Mon. Not. R. Astron. Soc. **000**, 1–17 (2016) Printed 30 May 2017 (MN LATEX style file v2.2)

An artificial neural network to discover Hypervelocity stars: Candidates in *Gaia* DR1/TGAS

T. Marchetti[1]⋆, E. M. Rossi[1], G. Kordopatis[2], A. G. A. Brown[1], A. Rimoldi[1], E. Starkenburg[3], K. Youakim[3] and R. Ashley[4]

[1]*Leiden Observatory, Leiden University, PO Box 9513 2300 RA Leiden, the Netherlands*
[2]*Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, France*
[3]*Leibniz-Institut fur Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany*
[4]*Department of Physics, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, UK*

30 May 2017

ABSTRACT

The paucity of hypervelocity stars (HVSs) known to date has severely hampered their potential to investigate the stellar population of the Galactic Centre and the Galactic Potential. The first *Gaia* data release (DR1, 2016 September 14) gives an opportunity to increase the current sample. The challenge is the disparity between the expected number of hypervelocity stars and that of bound background stars. We have applied a novel data mining algorithm based on machine learning techniques, an artificial neural network, to the Tycho-*Gaia* astrometric solution (TGAS) catalogue. With no pre-selection of data, we could exclude immediately ∼ 99% of the stars in the catalogue and find 80 candidates with more than 90% predicted probability to be HVSs, based only on their position, proper motions, and parallax. We have cross-checked our findings with other spectroscopic surveys, determining radial velocities for 30 and spectroscopic distances for 5 candidates. In addition, follow-up observations have been carried out at the Isaac Newton Telescope for 22 stars, for which we obtained radial velocities and distance estimates. We discover 14 stars with a total velocity in the Galactic rest frame > 400 km s^{-1}, and 5 of these have a probability > 50% of being unbound from the Milky Way. Tracing back their orbits in different Galactic potential models we find one possible unbound HVS with $v \sim 520$ km s^{-1}, 5 bound HVSs, and, notably, 5 runaway stars with median velocity between 400 and 780 km s^{-1}. At the moment, uncertainties in the distance estimates and ages are too large to confirm the nature of our candidates by narrowing down their ejection location, and we wait for future *Gaia* releases to validate the quality of our sample. This test successfully demonstrates the feasibility of our new data mining routine.

Key words: Spectroscopic surveys, Galaxy: Centre, Galaxy: kinematics and dynamics, Galaxy: stellar

1 INTRODUCTION

Observationally, hypervelocity stars (HVSs) are stars that can reach radial velocities in excess of the Galactic escape speed at their location, and whose trajectories are consistent with a Galactic Centre (GC) origin (Brown et al. 2005). Currently, about ∼ 20 unbound stars have been discovered (Brown et al. 2014): most of them are late B-type stars ($\sim 2.5-4$ M$_\odot$) detected in the outer halo (but note Zheng et al. 2014) with velocities between $\sim 300-700$ km s^{-1} (see Brown 2015, for a review). These stars are in principle unique tools to gather information on the Galactic Centre stellar population and dynamics (Madigan et al. 2014; Zhang et al. 2013, e.g.) and on the Galactic potential (e.g. Gnedin et al. 2005; Yu & Madau 2007; Perets et al. 2009). Using current data, a first proof of principle of how to get joint constraints on both environments was published in Rossi et al. (2017), and attempts to constrain the dark matter halo alone were performed by Sesana et al. (2007) and Fragione & Loeb (2016)[1]. These analyses however are severely hampered by the quality and quantity of the current small and rather biased sample.

So far the most successful observational strategy has been to spectroscopically select late B-type stars in the outer halo. Since the stellar halo is dominated by an old stellar population, young stars

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

[1] See also Gnedin et al. (2010), who uses the velocity dispersion of halo stars from the hypervelocity star survey.

likely come from other star-forming regions in the Galaxy, and a late B-type star has a long enough life-time ($\sim$ 100 – 300 Myr) to be able to travel to the outer halo from the Galactic Centre if its velocity is hundreds km s^{-1}. Most of the confirmed unbound HVSs have only radial velocity measurements and uncertainties in their photometric distances are large. Proper motions have been acquired with the Hubble Space Telescope for 16 high velocity stars (Brown et al. 2015), but even if the GC origin was confirmed for 13 of these objects, uncertainties are still too large to precisely constrain their origin, and therefore to identify them as HVSs.

Recent years have seen an increasing effort to identify low mass HVSs in the inner Galactic halo. These searches use high proper motion or high radial velocity criteria, as it is not possible to spectroscopically single out these low mass stars in the halo, as is done for B-type HVSs. A few tens of candidates have been reported, but the large majority are bound and/or consistent with Galactic disc origin (e.g. Li et al. 2012; Palladino et al. 2014; Ziegerer et al. 2015; Vickers et al. 2015; Hawkins et al. 2015; Zhang et al. 2016; Ziegerer et al. 2017). Positive identification is prevented by large distance and proper motion uncertainties.

Major observational advancements in the field are therefore expected from the data taken by the ESA mission *Gaia*, launched on the 19th of December 2013 (Gaia Collaboration et al. 2016b,a). *Gaia* will attain an unparalleled astrometric measurement precision for a total of $\sim 10^9$ stars in the Galaxy. In the end-of-the-mission data release, we anticipate a few hundred (a few thousand) HVSs within 10 kpc from us, in the mass range $\sim 1-10\,M_\odot$, with relative error on total proper motion $< 1\%$ ($< 10\%$), and that radial velocities will be measured for a subsample of these (Marchetti et al. in preparation). For brighter HVSs, accurate *Gaia* parallaxes can eliminate the large distance uncertainties in the existing sample, and for fainter stars calibrated photometric distances may eventually be used.

The first data release (DR1) happened on September 14, 2016, and it contains the five-parameter astrometric solution (positions, parallaxes, and proper motions) for a subset of $\sim 2 \times 10^6$ stars in common between the Tycho-2 Catalogue and *Gaia* (TGAS catalogue, Michalik et al. 2015; Lindegren et al. 2016). Radial velocity information is notably missing. Our expectation is that between 0.1– and a few unbound HVSs may be expected to be present in the catalogue, depending on the unknown mass distribution and star formation history in the Galactic Centre (Marchetti et al. in preparation).

In this paper, we report a systematic search for HVSs in DR1. We use an artificial neural network (§2), which is first applied to the TGAS subset of the *Gaia* catalogue without any prior constraints placed on stellar properties to select HVS candidates (§3). We then cross check our sample of best candidates with published spectral catalogues to acquire radial velocity and spectroscopic distance information (§4). We further proceed to describe the radial velocity follow-up observations for candidates with no published radial velocity and observable by the Isaac Newton Telescope (INT) (§4.2). In §5 we describe our Bayesian approach to determine distances, and then in §6 we present our results for HVS candidates in terms of total velocity and ejection location. We sort and characterize candidates in §7, and discuss their implications in §8.

2 DATA MINING ALGORITHM

Hypervelocity stars are rare objects, that occur in the Galaxy at an uncertain rate roughly between $10^{-5}-10^{-4}$ yr^{-1} (Hills 1988; Perets et al. 2007; Zhang et al. 2013; Brown et al. 2014). Considering the magnitude limit of *Gaia* and different assumptions on the population of binaries in the GC, such a rate implies only $\sim 0.1-1$ HVSs for every 10^6 stars in the final *Gaia* catalogue (Marchetti et al. in preparation). In particular for the TGAS catalogue, we expect to find at most a few HVSs (Marchetti et al. in preparation), although a larger number of slower stars generated via the same mechanism (called *"bound HVSs"*) are also expected (Bromley et al. 2006; Kenyon et al. 2008). Thus, *Gaia* can deliver a HVS sample that represents a huge leap in data quality and quantity, but building it requires careful data mining, especially since radial velocity measurements are currently missing.

The TGAS subset of *Gaia* DR1 provides the five-parameter astrometric solution for roughly two million objects, therefore we choose to build a data mining routine based only on the astrometric properties of the stars: position on the sky (α, δ), parallax ϖ, and proper motions $\mu_{\alpha*}, \mu_\delta$. This approach allows us to not make any a priori assumption on the stellar nature of HVSs, avoiding photometric and metallicity cuts which might bias our search towards particular spectral types, and lead to a sample which may not reflect the properties of the binary population in the Galactic Centre. Recent studies have shown indeed how the GC is a complex environment in which different stellar populations coexist and interact, and many properties (mass function, metallicity, binarity) are missing or poorly constrained due to observational limitations (see Genzel et al. (2010) for an exhaustive review). The nuclear star cluster surrounding the central massive black hole has also undergone several star formation episodes throughout its lifetime, which might have changed and influenced the stellar population and mass function (Genzel et al. 2010; Pfuhl et al. 2011).

We have therefore chosen to build a data mining routine based on a machine learning algorithm, an *artificial neural network*. Our chosen approach is a *supervised learning* problem: we present the algorithm with examples and their desired output (*training set*), and we let the algorithm learn the best function mapping inputs into outputs. We decided for a binary classification problem: the desired output of the algorithm is 0 for a "normal" background star, and 1 for a HVS. When we apply the classification rule to a new unlabelled example we can then interpret its output as the probability of that star being a HVS (Saerens et al. 2002).

We now start introducing neural networks, with a brief summary on the main idea behind this algorithm. Next in §2.2 we discuss how we build our training set, and finally in §2.3 and §2.4 how we optimize and determine the performance of the network based on the results on subsets of the data which were not used for the training.

2.1 Artificial Neural Networks

Artificial neural networks have been largely used in different branches of science for their ability to provide highly non-linear mapping functions, and for their intrinsic capacity to generalize: to provide reasonable outputs for examples not encountered while training the algorithm (see Haykin (2009) for an exhaustive explanation of neural networks). This latter property is particularly important for our goal, since our training set consists of mock data (see §2.2), and therefore we want to be flexible enough to find HVSs even if the real population is not perfectly represented by our simulations, which necessarily rely on several hypotheses and assumptions (see §2.2).

We have developed from scratch an artificial neural network with five input units (the astrometric parameters), two hidden lay-

ers of neurons, and a single output neuron for binary classification. Each neuron of the network is a computational unit which outputs a non-linear function[2] $f(v)$, where v is a linear combination of the j-th input M-dimensional vector $\boldsymbol{x}^{(j)}$ with some weight vector $\boldsymbol{\omega}$:

$$v_j(\boldsymbol{x}^{(j)};\boldsymbol{\omega}) = x_0\omega_0 + \sum_{i=1}^{\mathrm{M}} x_i^{(j)}\omega_i, \tag{1}$$

where $x_0 \equiv 1$ is referred to as the *bias unit*. In analogy with the brain architecture, the components ω_i are usually referred to as *synaptic weights*. A typical choice for f is a sigmoid function. We choose:

$$f(v) = a\tanh(bv), \tag{2}$$

with $a = 1.7159$ and $b = 2/3$. This activation function outputs real numbers in the interval $[-a, a]$, and satisfies several useful properties: it is an odd function of its argument; $f(1) = 1$ and $f(-1) = -1$; its slope at the origin is close to unity; and its second derivative attains its maximum value at $x = 1$. This choice has been shown to yield faster convergence than the usual logistic function, avoiding driving the hidden neurons into saturation (LeCun 1993).

For neurons in the first hidden layer the input $\boldsymbol{x}^{(j)}$ is just the data vector containing the $\mathrm{M} = 5$ astrometric parameters for the j-th training example: $\boldsymbol{x}^{(j)} = (\alpha_j, \delta_j, \varpi_j, \mu_{\alpha*j}, \mu_{\delta j})$, therefore the summation in Equation 1 extends over $i = 1, \ldots, 5$. For neurons in the second layer the input $\boldsymbol{x}^{(j)}$ is the M_1-dimensional vector output by the first layer of M_1 neurons, and the summation extends to $\mathrm{M} = \mathrm{M}_1$. Finally, the single neuron in the output layer takes in input a M_2-dimensional vector, with M_2 equal to the number of neurons in the second hidden layer, and in summation $\mathrm{M} = \mathrm{M}_2$. We call $D_j(\boldsymbol{\omega}) \in \mathbb{R}$ the final output of the neural network for the j-th example.

The training process consists in finding the vector of synaptic weights $\boldsymbol{\omega}$ which minimizes the total cost function

$$J(\boldsymbol{\omega}) \propto \sum_{j=1}^{N} (D_j(\boldsymbol{\omega}) - y_j)^2, \tag{3}$$

which is just the sum over all the N examples of the squared difference between the output of the neural network $D_j(\boldsymbol{\omega})$ and the desired output y_j of the labelled training example. The value of each synaptic weight is initialized with a random number drawn from a uniform distribution in the interval $[-\sigma_\omega, \sigma_\omega]$, with $\sigma_\omega = m_*^{-1/2}$, where m_* is the number of connections feeding into the corresponding layer of neurons (LeCun et al. 2012). The weights optimization is then performed with an adaptive stochastic (online) gradient descent method, using a specific learning rate η_k for each synaptic weight: the AdaGrad implementation (Duchi et al. 2011). We use the following iterative rule for the t-th update of the k-th weight ω_k (Singh et al. 2015):

$$\Delta\omega_k(t) = -\eta_k(t)g_k(t) = -\frac{\eta_0}{\sqrt{\sum_{i=1}^{t}(g_k(i))^2}}g_k(t), \tag{4}$$

where $\eta_0 > 0$ is called the *global learning rate*, $\boldsymbol{g}$ is the gradient of the cost function in Equation 3 (derivatives with respect to the weight vector $\boldsymbol{\omega}$), and the denominator is the norm of all the gradients of the previous iterations. The adopted value for η_0 is discussed in §2.3, while the gradient of the cost function is estimated with a back-propagation algorithm (see LeCun et al. (2012) for tips on an efficient implementation, essential when dealing with large datasets).

[2] In the following, we will use superscripts in round brackets to refer to a particular vector, and subscripts to specify its components.

2.2 Building the Training Set

We train the artificial neural network on a simulated end-of-mission *Gaia* catalogue for the Galaxy: the *Gaia* Universe Model Snapshot (GUMS, Robin et al. 2012), where we inject *mock* HVS data with errors on all astrometric and photometric measurements. A detailed description of how we construct our mock HVS will be the focus of an upcoming paper, and here we only briefly summarize our procedure. In the following we will adopt the Hills mechanism for modelling our mock population of HVSs, involving the disruption of a binary star by the Massive Black Hole (MBH) at the centre of our Galaxy (Hills 1988).

We explore the space (l, b, d, M) to populate each position in Galactic coordinates on the sky (l, b) with stars in a mass range $M \in [0.1 - 9]\ M_\odot$ and in a distance range $d \in [0, 40]$ kpc from us. We adopt a step of $\sim 9°$ in Galactic longitude l, $\sim 4.5°$ in Galactic latitude b, ~ 1 kpc in distance r, and $\sim 0.2\ \mathrm{M}_\odot$ in mass. We draw velocities from an ejection velocity distribution which analytically depends on the properties of the original binary approaching the massive black hole (Sari et al. 2010; Kobayashi et al. 2012; Rossi et al. 2014)[3]:

$$v_{\mathrm{ej}} = \sqrt{\frac{2Gm_c}{a}}\left(\frac{M_\bullet}{m_T}\right)^{\frac{1}{6}}, \tag{5}$$

where m_c is the mass of the star that remains bound to the MBH after the binary is disrupted, $m_T = M + m_c$ is the total mass of the disrupted binary, and $M_\bullet = 4.0 \times 10^6\ \mathrm{M}_\odot$ is the mass of the MBH in our Galaxy (Ghez et al. 2008; Gillessen et al. 2009; Meyer et al. 2012). Following Rossi et al. (2014, 2017), we model binary distributions for semi-major axis a and mass ratio q as power-laws: $f_a \propto a^\alpha$, $f_q \propto q^\gamma$, with exponents $\alpha = -1$ (Öpik's law, Öpik 1924) and $\gamma = -3.5$. This combination has been shown to result in a good fit between the observed sample of late B type HVSs in Brown et al. (2014) and the prediction of the Hills mechanism for reasonable choices of Milky Way potentials (Rossi et al. 2017). The total velocity v of the HVS is then computed decelerating the star in a given Galactic potential (refer to §6.2, Equations 12-14 for details on the adopted fiducial Milky Way potential).

For each star we compute the combination of proper motions and radial velocity which are consistent with an object moving radially away from the Galactic Centre, and we correct those values for the motion of the Sun and of the local standard of rest (LSR) (Schönrich 2012). We then roughly estimate the flight time from the GC to the given position in Galactocentric coordinates r_{GC} as $t_F = r_{\mathrm{GC}}/v_F$, where v_F is an effective velocity equal to the arithmetic mean between the ejection velocity and the decelerated velocity at the star's position. The age of the star is then computed summing the flight time and the age of the star at its ejection. The latter is computed as a random fraction of its main sequence (MS) lifetime (Brown et al. 2014), and the time spent on the MS is computed using analytic formulae in Hurley et al. (2000). We

[3] Rigorously, there should be a numerical factor in front of Equation 5, depending on the detailed geometry of the three-body encounter. This factor has been shown to be ~ 1 when averaged over the binary's phase (Rossi et al. 2014).

assume a super-solar metallicity $[M/H] = 0.4$, which corresponds to the mean value of the distribution in the GC (Do et al. 2015). Each star is evolved up to its age using the fast parametric stellar evolution code SeBa (Portegies Zwart & Verbunt 1996; Portegies Zwart et al. 2009) to obtain its radius, effective temperature, and mass, which we use to identify the best-matched stellar spectrum from the BaSeL 3.1 stellar spectral energy distribution (SED) libraries (Westera & Buser 2003) via chi-squared minimization. For each position of the sky we assess dust extinction using a three-dimensional Galactic dust model (Drimmel et al. 2003), and integrating the reddened flux in the respective passbands we estimate the magnitudes in the *Gaia G* band and in the Johnson-Cousins V, I_c bands. We finally use the python toolkit PyGaia[4] to estimate the errors on the astrometry with which *Gaia* would observe these objects. The errors are functions of the magnitude of the star, its color index $V - I_c$, and the ecliptic latitude β, the latter determining the number of observations of the object according to the satellite's scanning strategy.

Parallax and proper motions of each source are then replaced by drawing a random number from a Gaussian distribution centred on the nominal value and with standard deviation equal to the estimated uncertainty. This approach has two main advantages: it allows us to obtain negative parallaxes (which are present in the real *Gaia* catalogue) for faint objects with non-negligible relative errors on parallax; and it helps us mitigate the effect of the spatial grid in distance used for generating mock stars, preventing the algorithm from driving the learning rule towards discrete, fixed values in parallax.

We can therefore build a mock catalogue of HVSs, which we use for the training of the artificial neural network. We combine mock positions, parallaxes and proper motions of HVSs and "normal" background stars randomly picked from the GUMS in a single stellar catalogue, consisting of a total of $\sim 2.5\times10^6$ objects ($\sim 25\%$ HVSs, label = 1; $\sim 75\%$ *Gaia* stars, label = 0). We randomly split stars of the catalogue into a *training* set ($\sim 60\%$ of the catalogue), a *cross-validation* set ($\sim 20\%$ of the catalogue), and a *test* set ($\sim 20\%$ of the catalogue). The training set consists of the examples the algorithm will learn from, the cross-validation set is used to optimize hyperparameters (see §2.3), while we use the test set to determine the performance of the neural network (see §2.4). The use of different examples for performing these tasks is extremely useful to prevent overfitting and to ensure generalization. All features (five parameters) of the complete catalogue have been scaled in such a way to have mean of 0 and variance of 1, to achieve a faster convergence of the stochastic gradient descent algorithm (LeCun et al. 2012).

2.3 Optimization of the Algorithm

The effectiveness of a neural network, as the majority of machine learning algorithms, critically depends on the choice of the so-called *hyperparameters*, several parameters that need to be carefully tuned in order to achieve the best compromise between the algorithm performance, the time needed for its training, and its ability to generalize to new input data. We identify three hyperparameters in our algorithm: the number of neurons in the first hidden layer M_1, the number of neurons in the second hidden layer M_2, and the global learning rate η_0 for the adaptive stochastic gradient descent (see Equation 4).

[4] `https://github.com/agabrown/PyGaia`

A systematic grid search in the hyperparameter space to determine the best combination is not feasible because of time limitations and computational power. We use the pyswarm[5] implementation of a Particle Swarm Optimization (PSO) algorithm (Kennedy & Eberhart 1995) to explore the space (M_1, M_2, η_0) with 20 test particles. The algorithm iteratively adjusts particles' positions towards the minimum value attained by the cost function, with a velocity proportional to the distance from this extremum. Since each iteration involves the full training of the algorithm in order to determine the value of the cost function, we choose to apply PSO to a limited sample of the training set (1000 random training examples), and then we select the combination of parameters which results in the best performance on the full cross-validation set, defined in terms of the Matthews correlation coefficient MCC (Matthews 1975, see next subsection). The PSO algorithm converges to the following values: $M_1 = 119$, $M_2 = 95$, $\eta_0 = 0.071$[6].

2.4 Performance of the Algorithm

As mentioned before, we choose a stochastic gradient descent optimization to minimize the global cost function. Because of the intrinsic randomness of this algorithm, we train the neural network several times on the complete training set, shuffling the order of the presented example units during each training. Plotting learning curves (the value of the cost function versus the number of training examples presented to the network), we find that 8 complete iterations are enough to reach a minimum in both the training and cross-validation cost functions, again confirming that overfitting is not an issue.

We determine the performance of the algorithm on the test set by computing two different error metrics: the *Matthews correlation coefficient* MCC (Matthews 1975) and the F_1 *score*. Calling TP and TN (FP and FN) respectively the number of true (false) positives and negatives of the confusion matrix on the test set, error metrics are computed as:

$$\mathrm{F_1} \equiv 2\frac{\mathrm{PR}}{\mathrm{P+R}}, \tag{6}$$

$$\mathrm{MCC} \equiv \frac{\mathrm{TP\ TN - FP\ FN}}{\sqrt{(\mathrm{TP+FP})(\mathrm{TP+FN})(\mathrm{TN+FP})(\mathrm{TN+FN})}}, \tag{7}$$

where P and R are called, respectively, *precision* and *recall*, and they are defined as $\mathrm{P} \equiv \mathrm{TP}/(\mathrm{TP+FP})$, $\mathrm{R} \equiv \mathrm{TP}/(\mathrm{TP+FN})$. The $\mathrm{F_1}$ score assumes values in $[0, 1]$ while the MCC in $[-1, 1]$, and in both cases a value of 1 corresponds to a perfect classifier (diagonal confusion matrix). At the end of the training, we obtain the following values on the test set: $\mathrm{F_1} \sim \mathrm{MCC} \simeq 0.95$.

3 APPLICATION TO *GAIA* DR1

Once we have fully trained the neural network on the training set, determining the optimal values for the synaptic weights, we apply the classification rule to real unlabelled data to search for HVS

[5] `https://github.com/tisimst/pyswarm/`

[6] We initially included the *regularization parameter* λ as a 4th hyperparameter, but due to time limitation with the PSO we decided to discard it, since several tests showed that it always converged to values close to zero. A value $\lambda \sim 0$ is an indication that the algorithm is not overfitting the training set.

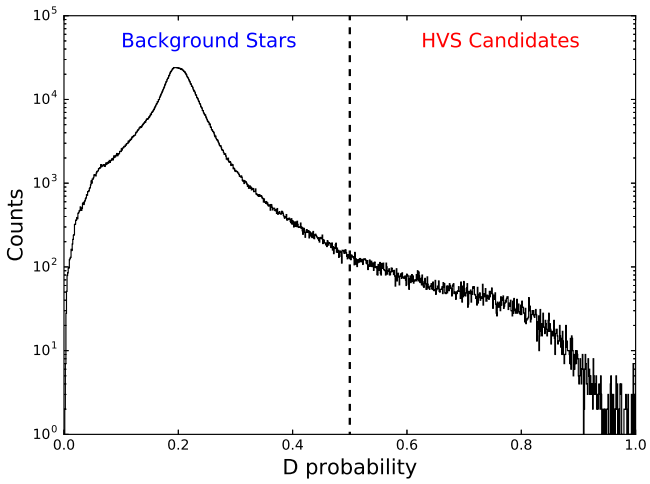

Figure 1. Histogram of the probability D of an object of being a HVS (output of the neural network), for all ∼ 2 million stars in the TGAS subset of *Gaia* DR1. A dashed vertical line marks the decision boundary $D = 0.5$.

candidates. The application of the neural network to the full TGAS subset of *Gaia* DR1 (2057050 sources) results in 22263 stars with a predicted probability > 50% of being a HVS, ∼ 1% of the original dataset. The histogram of the output probability D given by the neural network on the full TGAS catalogue is shown in Figure 1. To further reduce the sample of HVS candidates and to have reliable distance determinations, we filter out stars with a relative error on parallax $|\sigma_\varpi/\varpi| > 1$, obtaining a total of 8175 objects (∼ 0.4% of the original catalogue).

In these first cuts no information on the measured uncertainties is used to determine the probability of a star being a HVS. We subsequently include errors with a Monte Carlo (MC) simulation, randomly drawing one thousand realizations of the astrometry (parallax and proper motions) of each star from a Gaussian distribution centred on the nominal mean value and with a standard deviation equal to the corresponding quoted random uncertainty. This allows us to get for each star in TGAS a probability distribution of the output D of the neural network, which can then be characterized by its mean $\bar{D}$ and standard deviation σ_D. As a final cut, we select only stars with $\bar{D} - \sigma_D > 0.9$, for a total of 80 best HVS candidates, ∼ 0.004% of the original catalogue size.

We stress that all our cuts rely on the astrometry of the objects, without any prior assumption on the spectral type, photometry or more in general stellar properties of the selected best sample, and without any information on radial velocities.

4 ACQUIRING SPECTRAL INFORMATION

To confirm or reject a candidate in our quest for HVSs, a measure of the star *total* velocity is necessary. In the following, we will describe how we obtained reliable heliocentric radial velocities (HRVs) for 47 stars out of the 80 candidates.

4.1 Catalogue cross-matching

Our final sample has been cross-matched with several spectroscopic surveys of the Milky Way, covering both the Northern and Southern hemisphere[7]. We find a total of 30 stars in common: a subsample of these (5 stars) have both radial velocity and spectroscopic distance from the RAdial Velocity Experiment (RAVE) DR4 and/or DR5 (Kordopatis et al. 2013a; Kunder et al. 2017).

4.2 Follow-up observations with the INT

We successfully applied for director's discretionary time at the Isaac Newton Telescope (INT) in La Palma, Canary Islands, where we followed up spectroscopically 22 HVS candidates on the night of the 5th of October, 2016. We used the Intermediate Dispersion Spectrograph (IDS) with the RED+2 CCD, in combination with the R1200R grating, a 1.35" slit width, and the GG495 sorting order filter. This set-up provided an effective spectral range of ∼ 8000 – 9150 Å and a resolution at 7000 Å of 6731 over 2 pixels at the detector. We ensured that all observed spectra had a S/N of at least 50.

4.2.1 Spectra reduction

The spectra were reduced using the Image Reduction and Analysis Facility (IRAF, Tody 1986) software package. The reduction procedure included pre-processing (bias and flat field corrections), spectrum extraction, wavelength calibration, heliocentric radial velocity correction, and continuum normalisation.

4.2.2 Radial velocities, atmospheric parameters and spectroscopic distance determination

A first pass for radial velocity determination is performed by using the python routine pyasl.crosscorrRV, adopting a Solar template as reference, and errors in radial velocities are obtained following Zucker (2003). In order to obtain the effective temperature, surface gravity and metallicity of the stars, the same pipeline as the one used in RAVE (Kordopatis et al. 2011a, 2013a) has been applied to the spectra. This implies keeping only the wavelength range $\lambda\lambda = [8450.80-8746.55]$, removing the cores of the Calcium triplet lines (to avoid a mismatch between the synthetic templates used by the pipeline, computed assuming Local Thermodynamical Equilibrium, and the cores of the lines formed in Non LTE), and convolving the observations to a resolution of $R = 7500$. The output of the pipeline is then calibrated using the formulas presented in Kunder et al. (2017).

Our final radial velocities are obtained through the cross-correlation of a synthetic spectrum of the best-fit parameters to the observed spectrum. This cross-correlation is done with the package fxcor in IRAF (Tody 1986). Both the observed and synthesized spectrum are continuum normalized before cross-correlation and we use a Gaussian fit to all points with a correlation of 0.5 or higher to determine the radial velocity and its corresponding measurement uncertainty. During the observations a sample of 14 radial velocity standard stars from Soubiran et al. (2013) were observed with the same setup and matched closely in sky position to our program targets to check the accuracy of our determined radial velocities. We find that there is a good agreement between the literature values and our radial velocities. A mean offset of ∼ 0.1 km s^{-1}assures us

[7] RAVE DR4 and DR5 (Kordopatis et al. 2013a; Kunder et al. 2017), *Gaia*-ESO DR2 (Gilmore et al. 2012; Randich et al. 2013), LAMOST DR1 and DR2 (Cui et al. 2012), GALAH (Martell et al. 2017), APOGEE DR13 (Zasowski et al. 2013).

that there are no significant systematic effects. However, the rms variance between the literature values and our radial velocity determinations of 2.7 km s^{-1}is significantly larger than the median measurement uncertainty in the cross-correlation alone, which is only 1.1 km s^{-1}. We thus adopt an uncertainty floor of 2.5 km s^{-1}and add this in quadrature to our measurement uncertainties. Although we believe the radial velocities derived in this second iteration to be more precise than the first pass radial velocities due to the use of a synthetic spectrum that fits the stellar parameters, we note that the results presented in this paper are robust to the use of either set of radial velocities.

To obtain the spectroscopic distances of the stars, the calibrated stellar parameters are projected on Padova isochrones spanning ages from 100 Myr to 13.5 Gyr, with a step of 0.1 Gyr and a metallicity range between −2.2 dex and +0.2 dex. This allows us to obtain the absolute magnitudes in several photometric bands as in Kordopatis et al. (2011b, 2013c, 2015), and an estimation of the age of the stars as in Kordopatis et al. (2016); Magrini et al. (2017). The distances are then obtained using the distance modulus in the J band, and assuming $A_J = 0.709\ E(B-V)$ (Schlafly & Finkbeiner 2011), where $E(B-V)$ are the Schlegel extinctions towards each line-of-sight.

Kinematic properties from *Gaia* TGAS, radial velocities and stellar parameters derived from spectra of observed HVS candidates are presented in Table 1. For a precise cross-match with future *Gaia* releases and other Milky Way surveys, in Appendix A we report the *Gaia* and Hipparcos identifier of all the observed sources. We note that for 4 stars out of 22, the pipeline has not converged (quality flag $F = 1$, see Table 1) and therefore are excluded from the following analysis. Furthermore, visual inspection of TYC 2292-1267-1 (quality flag $F = 3$), shows a clear mismatch between the observed spectrum and the fitted template, and therefore was discarded as well.

The metallicity and mass distribution are shown, respectively, in Figure 2 and 3. The mean metallicity of our sample is −1.2 dex, consistent with the inner Galactic halo distribution, dashed (Chiba & Beers 2000) and dot-dashed (Kordopatis et al. 2013b) lines, but a total of 6 stars have $[M/H] > -0.5$ dex, and one candidate, TYC 3945-1023-1, has $[M/H] = -0.02 \pm 0.12$ dex. Most of the stars have masses slightly below the Solar value, with a peak of the distribution at $M \sim 0.85\ \mathrm{M}_\odot$, and a single star with $M \sim 2\ \mathrm{M}_\odot$: TYC 4032-1542-1. We can see that our sample is very different from the late B-type HVS candidates discovered in Brown et al. (2014). Considering the age estimates in Table 1, we note that the peak of the mass distribution is at the main-sequence turn-off of the stellar halo. Stars of this type have been used to trace the stellar halo because of their luminosity (e.g. Cignoni et al. 2007).

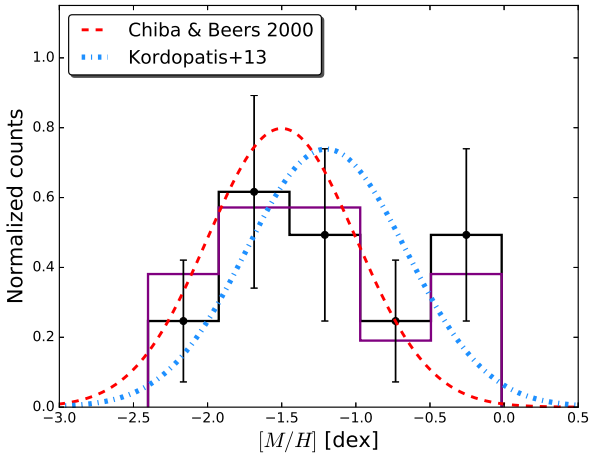

Figure 2. Normalized $[M/H]$ distribution for the observed HVS candidates, with error bars computed assuming Poisson noise. For a visual comparison, we overplot with a red dashed (blue dot-dashed) line the inner stellar halo metallicity, modelled as Gaussian with mean and standard deviation from Chiba & Beers (2000) (Kordopatis et al. (2013b)). Purple line shows the normalized $[M/H]$ distribution for high-velocity candidates (see Table 2).

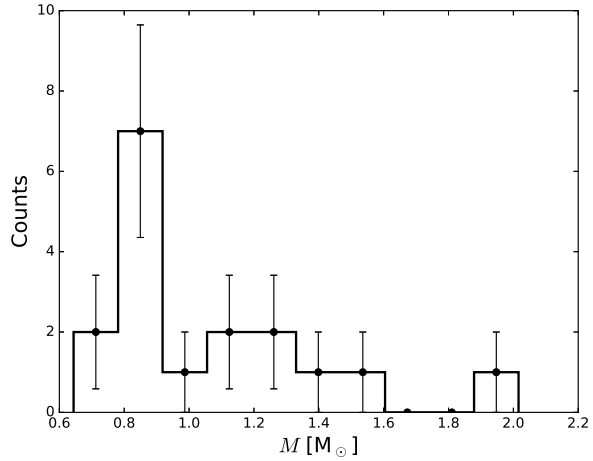

Figure 3. Mass distribution for the observed HVS candidates, with error bars computed assuming Poisson noise. The peak of the distribution is ∼ 0.85 $\mathrm{M}_\odot$.

5 DISTANCE ESTIMATION

Most of the stars in *Gaia* DR1 have non-negligible parallax errors. Therefore simply estimating distances as the inverse of parallax leads to biased results due to this highly non-linear transformation (Bailer-Jones 2015; Astraatmadja & Bailer-Jones 2016a). Additionally it can not be applied to negative parallaxes, which are present in our sample. In order to correctly take into account correlations between astrometric parameters supplied by the *Gaia* catalogue (parameter correlations may have an important impact on our results since we are implementing Monte Carlo simulations), we choose not to use the distance catalogue presented in Astraatmadja & Bailer-Jones (2016b), but to implement our own Bayesian approach, generalizing their method and considering covariances.

Assuming Gaussian noise for astrometric parameters, we model the likelihood for the triplet $\{\mu_{\alpha*}, \mu_\delta, \varpi\}$ as a multivariate normal distribution with mean vector:

$$\bar{x} = (\mu_{\alpha*},\ \mu_\delta,\ 1/d), \tag{8}$$

and with covariance matrix:

$$\Sigma = \begin{pmatrix} \sigma^2_{\mu_{\alpha*}} & \sigma_{\mu_{\alpha*}}\sigma_{\mu_\delta}\rho_{\mu_{\alpha*},\mu_\delta} & \sigma_{\mu_{\alpha*}}\sigma_{\varpi}\rho_{\mu_{\alpha*},\varpi} \\ \sigma_{\mu_{\alpha*}}\sigma_{\mu_\delta}\rho_{\mu_{\alpha*},\mu_\delta} & \sigma^2_{\mu_\delta} & \sigma_{\mu_\delta}\sigma_{\varpi}\rho_{\mu_\delta,\varpi} \\ \sigma_{\mu_{\alpha*}}\sigma_{\varpi}\rho_{\mu_{\alpha*},\varpi} & \sigma_{\mu_\delta}\sigma_{\varpi}\rho_{\mu_\delta,\mu_\varpi} & \sigma^2_{\varpi} \end{pmatrix}, \tag{9}$$

Table 1. Kinematic and observational properties of 22 HVS candidates spectroscopically followed-up with the INT telescope.

Tycho 2 ID	(RA, dec) (deg)	ϖ (mas)	$\mu_{\alpha*}$ (mas yr^{-1})	μ_δ (mas yr^{-1})	HRV (km s^{-1})	$T_{\rm eff}$ (K)	log g (cm s^{-2})	$[M/H]$ (dex)	$d_{\rm spec}$ (pc)	M ($M_\odot$)	$t_{\rm age}$ (Gyr)	F
2282-208-1	(16.81855, 33.66159)	2.17 ± 0.31	202.643 ± 1.213	−62.458 ± 0.398	−0.61 ± 1.29	5936 ± 136	3.8 ± 0.2	−1.35 ± 0.19	606 ± 152	0.92 ± 0.17	10.4 ± 3.8	1
2292-1267-1	(20.86832, 31.78668)	1.78 ± 0.35	90.782 ± 0.969	−15.275 ± 0.644	158.93 ± 5.99	7861 ± 83	4.0 ± 0.2	−0.20 ± 0.12	340 ± 69	1.70 ± 0.14	0.9 ± 0.2	3
2298-66-1	(25.30039, 33.51859)	2.45 ± 0.34	178.060 ± 1.213	−19.060 ± 0.319	−31.66 ± 2.78	5925 ± 328	3.8 ± 0.5	−2.08 ± 0.26	754 ± 569	0.95 ± 0.23	8.2 ± 4.5	0
2320-470-1	(31.29, 35.6289)	2.06 ± 0.27	106.443 ± 0.967	6.138 ± 0.290	−43.08 ± 1.32	5730 ± 214	3.4 ± 0.5	−3.29 ± 0.27	1240 ± 650	1.00 ± 0.21	6.9 ± 4.0	1
2376-691-1	(66.43652, 33.59088)	1.17 ± 0.29	62.060 ± 2.077	−9.137 ± 1.547	22.02 ± 1.63	5260 ± 74	3.5 ± 0.2	−0.67 ± 0.11	249 ± 64	1.22 ± 0.19	4.8 ± 3.8	2
2393-1001-1[8]	(78.45391, 32.03592)	2.21 ± 0.28	121.797 ± 1.710	−46.605 ± 1.158	−106.50 ± 0.94	4651 ± 1.158	0.6 ± 0.2	−2.40 ± 0.14	3036 ± 462	0.85 ± 0.26	7.6 ± 2.2	0
2818-556-1	(23.79684, 40.43319)	2.56 ± 0.37	147.979 ± 1.369	−41.076 ± 0.468	−92.17 ± 1.42	5734 ± 63	3.4 ± 0.2	−0.98 ± 0.17	686 ± 153	1.30 ± 0.18	3.4 ± 2.9	2
2822-1194-1	(23.14799, 42.03068)	1.85 ± 0.64	88.644 ± 1.849	2.063 ± 0.496	−23.19 ± 1.87	6403 ± 116	4.2 ± 0.2	−0.48 ± 0.12	532 ± 160	1.10 ± 0.09	1.8 ± 2.5	0
3163-1181-1	(303.97045, 44.18376)	2.30 ± 0.25	156.232 ± 1.116	67.079 ± 1.026	−194.08 ± 1.61	5570 ± 74	3.4 ± 0.2	−0.30 ± 0.11	463 ± 85	1.59 ± 0.17	2.0 ± 1.1	1
3263-733-1	(15.00873, 45.13101)	1.83 ± 0.34	95.576 ± 1.290	−3.277 ± 0.425	14.91 ± 1.46	5425 ± 89	3.8 ± 0.1	−0.81 ± 0.16	517 ± 55	0.88 ± 0.09	12.3 ± 2.2	0
3285-1422-1	(32.53176, 47.41257)	1.10 ± 0.29	75.04 ± 1.682	−31.531 ± 0.505	25.43 ± 1.54	5214 ± 89	4.1 ± 0.1	−1.58 ± 0.16	143 ± 87	0.64 ± 0.08	10.9 ± 1.3	2
3330-120-1	(56.71171, 48.53692)	2.61 ± 0.30	194.055 ± 0.323	−123.109 ± 0.255	−24.12 ± 1.26	5735 ± 89	3.8 ± 0.1	−1.55 ± 0.16	571 ± 30	0.83 ± 0.03	12.5 ± 0.9	0
3661-974-1	(4.55758, 57.6662)	3.49 ± 0.651	180.078 ± 1.110	104.039 ± 0.651	−154.53 ± 2.02	6507 ± 100	4.1 ± 0.2	−0.99 ± 0.16	397 ± 83	0.87 ± 0.09	10.2 ± 2.7	1
3744-1546-1	(67.80849, 58.96855)	1.81 ± 0.42	143.706 ± 1.923	−38.217 ± 1.272	8.72 ± 1.49	6232 ± 174	4.3 ± 0.3	−1.68 ± 0.20	294 ± 78	0.78 ± 0.05	9.9 ± 4.0	2
3945-1023-1	(304.24414, 56.57186)	−6.07 ± 0.89	−6.097 ± 1.826	−1.265 ± 1.879	−18.79 ± 1.80	6239 ± 83	3.8 ± 0.2	−0.02 ± 0.12	1185 ± 150	1.54 ± 0.11	2.3 ± 0.5	0
3983-1873-1	(338.34366, 52.68866)	1.84 ± 0.23	133.342 ± 0.094	72.34 ± 0.082	−165.28 ± 0.86	4832 ± 68	2.0 ± 0.2	−1.27 ± 0.14	1096 ± 151	1.06 ± 0.19	5.4 ± 2.5	0
4032-1542-1	(26.42901, 60.39286)	0.74 ± 0.40	68.109 ± 0.761	−13.725 ± 0.73	−115.48 ± 7.15	7600 ± 83	3.7 ± 0.2	−0.23 ± 0.12	1009 ± 187	2.02 ± 0.16	0.9 ± 0.2	0
4307-1106-1	(8.16184, 74.08742)	2.31 ± 0.52	72.556 ± 1.141	15.474 ± 1.291	45.88 ± 1.79	5517 ± 74	3.5 ± 0.2	−0.45 ± 0.11	844 ± 193	1.41 ± 0.20	3.1 ± 2.4	0
4507-1461-1	(33.29978, 82.01739)	2.52 ± 0.31	85.192 ± 0.661	0.366 ± 0.836	−384.65 ± 2.22	6516 ± 100	4.2 ± 0.2	−1.24 ± 0.16	331 ± 30	0.82 ± 0.02	11.8 ± 1.6	0
4509-1013-1	(58.91556, 75.28116)	2.15 ± 0.24	97.297 ± 0.886	−29.216 ± 0.758	−155.52 ± 1.55	5890 ± 89	3.8 ± 0.1	−1.71 ± 0.16	549 ± 69	0.83 ± 0.08	12.0 ± 1.9	0
4515-1197-1	(79.71826, 77.83392)	1.28 ± 0.28	96.148 ± 0.892	45.051 ± 1.045	−198.41 ± 1.09	5398 ± 63	3.4 ± 0.2	−1.63 ± 0.17	902 ± 170	0.88 ± 0.15	11.4 ± 3.5	0
4521-322-1	(55.43942, 81.069)	3.22 ± 0.35	160.469 ± 0.536	1.117 ± 0.768	−129.92 ± 1.19	5872 ± 89	4.0 ± 0.1	−1.38 ± 0.16	428 ± 29	0.83 ± 0.02	12.4 ± 0.6	0

[8] This star has a very low log g, making the position of the isochrones uncertain. Furthermore, its metallicity is outside of the range of our isochrones, therefore distance, mass, and age could be biased or offset.

Notes: Hipparcos and *Gaia* identifiers for these stars are given in Table A1 in Appendix A. Proper motions and parallaxes are from *Gaia* TGAS, while stellar parameters have been derived using the RAVE pipeline. The 2.5 km s^{-1}uncertainty floor is *not* included in the quoted HRV errors, see discussion in §4.2.2. F = flag for the stellar parameter pipeline: 0 = converged; 1 = not converged; 2 = the pipeline oscillated between two solutions and the mean has been performed; 3 = bookkeping flag, the pipeline has converged.

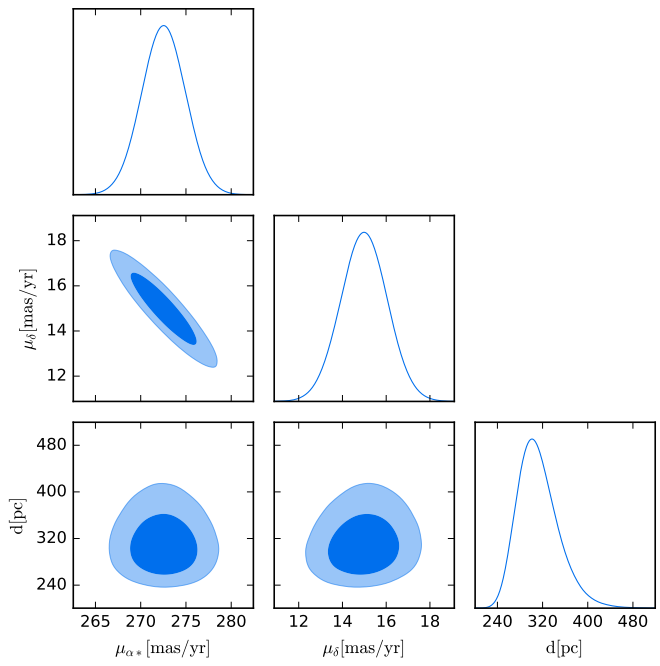

Figure 4. Proper motions and distance posterior distributions for the candidate TYC 49-1326-1 as obtained from the MCMC. Correlations from TGAS are $\rho_{\mu_{\alpha*},\mu_\delta} = -0.909$, $\rho_{\mu_{\alpha*},\varpi} = 0.023$, $\rho_{\mu_\delta,\mu_\varpi} = -0.103$. Dark (light) blue regions indicate the extent of the 1σ (2σ) credible intervals.

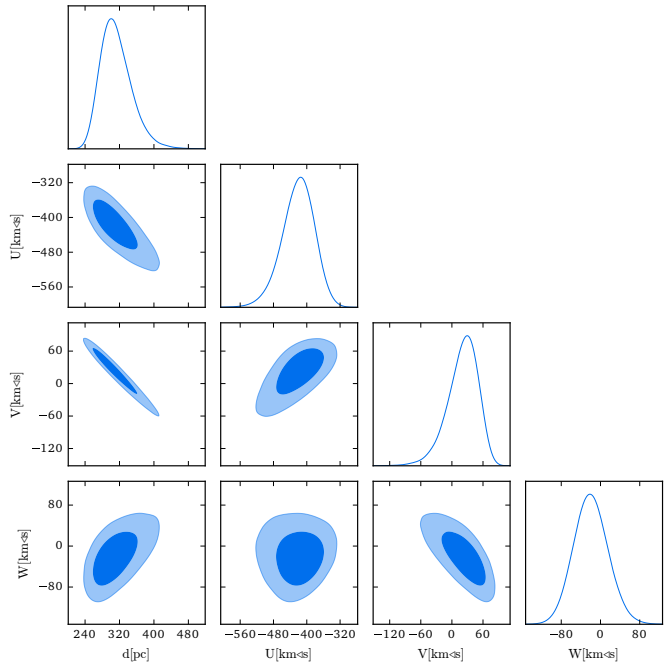

Figure 5. Distance and Galactic rectangular velocities U, V, W posterior distributions for TYC 49-1326-1 as obtained from the sampling of the astrometry shown in Figure 4. Dark (light) blue regions indicate the extent of the 1σ (2σ) credible intervals. The total galacto-centric velocity is $v_{\rm GC} = 419^{+38}_{-35}$ km s^{-1}.

where $\rho_{i,j}$ is the correlation between the parameters i and j, as given in TGAS. We model the prior probability on distances following the "Milky Way prior" approach presented in Astraatmadja & Bailer-Jones (2016a). We consider a three-dimensional density model for our Galaxy, that takes into account selection effects of the *Gaia* survey:

$$P_{\rm MW}(d, l, b) = d^2 \rho_{\rm MW}(d, l, b)\, p_{\rm obs}(d, l, b). \quad (10)$$

The stellar number density of the Milky Way $\rho_{\rm MW}(d, l, b)$ is modelled as the sum of three components (see Appendix A in Astraatmadja & Bailer-Jones (2016a) for details), while $p_{\rm obs}(d, l, b)$ describes the fraction of observable stars in a given sky position (Equation (4) in Astraatmadja & Bailer-Jones (2016a)). We choose this prior in our analysis because it gives the best results when comparing distances with a sample of known Cepheids (Astraatmadja & Bailer-Jones 2016b). The impact of assuming different priors on distance is discussed in Appendix B: except at distances > 800 pc, where errors are large, different priors give similar results. We assume uniform priors on proper motions. By means of Bayes' theorem we draw random samples of proper motions and distances from the resulting posterior distribution with an affine invariant ensemble Markov Chain Monte Carlo (MCMC) sampler (Goodman & Weare 2010), using the *emcee* implementation (Foreman-Mackey et al. 2013). We run the chain with 32 walkers and 4000 steps per walker, for a total of 128000 points drawn from the resulting posterior probability distribution. We check the convergence of the chain in terms of both the mean acceptance fraction and the autocorrelation time.

An example of a cornerplot showing Bayesian posterior distributions and correlations between the astrometric parameters for the candidate TYC 49-1326-1 is shown in Figure 4.

For the subset of 22 stars with a spectroscopic distance estimate we simply draw proper motions from a bivariate Gaussian distribution using the 2×2 covariance matrix provided by TGAS, and distances from a Gaussian with standard deviation equal to the estimated random uncertainty on distance.

If parallax-inferred and spectroscopic distance estimates are consistent within the errors, we expect the difference between the two divided by combined uncertainties to be distributed as a Gaussian with mean of zero and standard deviation of one. If we compute a Kolmogorov-Smirnov test to check whether these two distributions are consistent, we find that the null hypothesis cannot be rejected at a 5% level of significance. This is due to large uncertainties in distances, especially when adopting TGAS parallaxes. Since the two estimates can be remarkably different for individual stars, in the following we will present and discuss results assuming both distances.

6 RESULTS

Exploiting archival and new data we have assembled a total of 47 candidates with 3D position and velocity. A positive identification of a HVS requires both a radial trajectory from the Galactic Centre and a total velocity above the local escape speed. A star with the latter property but a trajectory that originates from the stellar disc will be called an *hyper runaway star*. Finally, *bound HVSs* (BHVSs) have Galactic Centre origin but velocity below the escape speed.

6.1 Total Galactocentric velocity

In order to identify HVSs, we compute the total velocity in the Galactic rest frame v_{GC} for the 47 candidates with a reliable ra-

dial velocity measurement. We start correcting radial velocities and proper motions for solar and LSR motion, assuming a three-dimensional Sun's velocity vector and LSR velocity (Schönrich 2012). We then calculate Galactic rectangular velocities U, V, and W with the following convention: U is positive if pointing towards the GC, V is positive along the direction of Galactic rotation, and W is positive towards the North Galactic Pole (Johnson & Soderblom 1987). The total velocity in the Galactic rest-frame is then simply computed summing in quadrature these three velocity components. We estimate uncertainties in the velocity vector via MC simulations, using the sampling in proper motions and distance described in §5. An example of posterior distributions for rectangular velocities is shown in Figure 5 for the candidate TYC 49-1326-1, obtained using posterior distributions shown in Figure 4.

For each star we draw 10^5 random realizations of its astrometric parameters, and the resulting total velocities are plotted in the first column of Figure 6 as a function of Galactocentric distance. We quote our results in terms of the median of the distribution, and errors are derived from the 16th and 84th percentiles. We overplot the median escape speed from the Milky Way derived in Williams et al. (2017) using a dashed line, with corresponding 68% (95%) credible intervals shown as a dark (light) blue region. This shows how the algorithm succeeded in finding high-velocity stars: 45 out of 47 candidates have a median Galactic rest frame velocity > 150 km s^{-1}, which is the typical velocity dispersion of stars in the halo (Smith et al. 2009; Evans et al. 2016). Considering parallax-inferred distances, first row, 11 objects are compatible within their uncertainties to be unbound from the Milky Way. If we use spectroscopic estimates, we find 3 stars with a total velocity consistent with being greater than the median escape speed at their position. Discussion of individual objects is postponed to §7.

Total velocities and distances are presented in Table 2 for the 15 stars with a median Galactic rest-frame velocity > 350 km s^{-1}obtained with at least one of the distance estimation methods. The *Gaia* and Hipparcos identifier of these high velocity candidates is presented in Appendix A. We assign to each star its probability of being unbound from the Galaxy, $P^{\rm u}$. From the posterior probability on distance d, we can compute the escape velocity from the Galaxy in each realization of the star's position using the analytic fit in Williams et al. (2017). We define $P^{\rm u}$ as the fraction of Monte Carlo realizations with $v_{\rm GC}(d) > v_{\rm esc}(d)$.

In the right panels of Figure 6 we present Toomre diagrams in the LSR frame for our candidates. In a Toomre's diagram one can identify three regions (separated by two solid black lines), corresponding to stars in the thin, thick disc, and halo (Venn et al. 2004; Hawkins et al. 2015). In the stellar halo kinematic region we report the local escape speed with associated errors (blue stripe, Williams et al. 2017)[9]. The two panels correspond to different distance determinations. Most of our candidates are consistent, from a kinematic point of view, with being halo stars. A total of 12 objects are consistent with being thin/thick disc stars considering parallax-inferred distances, and therefore will not be furthermore discussed.

6.2 Orbital traceback

We now proceed to establish the star candidate's origin by tracing back its trajectory in different models for the Galactic potential. We decide to perform the full orbit integration only for the most promising high-velocity stars in our sample, imposing the cut $\max(v_{\rm GC},\ v_{\rm GCspec}) > 350$ km s^{-1}, where quoted values denote the median of the distribution. A total of 15 objects passes this cut (see Table 2).

We use the publicly available python package *galpy*[10] (Bovy 2015) to integrate the orbit of each object in the Milky Way. We run 10^5 MC realizations of the star's orbit, using as initial conditions the position, distance, and U, V, W velocities previously randomly sampled from the posterior distributions. We use a four components Galactic potential, and we study the impact of our results depending on the choice of its parameters.

Our fiducial model consists of a point mass black hole potential:

$$\phi_{BH}(r) = -\frac{GM_\bullet}{r}, \tag{11}$$

a spherically symmetric bulge modelled as a Hernquist spheroid (Hernquist 1990):

$$\phi_b(r) = -\frac{GM_b}{r + r_b}, \tag{12}$$

a Miyamoto-Nagai disc in cylindrical coordinates (R, z) (Miyamoto & Nagai 1975):

$$\phi_d(R, z) = -\frac{GM_d}{\sqrt{R^2 + \left(a_d + \sqrt{z^2 + b_d^2}\right)^2}}, \tag{13}$$

and a Navarro-Frenk-White (NFW) profile for the dark matter halo (Navarro et al. 1996):

$$\phi_h(r) = -\frac{GM_h}{r} \ln\left(1 + \frac{r}{r_s}\right). \tag{14}$$

We adopt the following values for the potential parameters: $M_b = 3.4 \times 10^{10}$ M$_\odot$, $r_b = 0.7$ kpc, $M_d = 1.0 \times 10^{11}$ M$_\odot$, $a_d = 6.5$ kpc, $b_d = 0.26$ kpc (Johnston et al. 1995; Price-Whelan et al. 2014; Hawkins et al. 2015), $M_h = 0.76 \times 10^{12}$ M$_\odot$, $r_s = 24.8$ kpc (Rossi et al. 2017). This potential gives a local escape speed ~ 580 km s^{-1}, in agreement with results in Piffl et al. (2014), and, using data presented in Huang et al. (2016), provides a good fit to the rotation curve of the Milky Way out to ~ 100 kpc (see Appendix A, Figure A1, in Rossi et al. 2017).

For those stars for which we do not have a spectroscopic estimate of the age, we trace the orbit back in time for a fiducial time of 10 Gyr, motivated by the typical age and mass of the observed sample (see Table 1 and Figure 3). We integrate each orbit with a time resolution of 0.5 Myr, keeping track of each disc crossing (Galactic latitude $b = 0$).

If a star is ejected via the Hills mechanism but it is still gravitationally bound to the Milky Way, after the turn-around (maximum distance from the GC) it might cross multiple time the disc before being observed. This is supported by the fact that INT observations suggest that the majority of our stars have ages much larger than typical flight times from the stellar disc to the observed position, the latter being of the order of hundreds of Myr. An example of such a bound orbit is shown in Figure 7. Thus it is not trivial to determine which disc crossing should be assigned in order to understand whether or not our candidates effectively originate from the GC. Zhang et al. (2016), searching for nearby low mass high velocity stars, assume the most-recent disc crossing to be the ejection location of the star in the Galaxy. Given the complexity of bound

[9] We choose for simplicity to plot the local value.

[10] `http://github.com/jobovy/galpy`

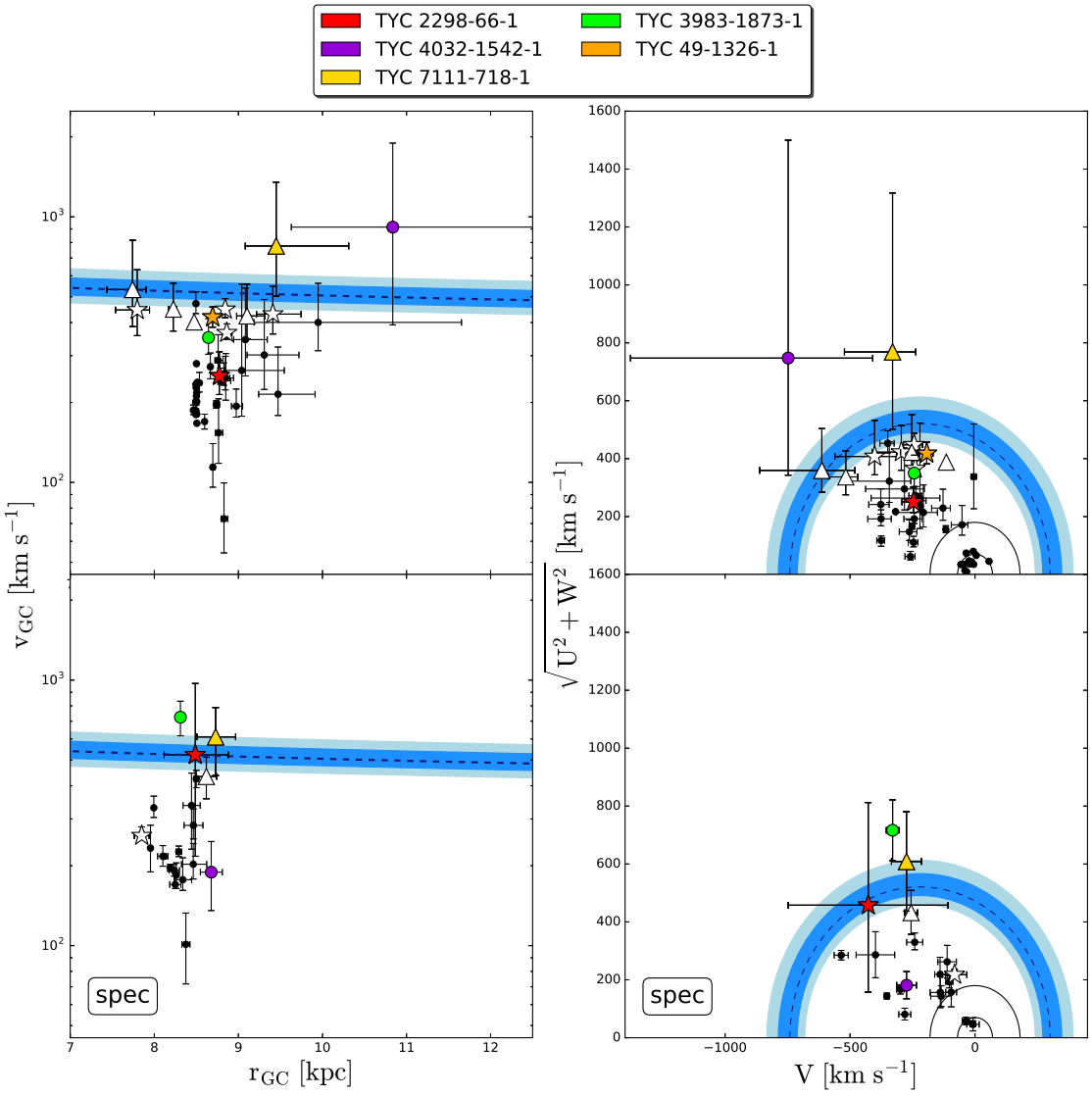

Figure 6. *First column*: Total Galactic rest frame velocity versus Galactocentric distance for those HVS candidates with a reliable radial velocity measurement. *Second column*: Toomre diagrams (in the LSR frame) for the same candidates. The two black rings in the bottom-right corner refer to the boundaries of the thin and thick disk, respectively at a constant velocity of 70 and 180 km s^{-1}(Venn et al. 2004). Most of our candidates lie in the kinematic region corresponding to halo stars. *First row*: velocities computed using distances inferred from parallax, using the MW prior. *Second row*: velocities computed using a spectroscopic distance estimate, when available. *All plots*: The dashed line is the median posterior escape speed (as a function of radius in the first column, and the local 521^{+46}_{-30} km s^{-1}in the second one) from Williams et al. (2017) with the 68% (94%) credible interval shown as a dark (light) blue band. Stars mark HVS/BHVS candidates in Table 2. Triangles mark runaway star candidates in Table 2. 11 objects are consistent with being unbound from the Milky Way in the first row, and 3 if we adopt spectroscopic distances.

orbits, we simply check the consistency of the GC origin hypothesis for our candidates by recording the closest disc crossing to the GC. This approach allows us to directly exclude stars that are not HVSs, since it is a necessary condition for a HVS that this method results in a density contour level containing the GC.

We find 8 stars to have orbits consistent with coming from the Galactic Centre using parallax-inferred distances. Within the sample of stars with spectroscopic distances we find 3 candidates, and all of them originate from the GC also when parallax-inferred distances are used.

We check the robustness of this conclusion integrating trajectories in different Milky Way potentials. Our choice for the mass of the bulge is significantly higher compared to the latest observational constraints (Bland-Hawthorn & Gerhard 2016; McMillan 2017), therefore we integrate each candidate assuming a bulge mass equal to half the previous adopted value: $M_b = 1.7 \times 10^{10}$ M$_\odot$, keeping fixed all the other parameters. As a second test, we adopt the potential in Kenyon et al. (2014), commonly adopted in HVS papers, which has a less massive bulge and stellar disc (but different scale parameters). In both cases we find the same candidates to

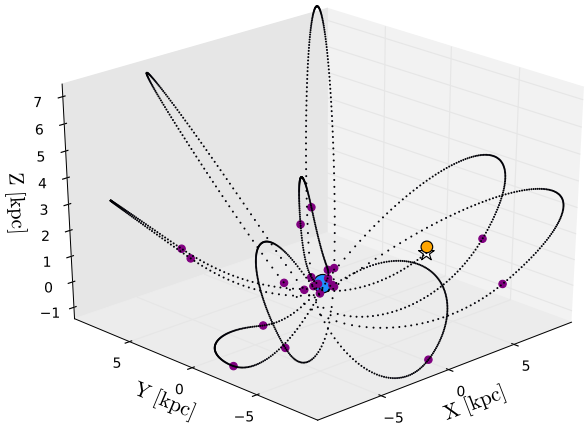

Figure 7. Example MC realization of a single bound orbit of TYC 2298-66-1 using the spectroscopic distance estimate. The blue (orange) circle marks the position of the GC (Sun), and the white star corresponds to the observed position of the star. Purple dots mark the disc crossings of the star prior to, and including the one happening closest to the GC. The initial conditions are $d_0 = 1018$ pc, $v_{\rm GC} = 225$ km s^{-1}, the eccentricity is $e \sim 0.96$, and the estimated flight time from the assigned ejection location to the observed position is $t_{\rm f} = 1.3$ Gyr $\ll t_{\rm age} = 8.2$ Gyr. For this particular orbit, the closest disc crossing is at ~ 260 pc from the Galactic Centre.

be consistent with coming from the GC. As a final test, we study the impact of assuming a triaxial profile for the bulge, which might influence the orbital traceback in the inner regions of the Galaxy. Results from star counts recently revealed that the Milky Way bulge has a boxy/peanut shape (McWilliam & Zoccali 2010; Wegg & Gerhard 2013), which can be characterized by an axis ratio from top $(b/a) \sim 0.5$, and an edge-on axis ratio $(c/a) \sim 0.26$ (Bland-Hawthorn & Gerhard 2016). Adopting the same mass and scale radius as in our fiducial potential and using a triaxial Hernquist profile to model the bulge, we find the shape of the density contour to change considerably, but the assumption of consistency with coming from the GC is solid.

Figure 8 shows example probability density functions of the disc crossing locations in the Galactic plane (rotating anticlockwise) for two candidates which will be further discussed in next sections, assuming our fiducial model for the Galactic potential. TYC 49-1326-1, left panel, is consistent with coming from the GC, while for TYC 3983-1873-1, right panel, the GC origin is excluded.

7 DISCUSSION OF INDIVIDUAL CANDIDATES

We divide candidates in Table 2 in three major classes: HVS and BHVS candidates, runaway star candidates, and "uncertain" objects. To help the discussion, the metallicity distribution of these stars is shown with a purple line in Figure 2, where it is compared to typical metallicity distributions of stars in the inner Galactic halo. We will now discuss separately candidates from each class in detail, focusing on the most promising objects and on stars already present in literature. One additional candidate not included in Table 2, but known from literature, is discussed in §7.4.

7.1 HVS and BHVS Candidates

In addition to HVSs, the Hills mechanism naturally predicts a population of *bound HVSs*: stars having a velocity high enough to escape from the MBH's gravitational field at their ejection, but not sufficient to be unbound from the whole Milky Way. These stars, being decelerated and deflected by the Galactic potential, can cross the disc multiple times during their life, following a wide variety of highly-non-radial orbits, as previously shown in Figure 7. The identification of such objects is observationally particularly difficult. The probability of observing a star at a particular moment of its orbit is proportional to the residence time t_r in that orbit element: $p \propto t_r \propto v^{-1}$, therefore we expect most of these stars to be observed when they have low velocities, and they could thus be easily mistaken for halo stars.

Hypervelocity and bound hypervelocity star candidates are marked with a star symbol in Figure 6. Stars are classified as HVSs if (i) their velocity is > 350 km s^{-1} with at least one distance estimate, and (ii) if they are consistent with coming from the GC (within 2σ) when traced back in different Galactic potentials. We find a total of 6 stars satisfying both properties within their uncertainties: TYC 2298-66-1, TYC 8422-875-1, TYC 2456-2178-1, TYC 2348-333-1, TYC 49-1326-1, and TYC 5890-971-1. The consistency with the GC origin does not depend on the assumed distance. The further sub-classification as HVSs or BHVSs depends on the value of $P^{\rm u}$. All of these stars are on highly radial orbits, with median eccentricities > 0.9.

- TYC 2298-66-1 (LP 295-632) is a high proper motion metal-poor candidate, identified by a red symbol in Figure 6. It is the only star with a probability $> 50\%$ of being unbound from the Galaxy when using the spectroscopic distance estimate ($v \sim 530$ km s^{-1}, even if with large uncertainties), therefore it is a HVS candidate.
- TYC 8422-875-1 (HD 201484, V Ind) is a F0 V variable star of RR Lyrae type (Houk 1978). In the discussion of this candidate, we use Figure 9 to help us distinguish which distance estimate is more likely to be correct. This plot compares the position of the star in the parallax-distance modulus diagram to the analytical prediction computed assuming the Schlegel extinction towards the line-of-sight. The distance modulus is taken from RAVE DR5 (Kunder et al. 2017), and the resulting point is shown in black. The total velocity of TYC 8422-875-1 strongly depends on the distance assumption, but from Figure 9 we can see that parallax-inferred distance is more likely to be correct. Furthermore, since this star is a RR Lyrae, we can independently determine its distance modulus using a period-luminosity-metallicity (PLZ) relation (Leavitt 1908; Leavitt & Pickering 1912). Period, [Fe/H] metallicity, and mid-infrared [3.6] magnitude are taken from Monson et al. (2017), and we estimate the distance modulus using the PLZ relation in the *WISE W*1 band from Sesar et al. (2017). This results in a distance modulus ~ 9.3, consistent with the parallax measured by *Gaia*, as shown with a red star in Figure 9. We then conclude that V Ind is a BHVS candidate, with $v \sim 450$ km s^{-1} and a probability of $\sim 30\%$ of being unbound.
- TYC 2456-2178-1 is a BHVS candidate, with $v \sim 430$ km s^{-1} and a probability $\gtrsim 20\%$ of being unbound from the Galaxy.
- TYC 2348-333-1 (G 95-11) is a high proper motion and high velocity star which has been previously used to estimate the local Galactic escape speed together with other stars from the $uvby-\beta$ survey of high velocity and metal poor stars (García Cole et al. 1999). With a total velocity around 450 km s^{-1}, this star is most likely a BHVS. We note that our distance estimate is higher than

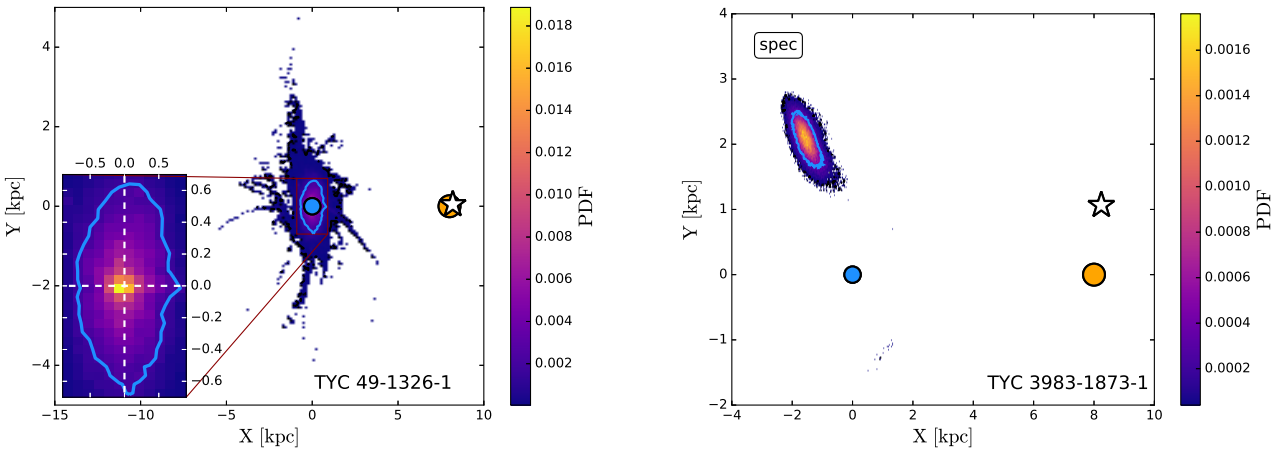

Figure 8. Normalised probability distribution function of Galactic disc crossings for the candidates TYC 49-1326-1, assuming the parallax-inferred distance (left panel), and TYC 3983-1873-1, using the spectroscopic distance (right panel). The blue line marks the 1σ contour, and the coloured region extends up to the 2σ contour. The MW rotates anticlockwise. The blue (orange) circle marks the position of the GC (Sun), while the white star corresponds to the median observed position of the candidate. The white dashed cross marks the position of the GC in the zoomed inset.

Table 2. Derived kinematic properties for the 15 HVS candidates with $\max(v_{\rm GC},\ v_{\rm GCspec}) > 350$ km s^{-1}, and interpretation.

Tycho 2 ID	HRV (km s^{-1})	$[M/H]$ (dex)	d (pc)	$d_{\rm spec}$ (pc)	$v_{\rm GC}$ (km s^{-1})	$v_{\rm GCspec}$ (km s^{-1})	$P^{\rm u}$	$P^{\rm u}_{\rm spec}$	Ref
HVS / BHVS candidates									
2298-66-1	-31.66 ± 2.78	-2.08 ± 0.26	431^{+78}_{-55}	754 ± 569	248^{+58}_{-38}	519^{+451}_{-307}	0.1%	50.3%	1
8422-875-1[11]	200.8 ± 0.8	-1.01 ± 0.07	1010^{+400}_{-218}	208 ± 124	446^{+186}_{-89}	259^{+21}_{-7}	29.1%	0.0%	2, 5
2456-2178-1	-243.08 ± 49.53	-2.25 ± 0.24	976^{+358}_{-207}		430^{+117}_{-68}		22.7%		3
2348-333-1	205.26 ± 0.34	-1.26 ± 0.40	407^{+51}_{-40}		448^{+44}_{-32}		7.6%		3, 4
49-1326-1	265.1 ± 37.6		304^{+38}_{-30}		419^{+38}_{-35}		1.2%		2, 5
5890-971-1	348.6 ± 0.8		550^{+93}_{-72}		366^{+29}_{-20}		0.2%		6, 7
Runaway star candidates									
7111-718-1	76.7 ± 1.2	-1.53 ± 0.17	1967^{+1413}_{-683}	1552 ± 430	776^{+576}_{-274}	611^{+176}_{-172}	82.2%	70.7%	2, 5
8374-757-1	71.8 ± 3.7		832^{+338}_{-179}		532^{+284}_{-147}		50.4%		8
1071-404-1	-267.12 ± 0.26	~ -0.5	439^{+91}_{-64}		449^{+113}_{-78}		23.7%		4
4515-1197-1	-198.41 ± 1.09	-1.63 ± 0.17	881^{+292}_{-175}	902 ± 170	423^{+137}_{-76}	433^{+78}_{-76}	23.5%	15.6%	1
9404-1260-1	-94.9 ± 0.6		$67.0^{+1.0}_{-0.9}$		402^{+4}_{-4}		0.0%		9
Uncertain candidates									
3983-1873-1	-165.28 ± 0.86	-1.27 ± 0.14	572^{+88}_{-67}	1096 ± 151	351^{+64}_{-47}	726^{+107}_{-108}	1.5%	97.2%	1
4032-1542-1	-115.48 ± 7.15	-0.23 ± 0.12	3216^{+2918}_{-1574}	1009 ± 187	918^{+979}_{-527}	183^{+59}_{-57}	75.7%	0.0%	1
3945-1023-1	-18.79 ± 1.80	-0.02 ± 0.12	4978^{+2802}_{-1686}	1185 ± 150	399^{+162}_{-87}	215^{+4}_{-4}	24.5%	0.0%	1
3330-120-1	-24.12 ± 1.26	-1.55 ± 0.16	401^{+56}_{-43}	571 ± 30	247^{+58}_{-44}	425^{+32}_{-32}	0.1%	0.3%	1

[11] The parallax-inferred distance d is more likely to be correct for this RR Lyrae star (see Figure 9), and is consistent with the value obtained using a PLZ relation (see discussion in §7.1).

Notes: Hipparcos and *Gaia* identifiers for these stars are given in Table A1 in Appendix A. The subscript "spec" refers to quantities computed using the spectroscopic distance (when available). For distances and Galactocentric velocities, results are quoted in terms of the median of the distribution with uncertainties derived from the 16th and 84th percentiles. The 2.5 km s^{-1} uncertainty floor (see discussion in §4.2.2) is *not* included in the quoted HRV errors.

References: (1) This paper, observations at the INT; (2) Kordopatis et al. (2013a); (3) (Cui et al. 2012); (4) Latham et al. (2002); (5) Kunder et al. (2017); (6) Przybylski (1978); (7) Barbier-Brossat et al. (1994); (8) Kharchenko et al. (2007); (9) Holmberg et al. (2007).

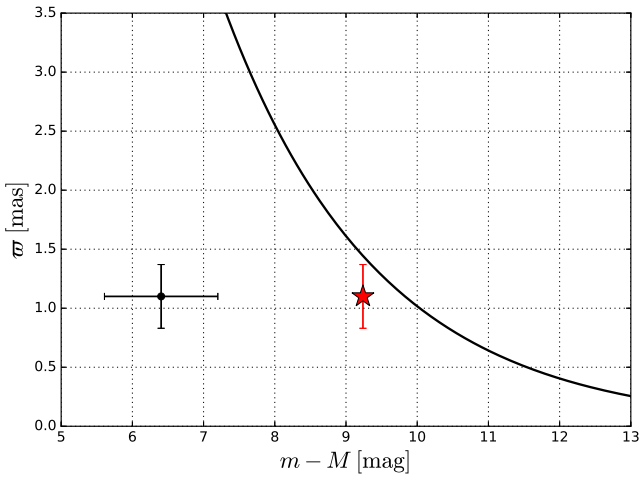

Figure 9. Parallax-distance modulus diagram for the RR Lyrae star TYC 8422-875-1 (HD 201484, V Ind), using the parallax from TGAS and the distance modulus from RAVE DR5 (black point). The line shows the analytic prediction assuming the Schlegel extinction towards the line-of-sight. The parallax-inferred distance estimate is clearly favoured. The red star corresponds to adopting the distance modulus obtained using the PLZ relation. Other candidates lie too close to the curve to have a clear preference towards one distance estimate.

the value $\sim$ 250 pc given in García Cole et al. (1999), resulting in a higher total velocity.

- TYC 49-1326-1 (G 75-29), marked with an orange star in Figure 6, is a BHVS candidate with a total velocity particularly well constrained of 419^{+38}_{-35} km s^{-1}.
- TYC 5890-971-1 (HD 27507), even if it has a total velocity lower than the other candidates, is worth mentioning because it is historically the first discovered HVS candidate. Przybylski (1978) discussed the possibility that HD 27507 is a star escaping from our Galaxy given its high velocity, and a following proper motion redetermination confirmed this conclusion (Clements et al. 1980). The authors found a total velocity $\sim$ 360 km s^{-1}, in good agreement with our results, but studies in the past decades substantially increased the value of the local escape speed (see Williams et al. (2017) for the latest constraints), making this star unlikely to be unbound from the Milky Way. Nevertheless, its orbit is consistent with coming from the GC, making TYC 5890-971-1 a bound HVS candidate.

7.2 Runaway Star Candidates

Runaway stars (RSs) are high velocity stars ejected in many-body dynamical encounters in dense stellar systems (Poveda et al. 1967; Portegies Zwart 2000) or by the explosion of a supernova in a binary system (Blaauw 1961; Tauris & Takens 1998). Tauris (2015) showed how it is possible to reach Galactic rest frame velocities up to $\sim$ 1280 km s^{-1}for the ejected companion star in a binary disrupted via an asymmetric supernova explosion. These extreme velocities can be achieved by low-mass G/K candidates in very compact presupernova binaries. High velocity runaway stars observed in the halo are most likely produced in the disc (Bromley et al. 2009; Duarte de Vasconcelos Silva 2012; Kenyon et al. 2014). Since most of our stars have masses slightly below the Solar value, this mechanism can possibly explain the notable velocity of our stars that do not originate from the GC.

With this classification rule we identify as runaway candidates 5 high-velocity stars: TYC 7111-718-1, TYC 8374-757-1, TYC 1071-404-1, TYC 4515-1197-1, and TYC 9404-1260-1. Regardless of the adopted distance, these stars always have median $v_{\rm GC} > 350$ km s^{-1}. In particular, 2 stars have a probability $> 50\%$ of being unbound from the Milky Way, and are therefore classified as *hyper runaway stars* (HRSs). Runaway star candidates are marked with a triangle symbol in Figure 6. In the following we discuss them individually.

- TYC 7111-718-1, marked in yellow in Figure 6, is a strong hyper-runaway star candidate, with a velocity $>$ 600 km s^{-1}, in excess of the local escape speed regardless of the adopted distance estimate. From a chemical point of view, it is consistent with the inner Galactic halo population.
- TYC 8374-757-1 (HD 176387, MT Tel) is a RR Lyrae variable star. It was previously discovered by Przybylski (1967), which discussed, despite large uncertainties in proper motions, its nature as a high velocity star. Because of large errors in distance we cannot strongly constrain its total velocity, which, with a median value $\sim$ 530 km s^{-1}, is nevertheless consistent with being greater than the escape speed, making MT Tel a hyper-runaway star candidate. We repeat the same approach discussed for TYC 8422-875-1 to determine the distance of MT Tel using the PLZ relation in Sesar et al. (2017) using data from Monson et al. (2017). We find a distance modulus $\sim$ 8.1, consistent with the parallax from *Gaia*, confirming our high-velocity determination.
- TYC 1071-404-1, TYC 4515-1197-1, and TYC 9404-1260-1 are RS candidates most likely bound to the MW, with a remarkably high total velocity $\gtrsim$ 400 km s^{-1}.

Another intriguing origin for these stars not originating from the GC is that they come from the Large Magellanic Cloud (LMC), either as runaway stars (Boubert et al. 2017), or by the extension of the Hills mechanism to a hypothetical MBH at the centre of the LMC (Boubert & Evans 2016). Uncertainties are at the moment too large to pinpoint their ejection location, and we dot not further expand on this possibility in this paper.

7.3 Uncertain Candidates

In our final sample (Table 2) there are 4 stars with uncertain interpretation: TYC 3983-1873-1, TYC 4032-1542-1, TYC 3945-1023-1, TYC 2393-1001-1, and TYC 3330-120-1. These objects have a debated nature, with velocities and origins highly dependent on the assumed distance indicator. We classify as runaway star (halo star) candidates that are not consistent with coming from the GC, and with a total velocity > 350 km s^{-1}(< 350 km s^{-1}).

- TYC 3983-1873-1 (BD+51 3413) is a high proper motion HVS candidate (green points in Figure 6). It is one of the few candidates with a spectroscopic distance higher than the parallax inferred one, which results in a total velocity of $\sim$ 725 km s^{-1}, more than 1σ above the median escape speed. Remarkably, if we assume a spectroscopic distance, this object is not consistent with coming from the GC, and should therefore be classified as a HRS, while it is a BHVS candidate ($v \sim 350$ km s^{-1}) if we adopt the parallax-inferred distance.
- TYC 4032-1542-1, marked in purple in Figure 6, suffers from a particularly poor distance determination. The spectroscopic distance gives a relatively low velocity of $\sim$ 190 km s^{-1}, consistent

with that of a high velocity halo star. Its velocity increases considerably if we rely on the much more uncertain parallax-inferred distance ($v \sim 900$ km s^{-1}). A point worth mentioning is that the metallicity is considerably higher than the mean value in the inner halo, making this object worth inspecting in order to constrain its nature and origin as kinematic and chemical outlier. Furthermore, TYC 4032-1542-1 is an A type star, more massive compared to the other candidates, therefore it is more difficult to explain its high velocity invoking the disruption of a close binary via supernova explosions (Tauris 2015, and see discussion in §7.2).

- TYC 3945-1023-1 is a RS ($v \sim 400$ km s^{-1}) or a halo star ($v \sim 200$ km s^{-1}) candidate, if we assume the parallax-inferred or the spectroscopic distance estimate respectively.
- TYC 3330-120-1 is a runaway star candidate ($v \sim 425$ km s^{-1}) if we adopt the spectroscopic distance, but behaves as a typical halo star ($v \sim 250$ km s^{-1}) if we infer distance from parallax.

7.4 HD 5223: Most Likely Not a HVS

In this subsection we present one additional star discovered with our data mining algorithm, TYC 1739-1500-1 (HD 5223). Even if it doesn't pass the velocity cut in Table 2, this star was previously known and discussed for its high velocity, which we now revisit using *Gaia*'s much more precise data.

HD 5223 is a carbon-enhanced metal-poor star presented in Pereira et al. (2012), which concluded that this object is a hypervelocity star with a total velocity in the Galactic frame of 713 km s^{-1}. Our velocity determination $v = 288^{+72}_{-46}$ km s^{-1}is considerably lower because of a substantial difference in the assumed distance: Pereira et al. (2012) determined $d = 1.2$ kpc, while our computation seems to suggest lower values: $d = 565^{+117}_{-80}$ pc. If our estimate is correct, HD 5223 is bound to the MW, and furthermore we find its orbit not to be consistent with coming from the GC.

8 DISCUSSION AND CONCLUSIONS

We successfully developed a new automatized method to extract high velocity stars, using a data-driven algorithm trained on mock populations of hypervelocity stars. Our data mining routine, an artificial neural network, is optimized for the very unbalanced search of rare objects in a large dataset. This approach avoids a bias towards particular spectral types or stellar properties, making as few assumptions as possible on the stellar nature of stars coming from the Galactic Centre. Applying the algorithm to the TGAS subset of the first release of the *Gaia* satellite, we have identified a total of 80 objects with a predicted probability > 90% of being a HVS, and for 30 of those we were able to find a radial velocity measurement from literature. We followed up spectroscopically 22 candidates at the Isaac Newton Telescope, for a total of 47 stars with a reliable radial velocity determination. Our stars show a uniform distribution across the sky, showing that the algorithm is not selecting sources in a preferential direction.

With a Bayesian approach we inferred distances from parallax for all our candidates, and total velocities in the Galactic rest frame were computed in order to establish their nature and origin. Without pre-selection of data we were able to recover several objects already noted and discussed in literature because of their remarkably high velocities. We found 45 candidates with a median rest frame velocity > 150 km s^{-1}, 14 of them having $v > 400$ km s^{-1}, and a subset of 5 stars has a probability > 50% of being unbound from the Milky Way, with median velocities up to ~ 900 km s^{-1}.

Tracing back orbits with Monte Carlo simulations in different Galactic potentials we found:

- 6 stars being consistent with coming from the Galactic Center. One of these stars, with a velocity of ~ 520 km s^{-1}, has a probability > 50% of being unbound from the Galaxy (HVS), while the others are bound hypervelocity star candidates, with velocities > 360 km s^{-1};
- 5 stars with high velocities but trajectories not consistent with coming from the Galactic Centre: these stars are runaway star candidates. Two of these stars have probabilities > 50% of being unbound from the Milky Way, and are therefore classified as hyper runaway stars. The explosion of a supernova in a binary system is a plausible mechanism for having accelerated these stars to such high velocities. It is remarkable that a good fraction of our RS candidates have velocities consistent with being higher than the escape velocity from the Galaxy, since these stars are thought to be extremely rare: approximately 1 for every 100 HVSs (Bromley et al. 2009; Perets & Šubr 2012; Kenyon et al. 2014; Brown 2015);
- 4 stars with a velocity and origin highly dependent on the assumed distance estimate. Two of these stars have a high probability of being unbound from the Milky Way.

At the moment, positive identifications are strongly hampered by large uncertainties in distance and limited information on the age and flight time of our sources. The advent of future *Gaia* releases will dramatically increase the number of HVSs we expect to find. The more accurate parallax determination, less affected by systematics, will allow us to decrease error bars and to identify in a clearer way the most interesting objects, narrowing down their ejection location. The brightest stars in the catalogue will also have a radial velocity measurement, allowing us to train the neural network adding this precious information as an extra feature to the astrometric solution.

We are currently working to increase the quality of the training set of mock HVSs, considering not only radial trajectories, but modelling orbits of bound stars and including deviations due to the disc and to a possible triaxiality of the bulge (e.g. McWilliam & Zoccali 2010) and/or the halo (e.g. Bullock 2002; Helmi 2004). Another natural advancement would be to model runaway and halo stars to create mock populations, and then to perform a multiclass classification analysis in order to decrease the number of false positives and achieve a more precise classifier.

ACKNOWLEDGEMENTS

We thank J. Brinchmann and A. Patruno for useful discussion and comments, U. Bastian and L. Lindegren for suggestions and advice on the use of *Gaia* astrometric data and on distances determination, and T. Astraatmadja and C. Bailer-Jones for the implementation of the Milky Way Prior. We also thank Warren Brown for the careful reading of the manuscript and his useful comments. TM and EMR acknowledge support from NWO TOP grant Module 2, project number 614.001.401. ES and KY gratefully acknowledge funding by the Emmy Noether program from the Deutsche Forschungsgemeinschaft (DFG). This work has made use of data from the European Space Agency (ESA) mission *Gaia* (`http://www.cosmos.esa.int/gaia`), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, `http://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. The

Isaac Newton Telescope is operated on the island of La Palma by the Isaac Newton Group of Telescopes in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias. 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).

REFERENCES

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
Barbier-Brossat M., Petit M., Figon P., 1994, A&AS, 108
Blaauw A., 1961, Bull. Astron. Inst. Netherlands, 15, 265
Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529
Boubert D., Evans N. W., 2016, ApJ, 825, L6
Boubert D., Erkal D., Evans N. W., Izzard R. G., 2017, preprint, (arXiv:1704.01373)
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
Bromley B. C., Kenyon S. J., Brown W. R., Geller M. J., 2009, ApJ, 706, 925
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
Bullock J. S., 2002, in Natarajan P., ed., The Shapes of Galaxies and their Dark Halos. pp 109–113 (arXiv:astro-ph/0106380), doi:10.1142/9789812778017_0018
Chiba M., Beers T. C., 2000, AJ, 119, 2843
Cignoni M., Ripepi V., Marconi M., Alcalá J. M., Capaccioli M., Pannella M., Silvotti R., 2007, A&A, 463, 975
Clements E. D., Swifte R. H. D., Alexander J. B., 1980, The Observatory, 100, 5
Cui X.-Q., et al., 2012, Research in Astronomy and Astrophysics, 12, 1197
Do T., Kerzendorf W., Winsor N., Støstad M., Morris M. R., Lu J. R., Ghez A. M., 2015, ApJ, 809, 143
Drimmel R., Cabrera-Lavers A., López-Corredoira M., 2003, A&A, 409, 205
Duarte de Vasconcelos Silva M., 2012, PhD thesis, PhD Theses Collection, 2299, 8115
Duchi J., Hazan E., Singer Y., 2011, J. Mach. Learn. Res., 12, 2121
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., Loeb A., 2016, preprint, (arXiv:1608.01517)
Gaia Collaboration et al., 2016a, A&A, 595, A1
Gaia Collaboration et al., 2016b, A&A, 595, A2
García Cole A., Schuster W. J., Parrao L., Moreno E., 1999, Rev. Mexicana Astron. Astrofis., 35, 111
Genzel R., Eisenhauer F., Gillessen S., 2010, Reviews of Modern Physics, 82, 3121
Ghez A. M., et al., 2008, ApJ, 689, 1044
Gillessen S., Eisenhauer F., Trippe S., Alexander T., Genzel R., Martins F., Ott T., 2009, ApJ, 692, 1075
Gilmore G., et al., 2012, The Messenger, 147, 25
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. App. Math. Comp. Sci., 5, 65
Hawkins K., et al., 2015, MNRAS, 447, 2046
Haykin S., 2009, Neural Networks and Learning Machines. No. v. 10 in Neural networks and learning machines, Prentice Hall, https://books.google.nl/books?id=K7P36lKzI_QC
Helmi A., 2004, MNRAS, 351, 643
Hernquist L., 1990, ApJ, 356, 359
Hills J. G., 1988, Nature, 331, 687
Holmberg J., Nordström B., Andersen J., 2007, A&A, 475, 519
Houk N., 1978, Michigan catalogue of two-dimensional spectral types for the HD stars
Huang Y., et al., 2016, MNRAS, 463, 2623
Hunter J. D., 2007, Computing In Science & Engineering, 9, 90
Hurley J. R., Pols O. R., Tout C. A., 2000, MNRAS, 315, 543
Johnson D. R. H., Soderblom D. R., 1987, AJ, 93, 864
Johnston K. V., Spergel D. N., Hernquist L., 1995, ApJ, 451, 598
Kennedy J., Eberhart R., 1995, in Neural Networks, 1995. Proceedings., IEEE International Conference on. pp 1942–1948 vol.4, doi:10.1109/ICNN.1995.488968
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
Kharchenko N. V., Scholz R.-D., Piskunov A. E., Röser S., Schilbach E., 2007, Astronomische Nachrichten, 328, 889
Kobayashi S., Hainick Y., Sari R., Rossi E. M., 2012, ApJ, 748, 105
Kordopatis G., Recio-Blanco A., de Laverny P., Bijaoui A., Hill V., Gilmore G., Wyse R. F. G., Ordenovic C., 2011a, A&A, 535, A106
Kordopatis G., et al., 2011b, A&A, 535, A107
Kordopatis G., et al., 2013a, AJ, 146, 134
Kordopatis G., et al., 2013b, MNRAS, 436, 3231
Kordopatis G., et al., 2013c, A&A, 555, A12
Kordopatis G., et al., 2015, A&A, 582, A122
Kordopatis G., Amorisco N. C., Evans N. W., Gilmore G., Koposov S. E., 2016, MNRAS, 457, 1299
Kunder A., et al., 2017, AJ, 153, 75
Latham D. W., Stefanik R. P., Torres G., Davis R. J., Mazeh T., Carney B. W., Laird J. B., Morse J. A., 2002, AJ, 124, 1144
LeCun Y., 1993, in Tutorial presented at Neural Information Processing Systems. p. 49
LeCun Y. A., Bottou L., Orr G. B., Müller K.-R., 2012, Efficient BackProp. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 9–48, doi:10.1007/978-3-642-35289-8_3, http://dx.doi.org/10.1007/978-3-642-35289-8_3
Leavitt H. S., 1908, Annals of Harvard College Observatory, 60, 87
Leavitt H. S., Pickering E. C., 1912, Harvard College Observatory Circular, 173, 1
Li Y., Luo A., Zhao G., Lu Y., Ren J., Zuo F., 2012, ApJ, 744, L24
Lindegren L., et al., 2016, A&A, 595, A4

Madigan A.-M., Pfuhl O., Levin Y., Gillessen S., Genzel R., Perets H. B., 2014, ApJ, 784, 23
Magrini L., et al., 2017, preprint, (arXiv:1703.00762)
Martell S. L., et al., 2017, MNRAS, 465, 3203
Matthews B., 1975, Biochimica et Biophysica Acta (BBA) - Protein Structure, 405, 442
McMillan P. J., 2017, MNRAS, 465, 76
McWilliam A., Zoccali M., 2010, ApJ, 724, 1491
Meyer L., et al., 2012, Science, 338, 84
Michalik D., Lindegren L., Hobbs D., 2015, A&A, 574, A115
Miyamoto M., Nagai R., 1975, PASJ, 27, 533
Monson A. J., et al., 2017, AJ, 153, 96
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Öpik E., 1924, Publications of the Tartu Astrofizica Observatory, 25
Palladino L. E., Schlesinger K. J., Holley-Bockelmann K., Allende Prieto C., Beers T. C., Lee Y. S., Schneider D. P., 2014, ApJ, 780, 7
Pereira C. B., Jilinski E., Drake N. A., de Castro D. B., Ortega V. G., Chavero C., Roig F., 2012, A&A, 543, A58
Perets H. B., Šubr L., 2012, ApJ, 751, 133
Perets H. B., Hopman C., Alexander T., 2007, ApJ, 656, 709
Perets H. B., Wu X., Zhao H. S., Famaey B., Gentile G., Alexander T., 2009, ApJ, 697, 2096
Pfuhl O., et al., 2011, ApJ, 741, 108
Piffl T., et al., 2014, A&A, 562, A91
Portegies Zwart S. F., 2000, ApJ, 544, 437
Portegies Zwart S. F., Verbunt F., 1996, A&A, 309, 179
Portegies Zwart S., et al., 2009, New A, 14, 369
Poveda A., Ruiz J., Allen C., 1967, Boletin de los Observatorios Tonantzintla y Tacubaya, 4, 86
Price-Whelan A. M., Hogg D. W., Johnston K. V., Hendel D., 2014, ApJ, 794, 4
Przybylski A., 1967, MNRAS, 136, 185
Przybylski A., 1978, PASP, 90, 451
Randich S., Gilmore G., Gaia-ESO Consortium 2013, The Messenger, 154, 47
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
Saerens M., Latinne P., Decaestecker C., 2002, IEEE Trans. Neural Networks, 13, 1204
Sari R., Kobayashi S., Rossi E. M., 2010, ApJ, 708, 605
Schlafly E. F., Finkbeiner D. P., 2011, ApJ, 737, 103
Schönrich R., 2012, MNRAS, 427, 274
Sesana A., Haardt F., Madau P., 2007, MNRAS, 379, L45
Sesar B., Fouesneau M., Price-Whelan A. M., Bailer-Jones C. A. L., Gould A., Rix H.-W., 2017, ApJ, 838, 107
Singh B., De S., Zhang Y., Goldstein T., Taylor G., 2015, CoRR, abs/1510.04609
Smith M. C., et al., 2009, MNRAS, 399, 1223
Soubiran C., Jasniewicz G., Chemin L., Crifo F., Udry S., Hestroffer D., Katz D., 2013, A&A, 552, A64
Tauris T. M., 2015, MNRAS, 448, L6
Tauris T. M., Takens R. J., 1998, A&A, 330, 1047
Tody D., 1986, in Crawford D. L., ed., Proc. SPIEVol. 627, Instrumentation in astronomy VI. p. 733, doi:10.1117/12.968154
Venn K. A., Irwin M., Shetrone M. D., Tout C. A., Hill V., Tolstoy E., 2004, AJ, 128, 1177
Vickers J. J., Smith M. C., Grebel E. K., 2015, AJ, 150, 77
Wegg C., Gerhard O., 2013, MNRAS, 435, 1874

Table A1. Tycho 2, Hipparcos, and *Gaia* identifiers of stars observed at the INT and of high velocity candidates.

Tycho 2 ID	Hipparcos ID	*Gaia* ID
1071-404-1	98492	4299974437593772672
2282-208-1		314799593600582656
2292-1267-1		316401685121779712
2298-66-1		317585859144818688
2320-470-1		329685915888890880
2348-333-1		137859551029399040
2376-691-1		172747742173867904
2393-1001-1		180650104040989568
2456-2178-1		893048667206860800
2818-556-1		347908809291960832
2822-1194-1		348293878879518848
3163-1181-1		2081319505008076416
3263-733-1		377741720849393920
3285-1422-1		353451584846863104
3330-120-1	17648	248695099116287872
3661-974-1		422054582068454016
3744-1546-1		470781741956237696
3945-1023-1		2187713404073484288
3983-1873-1	111334	2000722382112691456
4032-1542-1		509654254003883776
4307-1106-1		539315160710386944
4507-1461-1		569097391651702656
4509-1013-1		550795677011227648
4515-1197-1		552553933541803008
4521-322-1		568189573004745472
49-1326-1		2503868695508755840
5890-971-1	20214	3172032703298013696
7111-718-1		5590900663125136000
8374-757-1	93476	6662886601414152448
8422-875-1	104613	6483680327939151488
9404-1260-1	46120	5195968559017084160

Westera P., Buser R., 2003, in Piotto G., Meylan G., Djorgovski S. G., Riello M., eds, Astronomical Society of the Pacific Conference Series Vol. 296, New Horizons in Globular Cluster Astronomy. p. 238
Williams A. A., Belokurov V., Casey A. R., Evans N. W., 2017, MNRAS, 468, 2359
Yu Q., Madau P., 2007, MNRAS, 379, 1293
Zasowski G., et al., 2013, AJ, 146, 81
Zhang F., Lu Y., Yu Q., 2013, ApJ, 768, 153
Zhang Y., Smith M. C., Carlin J. L., 2016, ApJ, 832, 10
Zheng Z., et al., 2014, ApJ, 785, L23
Ziegerer E., Volkert M., Heber U., Irrgang A., Gänsicke B. T., Geier S., 2015, A&A, 576, L14
Ziegerer E., Heber U., Geier S., Irrgang A., Kupfer T., Fürst F., Schaffenroth J., 2017, A&A, 601, A58
Zucker S., 2003, MNRAS, 342, 1291

APPENDIX A: *GAIA* IDENTIFIERS

In Table A1 we present Tycho 2, Hipparcos, and *Gaia* identifiers for the candidates observed at the INT (Table 1) and for the stars with $v > 350$ km s^{-1}(Table 2).

APPENDIX B: ASSUMING DIFFERENT PRIORS ON DISTANCE

One could argue that assuming a three-components stellar density (bulge + disc + halo) for our Galaxy $\rho_{\rm MW}(\boldsymbol{d})$, as in Equation 10, is not appropriate to model the spatial distribution of HVSs, a population of stars that, by definition, is not distributed according to the density profile of the Milky Way. Therefore in this appendix we discuss the implication of assuming different priors on distances $P(\boldsymbol{d})$ in the MCMC sampling described in §5. In practice we adopt two different priors and we test the impact of these choices on our results: an exponential decreasing prior $P_{\rm exp}(d)$, and a prior specifically tailored for HVSs, the *HVS prior* $P_{\rm HVS}(\boldsymbol{d})$, that we introduce in this paper.

Astraatmadja & Bailer-Jones (2016a) show that an exponential decreasing prior

$$P_{\rm exp}(d) \propto d^2 \exp\left(-\frac{d}{L}\right) \tag{B1}$$

with $L = 1.35$ kpc gives a better performance in terms of RMS errors compared to the MW prior, when resulting distance estimates are compared with GUMS simulated data. This choice assumes that the disc has the same scale-height as the scale-length, and clearly it is not an accurate description of the MW. We find that this prior overestimates distances for the majority of our candidates, with values well above the spectroscopic ones. This is evident in top panel of Figure B1, where for distances greater than ~ 600 pc we can see that median values obtained with the exponential prior are always higher than the ones derived with the MW prior. This is due to the choice of L, which sets the exponential cut-off of the distribution. Since $L = 1.35$ kpc is higher than the typical distance of stars in the TGAS calatogue, this prior biases our candidates towards greater distances, and thus towards higher total velocities, proper motions and radial velocities being equal.

Assuming a continuous and isotropic ejection of HVSs from the Galactic Centre, the number density of HVSs goes approximately as $1/r^2$, where r is the galactocentric radius (Brown 2015). Following Equation 10 we therefore construct the HVS prior as:

$$P_{\rm HVS}(d, l, b) \propto \left(\frac{d}{r(d, l, b)}\right)^2 p_{\rm obs}(d, l, b), \tag{B2}$$

with $r(d, l, b) = \sqrt{d^2 + d_\odot^2 - 2dd_\odot \cos(l)\cos(b)}$ and $d_\odot = 8$ kpc. When deriving distances and total velocities with this prior, we find again results to be consistent with the ones derived using the MW prior, but uncertainties are considerably larger, and this prior overestimates distances for further stars, as shown in bottom panel of Figure B1.

In the end, we choose to adopt the MW prior for presenting our results since it allows us to maintain a conservative approach: because of large uncertainties, we only interpret our candidates as HVSs at the end of the kinematic analysis, without biasing our distances and velocities using that assumption.

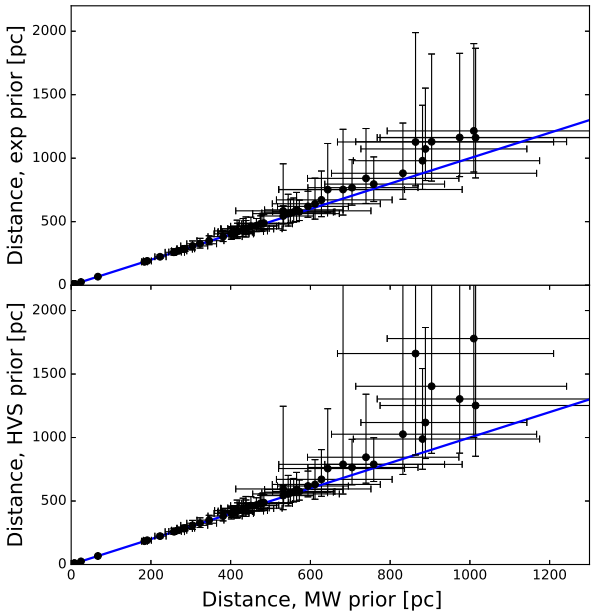

Figure B1. Comparison of distances obtained using the MW prior, on the x-axis, and the exponential decreasing (HVS) prior, y-axis on the top (bottom) panel. The blue line corresponds to equal estimates.