

Where did they come from? Where did they go? High-velocity stars in Gaia and SDSS

Nabeel Rehemtulla
(Northwestern University)

June 9, 2022

Abstract

High-velocity stars are uncommon remnants of a wide variety of astrophysical phenomena. Previous works have shown that they are influential to studies from Type Ia supernova progenitors to the Milky Way dark matter distribution. By the nature of their unusual kinematics, we start our investigation of them with obtaining their 6D phase space coordinates. Gaia provides proper motions, parallaxes, and sky positions, and SDSS provides spectra for use in inferring Earth-centric line-of-sight velocity. Once this information is present we perform orbital integration to trace their orbit backwards and forwards in time, and compute whether the star is bound or unbound to the Milky Way. We find a range of orbital properties even for a single star and its Monte Carlo samplings and only a single unbound star: Hivel15.

1 Introduction

The first hyper-velocity star SDSS J090745.0+024507 was discovered using data from the Sloan Digital Sky Survey (SDSS) and was observed to have a radial velocity of 853 km s^{-1} (Brown et al., 2005), the fastest velocity every recorded in the Milky Way (MW) at the time.

Identification of high-velocity and hyper-velocity stars is dependent on measuring a star’s space velocity. This has been made uniquely possible by recent observational efforts like Gaia whose data release 2 contained extremely precise proper motion measurements of more than one billion MW stars and relative to Earth line-of-sight velocities of more than seven million of those.

When a high-velocity star is discovered, one naturally asks questions to elucidate the context surrounding the star. We are most interested in where the star was born, where it is going, and how it was given such a large velocity.

One of the most interesting potential origins for a high-velocity star is an extragalactic source. A star can attain such an extreme velocity and leave their home galaxy after a near approach with its host galaxy’s supermassive black hole and later potentially make its way into the MW. These extragalactic stars are special tracers of the qualities of a region of their host galaxy (e.g. metallicity) well after leaving it. Marchetti (2021) and Marchetti et al. (2019) outline some of the fastest stars in the MW discovered in Gaia data, calculate the probability each is bound or unbound to the MW, and perform orbital integration to track where these stars came from to infer whether they may or may not be formed in the MW.

High-velocity stars can also provide clarity on the unresolved Type Ia supernova progenitor landscape. Shen et al. (2018) describe how three hyper-velocity white dwarfs found by Gaia are evidence defending the Dynamically driven double-degenerate double-detonation scenario. This is where two white dwarfs (WDs) orbit each other with mass transfer causing a loss of orbital energy until both He on the surface and the C/O in the core are ignited. Following this, there is the possibility that there is a supernova remnant along with a surviving white dwarf, which would be given an extreme velocity. Shen et al. (2018) note that further discoveries of hyper-velocity WDs will provide additional clarity towards Type Ia progenitors.

Once we have detected high-velocity stars we can use them to infer properties of the MW. When ejected from the Galactic center, we can make inferences on the MW’s dark matter distribution. Full 6D kinematics of SDSS J090745.0+024507 provided constraints to our dark matter halo’s triaxiality (Gnedin et al., 2005). Old metal-poor extreme velocity stars were used to produce a virial mass estimate of the MW through orbital integration (Hattori et al., 2018).

Each of these three domains has a great deal of information to be extracted from the star’s kinematics, metallicity, spectral type, age, and other properties. The key first steps to analyzing a HVS is performing orbital integration and understanding whether the star is bound or unbound to the MW. We do this by combining Gaia proper motions, sky positions, parallax distances, a line-of-sight velocity inferred from a spectrum, and an assumption of the MW potential. This analysis is more relevant than ever with Gaia’s full data release 3, publishing in June 2022, releasing more than 30 million radial velocities.

2 Data

Our study of high-velocity stars (HVSs) rests on the availability of their full 6D phase space coordinates. In observational terms, the coordinates with respect to Earth are the proper motion (in right ascension and declination), the line-of-sight velocity, the sky position (in right ascension and declination), and the line-of-sight distance.

We base our data set of HVSs off of a report of 591 HVS candidates with data from the Large Sky Area Multi-object Fiber Spectroscopic Telescope (LAMOST) and Gaia’s data release 2 (DR2) presented in [Li et al. \(2021\)](#) (henceforth Li21). We match these stars by sky position with sources in the SDSS database that have recorded spectra. The query is performed with the `Astroquery` ([Ginsburg et al., 2019](#)) library using a cone search with a 1 arcsecond radius centered on the Gaia DR2 sky position given in Li21. From these spectra we estimate the line-of-sight velocity of the matches with respect to Earth (see Sec. 3.2).

Gaia has made available extremely precise proper motions and sky positions for more than a billion Milky Way stars. A subset of these also have precise parallax measurements which we use to estimate distances to the HVSs with respect to Earth (see Sec. 3.1). We query Gaia’s early data release 3 (EDR3) for these 5 quantities and their uncertainties with `Astroquery` using a 1 arcsecond sided square search. We elect to not use the Gaia data provided in the Li21 list because it sources from DR2 while EDR3 data is now available with more precise measurements. Even with EDR3 data, there are some stars with imprecise parallax measurements, so we remove those matches with parallax measurements less than five times their parallax errors. This has the effect of removing stars with low signal-to-noise parallax measurements, including all listed with negative parallaxes.

With this matching and cutting, we build a list of 18 HVSs with Gaia astrometry and SDSS spectra for which we can obtain full 6D phase space coordinates.

Table 1: Inventory of high-velocity stars

Name	RA (deg)	Dec. (deg)	Parallax (mas)	Distance (kpc)	PM RA (mas yr ⁻¹)	PM Dec. (mas yr ⁻¹)	RV (km s ⁻¹)	Unbound?
Hivel15	250.95	43.61	0.231±0.042	4.330±0.784	-16.454±0.070	-36.536±0.106	-138.000±62.200	T
Hivel117	207.09	40.95	0.511±0.078	1.960±0.298	16.360±0.085	-54.615±0.100	69.100±62.400	F
Hivel141	234.15	17.89	0.180±0.035	5.550±1.080	-22.381±0.061	5.532±0.069	-0.084±62.300	F
Hivel216	197.21	50.02	0.196±0.036	5.090±0.922	0.939±0.043	-26.987±0.046	-138.000±62.400	F
Hivel217	201.94	40.68	0.423±0.045	2.360±0.250	-58.399±0.044	-10.270±0.051	-0.084±64.000	F
Hivel275	206.86	23.69	0.293±0.044	3.420±0.509	-35.313±0.079	-27.517±0.054	-207.000±62.500	F
Hivel321	198.11	31.21	0.276±0.043	3.620±0.559	-35.636±0.057	-3.181±0.054	-138.000±62.500	F
Hivel329	166.25	5.23	0.469±0.074	2.130±0.336	-27.349±0.105	-60.547±0.083	207.000±63.900	F
Hivel361	219.52	42.99	0.191±0.036	5.230±0.972	2.845±0.048	-22.444±0.060	-207.000±62.200	F
Hivel385	221.49	25.88	0.556±0.048	1.800±0.154	-57.245±0.083	20.637±0.090	-207.000±62.300	F
Hivel389	223.94	33.54	0.273±0.042	3.660±0.567	0.785±0.048	-33.304±0.065	-69.000±64.200	F
Hivel425	176.68	12.14	0.292±0.052	3.420±0.605	4.127±0.083	-37.862±0.068	207.000±62.200	F
Hivel432	191.67	13.43	0.648±0.093	1.540±0.222	27.116±0.173	-30.573±0.096	138.000±62.500	F
Hivel468	1.69	27.09	0.338±0.066	2.960±0.579	22.261±0.091	17.646±0.070	-0.084±62.500	F
Hivel500	224.69	4.45	0.442±0.070	2.260±0.356	-50.036±0.128	-10.018±0.118	207.000±62.500	F
Hivel563	168.28	48.02	0.340±0.067	2.940±0.578	-16.913±0.065	-4.763±0.107	-483.000±64.500	F
Hivel567	258.88	27.26	0.399±0.031	2.500±0.193	-48.039±0.046	-1.991±0.048	-276.000±62.400	F
Hivel588	210.74	37.81	0.365±0.055	2.740±0.409	0.532±0.050	-43.322±0.071	-69.000±64.100	F

Name from [Li et al. \(2021\)](#)

RA, Dec, Parallax, PM RA, PM Dec and uncertainties from Gaia EDR3

Distance inferred from Gaia parallax, see Sec. 3.1

RV inferred from SDSS spectrum, see Sec. 3.2

Unbound? Inferred from assumed Milky Way potential, see Sec. 3.4

3 Methods and Analysis

There are two main elements of analysis, orbital integration and bound/unbound classification. These are preceded by the computation of the two remaining kinematic components, line-of-sight distance and velocity. Once the full 6D phase space coordinates are assembled, we use `Astropy` ([Astropy Collaboration et al., 2013, 2018](#)) to transform to the Cartesian galactocentric reference frame. Both elements of analysis require an assumed Milky Way potential for which we use the `MilkyWayPotential`¹ model integrated in to `gala` ([Price-Whelan, 2017](#)), a galactic dynamics library. The `MilkyWayPotential` model contains 4 components: a nucleus, bulge, disk, and dark matter halo. Parameters for each of these are standard throughout the literature and are presented in the `gala` documentation.

3.1 Distance estimation from Gaia parallax

While there are more statistically robust methods for estimating distance from Gaia parallax measurements ([Bailer-Jones et al., 2018](#)), we use the basic parallax inversion method.

$$d \text{ [pc]} = 1''/\rho \quad (1)$$

Where d is the distance in parsecs, and ρ is the parallax in arcseconds. After removing the stars with large relative uncertainties, this parallax inversion method works well for our purposes. The uncertainty in the distance comes from propagation of the parallax uncertainty.

3.2 Line-of-sight velocity estimation from SDSS spectra

We employ cross correlation to quantify how well a manually shifted template aligns with a doppler shifted spectrum. We extract the spectral type of each star from the SDSS spectrum fits header, fetch the appropriate SDSS spectral template with `Astroquery`, and continuum normalize both the template and the observed spectrum. The cross correlation function tests the alignment of the template and source spectra for 10,000 individual line-of-sight velocities spaced between $-1,200 \text{ km s}^{-1}$ and $1,200 \text{ km s}^{-1}$. The shift value that produces the largest cross correlation response is chosen as the estimated line-of-sight velocity. The uncertainty on the line-of-sight velocity estimation is determined from the pixel-wavelength resolution of the spectrograph. We determine that each pixel in the SDSS spectra is approximately 1.4 Angstrom wide. Using the standard doppler shift formula, the line-of-sight velocity uncertainty evaluates to about 60 km s^{-1} at the location of $\text{H}\alpha$ (6563 Angstrom). This result does depend minutely on the wavelength chosen, although the cross correlation itself still considers the entire spectrum. $\text{H}\alpha$ is selected here because it is in all the spectra we use and is always one of the most prominent lines. Figure 1 shows an example output of this process for the star with the largest line-of-sight velocity magnitude in our sample, Hivel563. The cross correlation successfully aligns the $\text{H}\alpha$, and other, absorption lines in the observed spectrum with the template spectrum.

¹<http://gala.adrian.pw/en/latest/api/gala.potential.potential.MilkyWayPotential.html>

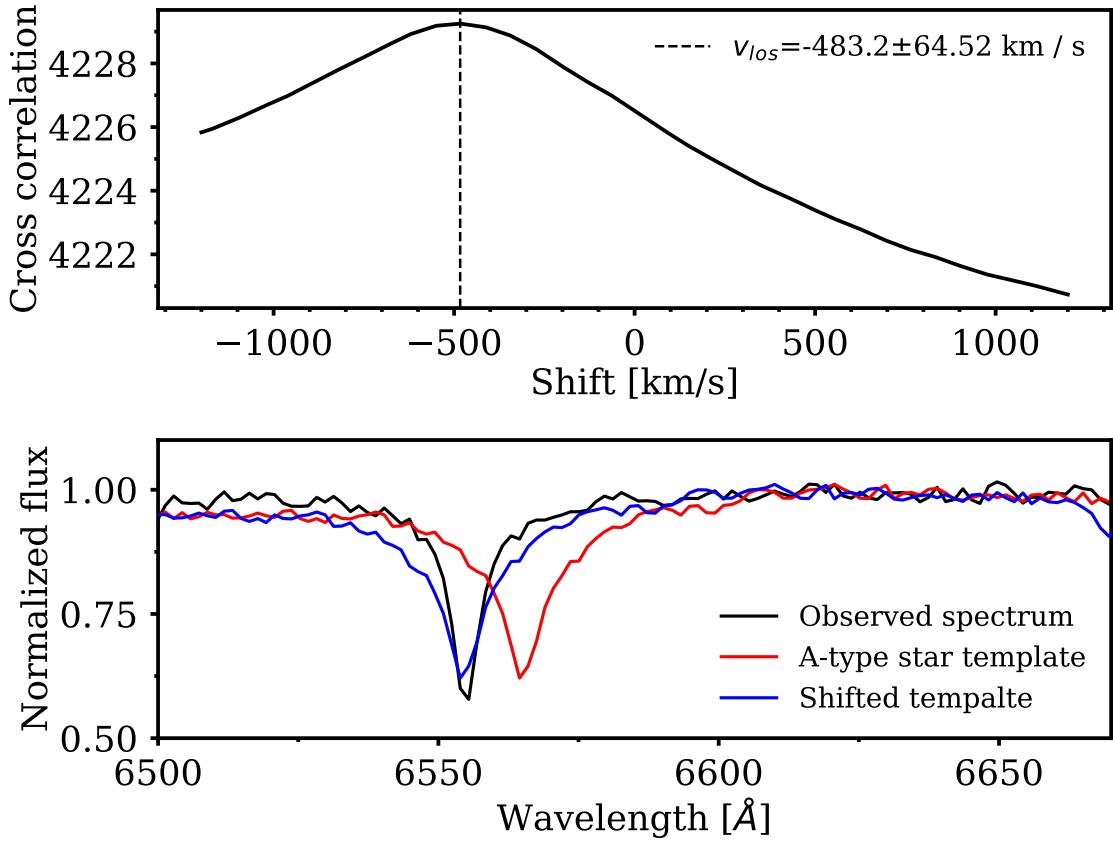


Figure 1: Doppler shifting template spectrum to match observed star spectrum and infer line-of-sight velocity. *Upper panel:* Cross correlation output for star Hivel563. The template aligns best with the observed spectrum when shifted by -483.2 km s^{-1} . Uncertainty comes from pixel-wavelength resolution of the spectrograph. *Lower panel:* Alignment of observed, unshifted template, and shifted template spectra in the region of H α .

3.3 Orbital integration

We integrate the orbit of each star forwards and backwards with `gala` for $N = 4000$ steps in either direction with $dt = 0.25$ Myr for a total of 2 Gyr of integration.

Uncertainties on the velocity components and the line-of-sight distance are considered using a Monte Carlo approach. We take $N = 100$ drawings from a gaussian centered at the original measurement with standard deviation equal to the uncertainty in the measurement. We then rerun the orbital integration for each sampling. Output of this analysis is shown for Hivel15, Hivel425, and Hivel567 in Figure 2 and the results for the remaining 15 are shown in the Appendix.

3.4 Bound/Unbound classification

We consider one of our high-velocity stars unbound to the Milky Way if its 3D velocity is greater than the escape velocity of the Galaxy at its position. We can easily compute the escape velocity of the Galaxy from the assumed potential model.

$$\Phi(\vec{r}) = \frac{v_{\text{esc}}^2(\vec{r})}{2} \quad (2)$$

Where Φ is the Galaxy's potential, and v_{esc} is the escape velocity.

Similarly to how uncertainties are considered for orbital integration in Sec. 3.3, we run $N = 250$ Monte Carlo samples of the uncertainty distributions on the kinematic components. Figure 3 shows the stars with their original measurements and their $N = 250$ Monte Carlo samplings in comparison to the escape velocity curve.

4 Results

4.1 Orbital integration

In the Monte Carlo samples shown by the faint orbital paths in Fig. 2, there is large variation in the shape of paths. This is chiefly due to the large line-of-sight velocity errors from the relatively low resolution SDSS spectra. The output for the three stars broadly demonstrates the three categories of orbital properties we observe. In the uppermost row, nearly all of Hivel15's Monte Carlo sampled orbits have the star making a close encounter with Sag A* and leaving the central region of the galaxy with a very large velocity. In the center row, some samples of Hivel425 have completed low-period orbits while others, similarly to Hivel15, show the star flying away from the center. In the lowermost row, nearly all of Hivel567's samples have it completing clearly bound, low-period orbits very near to Sag A*.

4.2 Bound/Unbound classification

Of our sample of 18 high-velocity stars, we classify only one star, Hivel15, as unbound to the Milky Way's gravitational potential. Most other stars have some, or even many, of their Monte Carlo samples above the escape velocity curve and thus would be considered unbound. The spread in the Monte Carlo samples is due almost entirely by uncertainties

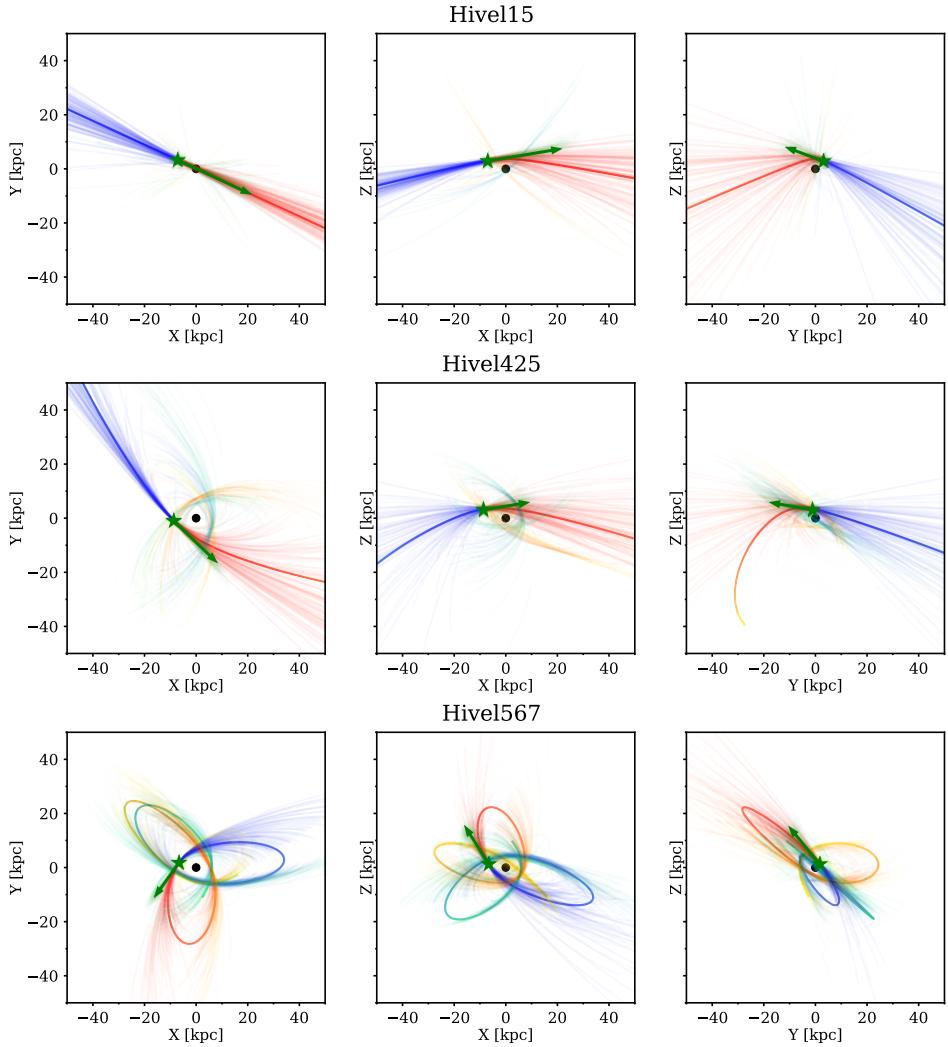


Figure 2: Orbital integration paths of three high-velocity stars in the Milky Way going forwards and backwards in time 1 Gyr each. The star's current position is shown by the green star and current velocity vector, scaled down for clarity, by the green arrow. Warm colored lines (red-orange-yellow) extending from the star show future path of star, and cold colored lines (blue-green-teal) show past path of the star; change in line color corresponds to passage of time. Sagittarius A* is shown as a black circle at (0,0). The faint lines represent $N = 100$ Monte Carlo samplings from the uncertainty distributions of the kinematic components. Due to large line-of-sight velocity errors, the behavior of the orbits vary dramatically over different Monte Carlo samples.

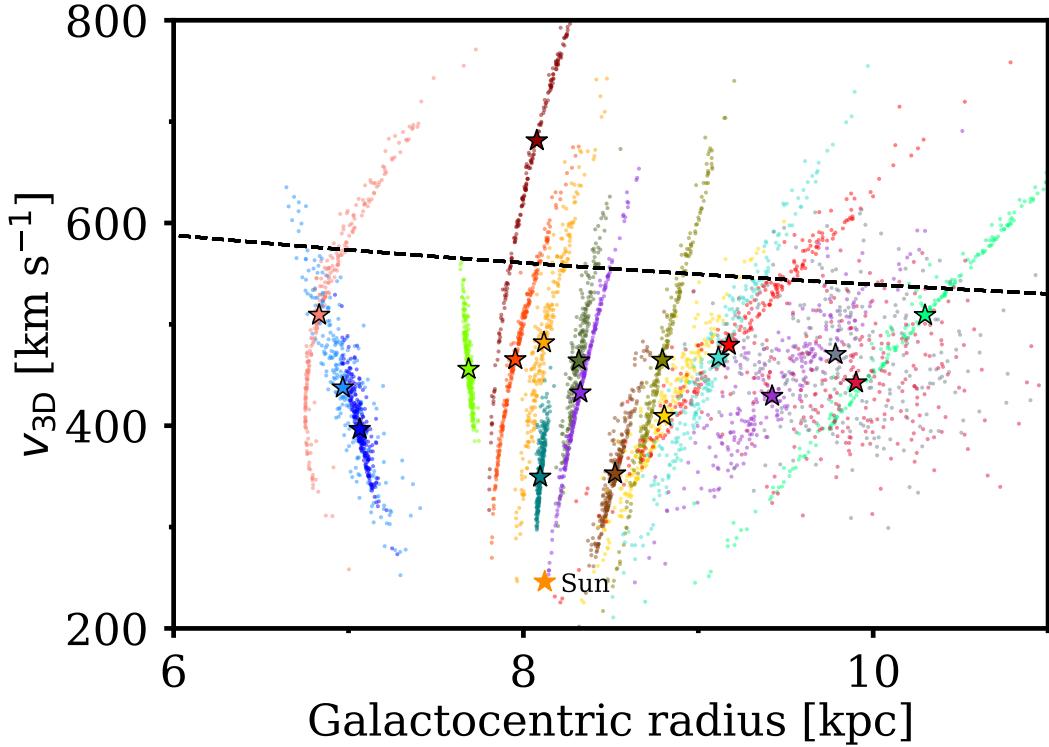


Figure 3: 3D velocities and radii of high-velocity stars in comparison to the escape velocity curve of an assumed Milky Way potential model. Colored stars with black outlines are high-velocity stars and colored points are Monte Carlo samplings of their uncertainty distributions. Stars above the escape velocity curve (dashed black line) are considered to be unbound to the potential. The maroon colored point with $v_{3D}(\sim 8 \text{ kpc}) = \sim 700 \text{ km s}^{-1}$ is Hivel15, the only unbound star in our sample.

on two quantities: the relative to Earth line-of-sight velocity and distance. We also consider errors on the proper motions but they are of dramatically smaller magnitude.

5 Conclusions

We match a list of high-velocity stars found in Gaia DR2 and followed up on by LAMOST (Li et al., 2021) with stars having recorded SDSS spectra. This yields a list of 18 high-velocity stars. We query SDSS for their spectra and Gaia EDR3 for the sky position, proper motions, and parallax. From the SDSS spectrum and Gaia parallax, we compute the line-of-sight velocity and distance respectively. At this point, we have obtained full 6D phase space coordinates for all of our high-velocity stars.

With this data we perform orbital integration and a bound/unbound classification. Both of these analyses are preceded by assuming a Milky Way potential, for which we adopt a simple 4-component (nucleus, bulge, disk, dark matter halo) model. Also done in both methods, we conduct Monte Carlo-like samplings from the uncertainty distributions of the star's position and velocity coordinates and rerun the analysis on each sample star.

The results of the orbital integration are dominated by the fact that the line-of-sight velocity uncertainties are much larger in fraction compared to the other components and their uncertainties. This causes most stars to exhibit a range of orbital properties over their many Monte Carlo samples. We identify 1 (Hivel15) of our 18 stars as unbound to the Milky Way's potential. We also find that, for most of our other stars, many of their Monte Carlo samples are unbound. These analyses, especially the past path of the orbital integration, can contribute to elucidating the mechanism by which these stars gain their unusually high-velocities.

In the future, this code base² could be applied to any Gaia SDSS star. When provided a seed RA and Dec, it will automatically search for matches across the SDSS and Gaia archives, estimate the line-of-sight velocity, distance, their respective uncertainties, perform orbital integration, and classify the star as bound or unbound. With minor adjustments it could be modified to function with any spectroscopic database and be used as a general purpose tool for basic analysis of high-velocity stars.

References

- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, [A&A, 558, A33](#)
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, [AJ, 156, 123](#)
- Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, [The Astronomical Journal, 156, 58](#), aDS Bibcode: 2018AJ....156...58B
- Brown, W. R., Geller, M. J., Kenyon, S. J., & Kurtz, M. J. 2005, [The Astrophysical Journal, 622, L33](#), aDS Bibcode: 2005ApJ...622L..33B
- Ginsburg, A., Sipőcz, B. M., Brasseur, C. E., et al. 2019, [AJ, 157, 98](#)

²Made publicly available at <https://github.com/nabeelre/gaia-sdss-hvs>

- Gnedin, O. Y., Gould, A., Miralda-Escudé, J., & Zentner, A. R. 2005, [The Astrophysical Journal](#), **634**, 344, aDS Bibcode: 2005ApJ...634..344G
- Hattori, K., Valluri, M., Bell, E. F., & Roederer, I. U. 2018, [The Astrophysical Journal](#), **866**, 121, aDS Bibcode: 2018ApJ...866..121H
- Li, Y.-B., Luo, A. L., Lu, Y.-J., et al. 2021, [The Astrophysical Journal Supplement Series](#), **252**, 3, aDS Bibcode: 2021ApJS..252....3L
- Marchetti, T. 2021, [Monthly Notices of the Royal Astronomical Society](#), **503**, 1374, aDS Bibcode: 2021MNRAS.503.1374M
- Marchetti, T., Rossi, E. M., & Brown, A. G. A. 2019, [Monthly Notices of the Royal Astronomical Society](#), **490**, 157, aDS Bibcode: 2019MNRAS.490..157M
- Price-Whelan, A. M. 2017, [The Journal of Open Source Software](#), **2**
- Shen, K. J., Boubert, D., Gänsicke, B. T., et al. 2018, [The Astrophysical Journal](#), **865**, 15, aDS Bibcode: 2018ApJ...865...15S

6 Appendix

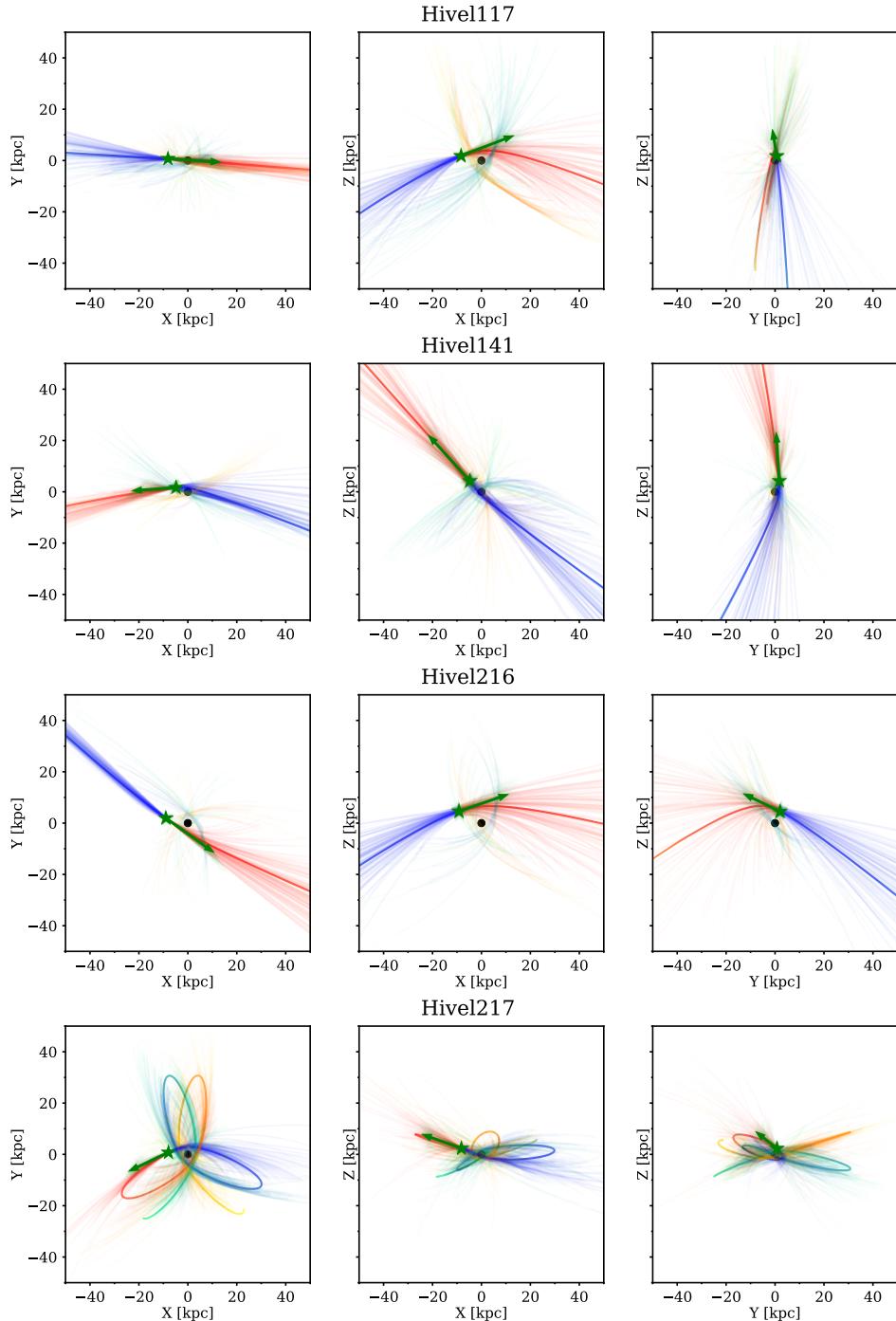


Figure 4: Same as Fig. 2 for Hivel117, Hivel141, Hivel216, Hivel217

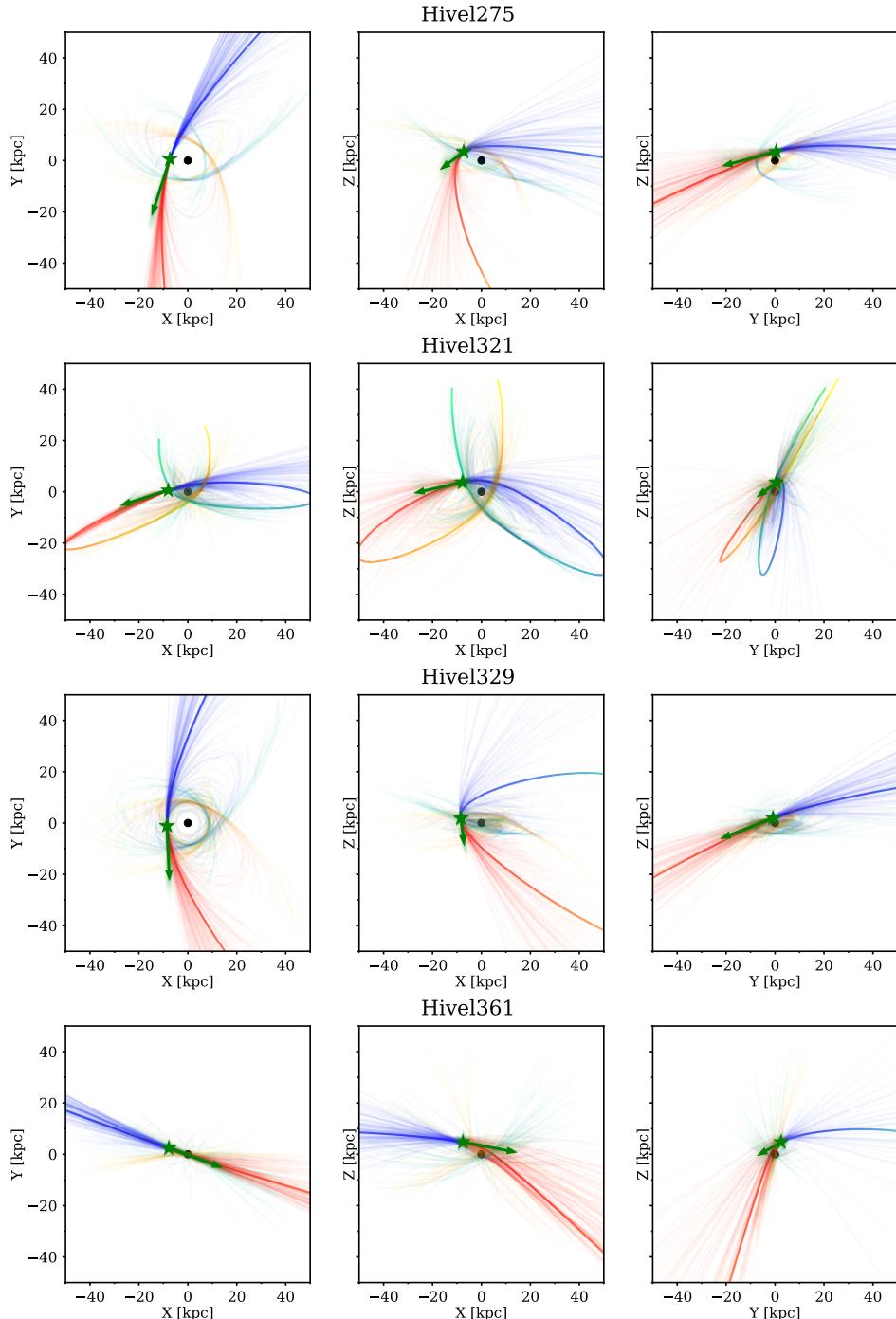


Figure 5: Same as Fig. 2 for Hivel275, Hivel321, Hivel329, Hivel361

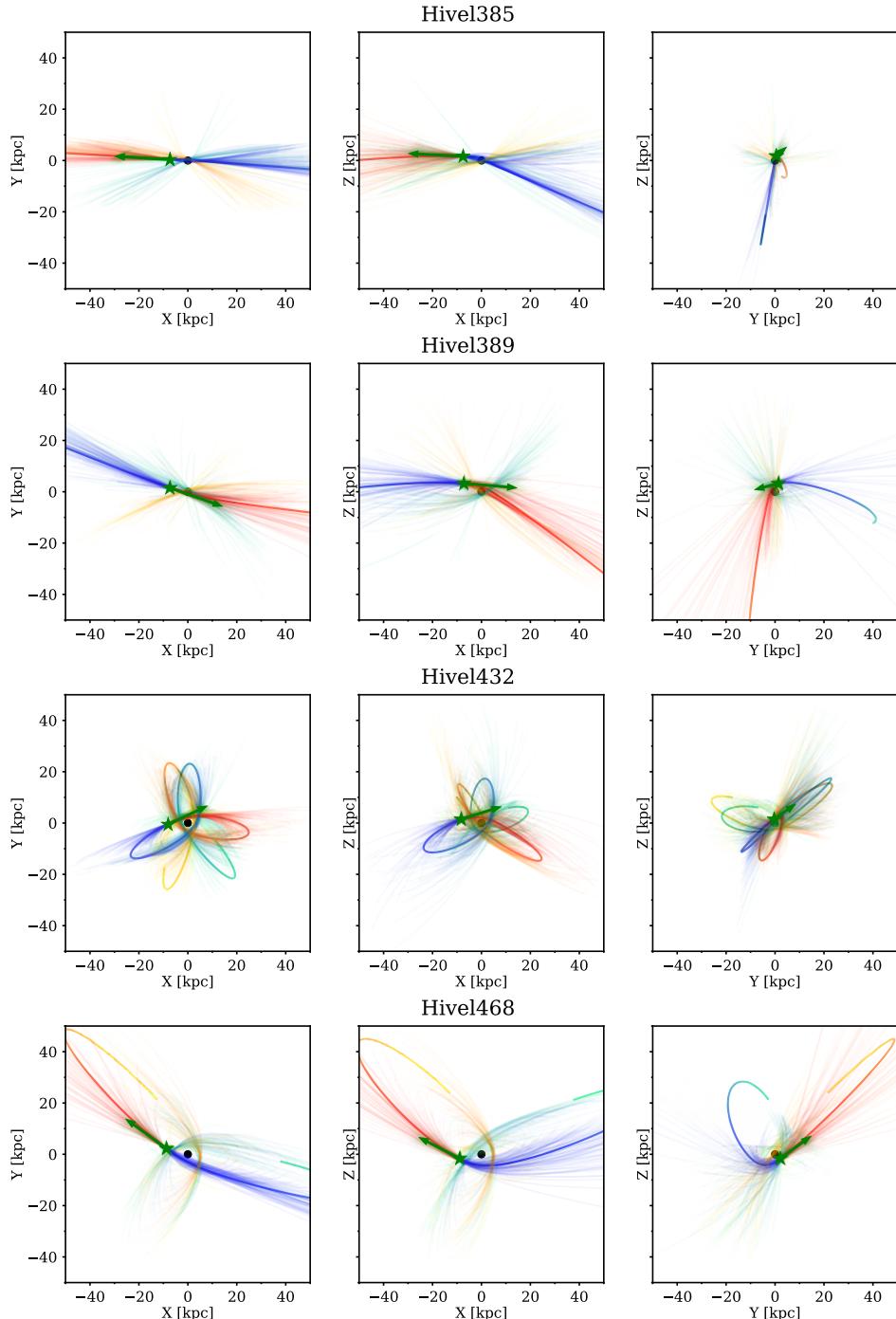


Figure 6: Same as Fig. 2 for Hivel385, Hivel389, Hivel432, Hivel468

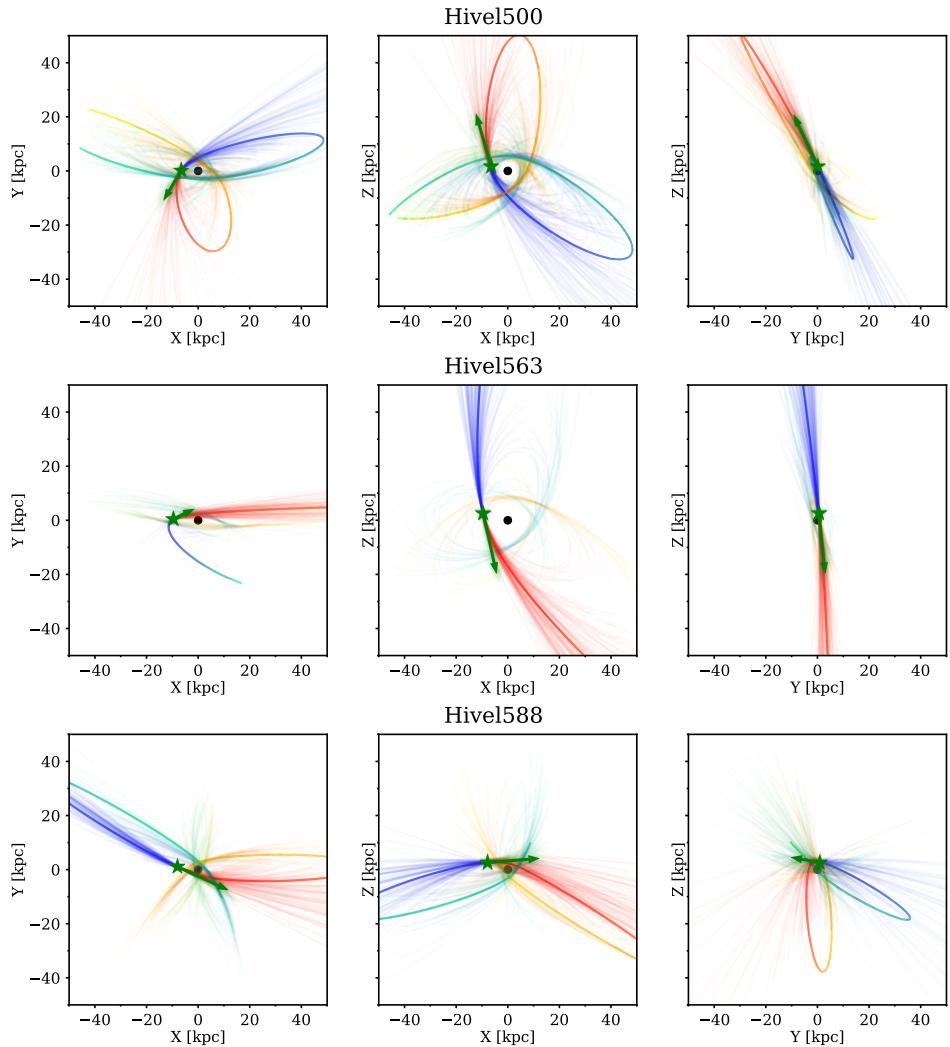


Figure 7: Same as Fig. 2 for Hivel500, Hivel563, Hivel588