Assessing distances and consistency of kinematics in *Gaia*/TGAS

Ralph Schönrich[★] and Michael Aumer

Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, Oxford OX1 3NP, UK

ABSTRACT

We apply the statistical methods by Schönrich, Binney & Asplund to assess the quality of distances and kinematics in the Radial Velocity Experiment (RAVE)–Tycho-Gaia Astrometric Solution (TGAS) and Large Sky Area Multiobject Fiber Spectroscopic Telescope (LAMOST)–TGAS samples of Solar neighbourhood stars. These methods yield a nominal distance accuracy of 1–2 per cent. Other than common tests on parallax accuracy, they directly test distance estimations including the effects of distance priors. We show how to construct these priors including the survey selection functions (SSFs) directly from the data. We demonstrate that neglecting the SSFs causes severe distance biases. Due to the decline of the SSFs in distance, the simple 1/parallax estimate only mildly underestimates distances. We test the accuracy of measured line-of-sight velocities (v_{los}) by binning the samples in the nominal v_{los} uncertainties. We find: (i) the LAMOST v_{los} have a $\sim -5\,\mathrm{km\,s^{-1}}$ offset; (ii) the average LAMOST measurement error for v_{los} is $\sim 7\,\mathrm{km\,s^{-1}}$, significantly smaller than, and nearly uncorrelated with the nominal LAMOST estimates. The RAVE sample shows either a moderate distance underestimate, or an unaccounted source of v_{los} dispersion ($e_{\parallel}$) from measurement errors and binary stars. For a subsample of suspected binary stars in RAVE, our methods indicate significant distance underestimates. Separating a sample in metallicity or kinematics to select thick-disc/halo stars, discriminates between distance bias and $e_{\parallel}$. For LAMOST, this separation yields consistency with pure v_{los} measurement errors. We find an anomaly near longitude $l \sim (300 \pm 60)^{\circ}$ and distance $s \sim (0.32 \pm 0.03)\,\mathrm{kpc}$ on both sides of the galactic plane, which could be explained by either a localized distance error or a breathing mode.

Key words: parallaxes – stars: distances – stars: kinematics and dynamics – Galaxy: kinematics and dynamics – solar neighbourhood.

1 INTRODUCTION

The *Gaia* satellite mission (Perryman 2001; Gaia Collaboration 2016, 2016b) is the most important survey of this decade for the fields of Galactic and stellar astronomy. It is retrieving precision astrometry for of order one billion stars throughout the Galaxy, coupled to spectroscopic observations that allow determination of stellar parameters and line-of-sight velocities for a large subset of these stars. In addition, this effort is coupled to a plethora of ground-based spectroscopic surveys, including Radial Velocity Experiment (RAVE), Gaia–ESO, APOGEE, SEGUE, GALAH, Large Sky Area Multiobject Fiber Spectroscopic Telescope (LAMOST), and WEAVE.

Since the first release of Tycho-Gaia Astrometric Solution (TGAS) data (Gaia Collaboration 2016; Lindegren et al. 2016), there has been a variety of papers examining the quality of the data, ranging from the original validation paper (Lindegren et al. 2016)

to studies with standard candles, like red giants and red giant clump stars (de Ridder et al. 2016; Davies et al. 2017). A thorough compilation of standard validations of TGAS astrometry is found in Arenou et al. (2017). They discuss significant deviations of the median parallax from 0 for remote objects in different regions of the sky, e.g. for objects in the Magellanic Clouds. Most such quality tests for parallaxes p rely on distant objects, which should have $p = 0$ arcsec (to within at most a few μas), i.e. quasars, galaxies, or objects in neighbouring galaxies. For such tests, the main caveats are potential contamination with Galactic sources, and to a minor extent their different geometry, and different spectral distribution.

For Milky Way (MW) studies, it is vital to have a direct validation also of larger parallax measurements. However, well-constrained standard candles like Cepheids are not numerous enough to allow for resolved, high-precision testing. Most other types of standard candles, like RR Lyrae stars, red clump stars, or blue horizontal branch stars are more distant than the local dwarf star samples. More importantly, their luminosities depend on additional parameters, like helium abundance, age, or metallicity. While relatively well controlled via the period dependence in pulsating stars, this
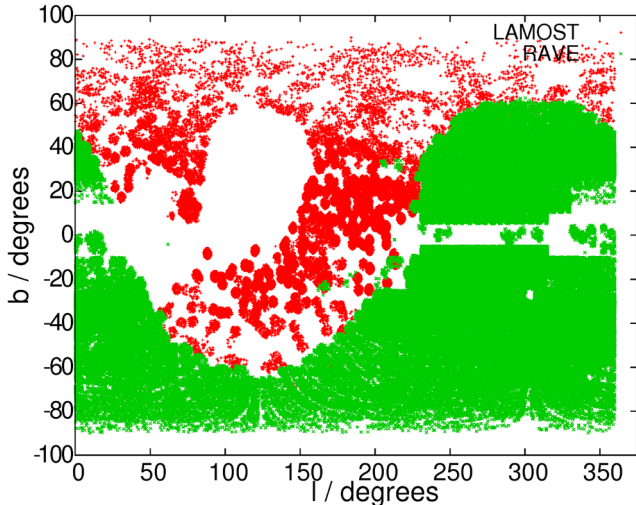
Figure 1. Positions of stars in the RAVE–TGAS (green) and LAMOST–TGAS (red) samples in galactic latitude b versus longitude l.

Figure 2. Schematic visualization of coordinates, velocity components in a local heliocentric Cartesian frame (U, V, W) versus local galactocentric Cartesian frame (U_g, V_g, W) and angle definitions for the observation of a star from the position of the Sun. In this plot, the Galactic disc rotates clockwise. The angle between U_g and U is α.

is a major issue for the more numerous red clump stars. Girardi & Salaris (2001) and Salaris & Girardi (2002) have found dependencies in excess of 0.1 mag, severely limiting their value to test *Gaia* parallaxes at high precision.

A key point for these validations is that even a perfectly unbiased parallax measurement does not yet imply unbiased distances. Each parallax measurement will still be uncertain, and so to estimate stellar distances, one has to translate the parallax error distribution into a distance distribution and combine the full information on this likelihood distribution (Strömberg 1927; Astraatmadja & Bailer-Jones 2016) with models for the density of stars in each region, and, (if stellar parameters are not directly used) with the selection function. Each of these factors can bias the derived stellar kinematics. They are frequently neglected in attempts to test and validate *Gaia* data on standard candles (see e.g. de Ridder et al. 2016; Davies et al. 2017). For everyone studying stellar kinematics, the main question is not if the parallaxes are unbiased, but if the derived stellar distances are unbiased. In this paper, we will apply a statistical distance estimator (Schönrich, Binney & Asplund 2012, hereafter SBA) that does exactly this.

The RAVE–TGAS sample, released with RAVE DR5 (Kunder et al. 2016), is the first major combined sample of ground-based spectroscopic data from the RAVE (Steinmetz et al. 2006) with combined astrometry from *Gaia* (Gaia Collaboration 2016; Lindegren et al. 2016), *Tycho-2* (Høg et al. 2000), and *Hipparcos* (Perryman 1997; van Leeuwen 2007). The RAVE–TGAS sample is also particularly interesting, since RAVE spectra have similar resolution and wavelength coverage to those of the *Gaia* spectrograph. The LAMOST survey (Wang et al. 1996; Liu et al. 2015; Luo et al. 2015) also overlaps with TGAS. As seen in Fig. 1, it complements RAVE on the northern sky and is of great use to study the outer disc regions of the MW.

The aim of this paper is to derive stellar distances from the *Gaia* parallaxes, and to validate the resulting kinematics. We make these data publicly available.[1] For the validation, we use the method of SBA, which employs correlations between either radial (U) or azimuthal (V) with vertical (W) velocities and combinations of galactic coordinates. The correlation between U and W velocity is affected by the vertical turn of the velocity ellipsoid above the plane, while the correlation between V and W velocities in disc samples can be affected by spiral breathing modes (Debattista 2014; Faure, Siebert & Famaey 2014), or the Galactic warp (Dehnen 1998; Poggio et al. 2017). That the MW has some warp and possible vertical waves is known both from observations in gas (Burke 1957; Kerr 1957) and stars (Djorgovski & Sosin 1989; Xu et al. 2015), but a possible imprint on local stellar kinematics demands scrutiny, since observational errors are known to cause similar apparent correlations between heliocentric velocities (SBA). We will discuss our finding of a warp signal in a second paper.

The structure of this paper is as follows: In Section 2, we describe the data sets we use, the merging of data from different sources, and our quality cuts. Section 3 provides a description how we calculate distances from parallaxes. In Section 4, we describe how our distance estimator, can be used to measure to high accuracy the mean distance bias in a sample and the v_{los} source dispersion $e_\parallel$. In Section 5, we apply this method to the RAVE–TGAS and LAMOST–TGAS samples. We show how a separation of the samples into subsets can be used to validate and measure the line-of-sight velocity uncertainties given in a pipeline. We also show how we can derive the effective spatial selection function for each sample and how our distance statistics can be used to validate the results. Our conclusions are found in Section 6.

2 DATA AND DEFINITIONS

2.1 Coordinate frame and definitions

For all subsequent discussions, please refer to Fig. 2 for definitions of our coordinate frame. We use either the heliocentric Cartesian frame with velocities (U, V, W) radially towards the Galactic Centre, azimuthally in the direction of Galactic rotation, and vertically out of the plane, or the galactocentric cylindrical coordinate frame, marked by indices g, naming the velocity vector (U_g, V_g, W). The galactocentric radius is termed R, the distance to the Galactic Centre r, Galactic longitude and latitude are l and b. The angle $\beta = \tan^{-1}(z/R)$ is the angle between the Galactic plane and the connection line star-Galactic Centre. The angle between the Sun-centre and the star-centre connection lines is called α. We use the Sun's galactocentric distance $R_0 = 8.27$ kpc, the Sun's total azimuthal

---

[1] Please find the data sets with distances and kinematics and our source code at http://www-thphys.physics.ox.ac.uk/people/RalphSchoenrich/data/tgasdist/data.tar.gz or request them directly from the authors.
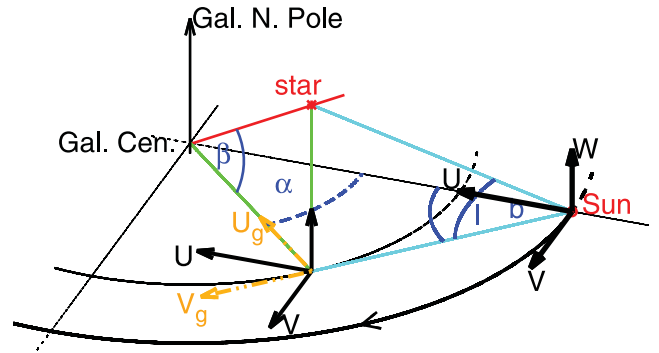
velocity $\boldsymbol{v}_\odot = V_c + V_\odot = 248 \, \text{km s}^{-1}$ in the Galactic rest frame, where V_c is the local circular speed of the disc, and the Solar motion with respect to the local standard of rest is $(U_\odot, V_\odot, W_\odot) = (13, 12.24, 7.24) \, \text{km s}^{-1}$. These values are chosen in concordance with Schönrich (2012), McMillan (2011, 2017), and Schönrich, Binney & Dehnen (2010). Following Joshi (2007), we place the Sun at $z = 20 \, \text{pc}$ above the Galactic mid-plane.

2.2 TGAS data

For this work, we cross-match TGAS data from *Gaia* DR1 (Gaia Collaboration 2016; Lindegren et al. 2016) both with RAVE DR5 (Kunder et al. 2016) and with DR2 of LAMOST (Luo et al. 2015). As can be seen in Fig. 1, the samples would together provide a full sky coverage. While LAMOST DR2 has far more stars, the catalogue covers significantly fainter stars giving a smaller common footprint with TGAS. More importantly, LAMOST's line-of-sight velocity determinations are about one order of magnitude less precise than in the RAVE survey, so on LAMOST our statistics will rather test their v_{los} measurements than the distances. To ensure sufficient quality of the data, we adopt a few general cuts:

(i) A derived distance $s > 30 \, \text{pc}$. Lindegren et al. (2016) report that TGAS is biased against stars with a proper motion larger than $\approx 3.5 \, \text{arcsec yr}^{-1}$. To prevent kinematic bias (a cut in proper motions might feign a distance underestimate), we drop the few stars with $s < 30 \, \text{pc}$, where a proper motion of 1 arcsec yr^{-1} would imply a transverse velocity $|v_\perp| \approx 140 \, \text{km s}^{-1}$.

(ii) Limiting $-100 < V_g/\text{km s}^{-1} < 310$: again this cut in heliocentric azimuthal velocity affects only a handful of stars, and we have tested that our results are robust against the precise choice. However, damaged data tend to assemble in the extreme wings, so we drop them.

(iii) Limiting the line-of-sight motion $|v_{\text{los}}| < 500 \, \text{km s}^{-1}$: excising another few objects is justified by indications of erroneous velocity determinations at extreme $|v_{\text{los}}|$ in RAVE.

2.3 RAVE

The RAVE DR5 sample contains 520 701 entries. However, about 15 per cent of the entries are multiple measurements of the same object. Of the 68 822 duplicate entries, we choose the one that has a smaller radial velocity (RV) error given by the RAVE pipeline, leaving us with 451 879 unique stars. To test the cross-match between the RAVE DR5 sample with TGAS, we have compared two different strategies: (i) using the already cross-matched positions in Kunder et al. (2016) and demanding a precise match in position to the *Gaia* catalogues, which gives us 210 415 cross-matched stars (by identifier we find 210 368 unique stars), and (ii) using the positions in RAVE versus TGAS positions and proper motions, demanding that the position and proper motion in TGAS predicts within 3 arcsec the original position in the RAVE catalogue. The latter exercise gives us 215 640 stars, of which a few dozen are apparent binaries in the *Gaia* catalogue. The *Gaia* data are provided at epoch 2015.0 and use the International Celestial Reference System (Arias et al. 1995; Feissel & Mignard 1998). To transform the celestial coordinates to Galactic coordinates l and b, we use the values $(\alpha_G, \delta_G, l_\Omega) = (192.85948, +27.12825, 32.93192)^\circ$ provided in ESA (1997).[2] We have compared our results from the different selections

[2] See also gaia.esac.esa.int Section 3.1.7 for reference.

and cross-matches, and since these choices do not significantly affect our results, we report here only results for the original RAVE DR5 cross-match. We further exclude all stars that are flagged as cluster members in the RAVE data set, or have a line-of-sight velocity error given as $\sigma_{\text{los,RAVE}} = 0 \, \text{km s}^{-1}$. If not explicitly stated otherwise, we use the quality cut $\sigma_{\text{los,RAVE}} < 5 \, \text{km s}^{-1}$, which will be justified in Section 5.2. A large number of stars in this cross-matched sample will fail our variable cut for the precision of the measured TGAS parallax. E.g. the RAVE subsample under our quality restrictions with parallaxes (p_0) better than 20 per cent, i.e. $p_0/\sigma_p > 5$, has only 88 464 members.

An important classification of RAVE spectra is given by the flags from Matijevič et al. (2012). These flags list the 20 closest matches to each RAVE spectrum for different classes of stars, e.g. normal stars (marked with 'n'), chromospherically active stars, or suspected binaries. Throughout this paper, we use these flags to define three subsets of RAVE: (i) all stars in RAVE irrespective of their classification, (ii) unflagged stars, i.e. stars that have all 20 flags set to 'n', and (iii) suspected binaries, which have at least one flag set to 'b'.

2.4 LAMOST

To cross-match the LAMOST sample with TGAS, we use the TGAS proper motions and positions to calculate back to the suspected position in LAMOST at epoch 2000. For a proper match, we demand $\sqrt{\Delta_{\text{RA}}^2 + \Delta_{\text{Dec.}}^2} < 1.2 \times 10^{-3} \text{deg}$ (or equivalently < 4.3 arcsec), between the predicted position and the position given in the LAMOST catalogue. Repeat observations in LAMOST are purged by taking for each object in TGAS the best match in position, or at equal match, the latest entry in the LAMOST catalogue. This leaves us with 107 663 stars in the cross-matched sample. To avoid kinematic biases in the selection function, which could affect our statistics, we remove the region of $0 < \text{RA deg}^{-1} < 67$ and $42 < \text{Dec. deg}^{-1} < 59$ from the LAMOST sample, since Luo et al. (2015) report that this region contains a subsample of plates that have a proper motion selection, limiting the stellar proper motions to smaller than 7 mas yr^{-1}. There might be minor contamination of in-plane fields with kinematic selections, but since the distance statistics we choose for this analysis are not sensitive to stars in the Galactic Plane, we judge the above exclusion region to be safe. Demanding a parallax uncertainty better than 20 per cent and a signal-to-noise ratio (S/N) > 30, the usual line-of-sight velocity cut $|v_{\text{los}}| < 500 \, \text{km s}^{-1}$ and a galactocentric azimuthal velocity of $-100 < V_g/\text{km s}^{-1} < 310$, we have a remaining sample of 34 384 stars. The S/N cut does not strongly affect the sample size, since the cross-match with *Gaia* includes mostly bright objects; with a cut at S/N > 5 the sample size would only increase to 38 834 stars.

We will show in Section 5.2 that line-of-sight velocities in LAMOST have to be corrected by adding $\delta v_{\text{los}} = 5 \, \text{km s}^{-1}$. Further, we will limit the sample to a nominal v_{los} measurement error $\sigma_{\text{los,LAM}} < 27 \, \text{km s}^{-1}$, and replace their estimate by using $\sigma_{\text{los}} = 7 \, \text{km s}^{-1}$, motivated by Section 5.5.

3 DISTANCE DETERMINATIONS

For the sake of well-defined error distributions in our distance determination, we will rely solely on parallaxes. Since the reported parallax errors are all larger than 0.2 mas, a 10 per cent (or 20 per cent) quality cut on the parallax will limit the sample to within ~0.5 kpc (1 kpc), with some weak dependence on the prior used. With parallax errors still a significant fraction of the measured parallax, it

is not advisable to use just the inverse parallax as a distance. For the derivation of distances from parallaxes, see Strömberg (1927), Schönrich & Bergemann (2014), or Astraatmadja & Bailer-Jones (2016). Since we are using exclusively the parallaxes to assess stellar distances, the magnitude-based selection function of the sample has to be written into the prior, so that the posterior distribution in distance is:

$$P(s) = N^{-1} s^2 G(p, p_0, \sigma_p) \rho(s(p), l, b) S(s(p)), \qquad (1)$$

where

$$N = \int ds \, s^2 G(p, p_0, \sigma_p) \rho(s(p), l, b) S(s(p)) \qquad (2)$$

is the normalization. $G(p, p_0, \sigma_p)$ is the observational likelihood of the parallax given the measurement, (p_0, σ_p), s^2 is the geometric factor, i.e. surface of the observational cone at fixed distance s, $\rho(s, l, b)$ is the supposed density of stars in the direction of galactic longitude and latitude (l, b), and S is the selection function, which we approximate as a function of distance s only. The expectation value for the distance of each star is then just:

$$\langle s \rangle = N^{-1} \int ds \, s^3 G(p, p_0, \sigma_p) \rho(s(p), l, b) S(s(p)). \qquad (3)$$

For the density prior ρ, we use the simple thin/thick disc + halo decomposition from equation (32) in Schönrich & Bergemann (2014), i.e. our density model simply reads as:

$$\rho(R, z) = e^{-\frac{(R-R_0)}{R_d}} e^{-\frac{|z|}{z_0}} + a_t e^{-\frac{(R-R_0)}{R_d}} e^{-\frac{|z|}{z_{0,t}}} + a_h \left(\frac{r}{R_0}\right)^{-2.5}, \qquad (4)$$

where R is the galactocentric radius in cylindrical coordinates, R_0 is the (cylindrical) galactocentric radius of the Sun, R_d is the scale-length of the disc, chosen at $R_d = 2.5 \, \text{kpc}$, z the altitude above the plane, $z_0 = 0.3 \, \text{kpc}$ and $z_{0,t} = 0.9 \, \text{kpc}$ are the scaleheights of the thin and thick discs, r is the galactic centre distance used for a simple spherical halo model, $a_t = 0.12$ is a local thick disc normalization set in concordance with Ivezić et al. (2008), $a_h = 0.001$ normalizes the halo component. We note that the placement of the Sun above the plane (we assume it at $z_\odot = 20 \, \text{pc}$ has virtually no influence on the distance estimates. When assuming a vertical position of the Sun either at $z_\odot = 0$ or $50 \, \text{pc}$, every single star in the entire RAVE sample experiences a relative distance change smaller than 0.01, the rms dispersion of the fractional distance differences is 3×10^{-4} and the mean offset is 8×10^{-5}. This is to some part due to the dominance of the selection function, to another part due to the inclination of most sightlines against the vertical direction.

Of great importance are, however, the selection function $S(s)$ and the choice, which statistical quantity we use. Fig. 3 shows the posterior distance distribution $P(s)$ according to equation (1) with a solid blue curve for a typical RAVE star with $\sigma_p/p \sim 0.2$. As a comparison, we show with a dashed red line the posterior distance distribution $P_0(s)$ when we would neglect the selection function $S(s)$ (depicted in arbitrary units with a dashed green line). The selection function favours at this distance more nearby stars, and so the posterior distribution loses most of its tail towards long distances. This strongly reduces the distance expectation value from $\langle s \rangle_0 \sim 0.6 \, \text{kpc}$ down to $\langle s \rangle \sim 0.47 \, \text{kpc}$. The median and mode (which is not a sensible statistical quantity due to its dependence on the axis scaling) of the distribution are shifted in the same direction, but react slightly less. Since the parallax probability distribution is symmetric, the resulting posterior probability in distance is strongly skew, and so the expectation value is a lot larger than the median and in particular the mode/maximum.
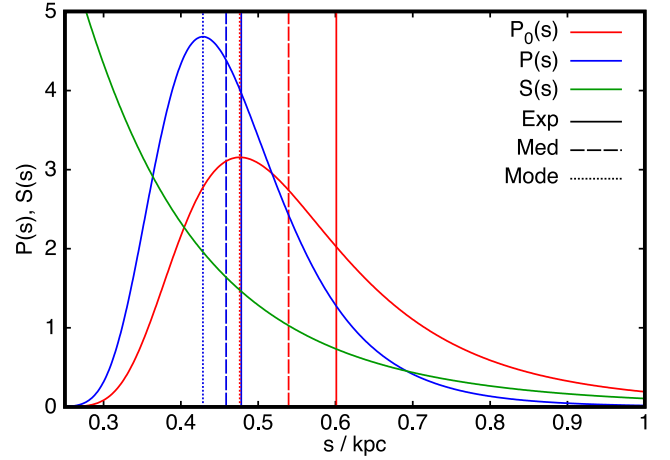
Figure 3. A typical posterior probability distribution $P(s)$ (solid blue line) for a star with $\sigma_p/p \sim 0.2$. In the same plot, we show the selection function $S(s)$ for the full RAVE sample (green) in arbitrary units, and the posterior probability distribution $P_0(s)$ (red), when we neglect $S(s)$. The vertical lines show expectation value (solid), median (dashed), and mode (dotted lines) for each distribution.

In short, we have to take two things from this: (i) the mode is not a sensible quantity to use and both median and mode underestimate the expectation value. (ii) The selection function strongly affects all distance estimates.

The selection function S would typically require the use of a full population synthesis model: it is apparent that the magnitude limits of the RAVE survey will introduce a strong bias towards nearby and more luminous stars, which ultimately favours younger stellar populations. Stellar metallicities affect brightness and colour, thus influencing the selection function as well. While S will hence in general not be isotropic, we approximate $S(s)$ as a function of distance s only. We discuss its derivation in Section 5.3.1.

4 ASSESSING DISTANCES AND LINE-OF-SIGHT VELOCITY DISTRIBUTION ERRORS

To assess distances, we make use of the statistical distance method of SBA, where a full derivation of all equations can be found. In addition to the theoretical justification, we have performed numerous validations on the original method. Since we are using a restricted approach in this paper, we present in Appendix A, a short validation on N-body galaxy simulations, which are structurally similar to the MW.

This method exploits position dependent correlations between different heliocentric velocity components. A stellar motion relative to the Sun (U_0, V_0, W_0) results in observable line-of-sight velocity v_{los} and proper motions μ_l and μ_b, depending on the distance s_0 to the star, calculated with the translation matrix M. Since this translation matrix is orthogonal, we have

$$\begin{pmatrix} U_0 \\ V_0 \\ W_0 \end{pmatrix} = M \begin{pmatrix} s\mu_l \\ s\mu_b \\ v_{los} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} s\mu_l \\ s\mu_b \\ v_{los} \end{pmatrix} = M^T \begin{pmatrix} U_0 \\ V_0 \\ W_0 \end{pmatrix} \qquad (5)$$

If our distance estimate s for the star is wrong by a fraction f, the terms containing the proper motions will be stretched by

$(1 + f) = s/s_0$, and our estimated velocity vector (U, V, W) for the star will be altered to

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = (\mathrm{I} + f\,\mathrm{T}) \begin{pmatrix} U_0 \\ V_0 \\ W_0 \end{pmatrix}, \qquad (6)$$

where $\mathrm{T} = \mathrm{MPM}^{\mathrm{T}}$, I is the identity matrix, and P is the identity for the proper motion terms, but zero in the v_{los} term. For the following discussion, we only need the terms of T connecting the vertical motion to all velocity components, i.e.

$$\begin{aligned} T_{\mathrm{UW}} &= \cos(l)\sin(b)\cos(b), \\ T_{\mathrm{VW}} &= \sin(l)\sin(b)\cos(b), \text{ and} \\ T_{\mathrm{WW}} &= \cos^2(b). \end{aligned} \qquad (7)$$

Take for example the motion of the Sun against a sample of halo stars. The Sun has an azimuthal velocity of roughly $V_\odot \sim 250\,\mathrm{km\,s^{-1}}$, while halo stars have no net azimuthal motion. T_{VW} is maximal when we look up out of the galactic plane at an angle $b \sim 45°$ in or against the direction of solar motion. To get a picture on how this term links vertical heliocentric velocities W with azimuthal velocities V, imagine approaching horizontally the wall of your room, while looking at the corners of the floor and the ceiling. Both corners will be blueshifted in their line-of sight velocity, but you also observe an angular motion. If your distance estimate to the corners is correct, the conclusion will be a zero net motion of the room. However, if the distance is overestimated, the angular motion of the upper corner will be mistakenly translated into a net upwards motion, and similarly, the lower corner will appear to move downwards. Similarly, if we overestimate stellar distances in our Galactic halo example, we will detect a net upwards motion (positive W) maximal near $(l, b) \sim (90°, 45°)$, and $(l, b) \sim (270°, -45°)$, and a downwards motion (negative W) around $(l, b) \sim (90°, -45°)$, and $(l, b) \sim (270°, 45°)$, and opposite bias for a distance underestimate. This motion will be proportional to the azimuthal velocity difference between the Sun and the stars.

As discussed in SBA, there are similar connections between each velocity pair. However, since we are dealing with a disc sample with a complicated selection function and strive for maximum accuracy, we will drop all terms that are to first order affected by galactic rotation and streaming (U versus V motion). Since proper motion errors are minimal, we can use the more precise non-linear estimator from SBA.

As a quick test for distance errors in a sample of stars, we can assume that (U, V, W) is sufficiently close to the real velocity vector (U_0, V_0, W_0) and fit a straight line to W versus $T_{\mathrm{VW}}V$, via the regression model

$$W_i = \gamma\, T_{\mathrm{VW,i}} V_i - W_\odot + \epsilon_i, \qquad (8)$$

where W_i and V_i are the measured vertical and azimuthal velocity components for the ith star, and $T_{\mathrm{VW, i}}$ is given in equation (7). The fit parameters $W_\odot$ and γ represent the vertical motion of the Sun, and the average fractional distance error, while ϵ_i represents the random term from stellar velocity dispersion and measurement errors, which should have zero expectation value. These fits are shown in Fig. 4 for different subsamples of RAVE–TGAS, and for LAMOST–TGAS. The dark red error bars show means of W when binning the sample in VT_{VW}, while the blue dashed lines show the fitted regression line from equation (8). It is evident that in particular LAMOST and the suspected binary stars in RAVE show a large slope γ, indicating either a major distance bias, or a problem with line-of-sight velocity determination.

We note that a robust linear estimator can be derived from equation (8) by replacing the individual azimuthal velocities with their expectation value, i.e. $V_i \to \langle V \rangle$. This estimator is particularly useful for remote samples of halo stars with a large mean heliocentric V. However, for disc stars, this estimator is about one order of magnitude less sensitive than our full estimator, and so for this work of little use compared to the full distance estimator, which we will now describe.

The regression in equation (8) can be formalized to an estimate of the fractional distance error f_{UV}. We adopt equation (19) from SBA:

$$f = \frac{\mathrm{Cov}(W, y)}{\mathrm{Var}(y) + \langle (T_{\mathrm{VW}}^2 + T_{\mathrm{UW}}^2)\,\sigma_{\mathrm{W}}^2 \rangle}. \qquad (9)$$

where $y = T_{\mathrm{UW}}U + T_{\mathrm{VW}}V - T_{\mathrm{WW}}W_\odot$ is the 'baseline' on which the rise in W is measured, Cov is the covariance, and Var(y) is the variance of y. To measure the distance bias in practice, we multiply all distances with the same factor, until the estimated f is zero. Error bars shown throughout the paper are derived by varying the distance correction, until f is at the 1σ confidence limit. The denominator of this term matters only for the error determination. The essential task is to find the distance correction factor, for which $\mathrm{Cov}(W, y) = 0$.

To construct an estimator that is more stable against outliers in W, we cap the vertical velocities at $|W| = 200\,\mathrm{km\,s^{-1}}$, i.e. assign all stars beyond that value the limit of $\pm 200\,\mathrm{km\,s^{-1}}$. In statistical terms, we are essentially using an M-estimator for our linear regression. We have checked by varying this ceiling between 150 and 400 $\mathrm{km\,s^{-1}}$ that our results are not affected by outliers.

In some situations, sample kinematics/measurements may be erroneous, or the data may be affected by streams. In this case, it is useful to look at the separate statistics from the $U \to W$ and $V \to W$ terms. These separate terms have analogously the estimators:

$$f_V = \frac{\mathrm{Cov}(W, T_{\mathrm{VW}}V)}{\mathrm{Var}(V\,T_{\mathrm{VW}}) + <T_{\mathrm{VW}}^2 \sigma_{\mathrm{W}}^2>}, \text{ and} \qquad (10)$$

$$f_U = \frac{\mathrm{Cov}(W, T_{\mathrm{UW}}U)}{\mathrm{Var}(U\,T_{\mathrm{UW}}) + <T_{\mathrm{UW}}^2 \sigma_{\mathrm{W}}^2>}. \qquad (11)$$

4.1 Bias corrections

There are three further contributions to these covariances: the tipping of the velocity ellipsoid outside the Galactic plane, observational uncertainties in proper motions, and both v_{los} measurement errors and intrinsic v_{los} dispersion from binaries. We summarize the last two items as v_{los} source variance $e_\parallel^2$. To get an unbiased distance estimator, we have to subtract these contribution from the measured covariance:

$$\mathrm{Cov}(W, T_{\mathrm{VW}}V) = \mathrm{Cov}_{\mathrm{measured}}(W, T_{\mathrm{VW}}V) - \mathrm{Cov}_{\mathrm{bias}}. \qquad (12)$$

We now separate the bias term $\mathrm{Cov}_{\mathrm{bias}}$ into its three sources: the bias from the tipping velocity ellipsoid $\mathrm{Cov}_{\mathrm{ve}}$, the measurement error on proper motions $\mathrm{Cov}_\perp$ and the $e_\parallel$ term $\mathrm{Cov}_\parallel$:

$$\mathrm{Cov}_{\mathrm{bias}}(W, T_{\mathrm{VW}}V) = \mathrm{Cov}_{\mathrm{ve}} + \mathrm{Cov}_\perp + \mathrm{Cov}_\parallel. \qquad (13)$$

Let us first discuss the minor contribution from the tipping of the velocity ellipsoid. As we can see in Fig. 2, some part of the galactocentric radial motion contributes to the heliocentric azimuthal velocity, when a star is observed away from the connecting line between Sun and Galactic Centre. This way, the inclination of the non-isotropic velocity ellipsoid in galactocentric coordinates translates into a minor correlation between the velocities in the
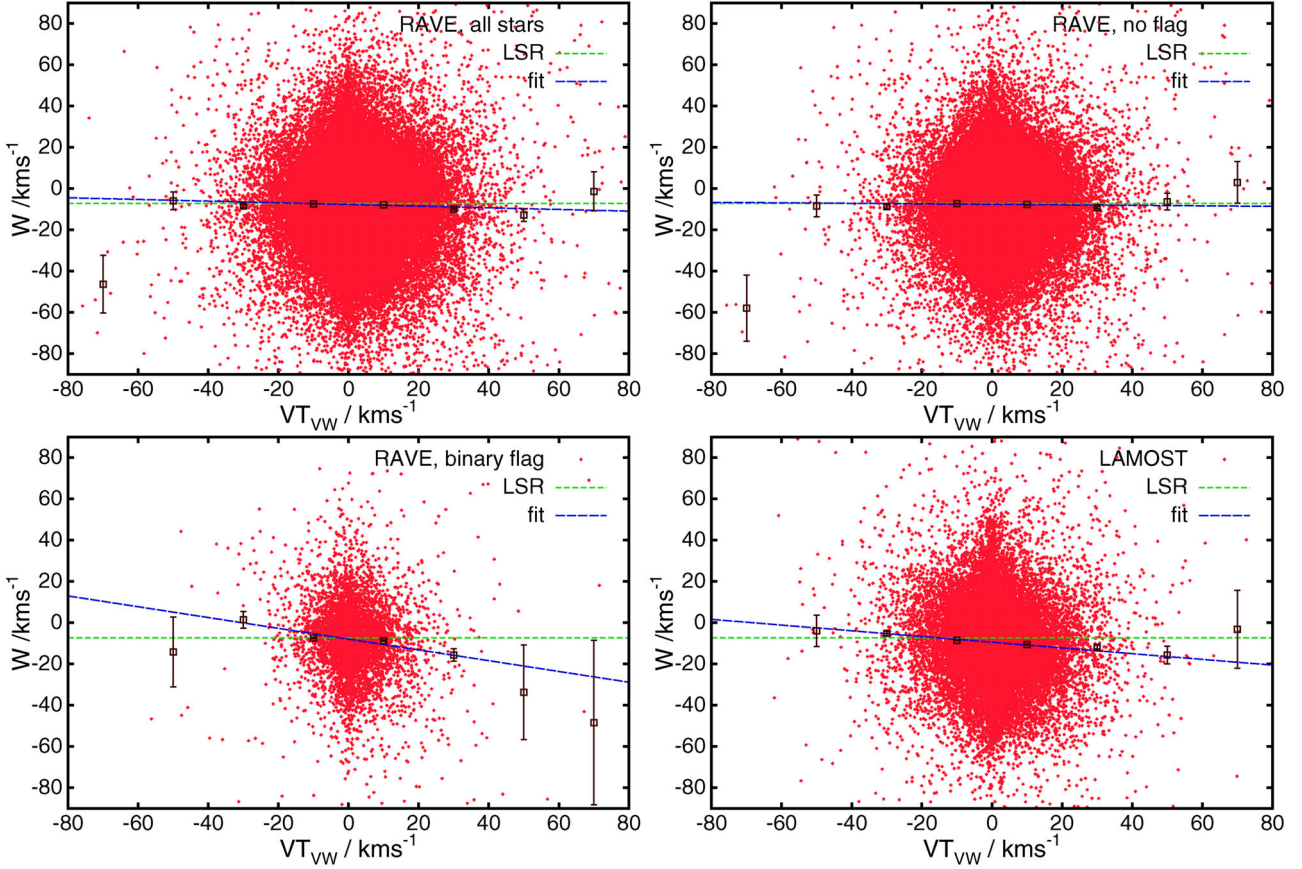
Figure 4. Trends in W versus VT_{VW} in all RAVE stars (top left), all unflagged RAVE stars (top right), all RAVE stars with at least one binary flag (bottom left), and all LAMOST stars with $\sigma_{los,LAM} < 25$ km s^{-1} (bottom right). The dark red error bars show the mean values of W when binning the sample along the x-axis. The green short dashed line marks the reflex motion of the Sun and the blue long-dashed line the linear fit to the data. Despite the curious outliers at extreme azimuthal velocities in the RAVE sample, the slopes in all four plots are significant and given in Table 1.

heliocentric frame. The velocity ellipsoid corrections from equations (40f) of Schönrich, Binney & Asplund (2012) for the $U \to W$ and $V \to W$ terms read:

$$\text{Cov}_{ve}(W, T_{UW}U) = \frac{1}{4}\left\langle \sin(2b)\cos(l)\cos(\alpha)\sin(2\beta) \right. $$
$$\left. \times \left(\sigma_U^2 - \sigma_W^2\right) \right\rangle,$$

$$\text{Cov}_{ve}(W, T_{VW}V) = -\frac{1}{4}\left\langle \sin(2b)\sin(l)\sin(\alpha)\sin(2\beta) \right.$$
$$\left. \times \left(\sigma_U^2 - \sigma_W^2\right) \right\rangle, \quad (14)$$

where σ_U^2 and σ_W^2, are the radial and vertical velocity dispersions. For definitions of the angles, see Fig. 2 and Section 2.1. Here, we assume that the velocity ellipsoid is pointing roughly towards the Galactic Centre (Siebert et al. 2008; Binney et al. 2014). σ_U and σ_W are measured directly from the data. Since our samples are local and close to the plane (small α and β), the resulting bias on the distance estimators is significantly less than 0.01 for f_U, and less than 0.002 for f_V.

Adapting equations (23ff) from SBA, we can calculate the bias in the covariance from proper motion errors:

$$\text{Cov}_\perp(W, T_{UW}U) = -\sum_i T_{UW,i}^2 e_{\perp,i}^2$$

$$\text{Cov}_\perp(W, T_{VW}V) = -\sum_i T_{VW,i}^2 e_{\perp,i}^2 \quad (15)$$

where again i is the index running over all stars in the sample. $e_{\perp,i}$ is the uncertainty in the proper motion of star i in the latitudinal (b) direction. If uncorrected, this term will feign a slight distance overestimate, but due to the excellent proper motions in TGAS, the resulting bias on the fractional distance error is below 0.3 per cent.

4.2 Line-of-sight velocity variance

The v_{los} errors require significantly more caution. Consider the effect on f_V, an analogous discussion will apply to f_U. Just where T_{VW} is maximal, i.e. at $l \sim 90°$, $270°$ and at an elevation of $b \sim 45°$ above the Galactic plane, an error in line-of-sight velocity determination will be evenly split into the derived V and W velocities, leading to a maximal correlation. This term will, if uncorrected, push the statistics towards a distance underestimate:

$$\text{Cov}_\parallel(W, T_{UW}U) = \sum_i T_{UW,i}^2 e_{\parallel,i}^2 \quad \text{and}$$

$$\text{Cov}_\parallel(W, T_{VW}V) = \sum_i T_{VW,i}^2 e_{\parallel,i}^2, \quad (16)$$

where $e_{\parallel,i}^2$ is the source v_{los} variance (we analogously call $e_\parallel$ the source v_{los} dispersion), i.e. all v_{los} variance that does not derive from the stellar velocity ellipsoid. For the full estimator f, both Cov terms have to be added. We conceptually split $e_\parallel$ into two contributions:

$$e_\parallel^2 = \sigma_{los}^2 + \sigma_{bin}^2, \quad (17)$$

where σ_{los} is the intrinsic measurement error in each survey, and σ_{bin} is the line-of-sight velocity dispersion from the orbital motion of multiple stellar systems (which we will address simply as 'binary' contamination). The source v_{los} variance is the most important bias in LAMOST, with a distance bias near 10 per cent. The impact of $\sigma_{\text{los,RAVE}}$ on the distance estimators is nearly negligible with less than 0.3 per cent, but σ_{bin} might be important. While the general problem of binary velocity dispersions has been known for a long-time (Olszewski, Pryor & Armandroff 1996; Hargreaves, Gilmore & Annan 1996), we could not locate a good prediction for the MW. A strong expected age dependence (Gieles, Sana & Portegies Zwart 2010), and probably some metallicity dependence, make it difficult to translate results from dwarf spheroidal galaxies to the MW. Depending on the binary fraction and distribution, the dispersion contributed to the line-of-sight velocities could be of order $3\,\text{km}\,\text{s}^{-1}$ with a large uncertainty (Vogt et al. 1995), though some open clusters appear to yield smaller values. We also note that our statistics will not pick up the full line-of-sight velocity dispersion, since the proper motions will also be affected to some extent. The difference is mostly in the long baseline in time for *Gaia* – short-period binary systems (with periods significantly shorter than the time elapsed between *Hipparcos* and *Gaia*) should fully affect f, whereas we expect that very long-period binary systems will have a vanishing effect. However, while long-period systems are more numerous, most of the contribution to the velocity distributions by binary systems is expected to stem from short-period systems (see e.g. de Rijcke & Dejonghe 2002).

This problem will be further examined in Sections 5.2 and 5.5.

Can we discriminate the contribution of $e_\parallel$ from a general distance error? To some extent, yes, with a sufficiently large number of stars in our sample. The bias from $e_\parallel$, $\text{Cov}_\parallel/N$, for any sample of N stars should be the same between subsamples with the same geometry (i.e. distribution in values of $T_{\text{VW},i}$), and same v_{los} measurement quality and binary contamination (i.e. similar $e_{\parallel,i}$). In contrast with this, the covariance term created by a distance error depends on the length of the baseline, which can be seen by isolating $\text{Cov}(W, T_{\text{VW}}V)$ or $\text{Cov}(W, y)$ in equation (9).

To summarize, these equations allow for a degenerate estimate between the mean distance error and an unknown bias in the line-of-sight velocities from binaries, but this degeneracy can be broken by selecting different subsamples with different baseline lengths ($\text{Var}(T_{\text{VW}}W)$), either by cuts in metallicity or (more dangerous) by enforcing cuts on the galactocentric azimuthal velocity V_{g}.[3]

To account for uncertainties in Cov_{bias}, we add 30 per cent of the distance bias from the tipping velocity ellipsoid and the proper motion errors as independent terms to the error budget in f. The geometric terms of the source v_{los} dispersion $e_\parallel$ are not uncertain, and so we add 10 per cent of this bias to the error budget.

4.3 Further biases and caveats

In a realistic galaxy, we expect further biases on our distance estimator from Galactic structure. The main features will be

(i) streams from accretion events,
(ii) the Galactic warp correlating V and W motions near the line of nodes,

(iii) (vertical) breathing modes from bar and spiral pattern, and
(iv) disc streaming motions near resonances.

We do not expect a major bias from disc streaming, because those motions are predominantly in the galactic plane (Dehnen 2000; Peréz-Villegas et al. 2017), and our distance estimator is only biased by vertical motions.

If a stellar stream through the sample, it may have a large vertical motion W. This does not yet imply even a local bias on our distance estimator, because even a correlation between W and V velocities is not enough. The stream has to correlate W significantly with $T_{\text{VW}}V$ to affect the f_V part. Even if the condition is fulfilled, good spatial coverage will cancel the bias to first order: If we have a bias at $b > 0$, an equivalent patch in the southern galactic hemisphere $b < 0$ will cancel the impact of the stream, because the stream's W and V velocities will be similar, but the geometric term $T_{\text{VW}}V$ reverses its sign. In addition, the f_U estimator will not be affected or react differently, and tend to cancel the effect. In short, streams in a sample with good spatial coverage will not be a problem, and in worst case be detected by internal discrepancy between f_U and f_V. Helmi et al. (2017) have already suggested that the TGAS sample has no dominant single stream or small number of identifiable streams.

Similarly, the Galactic warp may be a problem for the f_V term, but since the correlation between W and V will be similar in all directions (Dehnen 1998), f_V is only biased, if our sample looks predominantly into one quadrant of the sky. This is a major caveat, however, for dissecting samples in l and b.

Vertical breathing modes (Faure et al. 2014) are more dangerous, because their W motion reverses in parallel with a sign reversal of T_{VW} between $b > 0$ and $b < 0$. This can have some impact, if the vertical motion feature is not symmetric in the field of view. However, the effect can be detected by comparing f_V with f_U.

A summary of performance tests of our method on N-body simulations of MW-like galaxies, which reasonably match both Galactic structure and local kinematics, is discussed in Appendix A. On full sky samples measured at a solar-like position, the distance estimators perform at the level of their statistical error, i.e. the systematic effects discussed in this section must be below 1 per cent on the distance estimator. The worst case scenario that we could find was a simulation with an unrealistically long bar. In that case, the substructure introduced no concerning bias on the full sample, but a scatter of ~4 per cent in f_V and ~2 per cent in f when we restricted the sample to separate hemispheres in l.

We further note that (cf. the discussion in SBA) f is an intrinsically quadratic estimator. Stars with distance overestimates have a longer measured baseline in y, and so achieve a slightly larger weight in the distance estimator than distance underestimates. So, for a set of different fractional distance errors f_i for a set of stars, the fractional distance estimate is in fact $f = 1 - \sqrt{\sum_i <(1+f_i)^2>}$. While of some importance with spectrophotometric distances, the effect is not important here due to the good parallax quality and hence moderate random scatter of distance estimates in TGAS.

5 APPLICATION TO THE DATA

5.1 Mean distance error in the different subsamples

We start our discussion with an inventory of the mean distance biases in different subsamples, and will justify some of our assumptions in the later sections. For deriving distances in these samples, we use the full distance prior from equation (3) including the survey

---

[3] We iterate the warning that while it is feasible to cut the sample in V_{g}, cuts in heliocentric V are illicit, since they produce a biased cut in the actual velocity ellipsoid, biasing the sample in U_{g} and hence, via the inclination of the velocity ellipsoid in W.

Table 1. Distance correction factors assuming the given errors for different subsamples of RAVE and LAMOST stars. Column 2 shows the number of stars, γ is the slope in the simple linear regression from equation (8), f_V is the distance estimator from equation (10), Δf is the bias correction from equation (13) to f, and f is the full distance estimator from equation (9). The shorthands for the used samples are: RAVE: using RAVE stars with parallax errors better than 20 per cent; RAVE$_{10}$: RAVE stars with parallax errors better than 10 per cent; all: all stars; nf: no flags; nf,s: no flags and excluding the distance region $0.3 < s/\mathrm{kpc} < 0.35$; b: at least one binary flag; LAMe1: LAMOST with given line-of-sight uncertainties $\sigma_{\mathrm{los,LAM}} < 25\,\mathrm{km\,s^{-1}}$, and LAMe2: LAMOST with $\sigma_{\mathrm{los,LAM}} < 15\,\mathrm{km\,s^{-1}}$. A positive f indicates an average distance overestimate by a $1 + f$.

Sample	Number of stars	γ	f_V	Δf	f
RAVE$_{\mathrm{all}}$	88464	-0.041 ± 0.007	-0.027 ± 0.007	-0.009	-0.013 ± 0.005
RAVE$_{10,\mathrm{all}}$	40570	-0.063 ± 0.010	-0.041 ± 0.010	-0.004	-0.024 ± 0.007
RAVE$_{\mathrm{nf}}$	59923	-0.012 ± 0.008	-0.009 ± 0.009	-0.011	0.001 ± 0.006
RAVE$_{\mathrm{nf,s}}$	52865	-0.026 ± 0.009	-0.022 ± 0.009	-0.011	-0.006 ± 0.007
RAVE$_{10,\mathrm{nf}}$	22574	-0.030 ± 0.014	-0.014 ± 0.015	-0.005	-0.010 ± 0.009
RAVE$_{10,\mathrm{nf,s}}$	18710	-0.045 ± 0.015	-0.032 ± 0.016	-0.005	-0.017 ± 0.010
RAVE$_{\mathrm{b}}$	5929	-0.261 ± 0.027	-0.170 ± 0.024	-0.006	-0.123 ± 0.016
RAVE$_{10,\mathrm{b}}$	2803	-0.332 ± 0.041	-0.207 ± 0.037	-0.001	-0.114 ± 0.024
LAM$_{\mathrm{e1}}$	28475	-0.134 ± 0.012	0.131 ± 0.035	0.214	0.147 ± 0.031
LAM$_{\mathrm{e2}}$	6673	-0.141 ± 0.024	-0.013 ± 0.026	0.084	0.011 ± 0.019

selection functions, which we will derive in Section 5.3.1, and full bias corrections with the given errors according to equation (13).

Table 1 summarizes the distance statistics for different subsamples in RAVE and for LAMOST. The third column (γ) provides the slope in the simple linear fit of W motion against VT_{VW} given in equation (8). These simple fits are shown with blue dashed lines for four samples in Fig. 4: all RAVE stars (top left), unflagged RAVE stars (top right), suspected binaries in RAVE (bottom left), and LAMOST (bottom right). These slopes indicate distance underestimates or significant v_{los} source dispersions $e_\parallel$ in all samples, but in particular for the suspected binaries and LAMOST. The fourth and fifth columns of Table 1 show the two bias-corrected distance estimators f_V and f from equations (9) and (10). For the full RAVE sample, these terms both show a moderate, but significant distance underestimate of order 2 per cent.

Our first, naive, idea was to ascribe this distance underestimate to parallax errors in TGAS, as systematic effects are e.g. found in Arenou et al. (2017). However, when we select stars that have no flags according to Matijevič et al. (2012), i.e. are considered likely normal stars, the distance bias diminishes. In turn, when we test suspected binary systems, i.e. stars that have at least one 'b' flag, there is a very strong bias equivalent to more than 10 per cent distance underestimate. We can conclude from this that *Gaia* astrometry is generally of high quality, as long as we restrict the sample to apparently normal, single stars.

For the LAMOST sample, the situation is more complex. The simple linear regression test delivers a slope $|\gamma|$ in excess of 0.1, i.e. a very strong kinematic bias. However, LAMOST v_{los} determinations have nominal measurement errors $\sigma_{\mathrm{los,LAM}} > 10\,\mathrm{km\,s^{-1}}$. In the bias corrected statistics f and f_V, the expected impact by σ_{los} fully accounts for this correlation in the LAMOST subsample with $\sigma_{\mathrm{los,LAM}} < 15\,\mathrm{km\,s^{-1}}$, resulting in a near-zero f. When we ease this limit to $\sigma_{\mathrm{los,LAM}} < 25\,\mathrm{km\,s^{-1}}$, the bias correction apparently overshoots, giving strongly positive f. The only viable conclusion is that LAMOST greatly overestimates its v_{los} measurement error. We will test this hypothesis and estimate the real v_{los} measurement error of LAMOST in the next section.

5.2 Testing line-of-sight velocity errors

We can use our method to test line-of-sight velocity errors. Even if we do not fully trust the parallax distances, *Gaia*–TGAS astrometry and hence our distance estimates are independent of the stellar

parameters and line-of-sight velocity determinations. We can thus reasonably assume that the real distance bias in the sample should be nearly the same between samples of stars grouped according to their σ_{los} given by the spectroscopic surveys. A minor caveat is the potential correlations of line-of-sight velocity errors with stellar apparent magnitude, metallicities, and ages, which introduce a small bias to the selection function.

The top panel of Fig. 5 shows $1 + f$ for subsamples of RAVE grouped by $\sigma_{\mathrm{los,RAVE}}$ and binned to samples of 4500 stars, sliding the mask by 1500 stars each, so every third data point is independent. We use the given error estimates from the RAVE pipeline, i.e. we set $e_\parallel = \sigma_{\mathrm{los,RAVE}}$. The contribution to f from σ_{los} is small in this region, reaching $\delta f \sim 1.5$ per cent at $e_\parallel = 5\,\mathrm{km\,s^{-1}}$. Green error bars show the statistics for the entire RAVE sample, while the blue error bars depict the RAVE sample with no flags from Matijevič et al. (2012). There are some stars in the RAVE sample with given $\sigma_{\mathrm{los,RAVE}} > 10\,\mathrm{km\,s^{-1}}$, and the sharp decline in the $1 + f$ estimate seen in the full sample continues for those. Since their distance error should not be substantially different, this implies that beyond $\sigma_{\mathrm{los,RAVE}} > 5\,\mathrm{km\,s^{-1}}$, v_{los} estimates are significantly more uncertain than stated and should not be used at all. It is apparent that the unflagged sample shows excellent stability; the little aberration around $\sigma_{\mathrm{los,RAVE}} \sim 1.7\,\mathrm{km\,s^{-1}}$ is likely just a statistical fluctuation (about 10 error bars should have their 1σ confidence intervals not crossing 1.0). The entire sample, however, shows a clear abnormality for $\sigma_{\mathrm{los,RAVE}} > 2\,\mathrm{km\,s^{-1}}$. Apparently, the flagged stars contain line-of-sight velocity uncertainties beyond the provided errors, but the contamination is comparably small for $\sigma_{\mathrm{los,RAVE}} < 2\,\mathrm{km\,s^{-1}}$. There is some tentative indication for problems with the line-of-sight velocities of stars with very small errors $\sigma_{\mathrm{los,RAVE}} < 0.7\,\mathrm{km\,s^{-1}}$, possibly from some erroneous $\sigma_{\mathrm{los,RAVE}}$ determinations going in hand with v_{los} measurement errors.

The same analysis on LAMOST (bottom panel of Fig. 5) demonstrates the usefulness of the SBA method. Note that the x-axis in this plot just starts beyond the range of the top panel, i.e. in the entire sample, the correction of f for $e_\parallel$ and hence v_{los} errors is crucial. Here, we show the bias corrected estimates for $1 + f$ with red error bars, setting $e_\parallel = \sigma_{\mathrm{los,LAM}}$. These are compared to the same estimate without correcting for the $e_\parallel$ bias. If the pipeline estimates for $\sigma_{\mathrm{los,LAM}}$ were correct, the red error bars should lie in a horizontal line, while the uncorrected estimates (blue) should trend sharply downwards. However, we observe the opposite. While the uncorrected $1 + f < 1$ indicates a significant v_{los} source dispersion
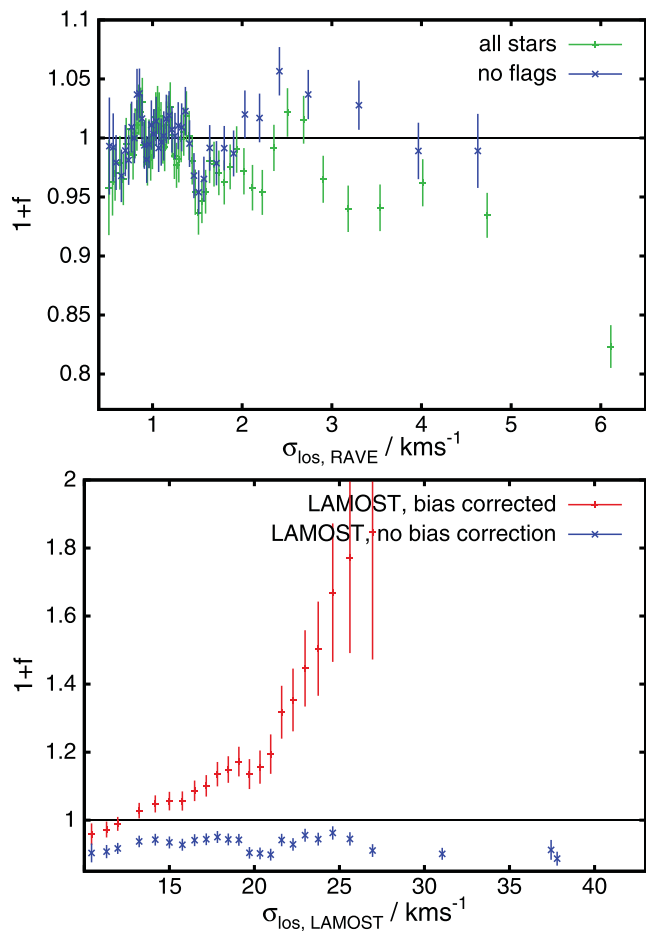
Figure 5. Testing line-of-sight velocity uncertainties σ_{los} as provided by the RAVE (top panel) and LAMOST surveys (bottom panel). Both samples are binned in σ_{los} using a mask of 4500 stars sliding by steps of 1500 stars, so every third data point is independent. The small v_{los} errors of RAVE result in small bias corrections, so here we show only corrected estimates for f. The entire RAVE sample is shown with green error bars, the subsample of unflagged stars is shown in blue. Note the better stability of the cleaned sample towards larger σ_{los}. Due to the larger errors in LAMOST, the bottom plot is on a different scale. Here, we plot the bias corrected distance statistics (red error bars) versus the uncorrected distance statistics (blue). The last data points for the corrected sample are above the figure margin.

$e_{\parallel}$, the points lie in a horizontal line, indicating that there is almost no correlation between the real $e_{\parallel}$ and the estimated $\sigma_{los,LAM}$. We conclude that $\sigma_{los,LAM}$ does not indicate the real v_{los} measurement error, and that the real v_{los} measurement errors must be far better than claimed by their pipeline. In accordance with this, the red error bars trend sharply upwards, because the bias correction increasingly overestimates the real v_{los} measurement uncertainty, explaining our findings in Table 1.

The top panel of Fig. 6 shows the measurement of $e_{\parallel}$ from the bias in f when ordering the sample in terms of $\sigma_{los,LAM}$. For this figure, we use our experience from RAVE that the distance bias should be negligible compared to the $e_{\parallel}$ bias correction, and hence determine $e_{\parallel}$ by demanding an estimate of $f = 0$ after the bias correction. The error bars are derived by varying $e_{\parallel}$ until the estimate for f becomes marginally positive or negative. Since the bias correction on f from equation (16) is quadratic in $e_{\parallel}$, the error bars on $e_{\parallel}$ are asymmetric. The plot clearly shows that $e_{\parallel}$ is nearly uncorrelated with $\sigma_{los,LAM}$, in fact it slightly declines. The sudden rise of $e_{\parallel}$ near
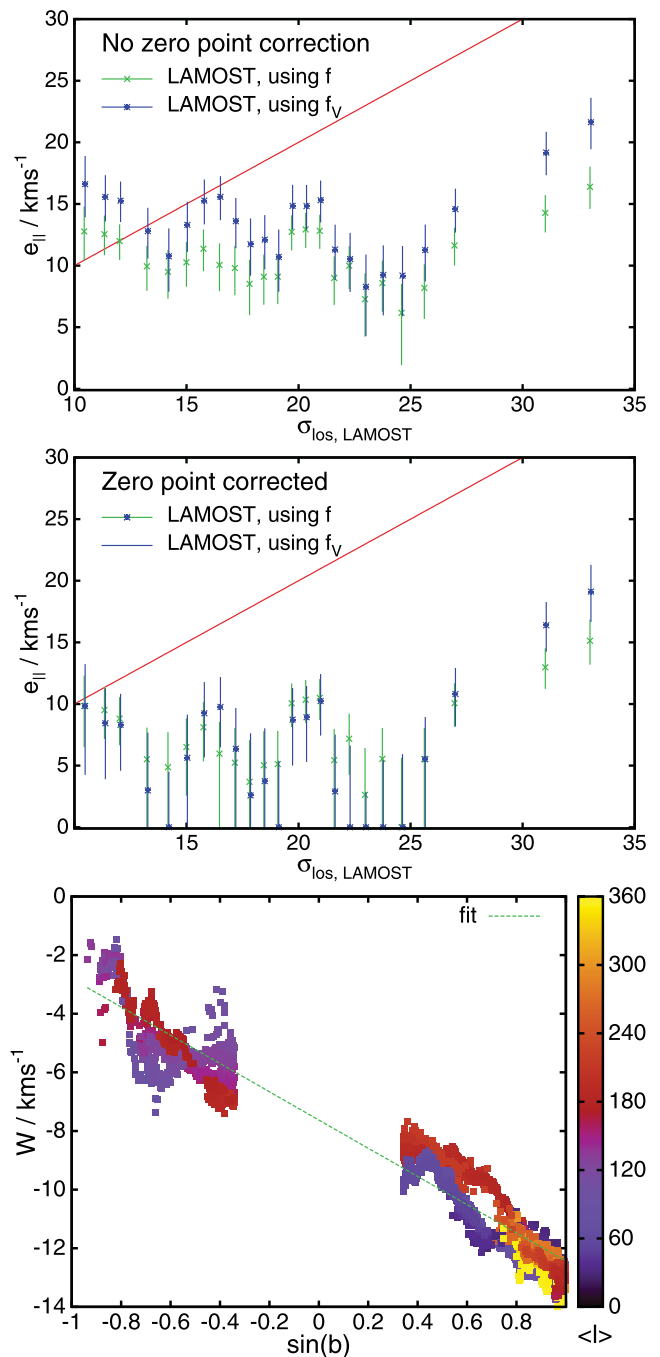

Figure 6. Top panel: measuring the v_{los} source dispersion $e_{\parallel}$ in subsamples of stars ordered by the measurement uncertainty as provided by LAMOST in each subsample. Since for these values $e_{\parallel}^2 \gg \sigma_{bin}$, the values on the y-axis can be directly compared to $\sigma_{los,LAM}$. We mark the identity line in red. The plot uses again a sliding mask, i.e. every third data point is independent. In the middle panel, we show the same statistics, but using a general correction of LAMOST v_{los} by the offset $\delta v_{los} = 5\,\mathrm{km\,s^{-1}}$. The bottom panel shows the average vertical velocity W versus average $\sin b$ when we take around each star all objects with $\sin b$ values within 0.1 and longitude l within $\pi/4$. The trend line (green dashed) was obtained by fitting equation (18) on the LAMOST sample; its slope indicates a $5\,\mathrm{km\,s^{-1}}$ global bias in LAMOST v_{los} measurements, which we corrected to create the middle panel.

Table 2. Fit parameters for testing the line-of-sight velocities via the dependence of vertical velocities on $\sin(b)$ as in equation (18). The slope ξ between W and $\sin(b)$ can be interpreted as an average v_{los} measurement offset δv_{los}.

Sample	$\delta v_{\mathrm{los}}/\mathrm{km\,s^{-1}}$	$-W_{\odot}/\mathrm{km\,s^{-1}}$
RAVE$_{\mathrm{all}}$	0.26 ± 0.13	-7.68 ± 0.09
RAVE$_{\mathrm{nf}}$	0.20 ± 0.16	-7.56 ± 0.10
RAVE$_{10,\mathrm{nf}}$	-0.09 ± 0.28	-8.12 ± 0.19
RAVE$_{\mathrm{b}}$	0.24 ± 0.47	-8.27 ± 0.31
LAMOST	-4.81 ± 0.21	-7.63 ± 0.15
LAMOST$_{\mathrm{e15}}$	-4.72 ± 0.40	-7.73 ± 0.29

$\sigma_{\mathrm{los,LAM}} = 27\,\mathrm{km\,s^{-1}}$ indicates problems with v_{los} estimates beyond this point, so that we from now own will adopt the quality cut $\sigma_{\mathrm{los,LAM}} < 27\,\mathrm{km\,s^{-1}}$. However, there is a problem: the estimates of $e_{\parallel}$ using the full set of terms in f are significantly smaller than the estimates of $e_{\parallel}$ when using only the azimuthal versus vertical motions in f_{V}. Something must be wrong beyond a random uncertainty.

The bottom panel of Fig. 6 solves this riddle. Around each star in LAMOST, we select all stars that have a similar value in galactic latitude, i.e. $\Delta \sin(b) < 0.1$ and lie within $\pi/4$ in l. We plot the average vertical velocity of each subsample versus its mean value in $\sin(b)$; the colour indicates the mean longitude $<l>$ in degrees. Apart from some statistical fluctuations, the sample lies almost perfectly along a line:

$$W = -W_{\odot} + \xi \sin(b). \tag{18}$$

However, $\sin(b)$ is just the projection factor of the v_{los} estimate into W, so the slope ξ can be interpreted as an average global offset in the v_{los} determinations: $\xi = \delta v_{\mathrm{los}}$. An alternative explanation would be a vertical breathing mode in the disc, but it cannot explain the perfect trend. Table 2 compares fits of equation (18) for different subsamples of RAVE and LAMOST. No subsample of RAVE shows an appreciable slope, ruling out a physical phenomenon. Note that the plotting technique leads to a strong autocorrelation between the plotted points (there would be about 20–30 independent samples with an error of $\sigma_{<\mathrm{w}>} \sim 1\,\mathrm{km\,s^{-1}}$), so to the eye, there might be an optical illusion of deviations from the plotted trend, which are, however, not significant. We conclude that LAMOST has an average zero-point offset $\delta v_{\mathrm{los}} \sim -5\,\mathrm{km\,s^{-1}}$, which we correct by adding $5\,\mathrm{km\,s^{-1}}$ to all LAMOST v_{los} estimates.

The middle panel of Fig. 6 confirms the benefit of this correction – the two different estimators for $e_{\parallel}$ are now in line. The estimated v_{los} source dispersion $e_{\parallel} \sim 7.1\,\mathrm{km\,s^{-1}}$ is significantly smaller than in the top panel, and shows that the LAMOST v_{los} determinations are far more precise than their estimate $\sigma_{\mathrm{los,LAM}}$. We also find a comparable dichotomy of mean radial stellar velocities inwards and outwards along the connection line between Sun and Galactic Centre, which confirms the v_{los} offset found on vertical motion. Such an offset in v_{los} is not unprecedented – SEGUE/Sloan Digital Sky Survey was plagued by a similar problem, and even after its correction, Schönrich (2012) found a residual bias, although smaller than this one. The fact that the v_{los} bias on SEGUE depends on colour and metallicity of stars as discussed by Schönrich (2012), indicates that future revisions of the LAMOST sample should examine those dependencies as well.

To summarize our findings in this section: we have shown that our method can be used to test the quality of σ_{los} estimates. v_{los} estimates and $\sigma_{\mathrm{los,RAVE}}$ estimates from the RAVE survey are of high quality, with minor issues for flagged stars. A strict cut of $\sigma_{\mathrm{los,RAVE}} < 5\,\mathrm{km\,s^{-1}}$ should be applied. The LAMOST v_{los} de-

terminations are far more precise than suggested by their nominal $\sigma_{\mathrm{los,LAM}}$. We adopt the quality cut $\sigma_{\mathrm{los,LAM}} < 27\,\mathrm{km\,s^{-1}}$. One should adopt a general error of $\sigma_{\mathrm{los}} \sim 7\,\mathrm{km\,s^{-1}}$, and correct all v_{los} estimates by $+5\,\mathrm{km\,s^{-1}}$.

5.3 Distance-dependent bias

5.3.1 Retrieving the prior/selection function

In this section, we will finally derive the selection function $S(s)$ required in the distance estimate in equation (3) and test it by measuring the distance bias f in samples binned in distance s.

A full a priori calculation of the distance priors and selection functions $S(s)$, which enter the distance estimation in equation (3), would demand a full chemodynamical model of the disc with all its assumptions. Population synthesis models can calculate $S(s, \tau,$ [Fe/H]) and give us some indication of the shape of the selection function. However, even with this knowledge, we still cannot reliably build a selection function for our samples, since we would need to know/assume the exact distribution of stars in age and metallicity at each point in the Solar neighbourhood.

For example, the RAVE selection function described in Wojno et al. (2016) can be roughly described as a flat selection in I-band magnitudes between about 9 and 11.3 mag, with some wings towards lower and higher magnitudes. Running a population synthesis model as used in Schönrich & Bergemann (2014) on this selection, demanding additionally $T_{\mathrm{eff}} > 4200\,\mathrm{K}$ for reasonable stellar parameter estimates, we obtain a relatively steep selection function $S(s, \tau)$ in distance and age: $S(s)$ at fixed metallicity and age falls off approximately exponentially with a scalelength of 0.12 kpc at $s > 0.2$ kpc, and with a flatter slope at $s > 0.5$ kpc. At fixed metallicity and distance, $S(\tau)$ behaves roughly like $1/(\tau + 1.5)$ Gyr.

Here, we will adopt a simpler, more direct approach: we can actually derive the $S(s)$ from the data. We know that with the correct prior, this should give stable values of f versus $\langle s \rangle$, just like we could demand a stable distance bias f when binning in σ_{los} in the previous section.

To obtain the full prior including the selection function, we start from the simple density model in equation (4). For the sake of simplicity, we assume that the selection function is solely a function of distance, i.e. $S(s)$, ignoring angle-dependent deviations from the simple density model, and more importantly from the age and metallicity structure of the disc. We can now measure $S(s)$ in a simple way: With an expected distribution of stars in distance $\rho_{\mathrm{S}}(s)$, we can estimate $S(s) \sim \rho_{\mathrm{m}}(s)/\rho_{\mathrm{S}}(s)$, where $\rho_{\mathrm{m}}(s)$ is the measured density of stars in distance. Since the distances of single stars are quite well determined, we approximate $\rho_{\mathrm{m}}(s)$ by simply binning the number of stars according to their distance expectation value $\langle s \rangle$. One caveat is that near the far end of the sample, the measurement will fall quickly below the real selection function, due to stars dropping out from the sample by the quality cut on parallax accuracy. The latter information is already used by the quality cut, and so must not be written into the prior.

To obtain $\rho_{\mathrm{S}}(s)$, we use the positions $(l_{\mathrm{i}}, b_{\mathrm{i}})$ on the sky of all stars i in the sample on our density model and sum up their normalized density distributions in distance s:

$$\rho_{\mathrm{S}}(s) = \sum_i N_{\mathrm{i}}' s^2 \rho(s, l_{\mathrm{i}}, b_{\mathrm{i}}), \tag{19}$$

where

$$N_{\mathrm{i}}'^{-1} = \int \mathrm{d}s \, s^2 \rho(s, l_{\mathrm{i}}, b_{\mathrm{i}}) \tag{20}$$

normalizes the integral in distance s over each of the individual distance distributions to 1, and $\rho(s, l_i, b_i)$ is calculated from equation (4). The problem is solved by an iteration. We start with a flat $S_0(s)$ in distance, calculate all distance expectation values from equation (3), bin them in distance to obtain $\rho_m(s)$ and calculate our first estimate for $S_1(s) = \rho_m(s)/\rho_S(s)$. We fit $S_i(s)$ with a simple analytic equation. A smooth fitting function is important, because a non-smooth function might pull data towards single points yielding a false convergence. We re-insert this analytic approximation to S into the distance calculation and repeat the procedure, until the selection function converges (typically within five iterations).

The top (linear) and middle (logarithmic plot) panel of Fig. 7 show the measurement and fits to the selection function from the entire RAVE sample. The blue points show the first estimate $S_1(s)$, and green points depict the final estimate for $S(s)$, which is the selection function that reproduces itself in further iterations. The precise parallaxes on the near end result in virtually no change of the shape of S at distances $s < 0.2\,\mathrm{kpc}$. On the long distance end, the selection function pulls stars closer, steepening $S(s)$.

As a fitting equation, we employ a lognormal distribution, plus a minor exponential term to smooth the edges:

$$S(s) = d_1(\exp(-((\ln(s/\,\mathrm{kpc}) - \ln(a))c_1)^2/2) + d_2 \exp(-sc_2), \quad (21)$$

where a, c_1, c_2, d_1, d_2 are free fitting parameters. The fit shown in Fig. 7 yielded $a = 0.112$, $c_1 = 1.37$, $d_1 = 9.52$, $d_2 = 1.3$, and $c_2 = 6.51\,\mathrm{kpc}^{-1}$, converging within five iterations. The fit quality is good; the discrepancy at large s is caused by the quality cut on parallaxes, which must not be part of $S(s)$.

5.3.2 Testing the selection function

Beyond this good convergence of $S(s)$, how can we judge if this selection function is reasonable? The key lies again in the distance statistics. The bottom panel in Fig. 7 shows $1 + f$ on the y-axis, using either a flat S (blue error bars), or the derived $S(s)$ from equation (21), depicted in green error bars. At the far end of the sample, where parallaxes are relatively uncertain, a flat selection function results in a massive overestimate of stellar distances, which in turn gets signalled by $1 + f \gg 1$ in the blue error bars. The last data points for the flat $S(s)$ are not shown, because they have $1 + f > 2$, outside the plotting range. The declining parallax precision hence results in a strong trend in f, while with the (relatively) correct selection function, we see no such trend. This is direct proof that the selection function is sound and has to be applied. Note that when binning in $1/p$, the breakdown in the distance statistics for a flat $S(s)$ would be found in almost all of the sample with $p \lesssim 5\,\mathrm{mas}$.

The left-hand column of Fig. 8 shows the same results for the RAVE–TGAS sample when selecting only unflagged stars (left-hand column). The fitting parameters for the RAVE selection function (equation 21) are $a = 0.21$, $c_1 = 2.19$, $d_1 = 5.43$, $d_2 = 0.86$, and $c_2 = 1.75\,\mathrm{kpc}^{-1}$. The exponential scalelength is fully degenerate with the normalization of the exponential term, and is hence undetermined by the fit. However, it has a significant impact on the long-range behaviour of $1 + f$: if we would use $e \lesssim 1.5$, i.e. a scalelength of the selection function larger than about $0.7\,\mathrm{kpc}$, $1 + f$ would become significantly larger than 1 for $s > 0.6\,\mathrm{kpc}$. The chosen value of $e = 1.75$ is in the range predicted by population synthesis models, and a compromise between our distance statistics and the direct fit for $S(s)$, and leads to a satisfactory result in the bottom left panel of Fig. 8.

On the right-hand side of Fig. 8, we analyse the selection function of the LAMOST survey. This selection function is not only limited
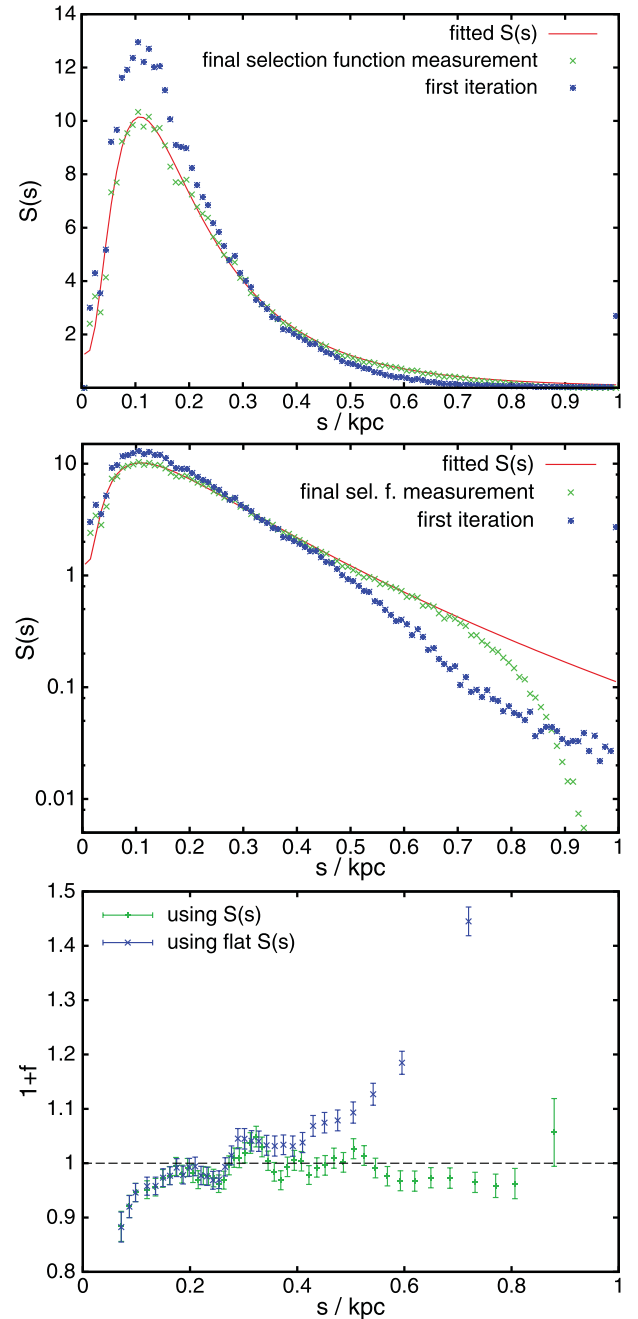

Figure 7. Derivation of the RAVE selection function using the entire RAVE–TGAS sample, the top panel in linear scale, and the middle panel in log scale. The blue points show the first measurement for $S(s)$ when starting with a trial flat selection function. The selection function $S(s)$ converges as shown within less than five iterations to its final value. The red line gives the smooth fitting function applied. The break down of the selection function on the right-hand end of the distribution is caused by the quality cuts on the parallax accuracy, not to be fitted. The bottom panel compares the distance statistics between using the final $S(s)$ (green) and a flat $S(s)$ (blue) binned in $\langle s \rangle$. The last (blue) data points in the flat $S(s)$ are not shown as they have $1 + f > 2$.

by the survey's magnitude cuts, but also by our applied signal-to-noise limit of S/N > 30 in the r band. The LAMOST selection function could not be fitted by the same lognormal distribution as RAVE, so we choose
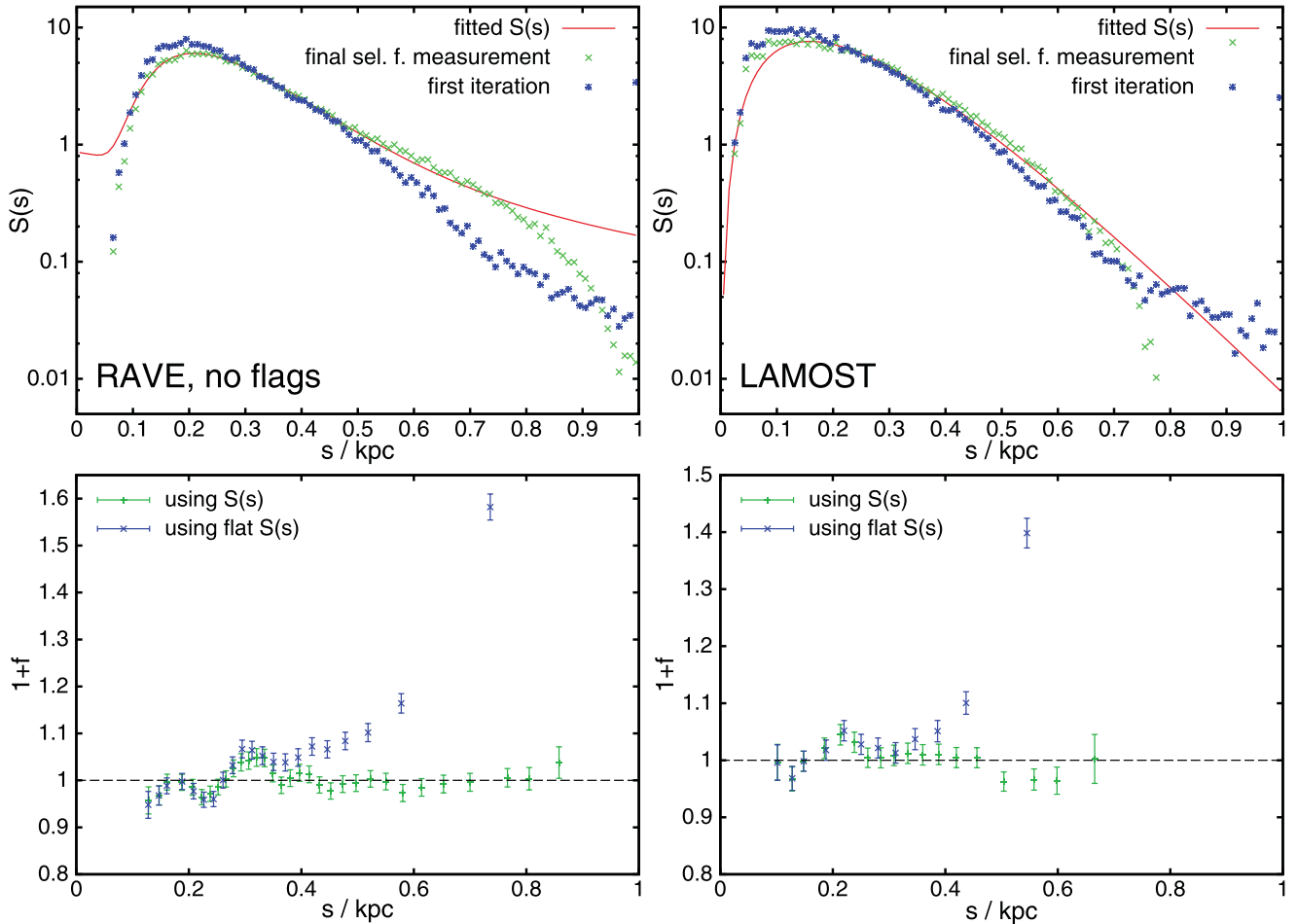
$$S_L(s) = as^2 \exp(-cs), \quad (22)$$

Figure 8. Same as Fig. 7, for RAVE–TGAS stars that are unflagged (left-hand column) and LAMOST (right-hand column). When calculating the distance statistics (bottom row), we apply all bias corrections in RAVE, but no bias corrections in LAMOST. The last data points in the flat $S(s)$ are not shown because they are far above the range of this plot.

where the free-fit parameters are determined as $a = 2230\,\mathrm{kpc}^{-2}$ and $c = 12.6\,\mathrm{kpc}^{-1}$. The fit is not good for small $s < 0.1\,\mathrm{kpc}$, since we made no attempt to fit the distribution in that range. As detailed above, S plays a negligible role for determining distances at $s < 0.1\,\mathrm{kpc}$ due to the high quality of TGAS parallaxes. On the right-hand side, there is a small bump in both the RAVE unflagged sample and in the LAMOST sample compared to our fits. Since the selection function only has to capture the rough shape of the distribution, this is acceptable. Also, the distance statistics in the bottom panel indicate that we cannot allow for a steeper decline in $S(s)$ at $s > 0.5\,\mathrm{kpc}$, since this would make the distance underestimate in this region significant.

5.3.3 Nature's bad joke: the simple parallax estimator

Fig. 9 shows the estimated $1 + f$ distance estimator versus s' for the RAVE–TGAS sample (top panel) and the LAMOST sample (middle panel), when we use the naive estimate $s' = 1/p$ instead of the full distance estimate from equation (3). These statistics look surprisingly good. While the general distance underestimate in RAVE gets slightly worse and there is a trend towards small $1 + f$ at large s', but the naive estimate obviously avoids the catastrophic failure we registered when neglecting the selection function, as shown in Figs 7 and 8.

Already Strömberg (1927) pointed out that one should be careful not to use $s' = 1/p$ blindly, but by a mood of nature, or rather the selection function, the resulting distance bias in these samples is quite small. We would expect a gradual increase of the bias towards larger distances, where the parallax quality decreases, just like it happened with $1 + f$ versus distance, when we used the flat selection function $S(s)$ in the previous section. Yet, for RAVE–TGAS, $1 + f$ is remarkably stable. This is achieved because the selection function in RAVE almost perfectly cancels the terms from the coordinate transformation and the density distribution (the selection function falls steeply in distance, acting in the same direction as the s' estimate). The LAMOST sample shows good stability out to about 0.5 kpc.

The cancellation of the distance biases is also evident from the bottom panel of Fig. 9, where we plot for the full RAVE sample with parallaxes better than 20 per cent the relative difference in the distance estimates $\Delta s/\langle s \rangle = (\langle s \rangle - s')/\langle s \rangle$ for each star. $\langle s \rangle$ is the full distance estimate using equation (3). The data points are coloured by their galactic latitude $|b|$. Since the density falls quicker towards higher latitudes, the distance underestimate by s' is stronger for stars at small latitude $|b|$. To understand the shape of this plot, we remember the behaviour of the RAVE selection function from equation (21) and the top panel of Fig. 7. At distances below $\sim 0.2\,\mathrm{kpc}$, the selection function is rising and $s^2 \rho_S(s)$ rises
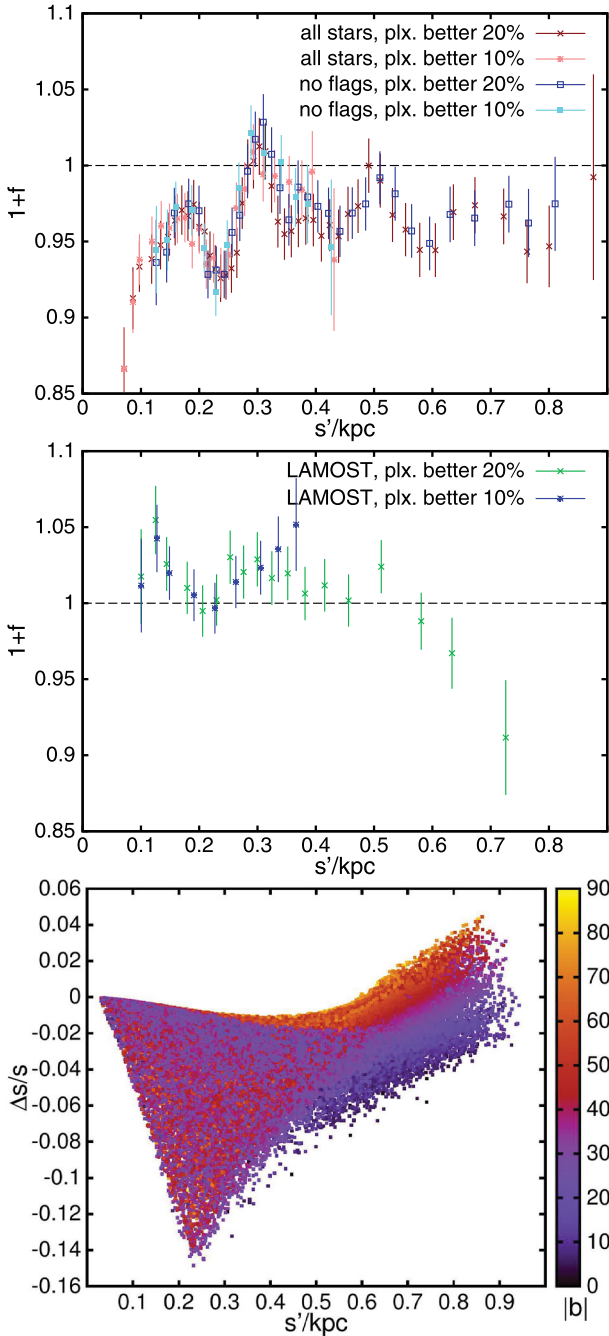
Figure 9. Distance statistics when using the simple $s' = 1/p$ distance estimator for RAVE (top panel) and LAMOST (middle panel). The distance statistics for LAMOST are balanced, since we use $e_\parallel = 7\,\mathrm{km\,s^{-1}}$. The bottom panel shows for RAVE stars with $\sigma_\mathrm{p}/p < 20$ per cent the relative difference $\Delta s/\langle s \rangle = (s' - \langle s \rangle)/\langle s \rangle$ between the full distance estimate $\langle s \rangle$ from equation (3) and $s' = 1/p$ coloured by the absolute $|b|$.

Figure 10. Comparison with the distance estimates from AB16. We could not find expectation values for their data set, so for the statistics f shown in the top panel, we follow their recommendation to use the mode of the MW-estimator. In the bottom panel, we show the fractional difference of their median distances $s_\mathrm{med,\,AB16}$ minus $s_\mathrm{med,\,SA17}$, with $s_\mathrm{med,\,SA17}$ on the ordinate, coloured in σ_p/p.

5.3.4 *Comparison with Astraatmadja & Bailer-Jones (2016)*

To understand how our distances compare to the estimates by Astraatmadja & Bailer-Jones (2016) (hereafter AB16), it is useful to revisit Fig. 3, which demonstrated two things: (i) the use of median and in particular mode of the distribution will strongly underestimate the true expectation value, since the posterior probability distribution is strongly skewed, and (ii) neglecting the selection function $S(s)$ will result in very strong distance overestimates.

Fig. 10 examines the distance estimates by AB16 on the full RAVE sample with parallaxes better than 20 per cent. The top panel shows our distance statistics applied to the distance mode of AB16, as they suggest that their comparison to other data suggest using their mode. The bottom panel directly compares our distances with theirs. It shows the fractional difference between the two distance median determinations. Towards large distances, the difference between the respective distance estimates diverges. We ascribe this to the lack of a selection function in AB16, which leads to increasing distance overestimates by them on the far end. Note that according to Fig. 3, the difference in expectation values would be even larger, since the shown median is less vulnerable to the long tail of the distributions. For nearby stars, the bright limit of the survey dominates, and the selection function rises, which is why their median distances

sharply as well, so that stars with even moderate parallax uncertainty show a strongly negative $\Delta s/\langle s \rangle$. However, beyond that point, the selection function and the spatial density fall steeply, bringing s' closer to the full estimate. As a result, the average $\Delta s/\langle s \rangle$ is about -2.4 per cent in the RAVE sample with parallaxes better than 20 per cent ($\sigma_\mathrm{p} < 0.2p$), and -1.4 per cent in the sample with parallaxes better than 10 per cent.

are shorter. We colour-coded the relative parallax error σ_p/p in the bottom panel, which shows that this is the main explaining variable for the scatter in this plot.

On the far end, the neglected selection function (leading to a distance overestimate), and the use of the mode (leading to a distance underestimate), thus nearly cancel each other, and so the distance statistics shown in the top panel of Fig. 10 show only a mild distance underestimate, just in line with Fig. 3, where for our example star also the expectation value for the correct posterior distribution and the mode of the posterior with neglected $S(s)$ coincide. For short distances/nearby stars, the selection function $S(s)$ rises, and so the two biases have the same sign, leading to comparably strong distance underestimates.

5.4 A bump in the distance statistics

The top panel of Fig. 11 shows the distance estimator $1 + f$ when binning stars in the RAVE sample according to their distance $\langle s \rangle$ in bins of 6000 stars each. Both sample selections of the RAVE–TGAS sample (red error bars: all stars and green error bars: unflagged stars) display a marked peak in $1 + f$ around $s \sim 0.3$ kpc. The feature is very narrow, and mostly limited by the distance accuracy and our ability to sample around it. It is stronger than we would expect to be possible for a statistical fluctuation, and we have checked that changes in the adopted selection function do not alter this result. In addition, an increase of the parallax accuracy does not remove the feature, even if we use $\sigma_p/p < 10$ per cent. We can exclude that this is a global issue with the parallax measurements, since the LAMOST survey in the opposite direction of the sky does not show this signal.

In the middle and bottom panels of Fig. 11, we try to dissect the sample further to find the origin of this feature. As we have discussed in Section 4.3, separating the different parts of the distance estimator, and in particular separating the sample in longitude or latitude make the analysis more prone to impacts by Galactic substructure. The middle panel compares the full estimator to its restricted counterparts from equation (10) which exclusively use the azimuthal velocity component (f_V, blue) or the RV component (f_U, green). The peak around $s = 0.3$ kpc is stronger in f_V than in f_U, though both estimators show a distance overestimate. The bottom panel attempts to dissect the sample in longitude l. This is not a safe selection, and likely the general downtrend seen in l is a consequence of the Galactic warp (the warp correlates W and V, since the sample is mostly at $b < 0$ and T_{VW} changes sign at $l = 180°$, f_V should change sign at this point). However, again the stars between $s = 0.3$ and 0.35 kpc show a markedly different behaviour. Attempts to further narrow the origin of this feature revealed that the signal is coming from around $l \sim 300°$ and latitudes both above and below the Galactic plane. This position corresponds roughly to RA $\sim 185°$ and Dec. $= -60°$. Detailed inspection of the kinematics did not reveal any evident stream at this position. Also, a stream passing through could correlate W and V motions, but the sign of T_{VW} changes with the sign change in latitude b, and so f should reverse sign as well, which it does not. Consistently, kinematic cuts to remove eventual outliers in W do not remove the feature. We have checked that the bias corrections are more than an order of magnitude too small to account for this deviation, i.e. even a large mistake in our assumptions cannot explain the feature. We have further drawn exclusion zones of $7°$ around the positions of the Magellanic clouds and found no difference in the values of f. A vertical breathing mode in the disc would be the only possibility that we could currently imagine to give rise to this feature, and the
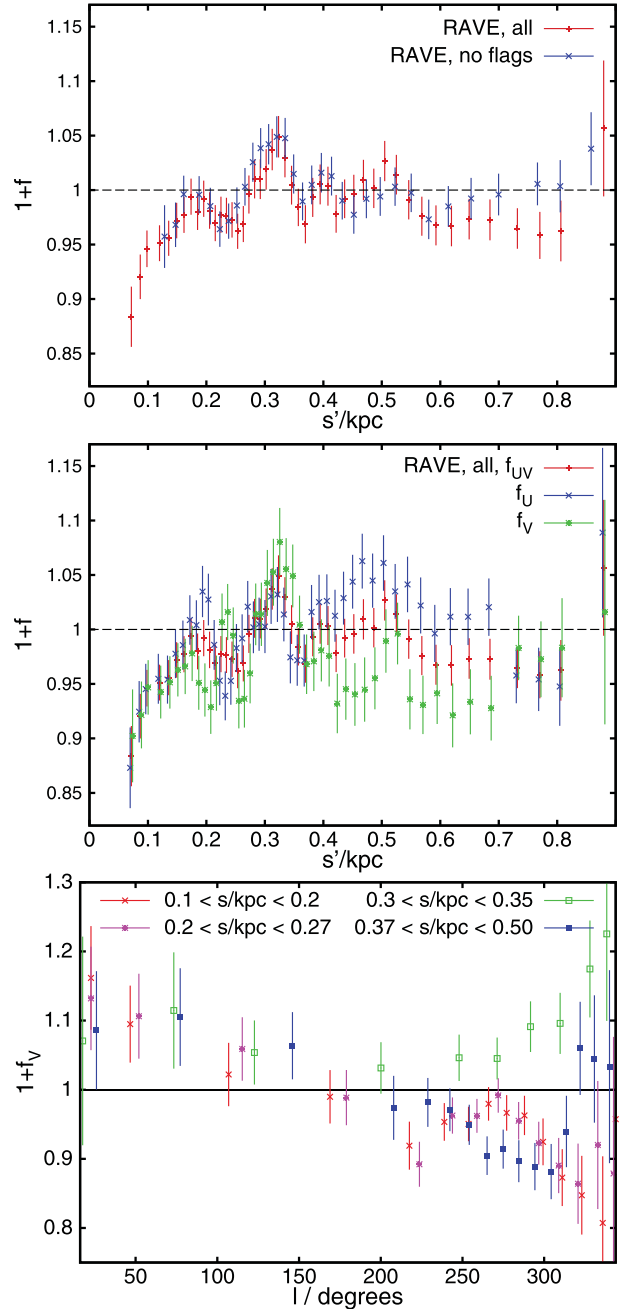

Figure 11. Top panel: comparing the distance statistics for RAVE stars, when binning in distance to samples of 6000 stars each, sliding the mask in steps of 2000 stars. Middle panel: comparing the full distance estimator $1 + f$ (red) to the restricted estimators using only the radial f_U (green), or the azimuthal f_V (blue) velocity components. Bottom panel: dissection of the distance statistics versus Galactic longitude l in the RAVE sample at different heliocentric distances s.

reversal of f_V to distances of $s > 0.4$ kpc could be interpreted as the location of a spiral arm or spur. However, there is no known spiral arm in this region. The next feature in this direction would be the Sagittarius/Carina arm, which is more than a kiloparsecs away (see e.g. Reid et al. 2014). In addition to the distance discrepancy, the explanation by a spiral arm/spur would also raise the question how this feature can be so narrow/sharp in distance. Another possible Galactic substructure in the solar vicinity, which could have a localized effect on stellar kinematics, is the Gould belt (Torra, Fernández

& Figueras 2000); however, shape, position, and distance distribution of the Gould belt stars (see e.g. Guillout et al. 1998) do not agree well with the observed feature.

We also note that all RAVE samples show a distance underestimate for stars with $s < 0.15\,$kpc. The parallaxes at this short distance are nominally too good to allow for such an effect by a wrong choice of the selection function $S(s)$. We think that this is either an extreme statistical outlier, or there is a problem in TGAS with large parallaxes and/or the bias against stars with large proper motions is significantly more severe than we take from the precaution in Lindegren et al. (2016). A definitive account is not possible from our side, since the bright limit of TGAS correlates with large proper motions, and so the fact that some nearby stars with large proper motions from *Hipparcos* are missing, can be interpreted in both ways.

To summarize this: the more likely conclusion is that there is a problem with TGAS parallaxes near $l \sim 300°$ and distances of $0.32 \pm 0.03\,$kpc ($p \sim 3\,$mas). This conclusion would be supported by Arenou et al. (2017), who report a problem possibly coinciding with the *Gaia* scanning regime change at a similar position on the sky (cf. their fig. 28). If it is not a glitch in the data, the feature is unlikely to be a stream. However, it might be related to disc structure. Since our estimate correlates azimuthal and vertical velocities, this cannot be a pure in-plane feature, though the location would rather point to a not yet established interarm, or spur. This feature warrants tests of the *Gaia* pipeline, and it should be investigated as a possible place for Galactic substructure.

5.5 Assessing the source v_{los} dispersion: LAMOST

We now return back to the question of the source v_{los} dispersion $e_\parallel$ and ways to differentiate it from parallax errors. As discussed above, the bias term in equations (13) and (16) can be used to measure $e_\parallel^2$, if the distance bias f in a sample is sufficiently well constrained.

This works very well e.g. for the LAMOST data set. In this case, the parallaxes and our method to retrieve the selection function and distance priors have been validated on the RAVE sample to an accuracy better than 2 per cent. The source v_{los} dispersion $e_\parallel$ can now be determined by plotting the corrected distance bias estimate $1 + f$, while varying the source v_{los} dispersion $e_\parallel$. This is done in Fig. 12, where the green lines show this variation for the LAMOST sample, dashed lines represent the 1σ error intervals. We use the quality cuts derived in Section 5.2, S/N > 30 and $\sigma_{los,LAM} < 27\,$km s^{-1}, however, do not use $\sigma_{los,LAM}$ in the bias correction of $1 + f$. Under the assumption that the distance estimates from TGAS parallaxes are unbiased, we obtain $e_\parallel = 7 \pm 1.0\,$km s^{-1}. This estimate contains a generous allowance for uncertainties in the velocity ellipsoid and proper motions. The systematic uncertainty from the distances is difficult to estimate, but from the previous indications it should be of order $2\,$km s^{-1}. $e_\parallel$ consists of the σ_{los} and additional source dispersion from binary systems. However, the latter should be of order $\sigma_{bin} \lesssim 3\,$km s^{-1}, so according to equation (17), their contribution to $e_\parallel$ is small.

5.5.1 Separating distance bias and line-of-sight velocity errors

Can we separate the line-of-sight velocity errors from a true distance error? In a larger sample, we can. The idea behind this is tested with the red and blue lines in Fig. 12. These lines show a separation of the LAMOST sample into subsamples with large and small V_g (top panel), or subsamples with large and small U_g (bottom panel).
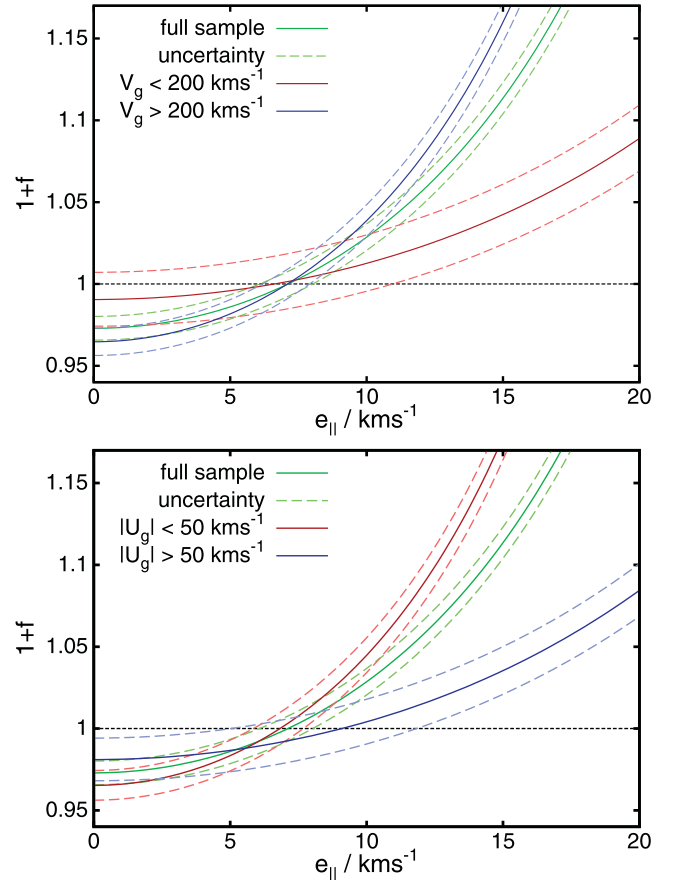

Figure 12. Both panels show the distance estimator $1 + f$ against the assumed v_{los} source dispersion $e_\parallel$ for the LAMOST sample. In these plots, $e_\parallel$ can be determined by demanding an unbiased distance estimator $1 + f \sim 1$. The full sample is shown with green lines, dashed lines show the 1σ error interval. The right combination of f and $e_\parallel$ can be tested by cutting the sample into stars with a large versus small baseline in VT_{VW} or UT_{UW} and demanding mutual agreement. In the top panel, we cut in galactocentric azimuthal velocity V_g, and in the bottom panel, we cut in the RV U_g. These subsamples are shown with red and blue lines in each panel.

We recall that f is estimated based on the correlation of vertical velocities W with the baseline (y) components $T_{VW}V$ and $T_{UW}U$, measured by the covariance $Cov(W, y)$ in equation (9). The selection in this plot target subsamples with a larger or smaller baseline y, i.e. while f is similar, both the covariance terms and the variance in the denominator of equation (9) vary. In contrast, the bias on the covariance from the v_{los} source dispersion is approximately the same, since it only depends on the sample geometry and $e_\parallel^2$. If the bias from $e_\parallel$ dominates the distance statistics, we hence expect the different sample selections in Fig. 12 to have the same $1 + f$ close to 1 at the correct value of $e_\parallel$, but the sample chosen to be closer to the Sun's velocity (blue lines in Fig. 12) should have a steeper relationship of $1 + f$ against $e_\parallel$ with larger curvature than its counterpart with larger heliocentric velocities (red lines). As a consequence, the low-velocity selection (blue) should also have a larger deviation of $1 + f$ from 1 at $e_\parallel = 0$, exactly what is observed in Fig. 12. In contrast, if the measured deviation is a distance error, $1 + f$ of the subsamples should be in agreement at small values of $e_\parallel$. In Fig. 12, both ways to separate the sample agree perfectly on $1 + f \sim 1$ at an $e_\parallel \sim 7\,$km s^{-1}. However, we also note that the still moderate sample size limits the statistical significance:

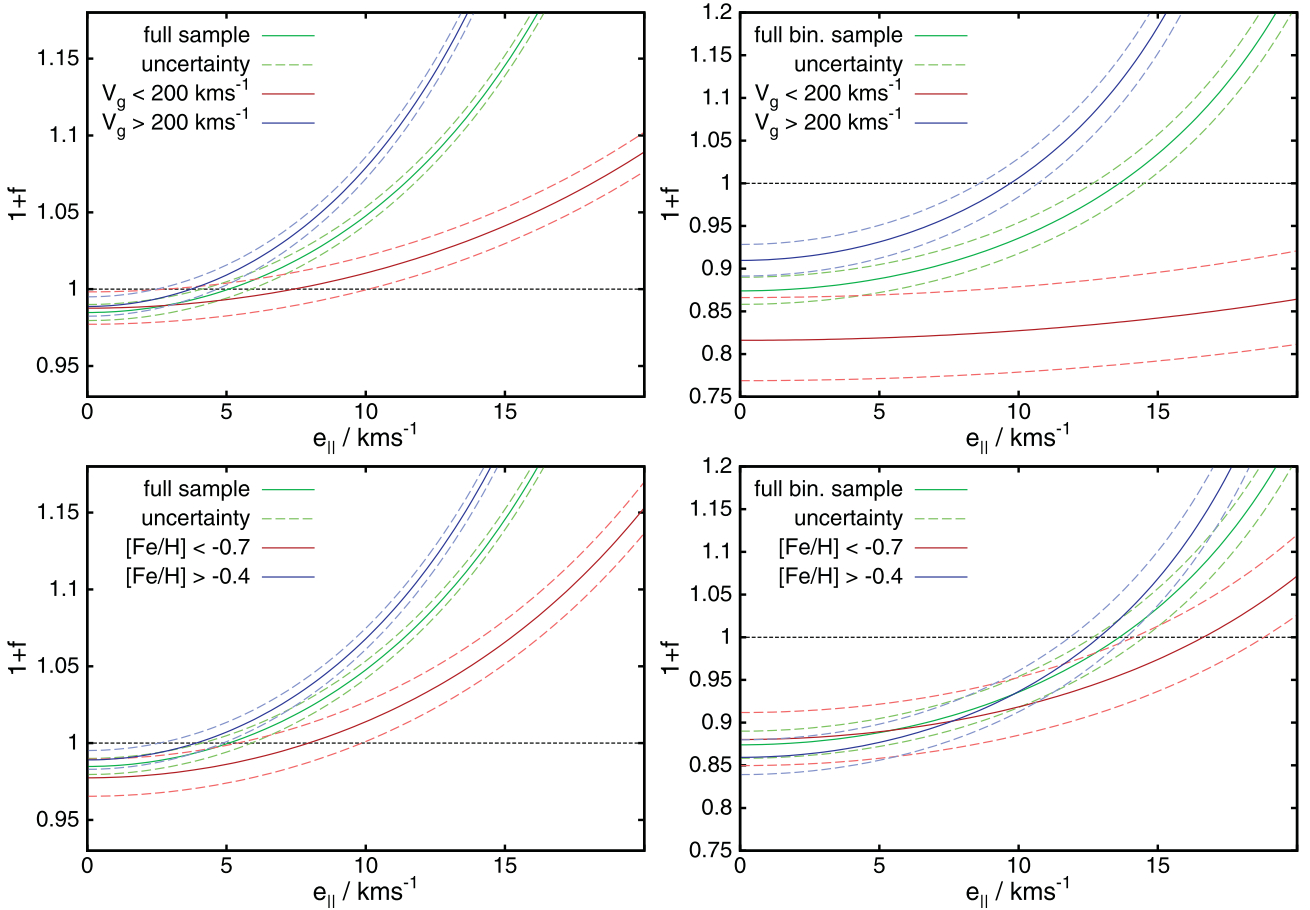**Figure 13.** Distance bias $1 + f$ versus v_{los} source dispersion $e_{\parallel}$ in RAVE for the entire sample (left-hand column) and for the subsample carrying binary flags (right-hand column). The plots are organized as in Fig. 12, the green lines show the full sample, blue lines the subsample with smaller heliocentric velocities, and red lines the subsample with larger heliocentric velocities. Here, the top panels show a separation at $V_g = 200$ km s^{-1}, the bottom row shows a separation in metallicity, selecting once stars with [Fe/H] < −0.7 dex and once with [Fe/H] > −0.4 dex. The top right plot should be taken with some caution, since the low-V_g binary sample only contains ∼400 stars.

the discrimination on LAMOST without our prior knowledge from RAVE only barely excludes $e_{\parallel} \sim 0$.

The cut can be done either in V_g (top) or U_g (bottom panel), but it has to be done in galactocentric velocities, since a heliocentric selection would slice through the velocity ellipsoid at an angle and hence correlate W and the horizontal components. Another strategy would be to select the sample at different values of T_{VW} and T_{UW}, since the bias term has a quadratic dependence on these, or to use e.g. a metallicity cut to select kinematically hotter stars that have a larger baseline y by their larger velocities.

Each of these options has their caveats: the kinematic cuts in U_g or V_g select to some extent for the errors we are measuring. Stars with larger velocity errors (either from distance or from velocity uncertainties) are pushed towards the tails of the velocity distribution. A good position to cut in a larger sample are hence the points in the velocity distribution where it has a small or vanishing slope. If one selects in metallicity, the major caveats are a possible difference in the binary fraction, or also a correlation of σ_{los} with the metallicity (typically σ_{los} increases for metal-poor, hot stars).

5.6 Assessing the line-of-sight velocity uncertainty in RAVE

While the main measurement for the LAMOST sample is the line-of-sight velocity errors, the situation in the RAVE sample is far less

clear. RAVE line-of-sight velocities have typically errors smaller than $\sigma_{los,RAVE} \lesssim 2$ km s^{-1}, i.e. an order of magnitude smaller than the LAMOST uncertainties. Consequently, we get a precise measurement on LAMOST line-of-sight velocity errors, but we just see that something is wrong with the RAVE sample, either stellar distances being too short, or some contamination with larger v_{los} errors, or binary systems contaminating the data.

Fig. 13 gives an overview of the trade-off between parallax/distance bias and $e_{\parallel}$, similar to Fig. 12. From the earlier analysis, we know that a major part of the problems can be ascribed to the suspected binary systems, so we show the full RAVE sample on the left-hand side and the suspected binary stars on the right-hand side. Note that the subsample of suspected binary stars with <6000 stars is very small, so the further separations (which have only 400 stars in the low-V_g component) should be taken with caution due to small number statistics.

As we can see from the entire sample printed in green on the left-hand column, the full sample has a significant average distance underestimate ($1 + f < 1$) if we assume $e_{\parallel} \lesssim 1$ km s^{-1} as suggested by the $\sigma_{los,RAVE}$ estimates of the RAVE pipeline, which tested fine in the statistics of Section 5.2. If we assumed that *Gaia*–TGAS and our distance derivation were completely bias-free, we would conclude that $e_{\parallel} \sim 5$ km s^{-1}. It would be very unlikely that this is caused by large v_{los} measurement error in light of our findings in

Section 5.2. Just like in LAMOST, we attempt to break the degeneracy by separating the sample in V_g (top panel) and in metallicity (bottom panel). This sample, however, is biased towards younger stars with small velocity dispersions and disc populations, so we have few useful halo stars. Both separations barely scratch statistical significance in the region of interest. Demanding agreement between the subsamples, we can only conclude that $e_\parallel < 5\,\mathrm{km\,s^{-1}}$, and that the statistics favour a problem with the stellar distances. We note that the values for $e_\parallel$ are at least at the upper end of what might be achievable by binary systems (Hargreaves et al. 1996; Olszewski et al. 1996). We tested that a change of the parallax quality to $\sigma_p/p < 10$ per cent yields no significant difference.

The two panels on the right-hand side of Fig. 13 attempt the same dissection on the subsample of suspected binary stars. While assuming unbiased distances for this subsample would yield a v_{los} source dispersion $e_\parallel \sim 14 \pm 2\,\mathrm{km\,s^{-1}}$, slicing the sample into subsets of either low and high azimuthal velocity (top panel) or low and high metallicity (bottom panel) fails to convince that a velocity dispersion σ_{bin} caused by binary stars is the sole explanation. As discussed above, the samples with large intrinsic baselines in VT_{VW} and UT_{UW} (i.e. the low-V_g sample, or the metal-poor subsample) should then start at a value of $1 + f$ closer to 1 at $e_\parallel = 0$, and intersect with their small-baseline counterpart at near-zero distance bias f. Here, we observe the opposite. The low-V_g subsamples (red lines) start already at more extreme $1 + f$ than the high-V_g subsample (blue lines). There is no agreement between the subsamples anywhere near the region of $1 + f \sim 1$. Both the top and the bottom panel suffer from the same discrepancy, though the bottom panel shows a milder version, while having different potential biases: the top panel possibly suffers from a potential direct kinematic bias, whereas the bottom panel could be impacted by binaries being wrongly evaluated as particularly metal-poor. This would be a reasonable suspicion in a low-resolution spectroscopic survey like RAVE. The unresolved binary broadens the lines, and while the equivalent width should not be strongly lowered, different line shapes might have this effect. In addition, the more metal-poor stars could have unaccounted v_{los} determination errors affecting both separations. A potential way out would be to claim that the Matijević et al. (2012) classification strongly depends on metallicity, but the most straightforward explanation is that TGAS astrometry fails for binary systems, in addition to some contribution by $e_\parallel$.

What we should take from this section is: Once we have a larger sample with certainly clean astrometry and a v_{los} determination at the quality level of RAVE and later *Gaia*, this method will be able to constrain even the binary velocity dispersion for different samples, and to differentiate well between $e_\parallel$ and distance uncertainties. The sizes of the current samples are just at the boundary of offering a full solution which breaks the degeneracy. And, we can clearly see from these statistics that stars with binary flags should be excluded from any analysis.

6 CONCLUSIONS

In this work, we have derived distance expectations and kinematics for stars in the RAVE–TGAS and LAMOST–TGAS samples.[4] We have applied the statistical methods developed by SBA to estimate the distance bias f in the RAVE–TGAS and the LAMOST–TGAS

samples, and provide a set of validated stellar distances and kinematics. While there are several papers testing the accuracy and precision of parallaxes from *Gaia*–TGAS, even perfectly unbiased parallaxes do not guarantee unbiased distances, since there is a large uncertainty in estimating the right priors and selection functions, which enter a Bayesian distance estimate. Our method uses correlations of derived velocities with position in the sky, and hence directly tests the derived distances. We have validated this method using a realistic Galaxy simulation, for which we know the exact distances, as shown in Appendix A.

We find that the entire RAVE–TGAS sample shows statistically a global distance underestimate by of order 2–3 per cent. However, more than half of this signal disappears when we restrict the sample to a subset of stars that are flagged as fully normal objects according to the classifications of Matijević et al. (2012), which detect binaries or peculiar stars, e.g. with emission lines. We have demonstrated that the objects with at least one binary flag show distance statistics equivalent to distance underestimates in excess of 15 per cent.

While it would be tempting to put all the blame of the distance bias on TGAS parallaxes, we have shown that distance statistics f are biased proportional to the source v_{los} variance $e_\parallel^2$. This comprises v_{los} measurement errors and contamination with binary stars, which act like a distance underestimate on f. On the one hand, this requires that we find a way to differentiate $e_\parallel$ from a true distance bias. On the other hand, it is a useful way to measure $e_\parallel$ and hence determine σ_{los} on samples where parallaxes are good and the $e_\parallel$ dominates the f estimates. We can use this dependence of f on $e_\parallel$ in two ways: (i) we can vary $e_\parallel$ with the condition $f \sim 0$, and thus measure $e_\parallel$; (ii) we can probe the accuracy of σ_{los} estimates in a pipeline by separating each sample in σ_{los} and looking for trends in f.

We have also examined the distances provided by Astraatmadja & Bailer-Jones (2016). The comparison underlines the importance of applying the right selection function. Neglecting this selection function leads to a relative distance overestimate in AB16. However, their choice to use the mode of their distributions (over-)compensates this effect, which leads to mildly short distance estimates. We advise to use expectation values, which are well-defined statistical quantities.

The LAMOST sample demonstrates this capability. Our method has revealed that the uncertainties given by the LAMOST pipeline have no detectable information content for uncertainties below $\sigma_{los,LAM} < 27\,\mathrm{km\,s^{-1}}$, while stars with larger $\sigma_{los,LAM}$ show a sudden increase in $e_\parallel$. More importantly, when comparing the full distance estimator f with its counterpart f_V that uses only the azimuthal V velocity versus the vertical motion W, these two estimators yield highly inconsistent results. We also find that the LAMOST sample shows a near-perfect relationship between W and $\sin(b)$, indicating a $5\,\mathrm{km\,s^{-1}}$ offset in the v_{los} measurements (a similar offset has been suggested in Tian et al. 2015), while RAVE has no significant correlation. When we correct LAMOST for this $\delta v_{los} = 5\,\mathrm{km\,s^{-1}}$, we measure $e_\parallel = 7.1 \pm 1\,\mathrm{km\,s^{-1}}$ with a systematic uncertainty of about $2\,\mathrm{km\,s^{-1}}$, i.e. the LAMOST v_{los} measurements are far more precise than indicated by their nominal errors.

In RAVE, the full sample starts showing signs of increased $e_\parallel$ above $\sigma_{los,RAVE} \geq 1.5\,\mathrm{km\,s^{-1}}$, and these reach an unacceptable level above $5\,\mathrm{km\,s^{-1}}$. This behaviour is not detectable when we restrict the sample to unflagged/normal stars, i.e. the flags indicate the most problematic v_{los} determinations.

We have demonstrated how to measure the selection function $S(s)$ and distance prior of each sample from the data directly in Section 5.3.1. We start from a flat $S_0(s)$, measure a new $S_1(s)$ by comparing the stellar distance distribution to the expectation in the prior,

---

[4] Please find the data sets with distances and kinematics and our source code at http://www-thphys.physics.ox.ac.uk/people/RalphSchoenrich/data/tgasdist/data.tar.gz, or request them directly from the authors.

and then iteratively insert this selection function into the distance determinations. This procedure converges within five iterations. By separating the sample in distance, we can show that neglect of this selection function [i.e. a flat $S(s)$] results in catastrophic distance bias towards larger distances, while the derived selection function provides unbiased distances within the measurement accuracy.

It is generally inadvisable to use the naive parallax estimator $s' = 1/p$ for stellar distances, because it neglects the proper transformation of the parallax error distribution into s, which gives a large distance underestimate against the true expectation value $\langle s \rangle$. However, to our surprise, we find that on both RAVE and LAMOST–TGAS, this estimator results only in a mild distance underestimate. The reason for this is that the selection function/distance prior falls steeply towards larger s, compensating partly for the missing transformation.

We note that the distance estimator f for RAVE stars with very short distances ($\langle s \rangle \lesssim 0.15$ kpc) consistently indicate distance underestimates. This can either be an extreme statistical fluctuation, or it might indicate either a problem with the largest TGAS parallaxes, or a stronger selection against stars with large proper motions than indicated in Lindegren et al. ([2016](#)).

We have detected an anomaly either in the stellar kinematics or in the parallaxes for the RAVE–TGAS sample around a galactic longitude $l \sim 300°$, and at a distance of $0.3 < s/\text{kpc} < 0.35$. The feature is found on both sides of the Galactic plane, and the correlation between vertical and azimuthal velocities is neither found in its foreground nor in its background. The same sign on both sides of the plane argues strongly against a (halo) stellar stream as source of this. The feature is also robust against outliers, and cannot be explained by eventual anomalies in the selection function. A likely explanation is substructure at this position in the Galaxy, e.g. a breathing mode near a spiral arm or spur, though we are not aware of such a feature at this position. Further analysis will be required to understand if this is a possible localized failure of TGAS astrometry, or in the other case, to unravel the exact nature of this structure.

We have further shown that the degeneracy between v_{los} source dispersion $e_{\parallel}$ and a true distance bias can be broken by selecting subsamples with different lengths of baseline in VT_{VW} or UT_{UW}. This is achieved by splitting the sample either with a metallicity selection or directly in stellar kinematics. The larger upcoming *Gaia* data releases will be sufficient to make this method work. In the current samples, this ability is just borderline significant. It shows that while we have a consistent picture for LAMOST being dominated by line-of-sight velocity uncertainties, the statistics suggest that the deviations in the RAVE survey, in particular on the binary subsample, derive from a combination of astrometric problems and intrinsic v_{los} dispersion.

ACKNOWLEDGEMENTS

REFERENCES

Arenou F. et al., 2017, A&A, 599, A50
Arias E. F., Charlot P., Feissel M., Lestrade J.-F., 1995, A&A, 303, 604
Astraatmadja T., Bailer-Jones C., 2016, ApJ, 832, 137 (AB16)
Aumer M., Binney J., 2017, MNRAS, 470, 2113
Aumer M., Binney J., Schönrich R., 2016, MNRAS, 459, 3326
Binney J. et al., 2014, MNRAS, 439, 1231
Burke B. F., 1957, AJ, 62, 90
Davies G. R. et al., 2017, A&A, 598, L4
De Ridder J., Molenberghs G., Eyer L., Aerts C., 2016, A&A, 598, L3
De Rijcke S., Dejonghe H., 1998, MNRAS, 298, 677
Debattista V., 2014, MNRAS, 443, 1
Dehnen W., 1998, AJ, 115, 2384
Dehnen W., 2000, AJ, 119, 800
Djorgovski S., Sosin C., 1989, ApJ, 341, 13
ESA, ed., 1997, ESA SP-1200: The Hipparcos and Tycho Catalogues. ESA, Noordwijk
Faure C., Siebert A., Famaey B., 2014, MNRAS, 440, 2564
Feissel M., Mignard F., 1998, A&A, 331, 33
Gaia Collaboration et al., 2016, A&A, 595, 2
Gaia Collaboration et al., 2016, A&A, 595, 1
Gieles M., Sana H., Portegies Zwart S., 2010, MNRAS, 402, 1750
Girardi L., Salaris M., 2001, MNRAS, 323, 109
Guillout P., Sterzik M. F., Schmitt J., Motch C., Neuhäuser R., 1998, A&A, 337, 113
Hargreaves J. C., Gilmore G., Annan J. D., 1996, MNRAS, 279, 108
Helmi A., Veljanoski J., Breddels M. A., Tian H., Sales L. V., 2017, A&A, 598, 58
Høg E. et al., 2000, A&A, 355, 27
Ivezić Z. et al., 2008, ApJ, 684, 287
Joshi Y. C., 2007, MNRAS, 378, 768
Kerr F. J., 1957 AJ, 62, 93
Kunder A. et al., 2016, AJ, 153, 75
Lindegren et al., 2016, A&A, 595, 4
Liu X.-W., Zhao G., Hou J.-L., 2015, Res. Astron. Astrophys., 15, 1089
Luo A.-Li et al., 2015, Res. Astron. Astrophys., 15, 1095
Matijevič G. et al., 2012, ApJS, 200, 14
McMillan P., 2011, MNRAS, 414, 2446
McMillan P., 2017, MNRAS, 465, 76
Olszewski E., Pryor C., Armandroff T., 1996, AJ, 111, 750
Pérez-Villegas A., Portail M., Wegg C., Gerhard O., 2017, ApJ, 840, L2
Perryman M., 1997, A&A, 323, 49
Perryman M., 2001, A&A, 369, 339
Poggio E., Drimmel R., Smart R. L., Spagna A., Lattanzi M. G., 2017, A&A, 601, 115
Reid M. J. et al., 2014, ApJ, 783, 130
Salaris M., Girardi L., 2002, MNRAS, 337, 332
Schönrich R., 2012, MNRAS, 427, 274
Schönrich R., Bergemann M., 2014, MNRAS, 443, 698
Schönrich R., Binney J., Dehnen W., 2010, MNRAS, 403, 1829
Schönrich R., Binney J., Asplund M., 2012, MNRAS, 420, 1281 (SBA)
Siebert A. et al., 2008, MNRAS, 391, 793
Steinmetz M. et al., 2006, AJ, 132, 1645
Strömberg G., 1927, ApJ, 65, 238
Tian H.-J. et al., 2015, ApJ, 809, 145
Torra J., Fernández D., Figueras F., 2000, A&A, 359, 82
van Leeuwen F., 2007, in Lilensten J., ed., Astrophysics and Space Science Library, Vol. 350, Space Weather : Research Towards Applications in

Europe 2nd European Space Weather Week. Springer-Verlag, Dordrecht, The Netherlands, p. 1

Vogt S., Mateo M., Olszewski E., Keane M., 1995, AJ, 109, 151

Wang S., Su D., Chu Y., Cui X., Wang Y., 1996, Appl. Opt., 35, 5155

Wojno J. et al., 2016, MNRAS, 468, 3368

Xu Y., Newberg H. J., Carlin J. L., Liu C., Deng L., Li J., Schönrich R., Yanny B., 2015, ApJ, 801, 105

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

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

APPENDIX A: VALIDATION ON GALAXY SIMULATIONS

To perform a sanity check on our method and to confirm its robustness against galactic substructure, we employ galaxy simulations that resemble the MW. The simulations are described in detail in Aumer, Binney & Schönrich (2016) and Aumer & Binney (2017). They feature a reasonable bar having appropriate thin and thick disc components and an active spiral pattern.

We focus on model Vα9s8$\lambda\zeta$. To build this model, we continuously add stellar particles on disc orbits to a model galaxy within a live dark matter halo over the course of 12 Gyr. The birth velocity dispersions of the stellar populations decline continuously in this model, from $>40\,\mathrm{km\,s^{-1}}$ during the first 1.5 Gyr to essentially near-circular orbits after 4 Gyr. This creates a disc galaxy with reasonable thin and thick disc components.

Moreover, the model grows inside–out with a declining star formation history. Vα9s8$\lambda\zeta$ has a final baryonic mass $M_{\mathrm{baryons}} = 6 \times 10^{10}\,\mathrm{M_\odot}$, and a live dark matter halo with mass $M_{\mathrm{dm}} = 1 \times 10^{12}\,\mathrm{M_\odot}$, which is set up with an initial concentration parameter of $c = 6.5$.

At our evaluation time 12 Gyr, the vertical profile is double-exponential. At $R = 8\,\mathrm{kpc}$, it can be fitted with exponential scale-heights $h_{z,\,\mathrm{thin}} \approx 285\,\mathrm{pc}$ and $h_{z,\,\mathrm{thick}} \approx 990\,\mathrm{pc}$. The thick disc contributes ~ 20 per cent to the local surface density. The final model has a bar of length $\sim 4\,\mathrm{kpc}$. Its $R = 8\,\mathrm{kpc}$ circular speed, dark matter density and baryonic surface density compare reasonably with observational constraints for the Solar neighbourhood. The level of radial migration in the model agrees with constraints from MW chemical evolution and its Solar neighbourhood velocity distributions show good agreement with RAVE+TGAS data if one corrects for selection effects. Overall, the model is well suited to create mock Solar neighbourhood samples.

To create these samples, we place the observer at $(R, z) = (8.3, 0)\,\mathrm{kpc}$ choosing 25 different positions, which are equally distributed in galactocentric azimuth Φ_{gal}. We then select all stars with distances $s < 1\,\mathrm{kpc}$, obtaining samples of about 15–20 000 stars each. In these samples, we calculate the stellar positions, proper motions, and line-of-sight velocities using a solar motion of $(U_\odot, \boldsymbol{v}_\odot, W_\odot) = (11, 250, 7)\,\mathrm{km\,s^{-1}}$ (we note that the only effect of the set solar motion is to alter the length of the baseline component $T_{\mathrm{VW}}V$ and thus the precision of f). Fig. A1 shows the tests of our distance statistics on these quite realistic samples. The top panel shows the full distance estimator $1 + f$ for each of the 25 observer positions in Φ_{gal}. The test statistics are consistent with a perfectly unbiased estimator. There are nine data points beyond 1σ, which is just the expected number of outliers.
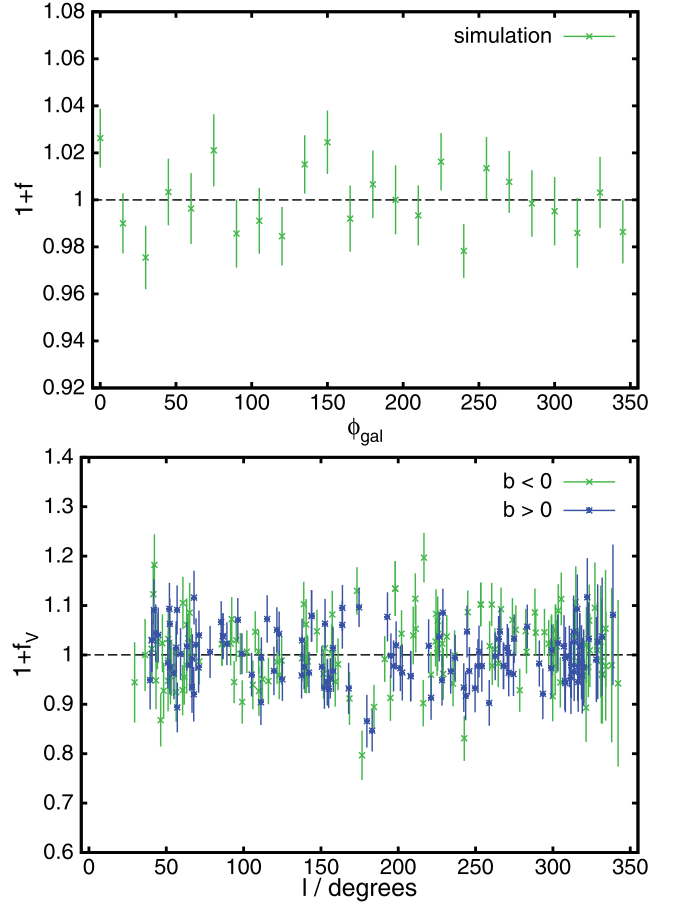

Figure A1. Testing the distance estimator on *N*-body galaxy model Vα9s8$\lambda\zeta$. The top panel shows the distance statistics for 1 kpc sized samples taken around simulated solar positions at $(R, z) = (8.3, 0)\,\mathrm{kpc}$, equally spaced in azimuth Φ_{gal} (*x*-axis). The bottom panel shows the restricted distance estimator f_V when we bin each sample in Galactic longitude l (*x*-axis) and separate the sample into positive (blue) and negative (green) galactic latitude b.

Table A1. Distance error estimates f in simulations with perfect distances. Each simulation has 25 samples around $R = 8\,\mathrm{kpc}$ at equidistantly spaced positions in azimuth. The first three columns give statistics for the full distance estimator f. $\langle f \rangle$ is the mean value of f from averaging over all 25 samples, σ_f is the residual dispersion of f, and $\sigma_{f,\mathrm{st}}$ is the average of the statistical errors derived with f. If the estimates are true, then σ_f should be close to $\sigma_{f,\mathrm{st}}$. Systematic biases in f depending on the sample position would show as σ_f significantly larger than $\sigma_{f,\mathrm{st}}$. The last three columns give the same statistics for f_V.

Model	$\langle f \rangle$	σ_f	$\sigma_{f,\mathrm{st}}$	f_V	$\sigma_{f,V}$	$\sigma_{f,V,\mathrm{st}}$
Vα9s8$\lambda\zeta$	-0.0004	0.014	0.014	0.0015	0.033	0.022
Vα5λ	-0.0073	0.018	0.013	-0.0064	0.024	0.019
Mα1ζ*	0.0026	0.019	0.017	-0.0013	0.030	0.027
P2	-0.0064	0.020	0.017	-0.0066	0.035	0.028
EHR2	-0.0072	0.010	0.008	0.0009	0.014	0.013

The mean distance estimator $<f>$ in each sample, the dispersion of the distance estimates σ_f, and the expected dispersion from shot noise and uncertainty in the correction of the velocity ellipsoid tilt $\sigma_{f,\mathrm{st}}$ for model Vα9s8$\lambda\zeta$ are summarized in the first line of Table A1. We also show the numbers for four additional MW-like models. Models Vα5λ and Mα1ζ* are of the same type as Vα9s8$\lambda\zeta$, but differ in model parameters and consequently in structural details

at final time. Models P2 and EHR2 also contain both thin and thick disc components, but the thick disc components are created as initial conditions and the thin disc is populated with stars born on near-circular orbits over 10 Gyr.

The values for all five models included in Table A1 demonstrate that, on average, the distance statistics show no bias at the levels needed for this work (generally less than 1 per cent). Moreover, the statistically expected scatter almost matches the observed scatter in $1 + f$, i.e. any impact by unaccounted systematics in the samples is significantly below 1 per cent. We note that the dispersion values are lowest in EHR2 as this model has a higher resolution and thus larger sample sizes than the other four. f_V is in principle more vulnerable to disc structure, but we see no clear indication for a significant additional scatter in its statistics provided in the three right-hand columns.

As detailed in Section 4.3, the fits become more risky when we separate samples in longitude and/or latitude, because effects caused by galactic structure do not cancel out any more to first order. This is particularly the case for simulation EHR2, which has a longer bar than the MW. As the bar thus has a larger impact at $R = 8$ kpc, in that simulation, cutting the test samples at $l = 180°$ results in a scatter of ~ 5 per cent in f_V. As expected, most of this gets compensated in the full estimator f, which has just 2 per cent dispersion (compared to about 1.2 per cent statistically expected dispersion). This simulation shows a discernible $m = 2$ pattern when the f_V statistics are plotted against Φ_{gal}. However, as we see in Table A1, even this simulation does not yield a discernible error if we use a full sky coverage.

The bottom panel in Fig. A1 looks for trends in the restricted distance estimator f_V when separating simulation Vα9s8$\lambda\zeta$ from the top panel in galactic longitude l, selecting subsamples of 3000 stars each. Before ordering the sample in l, we have additionally cut it into Northern and Southern hemispheres ($b > 0$ and $b < 0$). There is no detectable systematic deviation, apart from a minor issue with the error determination near $l = 180°$. At that position $T_{VW} \sim 0$, i.e. there is virtually no signal for f_V, which also affects the error determinations in a small region.

Similarly, the increase in uncertainties of f_V is due to the vanishing of T_{VW}. Overall, the statistics of this sample look fine – due to the cold stellar kinematics, and the small sample sizes, the statistically expected errors are ~ 12 per cent, close to the observed dispersion.

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