




On stellar migration from Andromeda to the Milky Way

Lukas Gülzow ^{1,2}★, Malcolm Fairbairn ³ and Dominik J. Schwarz ²

¹*Institut für Astroteilchenphysik (IAP), Karlsruher Institut für Technologie (KIT), D-76131 Karlsruhe, Germany*

²*Fakultät für Physik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany*

³*Department of Physics, King's College London, Strand, London WC2R 2LS, UK*

Accepted 2024 March 4. Received 2024 February 24; in original form 2023 June 22

ABSTRACT

Recent *Gaia* observations suggest that some hypervelocity stars (HVSs) might originate from outside the Galaxy. We ask whether these HVSs could come from as far as Andromeda. Therefore, we simulate HVSs originating in Andromeda with initial conditions based on attributes of high-velocity stars measured in the Milky Way and a simple model for the gravitational potential of Andromeda and the Milky Way. We evaluate the validity of this scenario based on the simulation results. While we expect that the vast majority of HVSs in our Galaxy will originate here, we expect the number of stars present from Andromeda at any one time to be between 12 and 3910, depending upon model assumptions. Further, we analyse the properties of HVSs that are able to reach the Milky Way. We discuss whether they could be detected experimentally based on recent constraints set on the ejection rate of HVSs from the Milky Way centre.

Key words: stars: kinematics and dynamics – Galaxy: kinematics and dynamics – galaxies: individual: M31 – galaxies: kinematics and dynamics – Local Group.

1 INTRODUCTION

Hypervelocity stars (HVSs) are some of the fastest objects in the Galaxy. Some of them exceed escape velocity and are unbound to the Milky Way gravity. Based on recent *Gaia* observations, a number of them could have extragalactic origins. In this paper, we investigate the Andromeda galaxy as a source of these HVSs.

HVSs are defined as stars that have velocities of the order of 1000 km s^{-1} . They were first predicted by Hills (1988). The first HVS in the Milky Way was discovered by Brown et al. (2005). Since then, the number of known HVSs in the Milky Way keeps increasing with every new probe of the Galaxy (Ginsburg, Brown & Wegner 2013). As the Milky Way escape velocity is of the same order of magnitude as the typical HVS velocity (Monari et al. 2018), it is easy to see that they can be unbound to the Milky Way gravitational potential. Many known HVSs move away from the Milky Way (Kreuzer, Irrgang & Heber 2020). A single HVS is confidently associated with the Milky Way centre (MWC; Koposov et al. 2020). The cause for such high kinetic energies is thought to be gravitational interactions between binary stars and the supermassive black hole in the MWC or other massive black holes in that region (Hills 1988; Brown et al. 2010). In this so-called Hills mechanism, one of the stars is captured by the black hole, while the other is ejected at a high velocity. In a recent series of analyses of the *Gaia* Data Release 3 (DR3) catalogue (Gaia Collaboration 2016, 2022; Fabricius et al. 2021), constraints have been set on the ejection rate of HVSs from the MWC caused by this mechanism (Evans, Marchetti & Rossi 2022a, b; Marchetti, Evans & Rossi 2022). Other possible origins include the ejection of one half of a binary, caused by the supernova explosion of the other half

(Wang & Han 2009), tidal tails from dwarf galaxies passing through the Milky Way (Abadi, Navarro & Steinmetz 2009; Piffl, Williams & Steinmetz 2011), and runaway stars from the Large Magellanic Cloud (Boubert et al. 2017; Erkal et al. 2019; Evans et al. 2021; Lin et al. 2023). Irrgang, Kreuzer & Heber (2018b) also found that additionally another ejection mechanism is likely at play. The known HVSs are main-sequence stars with masses of the order of a few solar masses (Irrgang et al. 2018a).

Montanari, Barrado & García-Bellido (2019) investigated multiple HVSs from the *Gaia* Data Release (DR2) catalogue (Gaia Collaboration 2016, 2018). They filtered the data according to Galactocentric velocity as well as the probability of being unbound from the Milky Way. The probabilities for the HVSs to be unbound were provided by Marchetti, Rossi & Brown (2019). They found 20 HVSs with a probability >80 percent of being unbound. 13 out of these 20 HVSs have trajectories that point towards the Milky Way disc. This indicates that their origin is not within the Milky Way, but outside of it. The trajectories of HVSs that originate inside the Milky Way characteristically point away from it. Montanari et al. (2019) argued that they may originate in globular clusters and dwarf galaxies surrounding the Milky Way, or even in the Andromeda Galaxy. While Montanari et al. (2019) focused on dwarf galaxies, this paper will focus on Andromeda as a possible origin. Previous work on this topic has been done by Sherwin, Loeb & O’Leary (2008).

We expect Andromeda to eject HVSs at a much higher rate than its satellites. Compared to the dwarf galaxies in the immediate neighbourhood of the Milky Way, Andromeda and its satellites are much further away. For this reason, we neglect Andromeda’s satellites and only consider HVSs ejected from the inner region of Andromeda.

We simulate a dynamical model of the gravitational system of Andromeda and the Milky Way in PYTHON (Astropy Collaboration 2013,

* E-mail: lukas.guelzow@kit.edu

2018; Virtanen et al. 2020). We randomly generate initial positions near the centre of Andromeda from which the HVS trajectories start. For the initial velocity vectors, we first take isotropically distributed directions. Then, we randomly generate velocity magnitudes within boundaries based on Andromeda’s escape velocity (Kaffe et al. 2018) and the distribution of star velocities in the Milky Way (Marchetti 2021). With this information, we calculate the trajectories of the ejected HVSs.

We found that it is possible for them to reach the Milky Way. We approximated the amount of Andromeda HVSs in the Milky Way at present time and analysed their position and velocity properties. In addition, we discuss whether it is possible to detect them.

In Section 2, we explain the utilized models for the simulation of the Milky Way, Andromeda, and the HVS trajectories in the system as well as the generation of initial conditions. Section 3 discusses the results of the simulation and the properties of the HVSs close to the Milky Way, and compares them with data from the *Gaia* DR3 catalogue. Finally, we draw our conclusions in Section 4.

2 MODEL AND SIMULATIONS

There are three main components to the simulation of HVS trajectories in the system of the Milky Way and Andromeda: First, we calculate the relative motion of the two galaxies in Section 2.1. Secondly, we model the gravitational potential of the entire system in Section 2.2. And finally, we generate initial conditions and use the equations of motions to simulate HVS trajectories in Section 2.3.

2.1 Andromeda trajectory

The system of the Milky Way and Andromeda is a dynamical one. Before we can determine any HVS trajectories, we calculate the trajectory of Andromeda in order to have an accurate dynamical model. Since the two galaxies do not collide during the relevant time frame for the propagation of HVSs, we describe both galaxies by point masses for the purpose of calculating Andromeda’s trajectory. Other gravitational influences in- and outside the Local Group are neglected (including cosmological expansion). In addition, we assume a present age of the Universe of $t_0 = 13.8$ Gyr.

We choose the coordinate system for this calculation as well as in the HVS trajectory simulation so that the MWC lies in the origin and is at rest. Andromeda lies on the x -axis and at $y = z = 0$ at present time. Additionally, the y -axis is chosen so that the Sun lies within the xy -plane to make the transformation to Galactic coordinates more convenient. The definition of the axis orientations only holds at present time; however, the coordinate system does not change in time.

The acceleration vector \mathbf{a}_{And} acting on Andromeda follows from the two-body problem,

$$\mathbf{a}_{\text{And}}(\mathbf{r}) = -G \frac{M_{\text{MW}} + M_{\text{And}}}{|\mathbf{r}|^3} \mathbf{r}, \quad (1)$$

with M_{MW} and M_{And} as the respective total masses of each galaxy and \mathbf{r} the position vector of Andromeda with respect to the MWC. Given the current position and velocity coordinates of Andromeda in the Galactocentric rest frame from van der Marel et al. (2012),

$$\begin{aligned} \mathbf{r}_{\text{M31}} &= \begin{pmatrix} -378.9 \\ 612.7 \\ -283.1 \end{pmatrix} \text{ kpc} \quad \text{and} \\ \mathbf{v}_{\text{M31}} &= \begin{pmatrix} 66.1 \pm 26.7 \\ -76.3 \pm 19.0 \\ 45.1 \pm 26.5 \end{pmatrix} \text{ km s}^{-1}, \end{aligned} \quad (2)$$

we can determine Andromeda’s trajectory at any point in time with equation (1) and the equations of motion established in Section 2.3.

2.2 Gravitational potential

The total acceleration acting on an object at any position in the system at any time is given by the gravitational potential. To find the total potential, we model the mass distribution of the two galaxies. We combine two different density profiles for the mass term of a single galaxy. The Navarro–Frenk–White (NFW) profile (Navarro, Frenk & White 1996) is used for the dark matter halo of a galaxy, making up the bulk of its mass. It is a spherically symmetric halo model based on N -body simulations in a cold dark matter scenario. It follows the mass density relation

$$\rho_{\text{NFW}}(r) = \frac{\rho}{\frac{r}{R_s} \left(1 + \frac{r}{R_s}\right)^2}. \quad (3)$$

Here, ρ is the characteristic density and R_s denotes the scale radius which both differ for every halo.

We add a second component of mass resulting from the Plummer density profile (Plummer 1911). The Plummer profile approximately describes the distribution of baryonic matter in a galaxy. It is a spherically symmetric model as well. That means it is at its most accurate near the centre of a galaxy. In the Plummer model, the mass density reads

$$\rho_{\text{P}}(r) = \frac{3M_{\text{B}}}{4\pi \cdot R_{\text{P}}^3} \left(1 + \frac{r^2}{R_{\text{P}}^2}\right)^{-\frac{5}{2}}. \quad (4)$$

Here, M_{B} is the total baryonic mass of the galaxy and R_{P} denotes the Plummer radius that determines the core radius of the galaxy.

The total acceleration \mathbf{a}_{tot} on an HVS caused by the gravitational influence of both galaxies follows by inserting these two mass components into Newton’s law of gravitation. \mathbf{a}_{tot} is defined by the three spacial components of $\mathbf{a}_{\text{tot}} = (a_{\text{tot},x}, a_{\text{tot},y}, a_{\text{tot},z})$ with

$$\begin{aligned} a_{\text{tot},i} &= -4\pi G \left[\rho_1 R_{s,1}^3 \left(\ln \left(\frac{R_{s,1} + |\mathbf{r}_1|}{R_{s,1}} \right) - \frac{|\mathbf{r}_1|}{R_{s,1} + |\mathbf{r}_1|} \right) \frac{r_{1,i}}{|\mathbf{r}_1|^3} \right. \\ &\quad + \rho_2 R_{s,2}^3 \left(\ln \left(\frac{R_{s,2} + |\mathbf{r}_2|}{R_{s,2}} \right) - \frac{|\mathbf{r}_2|}{R_{s,2} + |\mathbf{r}_2|} \right) \frac{r_{2,i}}{|\mathbf{r}_2|^3} \\ &\quad \left. + \frac{M_{\text{B},1}}{4\pi} \frac{r_{1,i}}{(|\mathbf{r}_1|^2 + R_{\text{P}}^2)^{3/2}} + \frac{M_{\text{B},2}}{4\pi} \frac{r_{2,i}}{(|\mathbf{r}_2|^2 + R_{\text{P}}^2)^{3/2}} \right]. \end{aligned} \quad (5)$$

Here, the top two lines correspond to the NFW profile and the bottom two lines correspond to the Plummer profile of the two galaxies. $r_{1,i}$ is the i -component of the position vector of the HVS with respect to the MWC \mathbf{r}_1 , and $r_{2,i}$ is the i -component of the position vector of the HVS with respect to the centre of Andromeda \mathbf{r}_2 . Analogously, the indices 1 and 2, respectively, refer to the Milky Way and Andromeda on characteristic density ρ , scale radius R_s , and M_{B} . ρ and M_{B} both depend on the total mass of the corresponding galaxy. R_s is set to 20 kpc for both galaxies since their virial radii are both of the order of $R_{\text{vir}} = 200$ kpc (Tamm et al. 2012; Nuza et al. 2014). The scale and virial radii are related by

$$R_s = \frac{R_{\text{vir}}}{c}. \quad (6)$$

c is a concentration parameter which has a typical value of $c = 10$ for galaxies like the Milky Way and Andromeda. We set $R_{\text{P}} = 4$ kpc for both galaxies because it results in a distribution of baryonic matter that is mainly concentrated within 4 kpc of their centres.

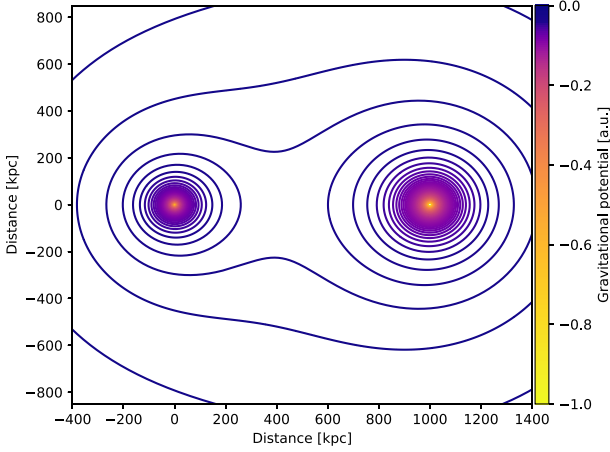


Figure 1. Contour plot of the combined potential of the Milky Way (at the origin of the coordinate system) and Andromeda galaxies in the half-mass scenario in arbitrary units at $t = 10$ Gyr after the big bang. The potential is computed according to our model for galaxy mass distribution, with the Plummer model for baryonic mass and the NFW model for the dark matter haloes. The potential has a saddle point between the two galaxies.

Due to varying values in the estimation of the total masses of both the Milky Way and Andromeda, we investigate two distinct scenarios in this paper:

- (i) $M_{\text{MW}} = M_{\text{And}} = 0.8 \times 10^{12} M_{\odot}$ (equal-mass scenario).
- (ii) $M_{\text{MW}} = 0.615 \times 10^{12} M_{\odot} \approx 0.54 M_{\text{And}}$ (half-mass scenario).

In scenario (i), we assume both galaxies to each have an approximate total mass of $0.8 \times 10^{12} M_{\odot}$. This is supported by Kafle et al. (2014, 2018), Karukes et al. (2020), Cautun et al. (2020), and Correa Magnus & Vasiliev (2021). In scenario (ii), we assume a mass ratio $M_{\text{MW}}/M_{\text{And}} \approx 0.54$ between the two galaxies. This scenario is supported by Watkins, Evans & An (2010), Peñarrubia et al. (2014), and Marchetti (2021). A 2D representation of the gravitational potential of the two galaxies in the half-mass scenario is shown in Fig. 1.

2.3 Modelling of hypervelocity star trajectories

To calculate a single HVS trajectory, we need to solve the equations of motion. In this case, the equations of motion are ordinary differential equations (ODEs) for the position \mathbf{r} and velocity \mathbf{v} of the star

$$\begin{aligned} \frac{d\mathbf{r}(t)}{dt} &= \dot{\mathbf{r}}(t) = \mathbf{v}(t), \\ \frac{d\mathbf{v}(t)}{dt} &= \dot{\mathbf{v}}(t, \mathbf{r}(t)) = \mathbf{a}(t, \mathbf{r}(t)). \end{aligned} \quad (7)$$

In three dimensions, we have a total of six equations, one for each of the three components of the position vector $\mathbf{r}(t)$ and velocity vector $\mathbf{v}(t)$. For each of the equations, we need a set of initial conditions with three components for both position and velocity. They are generated randomly for every HVS with weighting conditions determining the range in which they are likely to be found. In addition, we randomly generate an initial send-off time within the time interval [10, 13] Gyr after the big bang for each HVS. Outside of this interval, there is only a negligible amount of HVSs that is able to reach the Milky Way at present time.

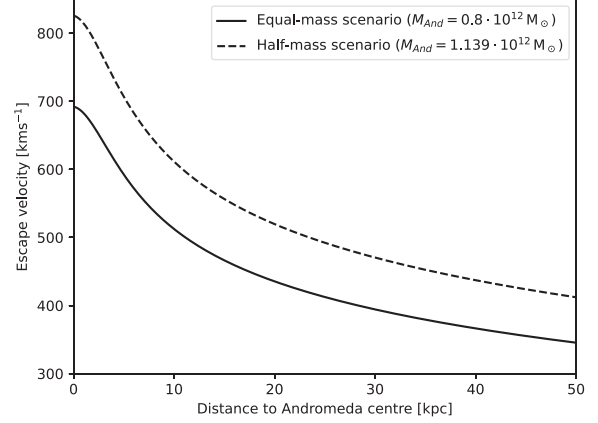


Figure 2. Escape velocity of the Andromeda galaxy versus distance to its centre for both mass scenarios. The graph is based on the gravitational potential arising from the NFW and Plummer models used in the simulation calculations in Section 2.2.

The initial radial distance from the centre of Andromeda r is determined by rearranging the Plummer model density relation in Equation (4). We find the radius–density relation

$$r(\rho_{\text{P},2}) = R_{\text{P},2} \sqrt{\left(\frac{3M_{\text{B},2}}{4\pi R_{\text{P},2}^3 \rho_{\text{P},2}} \right)^{\frac{2}{3}} - 1}. \quad (8)$$

Now, we randomly generate a density value between the Plummer model maximum density $\rho_{\text{P},2}(r=0)$ and $\rho_{\text{P},2}=0$ and substitute it in equation (8) to find the corresponding radius. This results in more initial radial coordinates towards the centre of Andromeda where more stars are located. The azimuthal angle φ and the inclination angle θ are randomly generated to give positions isotropically distributed around the centre of Andromeda.

In order to generate realistic initial velocity vectors for HVS, we employ a mathematical theorem from extreme value theory, called the Pickands–Balkema–de Haan theorem (Balkema & de Haan 1974; Pickands 1975). It deals with the statistical properties of extreme quantities. The theorem gives us access to the upper tail of an unknown distribution of a random variable if we assume that the distribution is continuous (which should be the case for the velocity distribution of stars). The theorem states that the tail above a fixed threshold is then well approximated by either an exponential distribution or a Pareto distribution. While the exponential distribution is specified by a scale and location parameter only, the Pareto distribution has an additional shape parameter. We show below that HVS in the Milky Way are indeed described very well by an exponential distribution (we also tested the Pareto distribution which improves the fit only marginally and thus we decided to go with the simpler set of assumptions). Based on this theorem, we characterize the velocity distribution of HVSs that exceed the escape velocity of Andromeda with an exponential distribution. At the same time, the extreme values are not sensitive to the exact form of the total distribution.

To find a reasonable exponential distribution, we use data from the analysis of star velocities in the Milky Way provided by Marchetti (2021). For the threshold of the distribution, we calculate the escape velocity of Andromeda depending on the distance to its centre according to our model. This is shown in Fig. 2.

We reproduce the plot from Marchetti (2021) showing the distribution of velocities of eligible stars in the *Gaia* eDR3 data catalogue

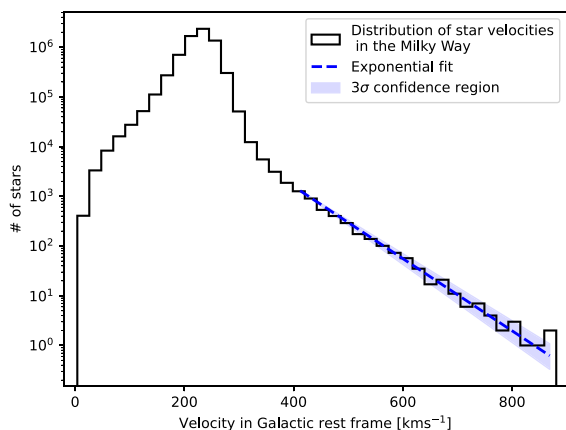


Figure 3. Velocity distribution of all Milky Way stars in *Gaia* eDR3 with sufficient information to perform velocity analysis (Marchetti 2021). Fitted with an exponential function $f(v)$ as a dashed line in the interval $v = [400, 900] \text{ km s}^{-1}$ in order to generate realistic initial conditions for HVSs near the centre of Andromeda. The 3σ confidence region is shown around the fit line and the best-fitting values are displayed in Table 1.

Table 1. Best-fitting parameters for the Milky Way velocity distribution in Fig. 3.

Velocity scale, σ_v (km s^{-1})	Amplitude, A (km s^{-1})	χ^2	d.o.f.
59.79 ± 1.27	94778.15 ± 1924.41	29.03	18

(Gaia Collaboration 2016, 2021; Fabricius et al. 2021) in Fig. 3. We fit an exponential function

$$f(v) = \frac{A}{\sigma_v} \exp\left(-\frac{v - v_0}{\sigma_v}\right) \quad (9)$$

to the high-velocity tail with the fit parameters velocity scale σ_v and amplitude A . The fit is shown in blue in Fig. 3. The fit result and the best-fitting parameters are displayed in Table 1. The function $P(v) = A^{-1}f(v)$ serves as the probability density function for a Milky Way star to have velocity v under the condition that $v > v_0 = 400 \text{ km s}^{-1}$. The cut-off v_0 is also called the location parameter of the distribution. Picking a different cut-off in the tail of the distribution does not alter the shape parameter σ_v , but merely changes the overall normalization A .

We now assume that we can adopt this present day velocity distribution of the Milky Way to Andromeda at earlier times. We believe that, at least for the equal-mass scenario, this is a sensible assumption.

We take the cumulative distribution function $F(v)$ corresponding to $P(v)$

$$F(v) = 1 - \exp\left(-\frac{v - v_0}{\sigma_v}\right) \quad \text{for } v \geq v_0. \quad (10)$$

Then, we rearrange it for v and substitute the escape velocity $v_{\text{esc}}(r)$ of Andromeda at the given position r for v_0

$$v[v_{\text{esc}}(r), F(v)] = v_{\text{esc}}(r) - \sigma_v [\ln(1 - F(v))]. \quad (11)$$

By generating a random number between 0 and 1 and substituting it for $F(v)$ in equation (11), we find the initial velocity magnitude v .

We note that the *Gaia* eDR3 high-velocity star selection from Marchetti (2021) may be biased and affected by spurious measurements. For a detailed discussion of the catalogue see Marchetti (2021). We also do not make the assumption that the velocities of

Table 2. Number and fraction of the total amount of simulated HVSs that reach a radius of r_{Filter} around the MWC for the two mass scenarios and result categories.

Mass scenario	Category	HVSs within r_{Filter}	Fraction
Equal	Min. distance	13 998	0.00078
Equal	Present time	2346	0.00013
Half	Min. distance	13 788	0.00077
Half	Present time	1891	0.00011

HVSs, ejected from near the centre of Andromeda, intrinsically follow the distribution of high-velocity stars in Fig. 3. We only use the exponential fit to the distribution as a reasonable function to model the upper tail of the unknown and assumedly continuous distribution of HVS velocities according to the Pickands–Balkeema–de Haan theorem. Sherwin et al. (2008) use a power-law distribution for the HVS velocities, based on the ejection mechanism. This is a special case of the Pareto distribution and, as such, also falls under the Pickands–Balkema–de Haan theorem. Plugging their velocity distribution into our model does not change the results significantly which supports the validity of our method (see Section 3).

The angular components of the initial velocity vector are generated so that the directions are isotropically distributed. Since we generate the initial conditions with respect to the centre of Andromeda, we add Andromeda’s position and velocity at the time of ejection to the initial vectors to find the initial conditions with respect to the MWC.

We can now use the equations of motion (equation 7) to calculate the trajectory of an HVS by plugging in its initial position and velocity as well as the acceleration from equation (5).

We use PYTHON to model the gravitational system of the Milky Way and Andromeda galaxies. Specifically, we utilize the SCIPY (Virtanen et al. 2020), NUMPY (Harris et al. 2020), and ASTROPY (Astropy Collaboration 2013, 2018) libraries. The equations of motion are solved in the simulation code by the ODEINT function from the SCIPY library. It integrates provided acceleration and velocity components to find position and velocity components, respectively. ODEINT utilizes the Runge–Kutta method (Butcher 1987; Hairer, Nørsett & Wanner 2000).

3 RESULT AND DISCUSSION

We structure the discussion in the following way: First, we analyse the properties of the simulated HVS in Section 3.1. Then, we compare these properties with *Gaia* observations in Section 3.2. At the end, we estimate the amount of Andromeda HVSs in the Milky Way at present time and discuss if it is possible to observe them in Section 3.3.

3.1 Properties of hypervelocity stars from Andromeda

For every HVS, the simulation produces data at the end of the calculated trajectory at present time $t_0 = 13.8 \text{ Gyr}$. We also store kinematic data at its minimum distance to the MWC. This data set includes the initial send-off time as well as the two positions in Galactic coordinates. Due to the isotropic distribution of initial velocity directions, the majority of simulated HVSs naturally move away from the Galaxy. We need to calculate a large number of trajectories to find a significant amount of HVSs with trajectories pointing towards the Milky Way.

It is important to note that the amount of simulated HVS is not to be interpreted as representative of the true amount of HVSs originating in Andromeda. Rather, it is a sufficiently large amount of data to

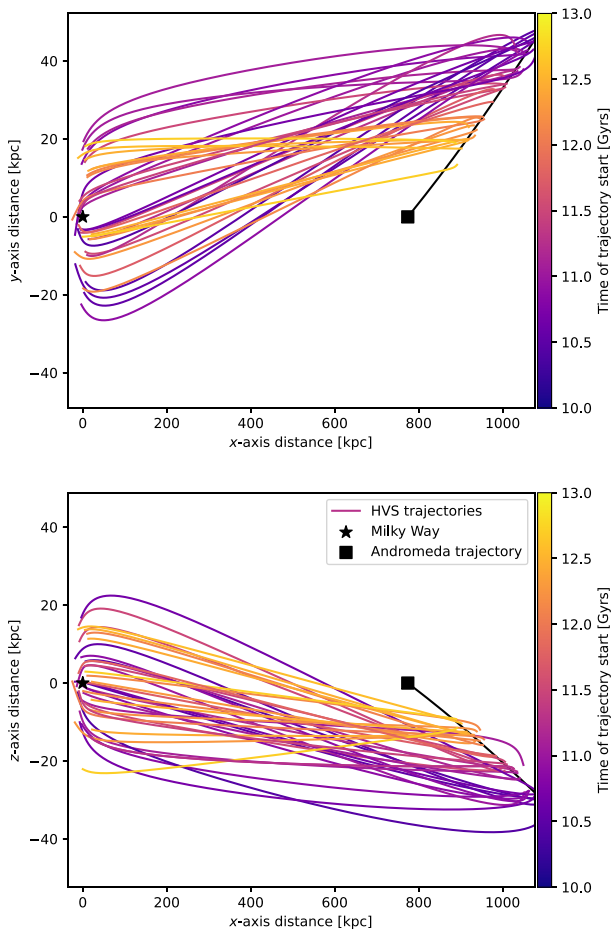


Figure 4. Projection plots of 38 full HVS trajectories as calculated by the simulation in the equal-mass scenario for the purpose of visualization. Chosen by hand with the condition of being within 25 kpc of the MWC at present time $t_0 = 13.8$ Gyrs. Colourmapped according to the time of the start of the trajectory. Coordinate system is chosen so that, at present time, Andromeda lies on the x -axis and at $y = z = 0$ and that the MWC and the Sun lie in the xy -plane. **Top:** Projection on to xy -plane. **Bottom:** Projection on to xz -plane.

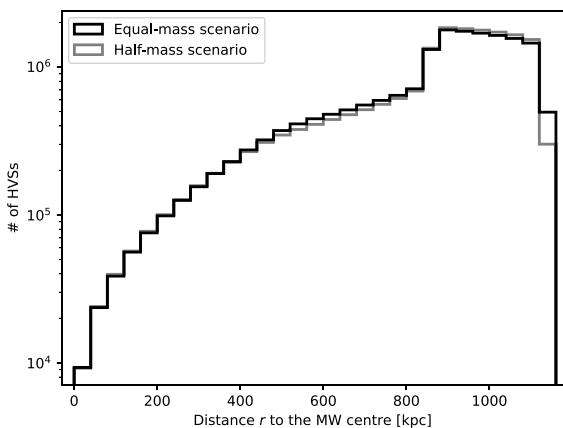


Figure 5. Distribution of minimum distance to the MWC of all simulated trajectories without any filter criteria. The star counts are displayed on a logarithmic scale. The flat peak to the right represents the HVSs that are sent off in the opposite direction to the Milky Way. Approximately 0.08 per cent of all simulated trajectories come within 50 kpc of the MWC at any point during their travel time. The equal-mass scenario is shown in black, the half-mass scenario in grey. There are only small differences between the scenarios. *MNRAS* **529**, 3816–3827 (2024)

analyse. The various assumptions and approximations made in our model are certainly the most important source of uncertainty for the results. Thus, we refrain from a detailed analysis of numerical errors, which are certainly much smaller, as we are relying on well established algorithms.

For each mass scenario, we calculate a total of 1.8×10^7 trajectories with randomly generated send-off times. We apply a filter to only consider HVSs in our analysis that reach a radius

$$r \leq r_{\text{Filter}} = 50 \text{ kpc} \quad (12)$$

around the MWC. We find that about 0.08 per cent of simulated trajectories fulfil this criterion at any point during their travel time. The amount of results for simulated HVSs within r_{Filter} of the MWC at present time is about one order of magnitude lower at, respectively, 0.013 and 0.011 per cent for the equal- and half-mass scenarios. The exact amount of results for each mass scenario is presented in Table 2. The reason for the significant difference in the amount of present time results between the two scenarios is Andromeda’s mass. In the half-mass scenario, HVSs require a higher escape velocity due to Andromeda’s larger mass. This causes some of the slowest HVSs that are sent off at the earliest possible times to still be too fast to remain in the Milky Way at present time. In addition, the Milky Way’s larger mass in the equal-mass scenario attracts the HVSs more strongly than in the half-mass scenario.

Fig. 4 visualizes a small, handpicked number of HVS trajectories that come especially close to the MWC in projection graphs in the xy - and xz -planes of the coordinate system used in the simulation (see Section 2.1). The y -axis is strongly compressed in order to show the differences between the trajectories in more detail. The different HVS trajectory start times are visible as a colour gradient along the trajectory of Andromeda. All shown HVS trajectories as well as Andromeda’s end at present time.

The distribution of minimum distances to the MWC of all simulated HVSs in both mass scenarios is displayed in Fig. 5 on a logarithmic scale. The graph shows a flat peak at the interval of distances where Andromeda is located during the [10, 13] Gyr time interval. The peak shows the HVSs that travel in directions pointing away from the Milky Way. For smaller distances to the Milky Way, the number of HVSs decreases rapidly. Differences between the two mass scenarios are negligible.

Fig. 6 displays the distance distribution of those HVSs that reach r_{Filter} in more detail around the MW centre and for both mass scenarios. The top panel again shows the minimum distances while the bottom panel has the distances of HVSs within r_{Filter} at present time. For the overall minimum distances, both mass scenarios show the same behaviour with a linear increase of HVS counts with increasing distance to the MWC. The linear slope indicates much lower HVS number density at higher radii. The absolute numbers of HVSs at minimum distance for both scenarios are similar as well.

The distribution of HVSs within r_{Filter} at present time, displayed in the bottom panel of Fig. 6, has an overall higher HVS count for the equal-mass scenario as was explained previously. The distributions in both scenarios show an increase in HVS counts with increasing distance to the MWC. It is consistent with a quadratic fit $f(r) = a r^2$ in each mass scenario. This indicates constant HVS number density within the considered r_{Filter} sphere. Additionally, we expect a dramatic increase of the population fraction of Andromeda HVSs with increasing distance, assuming the mass density of the Milky Way decreases with r^{-2} .

Next, we analyse the total velocity distributions of the simulated HVSs within r_{Filter} at t_0 for both mass scenarios in Fig. 7. Both distributions are fitted with the same exponential we fitted to the

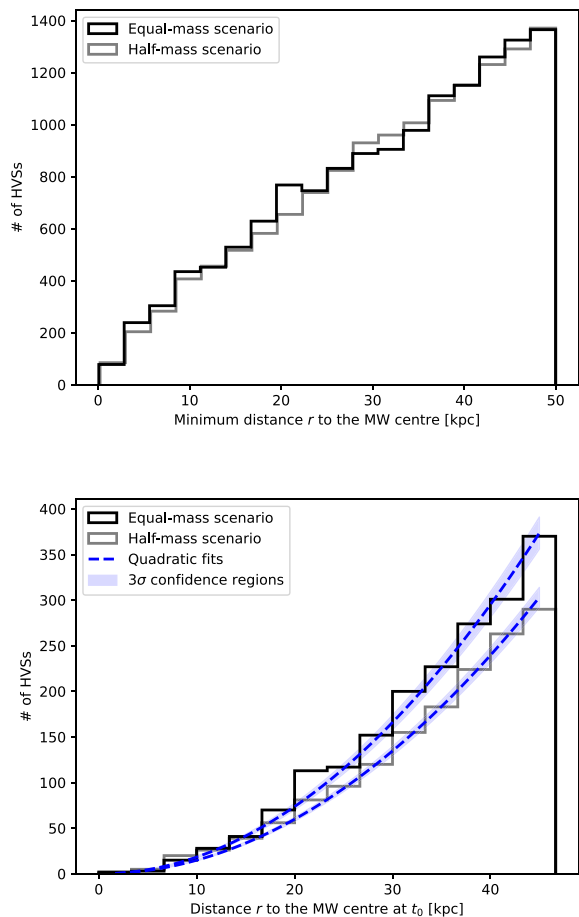


Figure 6. Distribution of distance to the MWC of simulation HVSSs within r_{Filter} . The black histograms show the equal-mass scenario. The grey histograms show the half-mass scenario. **Top:** Distances at minimum distance to the MWC. **Bottom:** Distances at present time $t_0 = 13.8$ Gyr. Each scenario is fitted with a quadratic function $f(r) = a r^2$. Fits are shown as dashed lines. Best-fitting values: Equal-mass ($a = 0.182 \pm 0.003$); half-mass ($a = 0.141 \pm 0.003$). 3σ confidence regions are shown around each fit.

Milky Way velocity distribution in Fig. 3 and which we used to generate the initial HVSS velocities. The best-fitting values are reported in Table 3. The highest velocity bins are excluded from each fit due to low statistics. In addition, the first bin is not considered either due to incomplete information.

To start with, we find that many HVSSs at lower velocities are not fast enough to be able to escape the Milky Way again. Thus, a significant fraction of the HVSSs arriving from Andromeda remain bound to the Milky Way gravity. Overall, the HVSSs still approximately follow the same exponential distribution even after the long journey to the Milky Way. While the slopes of the result distributions seem very close to the slope of the Milky Way distribution in the logarithmic representation, a comparison with the fit values in Table 1 shows differences for each fit parameter. However, the differences are smaller than an order of magnitude. We note that significantly fewer data points were used for the computation of the fits of the simulation results which causes the larger uncertainties. We also note that the simulated distributions have a similar amount of total stars in the shown bins as the initial distribution, shown as a dotted line. This

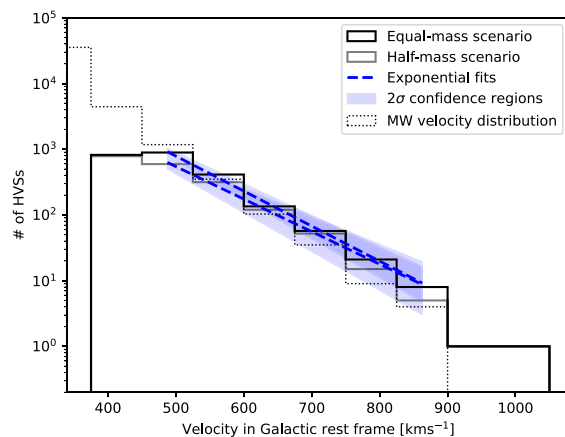


Figure 7. Total velocity distribution of the simulated HVSSs within r_{Filter} at present time. The equal-mass scenario is shown in black, the half-mass scenario in grey. Both histograms are fitted with the exponential function from equation (9) which was also fitted to the Milky Way distribution in Fig. 3. This distribution is shown here as the dotted line. For both fits, the first and the last two bins are excluded due to incomplete information and low statistics, respectively. The fits are displayed as dashed lines. 2σ confidence regions are shown around each fit.

Table 3. Best-fitting parameters for the fitted velocity distributions of the two simulation scenarios in Fig. 7.

Equal-mass scenario		χ^2	d.o.f.
Velocity scale, σ_v (km s^{-1})	Amplitude, A (km s^{-1})		
81.52 ± 3.05	218428 ± 9295	7.28	4
Half-mass scenario		χ^2	d.o.f.
Velocity scale, σ_v (km s^{-1})	Amplitude, A (km s^{-1})		
87.72 ± 5.34	149379 ± 10255	14.68	4

is coincidental and will change drastically based on the number of simulated HVSSs in the analysis.

In Fig. 8, we display the flight times of the HVSSs from Andromeda to the Milky Way, from their send-off to present time. Up to about 3000 Myr, the distributions are similar between the two scenarios. While the equal-mass scenario shows a large amount of HVSSs with flight times up to about 3600 Myr, the distribution in the half-mass scenario decays more quickly and only goes up to about 3300 Myr. The higher average initial velocity explains the lack of long flight times in the half-mass scenario. Under the assumption that the HVSSs are main sequence stars, the flight times account for a significant part of their life spans. Particularly for the flight times of multiple Gyrs, a significant fraction of HVSSs might evolve off the main sequence and not survive the journey to the Milky Way.

Figs 9 and 10 show sky maps of the distribution of the simulated HVSS positions in Galactic coordinates (l, b) for both mass scenarios. The Sun serves as the coordinate origin since this is how the HVSSs would be measured from Earth. Only HVSSs within a sphere of radius 40 kpc, that is wholly contained in the r_{Filter} sphere, are displayed to conserve symmetry. They are colourmapped according to their distance to the Sun. The minimum distance distributions in Fig. 9 form discs centred near the MWC at about $(-45^\circ, 25^\circ)$ on the opposite side of the direction of Andromeda at $(121.2^\circ, -21.6^\circ)$. The

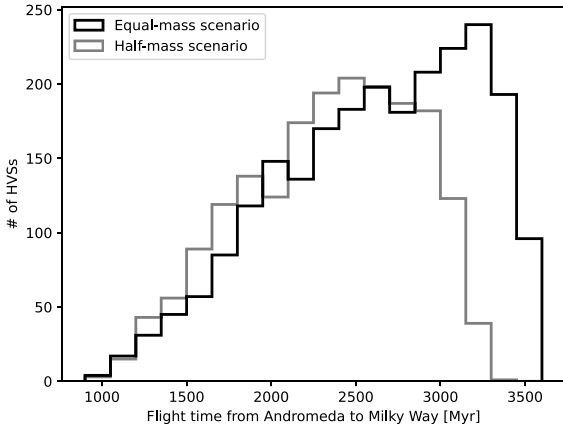


Figure 8. Distributions of flight times from Andromeda to the MWC of simulation HVSS within r_{Filter} at present time. The black histogram shows the equal-mass scenario. The grey histogram shows the half-mass scenario. Flight times are overall higher in the equal-mass scenario.

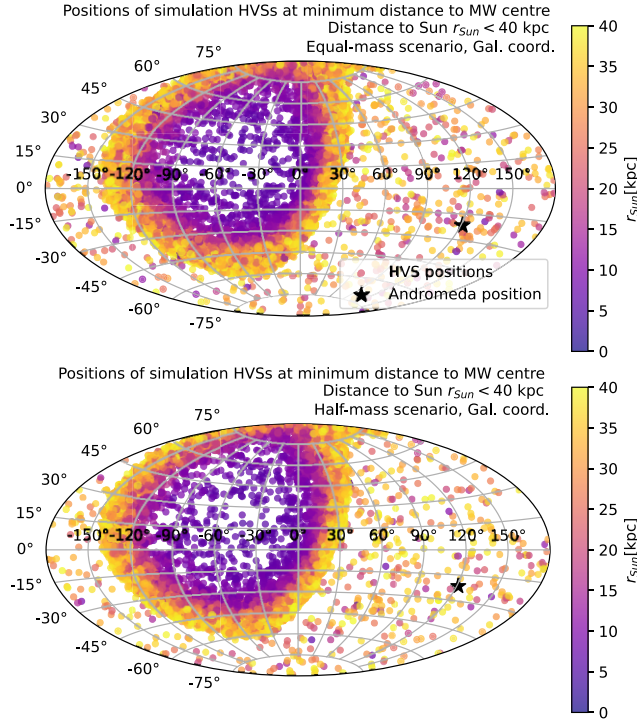


Figure 9. Sky maps showing the positions of the simulated HVSS within a sphere of radius $r = 40$ kpc around the Sun in Galactic coordinates at their minimum distance to the MWC. Colourmapped by distance to the Sun r_{Sun} . The position of Andromeda on the sky is marked by a black star. The distribution of HVSS is more concentrated in the equal-mass scenario due to the larger mass and gravitational pull of the MWC. Otherwise, the distributions are approximately identical. **Top:** Equal mass scenario. **Bottom:** Half mass scenario.

discs are oriented perpendicularly to Andromeda’s position vector. Evidently, HVSS typically reach the minimum distance to the MWC just after passing it. The HVSS that are not part of the discs do not pass the MWC before present time when the simulation ends. As such, their minimum distance positions are the same as their present time positions.

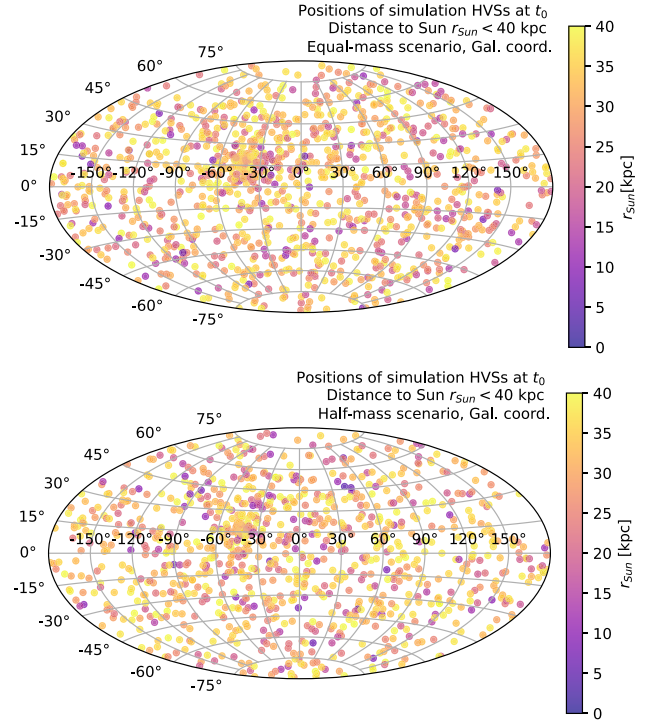


Figure 10. Sky maps showing the positions of the simulated HVSS in Galactic coordinates at present time $t_0 = 13.8$ Gyr within a sphere of radius $r = 40$ kpc around the Sun. Colourmapped by distance to the Sun r_{Sun} . The distribution is approximately homogeneous for both mass scenarios. **Top:** Equal mass scenario. **Bottom:** Half mass scenario.

The present time positions in Fig. 10 show almost isotropic distributions. Only near the Galactic centre at about $(-30^\circ, 15^\circ)$, near the position of the disc centres from Fig. 9, there is a visibly higher concentration of HVSS. This is expected due to the gravitational attraction of the MWC. Together with the constant number density from the quadratic fits in the bottom panel of Fig. 6, this result implies an approximately homogeneous and isotropic distribution of the simulated HVSS within the 40 kpc sphere around the Sun. For both minimum distance and present time positions, the two mass scenarios do not show a significant difference.

3.2 Comparison with *Gaia* DR3 data

To assess whether HVSS from Andromeda reaching the Milky Way is ruled out by observations, we compare the simulated HVSS at present time with star data from the *Gaia* DR3 (Fabricius et al. 2021; Gaia Collaboration 2016, 2022). We use the same radius of 40 kpc around the Sun as before and a velocity filter of $v > 500$ km s $^{-1}$ in the rest frame of the MWC. It is important to note that only objects with measured radial velocity, full astrometric solution (sky position, parallax, and proper motions) and sufficient measurement accuracy are eligible for this analysis. Only 10 721 objects of the total *Gaia* catalogue fulfil these criteria. Our methods of selecting data from the *Gaia* catalogue, and the velocity coordinate transformations are detailed in Appendices A and B, respectively.

In Fig. 11, we compare the positional data. The top panel shows the star positions, projected on to the Galactic plane with their elevation above it shown as colour. Almost all shown stars are close to zero elevation. The majority of stars is contained within an arc shape

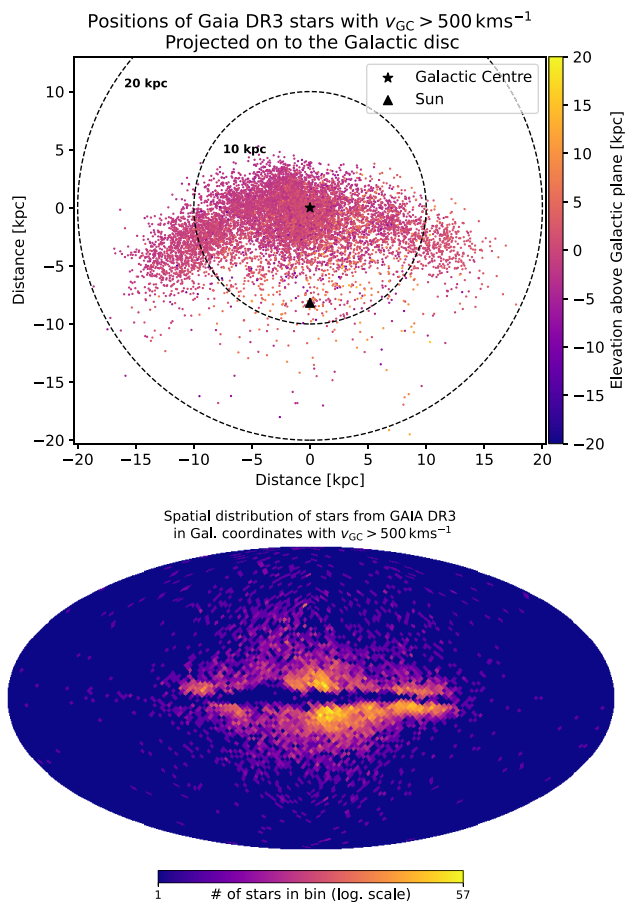


Figure 11. Positions of 10721 high-velocity stars in Galactic coordinates at present time $t_0 = 13.8$ Gyrs within a sphere of radius $r = 40$ kpc around the Sun. Data are taken from DR3 of the *Gaia* satellite mission. **Top:** Star positions projected on to the Galactic plane and colourmapped by their elevation above the disc. **Bottom:** Spatial distribution of *Gaia* stars as a population density sky map.

that is densest around the MWC. The area outside the arc is only sparsely populated. We expect a large concentration of high-velocity stars around the MWC and the Galactic disc due to the high stellar density in this region. In the bottom panel, the population density on the sky is displayed. We notice a prominent feature directly on the Galactic disc where the number of stars seems unexpectedly low. This is due to inaccurate measurements caused by dust attenuation along the line of sight, and the corresponding data cuts. Comparing this panel with Fig. 10, we see that *Gaia* observes stars at the region opposite Andromeda where we expect a concentration of HVSs from the simulation results. Overall, the distribution is very different from the isotropic scatter of the simulated HVSs.

Further, we directly compare the HVS velocity directions from the simulation results in the MWC rest frame with the corresponding *Gaia* data in Fig. 12. The simulation results in the top panel show a narrow range of possible directions which, expectedly, lie opposite to the position of Andromeda on the sky. The distribution of HVS directions in the equal-mass scenario is broader than in the half-mass scenario while also being less concentrated in the centre. This is caused by stronger gravitational focusing of the trajectories in the equal-mass scenario, meaning the trajectories that reach the Milky

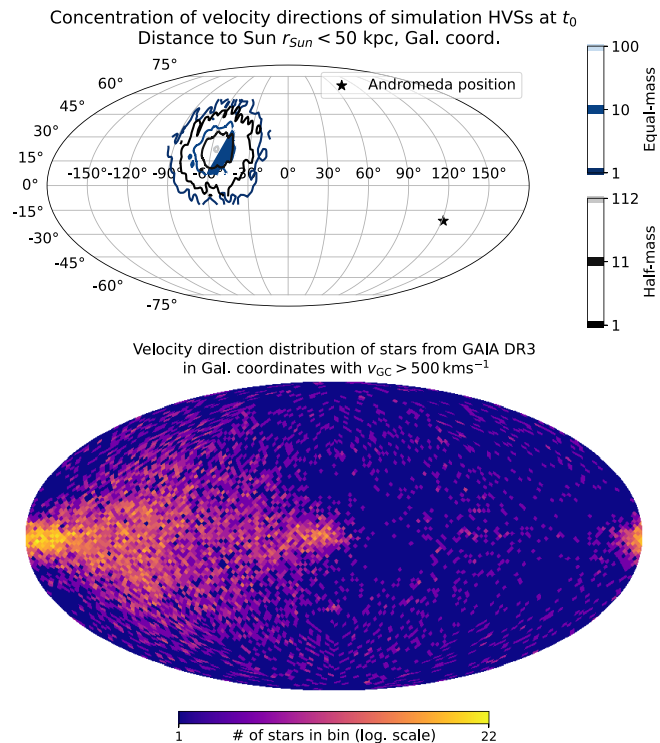


Figure 12. Comparison of velocity directions from simulation results and *Gaia* DR3 data. Directions are shown as if the velocity vector passed through the coordinate origin. Full 50 kpc radius around the MWC is considered. **Top:** Distribution of velocity directions of simulated HVSs in the MWC rest frame as contour plots. Blue contours show the broader distribution in the equal-mass scenario, grey-scale contours the distribution in the half-mass scenario. The velocity directions are more densely concentrated in the half-mass scenario. **Bottom:** Velocity directions of *Gaia* stars with velocity in the MWC rest frame $v_{\text{MWC}} > 500 \text{ km s}^{-1}$ as a population density sky map.

Way in the half-mass scenario, on average, point more directly from Andromeda at the Milky Way.

In comparison, the *Gaia* HVS velocity vectors in the bottom panel, shown as a logarithmic population density map, have strong preference along the position vector of the MWC ($0^\circ, 0^\circ$) and the opposite direction ($180^\circ, 0^\circ$). Further, we note a less populated preference in the directions of negative Galactic longitude, and Galactic latitude angles between 60° and -60° . This preference is denser within $([-150, -90]^\circ, [-30, 30]^\circ)$. The centre of this area lies opposite to the movement direction of the Sun. It also overlaps with the directions of the simulated HVSs. While we cannot observe clustering around the direction pointing away from Andromeda as in the simulations, we cannot exclude that some of the observed HVSs are indeed coming from Andromeda.

From these two comparisons, we conclude that the Andromeda HVS scenario is not excluded by *Gaia* data. In addition, the strongly focused velocity directions of HVSs from Andromeda might allow for their detection if the HVS numbers are large enough to cause a feature in a plot like the bottom panel of Fig. 12.

3.3 How many Andromeda HVSs are in the Milky Way?

Using the constraints of the ejection rate of HVSs from the MWC from Evans et al. (2022b), we can estimate the number of HVSs from

Table 4. Estimations of the amount of Andromeda HVSs expected to be in the Milky Way at present time based on constraints on the ejection rate of HVSs from the MWC from Evans et al. (2022b). Only the equal-mass scenario is shown. The top line displays the lower bound, and the bottom line the upper bound.

HVS ejection rate (yr^{-1})	Andromeda HVSs expected in the Milky Way
$10^{-4.5}$	12
10^{-2}	3910

Table 5. Number of Andromeda-ejected HVSs presently in Milky Way according to the power-law initial velocity distribution of Sherwin et al. (2008). The top line displays the lower bound, and the bottom line the upper bound.

HVS ejection rate (yr^{-1})	HVS amount predicted by power law
$10^{-4.5}$	7
10^{-2}	2341

Andromeda we expect in the Milky Way at present time. However, we need to assume the equal-mass scenario and that the ejection rate stays approximately the same between 10 and 13 Gyr after the big bang. We simply multiply the ejection rate per year with this time interval and the fraction of HVSs we expect from Andromeda (see Table 2). The results of the estimation are shown in Table 4. We acknowledge that these numbers might be lower since not all HVSs survive the travel time to the Milky Way. The significant differences between the two estimates are due to model dependences that are discussed in Evans et al. (2022b). Even for the lower bound of the ejection rate, we expect an amount of HVSs that is greater than 10. Such an amount is unlikely to be detected since *Gaia* DR3 only has full position and velocity information for a small fraction of its catalogue. The large amount of HVSs in the upper bound has a much higher likelihood to appear in measurements. However, a more detailed search for Andromeda HVSs in measurements of Milky Way stars is beyond the scope of this work.

For the Hills mechanism, Sherwin et al. (2008) expect there to be of the order of 1000 HVSs within the virialized Milky Way halo. We consider that this work does not take the entire Milky Way halo into account, but only about 1.43 per cent of the halo of radius $R_{M,200} = 206$ kpc assumed in Sherwin et al. (2008). Compared to the result obtained in this work, the lower bound and values up to about 150 HVSs are compatible. Notable differences between the two approaches include the models for the mass distribution and dynamics of the Local Group. Further, Sherwin et al. (2008) consider the stellar masses of the ejected HVSs, which we neglect.

As we briefly discussed in Section 2.3, we also considered the power-law HVS velocity distribution from Sherwin et al. (2008) for our model. For this comparison, we again simulated 1.8×10^7 trajectories in the equal-mass scenario, now with initial velocities generated from this power-law velocity distribution. We find that we expect about 0.0078 per cent of HVSs to reach the 50 kpc radius around the MWC. The resulting lower and upper bounds for the total amount of HVSs are presented in Table 5. The amount of HVSs expected in the Milky Way using the power-law distribution remains within the same order of magnitude for both the upper and lower bound. As such, the conclusion does not change.

4 CONCLUSIONS

To find out whether HVSs from Andromeda can reach the Milky Way, we have developed a simulation model of the gravitational system of the two galaxies. Using this model, we have calculated the trajectories of 1.8×10^7 HVSs each for two different scenarios. We have considered two mass scenarios with the Milky Way and Andromeda having equal mass, and the Milky Way having about half of Andromeda's mass, respectively.

The HVS initial positions have been randomly generated near the centre of Andromeda and ejected in random directions with initial velocity magnitudes based on Milky Way star behaviour and Andromeda's escape velocity. We have found that 0.013 and 0.011 per cent of simulated HVSs are within a radius of 50 kpc around the MWC at present time.

We have analysed the distance distributions within this 50 kpc filter radius. The minimum distances to the MWC follow a linear distribution. This means that the number density decreases significantly at large radii. The mass scenarios show no significant differences in this regard. In contrast, for the present time results, the equal-mass scenario shows a higher count of HVSs due to the stronger gravitation of the Milky Way and the initial HVS velocities in this scenario. The present time distance distributions show parabolic behaviour and have been fitted with a quadratic function, indicating homogeneous HVS number density within the considered sphere.

We have compared the velocity distributions of the present time data sets in both mass scenarios to the data set we used for the generation of initial velocities. We have found that, even after the long journey to the Milky Way, the HVSs still approximately follow the initial velocity distribution. Some HVSs from Andromeda slow down so much that they end up bound to the Milky Way. Additionally, we have considered the HVS flight times and found that a significant fraction of the ejected HVSs might evolve off the main sequence during the journey to the Milky Way.

From the simulation results, we have displayed as sky maps all HVSs arriving within 40 kpc of the Sun at present time t_0 as well as all HVSs that reach this minimum distance at any point during their trajectory. The positions of HVSs at their minimum distance form a disc centred on the MWC, opposite of Andromeda. The present time positions are distributed nearly isotropically and homogeneously within the considered sphere around the Sun. They are, expectedly, slightly more concentrated around the MWC due to its gravitational attraction. The mass scenarios show negligible differences for both minimum distance and present time sky maps.

We have compared this result with high-velocity star positions from the *Gaia* DR3. We have found that the simulated HVS star positions are consistent with the spatial distribution of observed HVSs, taking into account that HVSs should also originate from the MW itself, an effect that is not included in our simulations. We have compared the velocity directions from the simulation results with the *Gaia* data to find an approximately homogeneous population of velocity directions that overlaps with the simulated velocity directions.

In addition, we have estimated the expected amount of Andromeda HVSs in the Milky Way at present time using constraints on the ejection rate of HVSs from the MWC. We have found that even for the lower bound, we expect an amount greater than 10. The upper bound results in a large number of Andromeda HVSs that might be detectable in a more detailed analysis. We have compared these results to earlier work and found that they are compatible. In addition, we have tested our model with initial velocities generated from the HVS velocity distribution from said earlier work. As a result, the

expected number of Andromeda HVSs only changes within an order of magnitude.

We conclude that it is possible for HVSs from Andromeda to travel towards and reach the Milky Way. They are expected to form an approximately isotropic and homogeneous distribution around the MWC. Their expected velocity directions point away from Andromeda. Only a small fraction of HVSs ejected from Andromeda are expected to migrate to the Milky Way. However, it might be possible to detect them based on their velocity and trajectory orientation. Further, it would be interesting to extend this analysis to include estimates of the stellar ages and other astrophysical information of HVSs in order to gain further insight into their origin and migration history.

ACKNOWLEDGEMENTS

We acknowledge valuable discussions with Marius Neumann, Ferdinand Jünemann, David Clarke, Jelena Köhler, and Pranav Sampathkumar during the preparation of this paper.

The computations in this work were performed on the GPU cluster at Bielefeld University. We thank the Bielefeld *HPC.NRW* team for their support.

Malcolm Fairbairn was supported by the United Kingdom STFC grant ST/X000753/1.

This work was co-funded by the Erasmus+ programme of the European Union.

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 work made use of *ASTROPY* (<http://www.astropy.org>), a community-developed core PYTHON package for astronomy (Astropy Collaboration 2013, 2018), *NUMPY* (Harris et al. 2020), and *SCIPY* (Virtanen et al. 2020). All figures in this work were produced using *MATPLOTLIB* (Hunter 2007), *HEALPIX* (Gorski et al. 2005), and *TOPCAT* (Taylor 2005).

DATA AVAILABILITY

The simulation code and the simulation results are available as supplementary material and can also be found on GitHub (<https://github.com/lguelzow/On-Stellar-Migration-from-the-Andromeda-Galaxy>).

The data from Marchetti (2021) for the velocity analysis of Milky Way stars are available on T. Marchetti's website (<https://sites.google.com/view/tmarchetti/research>).

The acquisition of the *Gaia* DR3 data is described in Appendix A.

REFERENCES

- Abadi M. G., Navarro J. F., Steinmetz M., 2009, *ApJ*, 691, L63
 Astropy Collaboration, 2013, *A&A*, 558, A33
 Astropy Collaboration, 2018, *AJ*, 156, 123
 Bailer-Jones C. A. L., 2015, *PASP*, 127, 994
 Balkema A. A., de Haan L., 1974, *Ann. Probab.*, 2, 792
 Boubert D., Erkal D., Evans N. W., Izzard R. G., 2017, *MNRAS*, 469, 2151
 Brown W. R., Geller M. J., Kenyon S. J., Kurtz M. J., 2005, *ApJ*, 622, L33
 Brown W. R., Anderson J., Gnedin O. Y., Bond H. E., Geller M. J., Kenyon S. J., Livio M., 2010, *ApJ*, 719, L23
 Butcher J. C., 1987, *The Numerical Analysis of Ordinary Differential Equations: Runge-Kutta and General Linear Methods*. Wiley-Interscience, Hoboken, NJ

- Cautun M. et al., 2020, *MNRAS*, 494, 4291
 Correa Magnus L., Vasiliev E., 2021, *MNRAS*, 511, 2610
 Erkal D., Boubert D., Gualandris A., Evans N. W., Antonini F., 2019, *MNRAS*, 483, 2007
 Evans F., Marchetti T., Rossi E., Baggen J., Bloot S., 2021, *MNRAS*, 507, 4997
 Evans F. A., Marchetti T., Rossi E. M., 2022a, *MNRAS*, 512, 2350
 Evans F. A., Marchetti T., Rossi E. M., 2022b, *MNRAS*, 517, 3469
 Fabricius C. et al., 2021, *A&A*, 649, A5
 Gaia Collaboration et al., 2016, *A&A*, 595, A1
 Gaia Collaboration et al., 2018, *A&A*, 616, A11
 Gaia Collaboration et al., 2021, *A&A*, 649, A1
 Gaia Collaboration et al., 2023, *A&A*, 674, A1
 Ginsburg I., Brown W. R., Wegner G. A., 2013, preprint ([arXiv:1302.1899](https://arxiv.org/abs/1302.1899))
 Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, *ApJ*, 622, 759
 Groenewegen M. A. T., 2021, *A&A*, 645, A20
 Hairer E., Nørsett S., Wanner G., 2000, *Solving Ordinary Differential Equations I Nonstiff problems*, second edn. Springer, Berlin
 Harris C. R. et al., 2020, *Nature*, 585, 357
 Hills J. G., 1988, *Nature*, 331, 687
 Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
 Irrgang A., Kreuzer S., Heber U., Brown W., 2018a, *A&A*, 615, L5
 Irrgang A., Kreuzer S., Heber U., 2018b, *A&A*, 620, A48
 Kafle P. R., Sharma S., Lewis G. F., Bland-Hawthorn J., 2014, *ApJ*, 794, 59
 Kafle P. R., Sharma S., Lewis G. F., Robotham A. S. G., Driver S. P., 2018, *MNRAS*, 475, 4043
 Karukes E., Benito M., Iocco F., Trotta R., Geringer-Sameth A., 2020, *J. Cosmol. Astropart. Phys.*, 2020, 033
 Kopusov S. E. et al., 2020, *MNRAS*, 491, 2465
 Kreuzer S., Irrgang A., Heber U., 2020, *A&A*, 637, A53
 Lin Z., Xu Y., Hao C., Li Y., Liu D., Bian S., 2023, *ApJ*, 952, 64
 Lindegren L. et al., 2021, *A&A*, 649, A4
 Marchetti T., 2021, *MNRAS*, 503, 1374
 Marchetti T., Rossi E. M., Brown A. G. A., 2019, *MNRAS*, 490, 157
 Marchetti T., Evans F. A., Rossi E. M., 2022, *MNRAS*, 515, 767
 Monari G. et al., 2018, *A&A*, 616, L9
 Montanari F., Barrado D., García-Bellido J., 2019, *MNRAS*, 490, 5647
 Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
 Nuza S. E., Parisi F., Scannapieco C., Richter P., Gottlöber S., Steinmetz M., 2014, *MNRAS*, 441, 2593
 Peñarrubia J., Ma Y.-Z., Walker M. G., McConnachie A., 2014, *MNRAS*, 443, 2204
 Pickands J., 1975, *Ann. Stat.*, 3, 119
 Piffil T., Williams M., Steinmetz M., 2011, *A&A*, 535, A70
 Plummer H. C., 1911, *MNRAS*, 71, 460
 Sherwin B. D., Loeb A., O'Leary R. M., 2008, *MNRAS*, 386, 1179
 Tamm A., Tempel E., Tenjes P., Tihhonova O., Tuvikene T., 2012, *A&A*, 546, A4
 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
 Virtanen P. et al., 2020, *Nat. Methods*, 17, 261
 Wang B., Han Z., 2009, *A&A*, 508, L27
 Watkins L. L., Evans N. W., An J. H., 2010, *MNRAS*, 406, 264
 van der Marel R. P., Fardal M., Besla G., Beaton R. L., Sohn S. T., Anderson J., Brown T., Guhathakurta P., 2012, *ApJ*, 753, 8

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Stellar_Migration_from_Andromeda_Simulation_Code.tar

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: DATA SELECTION FROM GAIA DR3 CATALOGUE

The *Gaia* DR3 data are provided by the *Gaia* archive website at <https://gea.esac.esa.int/archive/>. We use the following prompt in the ADQL interface:

```
SELECT source_id, parallax, ra, dec, pmra, pmdec, radial_velocity, parallax_error, ra_error, dec_error, pmra_error, pmdec_error, radial_velocity_error, phot_g_mean_mag, nu_eff_used_in_astrometry, pseudocolour, ecl_lat, astrometric_params_solved
FROM gaiadr3.gaia_source
WHERE parallax >= 0.02
AND ruwe < 1.4
AND rv_nb_transits > 10
AND rv_expected_sig_to_noise >= 5
AND abs(parallax) > 2 * parallax_error
AND abs(pmra) > 5 * pmra_error
AND abs(pmdec) > 5 * pmdec_error
AND abs(radial_velocity) > 5 * radial_velocity_error
AND sqrt(power(radial_velocity, 2) + power((pmra/parallax) * 4.744213026, 2) + power((pmdec/parallax) * 4.744213026, 2)) >= 150
```

The SELECT line provides the columns of the data base used for the result table. The FROM line sets the data base from which the data are drawn. The WHERE and AND lines refer to further conditional statements for the data selection. Here, we use multiple conditions to ensure we obtain pure and accurate data from the *Gaia* catalogue. First, we require relative errors smaller than 20 percent for radial velocity and proper motion (pmra and pmdec). For parallax measurements, we initially use a more relaxed criterion of relative errors smaller than 50 percent. After this selection we remove the parallax offset present in the *Gaia* catalogue as detailed in Lindegren et al. (2021) and Groenewegen (2021). The parameters phot_g_mean_mag, nu_eff_used_in_astrometry, pseudocolor, ecl_lat and astrometric_params_solved are required for this. Once the parallax values are corrected, we apply the criterion of relative error smaller than 20 percent. This condition matches the parallax accuracy to the rest of the data and enables us to accurately estimate distances from parallaxes by simply inverting them (Bailer-Jones 2015). In addition to accuracy conditions on the parameters we use directly, we also select stars with criteria on the parameters ruwe (renormalized unit weight error), rv_nb_transits (number of epochs used to determine the radial velocity of each star), and rv_expected_sig_to_noise (expected signal-to-noise ratio in the combination of individual spectra used to determine the *Gaia* DR3 radial velocity). For details on these criteria, we refer to Marchetti et al. (2022) from where we have adopted them. Lastly, the final condition calculates the velocity magnitude of the object from its radial velocity and proper motion and only adds it to the result table if it is $\geq 150 \text{ km s}^{-1}$. Notably, this velocity magnitude is in the heliocentric rest frame. We choose this low threshold to make sure we do not miss any stars in the catalogue that have a velocity larger than 500 km s^{-1} in the Galactocentric rest frame.

APPENDIX B: COORDINATE TRANSFORMATION OF THE GAIA DR3 VELOCITIES

The data from the *Gaia* DR3 catalogue are given in the equatorial coordinate system. It utilizes the celestial sphere of the Earth. We use the ASTROPY function SKYCOORD to transform the star positions from the catalogue to Galactic coordinates. In order to obtain the distance to the Sun d_S and velocity vectors \mathbf{v}_{Gaia} of the measured HVSSs, we use the given parallax P and radial velocity v_{rad} as well as the proper motion velocities $\omega_{\text{RA}}, \omega_{\text{Dec}}$ in the direction of the equatorial coordinate components, right ascension (RA) and declination (Dec.). The parallax P in milliarcseconds (mas) is simply converted to the distance to the Sun d_S in kiloparsec (kpc) by taking the inverse.

To find the velocity vector of a given object in Galactic coordinates, we have to perform multiple coordinate transformations since we only have the previously mentioned measurement values. First, we determine the velocity vector of an HVS in equatorial Cartesian coordinates from the *Gaia* data using two coordinate rotations. This process is schematically displayed in Fig. B1. After these rotations and a transformation to spherical coordinates, we can again use the SKYCOORD function to convert the calculated velocity direction from equatorial coordinates to Galactic coordinates.

Initially, we construct the velocity vector \mathbf{v}_{Gaia} of a *Gaia* object in a Cartesian coordinate frame (x'', y'', z'') that uses the position vector \mathbf{r}_{Gaia} of the HVS in equatorial Cartesian coordinates as its x'' -axis and the equatorial coordinate directions RA and Dec. at the coordinates of the HVS as its y'' - and z'' -axes, respectively (see Fig. B1, top right). The velocity component $v_{\text{Gaia},x''}$ along the x'' -axis is simply the radial velocity of the HVS. We calculate the y'' - and z'' -components using the parallax and proper motion of the HVS

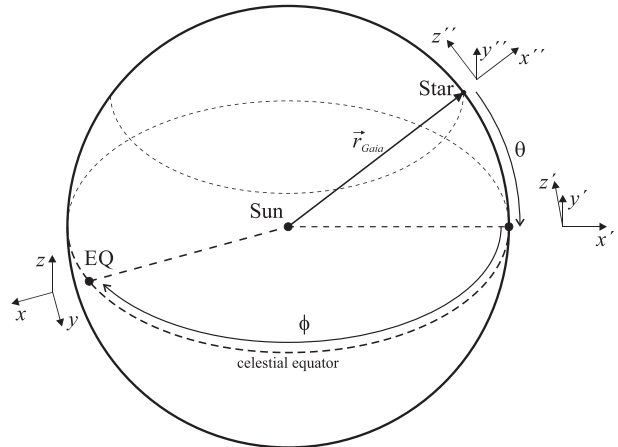


Figure B1. Schematic illustration of the transformation of the *Gaia* catalogue velocity data in hour angle equatorial coordinates. We transform the coordinate system (x'', y'', z'') determined by the position vector \mathbf{r}_{Gaia} of the HVS by rotating by $\theta = \text{Dec.}$ around the y'' -axis and by $\phi = -\text{RA}$ around the z'' -axis of the transitional coordinate system (x', y', z') . The z' -axis is slightly tilted for visual clarity. Now it matches the coordinate system (x, y, z) at the position $(0^\circ, 0^\circ)$, the vernal equinox (EQ). This coordinate system is aligned with the standard equatorial coordinate frame and gives us the velocity vector \mathbf{v}_{Gaia} .

with

$$\begin{aligned} v''_{\text{Gaia},y} &= \frac{\omega_{\text{RA}}}{P} \cdot \cos(\text{Dec.}) \cdot C, \\ v''_{\text{Gaia},z} &= \frac{\omega_{\text{Dec.}}}{P} \cdot C. \end{aligned} \quad (\text{B1})$$

We need to multiply by $\cos(\text{Dec.})$ in $v''_{\text{Gaia},y}$ since proper motion in the horizontal direction near the North pole of a spherical coordinate frame covers more real distance than near the equator of the coordinate frame in the same time frame. This is already included in the *Gaia* catalogue value of ω_{RA} . The constant $C \approx 4.744$ is needed for the conversion to km s^{-1} .

To find the velocity vector v_{Gaia} in the equatorial Cartesian frame (x, y, z) aligned with the standard right-handed equatorial coordinate frame, we rotate the coordinate frame (x'', y'', z'') to match (x, y, z) . This action is equivalent to rotating v''_{Gaia} with transposed rotation matrices. We use the transpose of the y -axis rotation matrix \mathbf{R}_y^T to rotate around the y'' -axis and the corresponding matrix \mathbf{R}_z^T for the rotation around the z' -axis of the resulting intermediate coordinate

frame (x', y', z') (see Fig. B1, centre right)

$$\begin{aligned} \mathbf{R}_y^T &= \begin{pmatrix} \cos \alpha & 0 & -\sin \alpha \\ 0 & 1 & 0 \\ \sin \alpha & 0 & \cos \alpha \end{pmatrix} \quad \text{and} \\ \mathbf{R}_z^T &= \begin{pmatrix} \cos \beta & \sin \beta & 0 \\ -\sin \beta & \cos \beta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \quad (\text{B2})$$

In this definition, α and β are arbitrary angles. We use the corresponding position coordinates $\mathbf{r}_{\text{Gaia}} = (\text{RA}, \text{Dec.})$ as rotation angles. We rotate first around the y'' -axis by the angle $\alpha = \text{Dec.}$ and subsequently around the z' -axis by the angle $\beta = -\text{RA}$. We need to use the negative of RA because the data are given in the hour angle system which is the left-handed equivalent to the standard equatorial frame.

Now, we can simply perform the transformation from Cartesian to spherical coordinates to find the point on the sphere in equatorial coordinates the velocity vector points at if placed in the coordinate origin. We transform the velocity direction into the Galactic coordinate frame using SKYCOORD.

This paper has been typeset from a $\text{T}_\text{E}\text{X}/\text{L}^{\text{A}}\text{T}_\text{E}\text{X}$ file prepared by the author.