DR2 ID	d (pc)	$v_{\rm GC}$ (km s^{-1})	$P_{\rm ub}$
5932173855446728064	2096^{+130}_{-117}	747^{+3}_{-3}	1.0
1383279090527227264	7144^{+809}_{-782}	745^{+105}_{-102}	0.98
6456587609813249536	7964^{+1297}_{-885}	660^{+135}_{-92}	0.82
5935868592404029184	$10\,010^{+2144}_{-1800}$	665^{+81}_{-67}	0.75
5831614858352694400	$17\,160^{+4736}_{-4055}$	600^{+101}_{-86}	0.73
5239334504523094784	$14\,426^{+3339}_{-2236}$	454^{+105}_{-66}	0.32
4395399303719163904	9934^{+2389}_{-1586}	535^{+112}_{-71}	0.37
1396963577886583296	$23\,038^{+5341}_{-3347}$	511^{+112}_{-68}	0.73
5593107043671135744	$32\,604^{+6740}_{-4982}$	511^{+79}_{-61}	0.9
5546986344820400512	$26\,048^{+6507}_{-4962}$	507^{+99}_{-74}	0.78
5257182876777912448	$21\,973^{+4863}_{-4292}$	515^{+106}_{-91}	0.66
4326973843264734208	4718^{+725}_{-580}	670^{+131}_{-104}	0.72
5298599521278293504	$24\,102^{+6820}_{-3800}$	489^{+140}_{-74}	0.63
6700075834174889472	$11\,382^{+4021}_{-2622}$	631^{+158}_{-98}	0.69
4073247619504712192	$11\,656^{+3234}_{-1949}$	601^{+101}_{-61}	0.47
6492391900301222656	7999^{+1457}_{-1042}	487^{+109}_{-73}	0.29
4596514892566325504	$10\,522^{+1717}_{-1145}$	436^{+83}_{-53}	0.14
5830109386395388544	$19\,057^{+4550}_{-3307}$	514^{+84}_{-59}	0.51
1990547230937629696	$13\,243^{+2851}_{-2563}$	456^{+71}_{-62}	0.37
5321157479786017280	$22\,613^{+5272}_{-4543}$	456^{+94}_{-81}	0.51

APPENDIX D: GLOBAL PARALLAX OFFSET

In this appendix, we discuss the impact of including the -0.029 mas global parallax zero-point mentioned in Lindegren et al. (2018a), derived from *Gaia*'s observations of distant quasars. Being a negative offset, the net effect is to lower the inferred distances, and therefore the resulting total velocities. We repeat the Bayesian analysis discussed in Section 2 to the 20 stars with $P_{\rm ub} > 80$ per cent. In this case, the likelihood probability is again a multivariate Gaussian distribution, but with mean vector (Bailer-Jones et al. 2018):

$$\boldsymbol{m} = [\mu_{\alpha*}, \mu_\delta, 1/d + \varpi_{\rm zp}], \tag{D1}$$

where $\varpi_{\rm zp} = -0.029$ mas. In Table D1, we report the updated values of the distance, total velocity, and probability of being unbound from the Galaxy for the 20 stars discussed in Section 5. We now find 14 candidates (70 per cent) to have an updated $P_{\rm ub} > 50$ per cent, and 4 stars (20 per cent) to have $P_{\rm ub} > 80$ per cent.

APPENDIX E: SYSTEMATIC ERRORS IN PARALLAX

Gaia DR2 uncertainties in parallax do not include the contribution from *systematic* errors, which might depend on the magnitude, position, colour, and other property of the source. The mean value of the systematic errors is the global parallax offset $\varpi_{\rm zp}$ already discussed in Appendix D. In this appendix, we discuss the impact of adding this contribution to the quoted values of the parallax uncertainties. To do that, we follow the advice and guidelines presented in Lindegren et al. (2018b). Internal uncertainties published in the *Gaia* DR2 catalogue can be artificially inflated to keep into account systematic errors (e.g. Lindegren et al. 2016):

$$\sigma_{\varpi,\rm ext} = \sqrt{k^2\sigma_\varpi + \sigma_{\rm s}^2}, \tag{E1}$$

where $k \gtrsim 1$ is a correction factor, and $\sigma_{\rm s}$ is the variance of the systematic error. These parameters need to be calibrated using external datasets. Lindegren et al. (2018b) suggest adopting $k = 1.08$, $\sigma_{\rm s} = 0.021$ mas ($k = 1.08$, $\sigma_{\rm s} = 0.043$ mas) for bright stars with $G \lesssim 13$ (faint stars with $G \gtrsim 13$). In Table E1, we report the updated values for distances, total velocities, and probability of being unbound from the Galaxy for the sample of 20 stars discussed in Section 5. All of the stars but one are classified as *faint* stars. 9 (5) stars out of 20 now have an updated probability $P_{\rm ub} > 0.5$ ($P_{\rm ub} > 0.8$). We want to stress that the adopted value for $\sigma_{\rm s}$ is likely overestimated for the typical magnitude of stars in our sample (Lindegren et al. 2018b), therefore this is a conservative approach, which underestimates distances (and therefore total velocities).

Table E1. Distances and total velocities in the Galactic rest frame for the 20 'clean' high-velocity star candidates with $P_{\rm ub} > 0.8$ presented in Table 2. Parallax uncertainties are inflated according to equation (E1). For comparison, stars are sorted as in Table 2.

Gaia DR2 ID	d (pc)	$v_{\rm GC}$ (km s^{-1})	$P_{\rm ub}$
5932173855446728064	2316^{+306}_{-265}	746^{+3}_{-3}	1.0
1383279090527227264	8577^{+3716}_{-2135}	931^{+484}_{-278}	0.94
6456587609813249536	9370^{+3917}_{-2262}	806^{+414}_{-234}	0.86
5935868592404029184	$10\,744^{+3489}_{-2685}$	694^{+131}_{-101}	0.8
5831614858352694400	$13\,924^{+5147}_{-3860}$	531^{+109}_{-79}	0.53
5239334504523094784	$12\,051^{+4102}_{-3115}$	384^{+124}_{-84}	0.22
4395399303719163904	$11\,019^{+3704}_{-3061}$	585^{+179}_{-138}	0.58
1396963577886583296	$15\,707^{+5086}_{-3944}$	372^{+92}_{-54}	0.21
5593107043671135744	$18\,643^{+5317}_{-4575}$	348^{+62}_{-53}	0.14
5546986344820400512	$16\,803^{+5056}_{-4307}$	371^{+77}_{-61}	0.2
5257182876777912448	$14\,545^{+4481}_{-3702}$	361^{+91}_{-65}	0.17
4326973843264734208	6032^{+2296}_{-1452}	909^{+426}_{-265}	0.91
5298599521278293504	$16\,316^{+5884}_{-4573}$	341^{+109}_{-77}	0.18
6700075834174889472	$12\,278^{+4717}_{-3393}$	667^{+187}_{-129}	0.75
4073247619504712192	$11\,462^{+4236}_{-2678}$	593^{+135}_{-83}	0.48
6492391900301222656	9897^{+4543}_{-2488}	630^{+359}_{-185}	0.69
4596514892566325504	$11\,421^{+4949}_{-2998}$	479^{+242}_{-141}	0.44
5830109386395388544	$14\,312^{+5392}_{-4022}$	430^{+97}_{-69}	0.27
1990547230937629696	$11\,614^{+4698}_{-2769}$	416^{+116}_{-66}	0.35
5321157479786017280	$15\,167^{+5524}_{-3904}$	328^{+95}_{-62}	0.13

This paper has been typeset from a TEX/LATEX file prepared by the author.