Planets Across Space and Time (PAST). I. Characterizing the Memberships of Galactic Components and Stellar Ages: Revisiting the Kinematic Methods and Applying to Planet Host Stars

Di-Chang Chen1,2, Ji-Wei Xie1,2, Ji-Lin Zhou1,2, Subo Dong3, Chao Liu4,5, Hai-Feng Wang6,10, Mao-Sheng Xiang7,8, Yang Huang6, Ali Luo7, and Zheng Zheng9

1 School of Astronomy and Space Science, Nanjing University, Nanjing 210023, People’s Republic of China; jwxie@nju.edu.cn
2 Key Laboratory of Modern Astronomy and Astrophysics, Ministry of Education, Nanjing 210023, People’s Republic of China
3 Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, People’s Republic of China
4 Key Laboratory of Space Astronomy and Technology, National Astronomical Observatories, CAS, Beijing 100101, People’s Republic of China
5 University of Chinese Academy of Sciences, Beijing 100049, People’s Republic of China
6 South-Western Institute for Astronomy Research, Yunnan University, Kunming 650500, People’s Republic of China
7 National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, People’s Republic of China
8 Max-Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany
9 Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112, USA

Received 2020 September 8; revised 2020 November 23; accepted 2020 December 11; published 2021 March 10

Abstract

Over 4000 exoplanets have been identified and thousands of candidates are to be confirmed. The relations between the characteristics of these planetary systems and the kinematics, Galactic components, and ages of their host stars have yet to be well explored. To address these questions, we conduct a research project, dubbed Planets Across Space and Time (PAST). To do this, one of the key steps is to accurately characterize the planet host stars. In this paper, Paper I of the PAST series, we revisit the kinematic method for classification of Galactic components and extend the applicable range of velocity ellipsoid from ~100 pc to ~1500 pc from the Sun in order to cover most known planet hosts. Furthermore, we revisit the age–velocity dispersion relation (AVR), which allows us to derive kinematic ages with a typical uncertainty of 10–20% for an ensemble of stars. Applying the above revised methods, we present a catalog of kinematic properties (i.e., Galactic positions, velocities, and the relative membership probabilities among the thin disk, thick disk, Hercules stream, and the halo) as well as other basic stellar parameters for 2174 host stars of 2872 planets by combining data from Gaia, LAMOST, APOGEE, RAVE, and the NASA exoplanet archive. The revised kinematic method and AVR, as well as the stellar catalog of kinematic properties and ages, lay the foundation for future studies of exoplanets in space and time in the Galactic context.

Unified Astronomy Thesaurus concepts: Exoplanets (498); Exoplanet catalogs (488); Stellar ages (1581); Stellar kinematics (1608); Milky Way Galaxy (1054)

Supporting material: machine-readable table

1. Introduction

It has been a quarter century since the discovery of the first exoplanet. To date, over 4000 exoplanets have been discovered and thousands of candidates are yet to be confirmed (NASA Exoplanet Archive, EA hereafter; Akeson et al. 2013). There is clear evidence (shown in Figure 1, data from http://exoplanet.eu) that our knowledge of exoplanets in the Galaxy is expanding. Before 2005, most known exoplanets were confined to the solar neighborhood, with distances less than ~100–200 pc. Now, the map of exoplanets is much wider with a large range of distances up to ~10,000 pc. Therefore, people began to study exoplanets in the context of the Galaxy. For example, there were continuous discussions on how to define the Galactic habitable zone (Gonzalez et al. 2001; Lineweaver et al. 2004; Sundin 2006; Gowanlock et al. 2011; Jiménez-Torres et al. 2013; Balbi & Tombesi 2017; Stojković et al. 2019), researches on the Galactic distribution of planets as a function of distance/population (e.g., Zhu et al. 2017), and studies on whether the planet occurrence rate depends on the Galactic velocity (e.g., Bashi & Zucker 2019; McTier & Kipping 2019).

One of fundamental questions in studying exoplanets in the Galactic context is: what are the differences in the properties of planetary systems at different positions in the Galaxy with different ages? The answer to this question will provide insights on the formation and evolution of the ubiquitous and diverse exoplanets in different Galactic environments. To address the question, in a series of papers from here on, we conduct statistical studies of planets at different positions in the Galaxy with different ages, a project that we dub Planets Across Space and Time (PAST).

To this end, the first step is to figure out where the exoplanets are in the Galaxy. Specifically, for a given exoplanet host star, we would like to know which Galactic component (i.e., the thin disk, the thick disk, or the halo) it belongs to. One of well-established methods to distinguish these Galactic components is the kinematic approach, as different components generally have different kinematic characteristics. For example, the thin disk has a smaller vertical scale height (Bovy et al. 2012a; Wu et al. 2018), but the thick disk is generally kinematically hotter with larger velocity dispersions (Gilmore et al. 1989; Reddy et al. 2003; Bensby et al. 2003, 2014; Buder et al. 2018). By comparing the kinematic properties of a given star to the typical kinematic characteristics of a Galactic
component, one may calculate the likelihood that the star belongs to this component (e.g., Bensby et al. 2003). However, the kinematic characteristics of this method were obtained with data in the solar neighborhood within ~100 pc (Bensby et al. 2003, 2014), and therefore this kinematic method is limited to a relatively small range of area. Thanks to the recent large-scale star surveys both from space (e.g., Gaia, Gaia Collaboration et al. 2016, 2018a, 2018b) and the ground (e.g., LAMOST, Wang et al. 1996; Su & Cui 2004; Cui et al. 2012; Zhao et al. 2012; Luo et al. 2012), we are now able to extend the kinematic method to beyond 1000 pc in order to characterize the majority of exoplanet host stars.

The second step is to obtain the ages of exoplanet host stars, since most exoplanet host stars have no (accurate) age estimates. Stellar ages can hardly be measured but only inferred or estimated indirectly through a number of techniques, which have their own strengths and weaknesses (Soderblom 2010). For example, the widely used isochrone placement method is applicable for estimating ages of a large range of stars, but it usually suffers from relatively large uncertainty (~50% typically) for main-sequence stars, which are the bulk of exoplanet hosts (e.g., Berger et al. 2020b). Asteroseismology is significantly better than any other age-dating method, and it can deliver age estimates for individual stars with uncertainties of ~10–20% (e.g., Gai et al. 2011; Chaplin et al. 2014). However, this method requires observations with sufficiently accurate, high-cadence photometric measurements and it is only applicable to stars in a limited range of spectral types that exhibit prominent oscillations. In addition, the carbon and nitrogen abundances have been suggested to be age indicators, but this is usually applicable to giant stars, and the reported age has achieved a precision of ~20–30% (Martig et al. 2016; Ness et al. 2016; Ho et al. 2017; Wu et al. 2018).

Stellar ages can also be estimated statistically from some empirical relationships. It has been known for decades that older stars have larger velocity dispersions, the so-called age–velocity dispersion relation (AVR; Strömgren 1946; Parenago 1950; Wielen 1977; Holmberg et al. 2009). To derive age from AVR, one generally just needs the stellar kinematics, and therefore the age is also called the kinematic age. Unlike the methods mentioned above, the kinematic method is only applicable to ensembles of stars (not individuals). Nevertheless, kinematic age is still meaningful from a statistical view given the fast rise of the exoplanet population. Furthermore, the strength of kinematic age is that it uses only the 3D space motions (i.e., astrometry and radial velocities) without involving stellar evolutionary models, and therefore it can apply to stars with a large range of parameters (Soderblom 2010). In recent years, the kinematic method has made possible a major opportunity to determine stellar ages, thanks to the high-quality astrometry and radial-velocity observations for millions of stars (including thousands of exoplanet hosts) provided mainly by Gaia and LAMOST.

In this paper, we revisit the methods to characterize stellar kinematic properties and apply them to over 2000 exoplanet host stars based on their astrometry and radial velocities provided mainly by Gaia and LAMOST. Specifically, in Section 2, we revisit the kinematic method to identify Galactic components (e.g., thin/thick disk). In Section 3, we revise the AVR to derive kinematic ages. Applying the revised kinematic method and AVR, in Section 4, we present a catalog of kinematic properties for 2174 planet host stars and conduct some analyses. In Section 5, we discuss our results and some future prospects. In Section 6, we provide some importing guidelines to the kinematic methods and planet host catalog and describe cautions and limitations to it. Finally, we summarize in Section 7.

2. Revisiting the Kinematic Method to Classify the Galactic Components

In this section, we revisit the kinematic method to classify stars into different Galactic components (e.g., thin/thick disk). The key is revising the characteristic kinematic parameters (Section 2.3), an extension from the solar neighborhood (~100 pc) to ~1500 pc, in order to cover most planet hosts, as shown in Figure 1.

2.1. Space Velocities and Galactic Orbits

We calculated the 3D Galactocentric cylindrical coordinates (R, θ, Z) by adopting a location of the Sun of R⊙ = 8.34 kpc (Reid et al. 2014) and Z⊙ = 27 pc (Chen et al. 2001). The Galactic rectangular velocities relative to the Sun (U, V, W) and their errors were calculated by the right-handed coordinate system based on the formulae and matrix equations presented in Johnson & Soderblom (1987). Here, U is positive when pointing to the direction of the Galactic center, V is positive along the direction of the Sun orbiting around the Galactic center, and W is positive when pointing toward the North Galactic Pole. Cylindrical velocities VR, Vθ, and VZ are defined as positive with increasing R, θ, and Z, with the latter toward the North Galactic Pole. To obtain the Galactic rectangular velocities relative to the local standard of rest (LSR); (ULSR, VLSR, WLSR), we adopted the solar peculiar motion [ULSR, VLSR, WLSR] = [9.58, 10.52, 7.01] km s−1 (Tian et al. 2015).

2.2. Classification of Galactic Components

We adopted the widely used kinematic approach as in Bensby et al. (2003, 2014) to classify the stars in our sample into different Galactic components, e.g., thin and thick disk stars. This method assumes that the Galactic velocities (ULSR, VLSR, WLSR) in different components (the thin disk, the thick disk, the halo, and the Hercules stream) follow a
multidimensional Gaussian distribution as
\[
    f(U, V, W) = k \times \exp \left( -\frac{(U_{\text{LSR}} - U_{\text{asym}})^2}{2\sigma_U^2} - \frac{(V_{\text{LSR}} - V_{\text{asym}})^2}{2\sigma_V^2} - \frac{W_{\text{LSR}}^2}{2\sigma_W^2} \right),
\]
where the normalization coefficient
\[
k = \frac{1}{(2\pi)^{3/2}\sigma_U\sigma_V\sigma_W}.
\]
Here, \(\sigma_U\), \(\sigma_V\), and \(\sigma_W\) are the characteristic velocity dispersions, and \(V_{\text{asym}}\) and \(U_{\text{asym}}\) are the asymmetric drifts.

For \(V_{\text{asym}}\), following Binney & Tremaine (2008) we adopted
\[
    V_{\text{asym}} = \bar{V}_\theta - V_c,
\]
where \(V_c\) is the circular speed of the LSR, 238 km s\(^{-1}\) (Schönrich 2012), and \(\bar{V}_\theta\) is the mean value of azimuthal velocities for a given component. For \(U_{\text{asym}}\), we adopted \(U_{\text{asym}} = 0\) for the disk and halo components and \(U_{\text{asym}} = -40\) km s\(^{-1}\) for the Hercules stream component (Binney & Tremaine 2008; Bensby et al. 2014).

The relative probabilities between two different components, i.e., the thick disk to thin disk (TD/D), thick disk to halo (TD/H), the Hercules to thin disk (Herc/D), and the Hercules to thick disk (Herc/TD) can be calculated as
\[
    \frac{TD}{D} = \frac{X_{\text{TD}}}{X_{\text{D}}} \cdot \frac{f_{\text{TD}}}{f_{\text{D}}}, \quad \frac{TD}{H} = \frac{X_{\text{TD}}}{X_{\text{H}}} \cdot \frac{f_{\text{TD}}}{f_{\text{H}}}, \quad \frac{\text{Herc}}{TD} = \frac{X_{\text{Herc}}}{X_{\text{TD}}} \cdot \frac{f_{\text{Herc}}}{f_{\text{TD}}},
\]
where \(X\) is the fraction of stars for a given component.

Then, for stars in our planet host sample, we calculated their above probabilities, and classified them into different Galactic components by adopting the same criteria as in Bensby et al. (2014), which are (1) thin disk: \(TD/D < 0.5\) and \(\text{Herc}/D < 0.5\), (2) thick disk: \(TD/D > 2\) and \(TD/H > 1\) and \(\text{Herc}/TD < 0.5\), (3) halo: \(TD/D > 2\) and \(TD/H < 1\) and \(\text{Herc}/TD < 0.5\), and (4) Hercules: \(\text{Herc}/D > 1\) and \(\text{Herc}/TD > 1\).

### 2.3. Revision of Characteristic Kinematic Parameters

One of the important steps during the above classification procedure is to obtain the characteristic kinematic parameters for each Galactic component, i.e., \(\sigma_U\), \(\sigma_V\), \(\sigma_W\), \(U_{\text{asym}}\), \(V_{\text{asym}}\), and \(X\). In the solar neighborhood within ~100 pc from the Sun, these parameters are available from Bensby et al. (2014) and listed in Table 1 here. However, as shown in Figure 1, the stars in our sample are located in a much wider zone (up to several kpc from the Sun). It has been found that the velocity ellipsoids change with the Galactic position (Williams et al. 2013). Therefore, the values of these characteristic kinematic parameters for each component should be revised and extended (Section 2.3) so that they are applicable for stars in a larger range of \(Z\) (e.g., \(|Z| < 1.5\) kpc) and \(R(7.5 < R < 10.0\) kpc).

### 2.3.1. Calibration Sample

To revise the values of characteristic kinematic parameters, we relied on a calibration sample based on the LAMOST and Gaia data. The LAMOST main-sequence turnoff and subgiant (MSTO-SG) star sample of Xiang et al. (2017a) provides the estimates of stellar age, mass, and radial velocity (RV) for 0.93 million Galactic disk stars from the LAMOST Galactic spectroscopic surveys. The typical uncertainty in age is 34%.

To construct the calibration sample, we first cross-matched the above LAMOST MSTO-SG catalog with the Gaia DR2 catalog. This was done by using the X-match service of CDS. We set a critical distance of 1.25\(\arcsec\) for position match. We carried out a magnitude cut, which was set as the G magnitude difference less than 2.3, to ensure our cross-matches were of similar brightness. The G magnitudes for the LAMOST MSTO-SG stars were calculated by using the XSTPS – GAC g, r, and i color–color polynomial fits in Table 7 of Jordi et al. (2010). For stars with multiple matches, we kept those with the smallest angular separations. After the above cross-match, we had 863,663 stars left.

We then applied the following filters to further clean the calibration sample.

1. Binary filter. We removed binary star systems because their kinematics contain additional motions (Dehnen & Binney 1998). This was done by choosing stars flagged as ‘Normal star’ (i.e., single and with spectral type of AFGKM) in the LAMOST MSTO-SG catalog (Xiang et al. 2017a).
2. Parallax precision filter. Following Dehnen & Binney (1998), we removed stars with relative parallax errors larger than 10% as reported in the Gaia DR2.
3. Age precision filter. We removed stars with ages older than 14 Gyr or with age errors larger than 25% as well as blue straggler stars (\(|Z| > 1.5\) kpc and age younger than 2 Gyr) in the LAMOST MSTO-SG catalog.
4. Distance filter (similar to Binney et al. 1997). The majority of the remaining stars are brighter than G mag = 16 where the parallax has a median error of 0.0649 mas. Recalling the above 10% parallax precision errors with a weight of \(\frac{1}{\text{error}}\) for feature extraction in the classification procedure.

### Table 1

Characteristics for Stellar Components in the Solar Neighborhood from Bensby et al. (2014)

| Component   | \(\sigma_U\) (km s\(^{-1}\)) | \(\sigma_V\) (km s\(^{-1}\)) | \(\sigma_W\) (km s\(^{-1}\)) | \(U_{\text{asym}}\) (km s\(^{-1}\)) | \(V_{\text{asym}}\) (km s\(^{-1}\)) | \(X\) |
|-------------|----------------------------|----------------------------|----------------------------|--------------------------------|--------------------------------|-----|
| Thin disk   | 35                         | 20                         | 16                         | 0                             | -15                            | 0.85|
| Thick disk  | 67                         | 38                         | 35                         | 0                             | -46                            | 0.09|
| Halo        | 160                        | 90                         | 90                         | 0                             | -220                           | 0.0015|
| Hercules    | 26                         | 9                          | 17                         | -40                           | -50                            | 0.06|

Note.
- \(\sigma_U\), \(\sigma_V\), and \(\sigma_W\) are the velocity dispersions for the different components; \(U_{\text{LSR}}\) and \(V_{\text{LSR}}\) are the asymmetric drifts in \(U\) and \(V\) relative to the LSR; and \(X\) is the normalization fractions for each component in the Solar neighborhood (in the Galactic plane). Values are taken from for the thin disk, thick disk, the stellar halo, and the Hercules stream (Bensby et al. 2007, 2014).
Figure 2. Top panel: Galactocentric radius ($R$) vs. height ($Z$) for the calibration sample. The position (8.34 kpc, 0.027 kpc) marks the location of the Sun. Bottom panel: Galactocentric radius ($R$) vs. angle ($\theta$) for the calibration sample, with (8.34 kpc, 0°) marking the position of the Sun.

requirement, it leads to a distance limit of $\sim 1/(0.0649/0.1) = 1.54$ kpc. We therefore removed stars with distances larger than this limit. Such a cut at 1.54 kpc also makes the distance distribution of the calibration sample closer to that of the planet host sample (Figure 13).

After applying the above filters, we were left with 130,403 stars. Figure 2 shows the location of stars in this calibration sample. As can be seen, these stars are mainly located at $7.5 < R < 10.0$ kpc and $|Z| < 1.5$ kpc, a region large enough to cover most known planet host stars (Figure 13).

Although most planet hosts are main-sequence stars while stars in the calibration sample are main-sequence turnoff stars and subgiants, it should not affect the calibration of kinematic properties. In fact, it has been shown that the velocity ellipsoid and AVR of main-sequence turnoff stars in the calibration sample are main-sequence turnoff stars and subgiants, it should not affect the calibration of kinematic properties. In fact, it has been shown that the velocity ellipsoid [Binney et al. 2014; Büdenbender et al. 2015; Everall et al. 2019] and the AVR (Wielen 1977; Holmberg et al. 2009; Yu & Liu 2018; Mackereth et al. 2019) are independent of stellar evolution stage and effective temperature (mass). Therefore, the revised characteristic kinematic parameters (Section 2.3) and AVR (Section 3) from the calibration sample is applicable to stars of different spectral types and evolutionary stages, including the planet host sample (Section 4).

2.3.2. Binning and Examining the Calibration Sample

In order to calculate the characteristic kinematic parameters for each Galactic component as a function of ($R$, $Z$) in the Galaxy, we binned the calibration sample as follows. For $|Z|$, we set eight bins with boundaries at $|Z| = 0, 0.1, 0.2, 0.3, 0.4, 0.55, 0.75, 1.0, and 1.5$ kpc, resulting in similar sizes ($\sim 20,000$ stars) for all of the bins except for the last two, whose sizes are $\sim 10,000$. For $R$, we set five bins with boundaries at $R = 7.5, 8.0, 8.5, 9.0, 9.5$, and 10 kpc. In total, there are $5 \times 8 = 40$ grids in the $R$–$Z$ plane. Two of the grids ($R$: 9.5–10.0 kpc, $|Z|$: 0.75–1.0 & 1.0–1.5 kpc) have too few stars (<400) and therefore are not considered hereafter.

Following Binney et al. (2000), we examined the kinematically isotropic homogeneity in each bin of the calibration sample. The kinematical homogeneity requires that the dispersion in proper motions, $S$, follows

$$S = \left( \frac{2}{3} \sigma_{\text{tot}}^2 \right)^{1/2} = \left[ \frac{2}{3} \left( \sigma_R^2 + \sigma_\theta^2 + \sigma_Z^2 \right) \right]^{1/2}. \tag{6}$$

The result of this examination is shown in Figure 3. As can be seen, the calibration sample, either as a whole or as divided into various grids, generally obeys the above relation, indicating the sample is kinematically unbiased.

2.3.3. Classifying the Calibration Sample

Next, we classified stars in the calibration sample into different Galactic components. Following Bonaca et al. (2017), we identified halo stars if $V_{\text{tot}} = (U_{\text{LSR}}^2 + V_{\text{LSR}}^2 + W_{\text{LSR}}^2)^{1/2} > 220$ km s$^{-1}$. Stars with $V_{\text{LSR}} \approx -50 \pm 9$ km s$^{-1}$ and $(U_{\text{LSR}}^2 + W_{\text{LSR}}^2)^{1/2} \approx 50 – 70$ km s$^{-1}$ were selected as the Hercules stream (Famaey et al. 2005; Bensby et al. 2007, 2014). We adopted the age-defined thin and thick disk components with a boundary at 8 Gyr (Fuhrmann 1998; Hou et al. 2013) to classify the rest of the sample into thin and thick disk stars. In Figure 4, we plot the Toomre diagram for the calibration sample. As expected, most stars with low velocities ($V_{\text{tot}} \lesssim 50$ km s$^{-1}$) are in the thin disk, while those with moderate velocities ($V_{\text{tot}} \approx 70–180$ km s$^{-1}$) are mainly in the thick disk (e.g., Feltzing et al. 2003; Adibekyan et al. 2013; Bensby et al. 2014).

2.3.4. Revising the Velocity Ellipsoid

We then revised the velocity ellipsoid of each Galactic component, namely calculating $\sigma_R$, $\sigma_\theta$, $\sigma_Z$, and $V_{\text{sym}}$ of each grid in the $R$–$Z$ plane. Here, we revised those values only for the thin and thick disk components. For the halo and Hercules components, their numbers in each grid were too low for the revision, therefore we adopted the velocity ellipsoid values as...
Dashed lines show constant values of the total Galactic velocity components. The diagram is color coded to represent different components. 

We therefore fit \( \sigma_U, \sigma_V, \sigma_W \) in this formula. To obtain the uncertainty of each fitting parameter, we assumed that the Galactic velocity follows the Gaussian distribution \( N(V, \text{err}_V) \), where \( \text{err}_V \) represents the corresponding uncertainty. Then, we resampled Galactic velocities based on these Gaussian distributions. After that, we refit the resampled data in this formula. To obtain the uncertainties (one-sigma interval) of the fitting parameters were set as the range of 50 ± 34.1 percentiles of these 1000 sets of

\[
\sigma = b_1 + b_2 \times \frac{R}{\text{kpc}} + b_3 \times \left(\frac{Z}{\text{kpc}}\right)^2 \text{km s}^{-1}.
\] (7)
best fits. Figure 5 shows the velocity dispersions \(\sigma_U, \sigma_V, \sigma_W\) as a function of \(Z\) in each \(R\) bin. The best fits are overplotted as the black solid curves. The values of fitting parameters and their one-sigma uncertainties are summarized in Table 3. As expected, velocity dispersions generally increase with \(Z\) for both thin and thick disks in all of the \(R\) bins.

For the asymmetric velocity, according to Robin et al. (2003) and Binney & Tremaine (2008), it generally follows the relation

\[
V_{\text{asym}} = -\sigma_U^2 \frac{\partial \ln \rho}{2V_{\text{LSR}}} + \sigma_V \frac{\partial \ln \sigma_U}{\partial \ln R} + \left(1 - \frac{\sigma_V}{\sigma_U}\right) \left(1 - \frac{\sigma_W}{\sigma_U}\right).
\]

Therefore, we used the following formula to calculate \(V_{\text{asym}}\), i.e.,

\[
V_{\text{asym}} = \frac{\sigma_U^2}{C_0}.
\]

Figure 6 shows \(V_{\text{asym}}\) as a function of \(\sigma_U^2\). The best fits are overplotted as the black solid lines. The values of \(C_0\) are \(-88.5^{+1.7}_{-1.9}\) km s\(^{-1}\) and \(-92.5^{+2.3}_{-2.1}\) km s\(^{-1}\) for the thin and thick disks, respectively, which are generally consistent with the theoretical estimate \((-82 \pm 6)\) km s\(^{-1}\); Binney & Tremaine (2008).

2.3.5. Revising the X Factor

As defined in Equation (5), \(X\) is the fraction of stars for a given component. For the halo and Hercules stream, their number density distributions and structures are not quite clear yet. Therefore, we set their \(X\) as the observed fractions, i.e.,

\[
X_H = \frac{N_H}{N_{\text{tot}}}, \quad X_{\text{Herc}} = \frac{N_{\text{Herc}}}{N_{\text{tot}}},
\]

\[
(9)
\]

Figure 5. The velocity dispersions as functions of position \((R, Z)\) in the Galaxy for the calibration sample. The black line in each panel denotes the result of the best fit of Equation (7) using the coefficients in Table 3.

Figure 6. The asymmetric velocity, \(V_{\text{asym}}\) as a function of \(\sigma_U^2\) for the thin disk (left panel) and thick disk (right panel). The black lines denote the results of the best fit using Equation (8).

Table 3
Fitting Parameters of the Velocity Dispersion as Functions of \((R, Z)\), i.e., Equation (7)

| \(b_1\) | \(b_2\) | \(b_3\) |
|-------|-------|-------|
| \(\sigma_U^2\) | \(63.4^{+1.3}_{-1.2}\) | \(-3.2^{+0.5}_{-0.3}\) | \(7.6^{+0.7}_{-0.5}\) |
| \(\sigma_V^2\) | \(41.6^{+2.3}_{-2.1}\) | \(-2.3^{+0.4}_{-0.2}\) | \(5.6^{+0.3}_{-0.2}\) |
| \(\sigma_W^2\) | \(27.3^{+3.4}_{-3.1}\) | \(-1.2^{+0.3}_{-0.2}\) | \(5.0^{+0.2}_{-0.1}\) |
| \(\sigma_U^{10}\) | \(58.4^{+4.0}_{-3.7}\) | \(0.6^{+0.1}_{-0.0}\) | \(4.1^{+0.1}_{-0.1}\) |
| \(\sigma_V^{10}\) | \(44.9^{+3.2}_{-3.0}\) | \(-0.7^{+0.1}_{-0.0}\) | \(5.2^{+0.2}_{-0.1}\) |
| \(\sigma_W^{10}\) | \(55.8^{+3.8}_{-3.1}\) | \(-2.2^{+0.2}_{-0.1}\) | \(6.1^{+0.3}_{-0.2}\) |
where \(N_{\text{H}}\), \(N_{\text{Herc}}\) and \(N_{\text{tot}}\) are the numbers of Hercules stream stars, halo stars, and total stars in each \((R-Z)\) grid.

For the thin and thick disks, the number density is modeled as the following formula (Chen et al. 2001; Binney & Tremaine 2008):

\[
n(R, Z) = n^0 \times \exp \left( -\frac{R - R_s}{h_R} \right) \exp \left( -\frac{|Z|}{h_Z} \right),
\]

where \(h_Z\) and \(h_R\) are the scale height and scale length of the disk, respectively. Here, we took \((h_R, h_Z)\) as \((3.4, 0.3)\) kpc for the thin disk and \((1.8, 1.0)\) kpc for the thick disk (Binney & Tremaine 2008; Cheng et al. 2012; Bovy et al. 2016). Then, the ratio of thick/thin disk star numbers in each \(R-Z\) grid can be calculated as

\[
X_{\text{TD/D}} = \frac{X_{\text{TD}}}{X_{\text{D}}} = \frac{\int_{R}^{\infty} \int_{Z}^{\infty} n_{\text{TD}}(R, Z) 2\pi R dR dZ}{\int_{R}^{\infty} \int_{Z}^{\infty} n_{\text{D}}(R, Z) 2\pi R dR dZ}.
\]

The results of these revised \(X\) values in all the \(R-Z\) grids are tabulated in Table 2. Figure 7 shows the \(X\) values of various Galactic components as functions of Galactic radius \(R\) and absolute value of height, \(|Z|\). As expected, \(X_{\text{D}}\) (\(X_{\text{TD}}\)) generally decrease (increase) with \(|Z|\) in all of the \(R\) bins.

![Figure 7](https://example.com/figure7.png)

**Figure 7.** The normalization fraction \(X\) of stars for each component as a function of Galactic radius \(R\) and absolute value of height \(|Z|\). The different colors denote subsamples of stars with different Galactic radii.

So far, we have calculated the characteristic parameters (i.e., \(\sigma_V\), \(\sigma_{\phi}\), \(U_{\text{sym}}\), and \(V_{\text{sym}}\)) as functions of \(R\) and \(|Z|\) (Table 2). In Table 4, we then compare our results in the solar neighborhood \(((R - R_s)^2 + Z^2)^{1/2} = 100\) pc, Bensby et al. 2014) to those of Bensby et al. (2014). As can be

\[
X_{\text{D}} = (1 - X_{\text{H}} - X_{\text{Herc}}) \times \frac{1}{1 + X_{\text{TD/D}}},
\]

\[
X_{\text{TD}} = (1 - X_{\text{H}} - X_{\text{Herc}}) \times \frac{X_{\text{TD/D}}}{1 + X_{\text{TD/D}}},
\]

The Astrophysical Journal, 909:115 (21pp), 2021 March 10 Chen et al.
seen, both our results and those of Bensby et al. (2014) provide very similar values for these characteristic parameters, demonstrating that our revision (Section 2) can also be applied to the solar neighborhood.

### 3. Revisiting the Age–Velocity Dispersion Relation to Derive Kinematic Ages

When the stars in the solar neighborhood are binned by age, the velocity dispersion of each bin increases with its age. This age–velocity relation (AVR) has been known and studied for decades (Strömberg 1946; Parenago 1950; Wielen 1977; Holmberg et al. 2009). A similar relationship has also been inferred for the external Galactic disk (Aumer et al. 2016; Robin et al. 2017). Here, we revisit the AVR with the calibration sample constructed in Section 2.3.1.

#### 3.1. Fitting AVR

In our study, we divided the foregoing calibration sample (Section 2.3.1) into 30 bins with approximately equal sizes (~4350 stars in each bin) according to their ages. Then, we calculated the total velocity dispersion for each bin, i.e.,

\[
\sigma_{\text{tot}} = \sigma_U^2 + \sigma_V^2 + \sigma_W^2. \tag{13}
\]

Figure 8 shows the velocity dispersion as a function of the median age of each bin. As can be seen, all of the components of velocity dispersion ($U_{\text{LSR}}, V_{\text{LSR}}, W_{\text{LSR}}$) and the total velocity dispersion ($V_{\text{tot}}$) increase with age. Following Holmberg et al. (2009) and Aumer et al. (2016), we fit the AVRs shown in Figure 8 by using a simple power-law formula, i.e.,

\[
\sigma = k \times \left( \frac{t}{\text{Gyr}} \right)^\beta \text{km s}^{-1}, \tag{14}
\]

where $t$ is stellar age, $\sigma$ is the velocity dispersion, and $k$ and $\beta$ are two fitting parameters.

We used the Levenberg–Marquardt algorithm (LMA) to find the best fit. To obtain the uncertainty of each fitting parameter, we assumed that the Galactic velocity and age follows the Gaussian distribution $N(V, \text{err}_V)$ and $N(t, \text{err}_t)$, where err represents the corresponding uncertainty. Then, we resampled Galactic velocities and stellar ages based on these Gaussian distributions. After that, we refit the AVR by using the resampled data. We repeated the above resampling process 1000 times and obtained 1000 sets of best fits. The uncertainties (one-sigma interval) of the fitting parameters

---

**Table 4**

The Kinematic Characteristics for Stellar Components in the Solar Neighborhood from Bensby et al. (2014) and This Work

|       | $X_D$ | $X_{TD}$ | $X_H$ | $X_{H\odot}$ |
|-------|-------|----------|-------|---------------|
| Bensby et al. (2014) | 0.85  | 0.09     | 0.0015 | 0.06          |
| This work   | 0.84  | 0.10     | 0.0013 | 0.06          |

|       | $\sigma_U, D$ | $\sigma_V, D$ | $\sigma_W, D$ | $V_{\text{sym}, D}$ |
|-------|---------------|---------------|---------------|---------------------|
| (km s$^{-1}$) | (km s$^{-1}$) | (km s$^{-1}$) | (km s$^{-1}$) |
| Bensby et al. (2014) | 35  | 20  | 16  | $-$15             |
| This work   | 34  | 21  | 16  | $-$14             |

|       | $\sigma_U, TD$ | $\sigma_V, TD$ | $\sigma_W, TD$ | $V_{\text{sym}, TD}$ |
|-------|---------------|---------------|---------------|---------------------|
| (km s$^{-1}$) | (km s$^{-1}$) | (km s$^{-1}$) | (km s$^{-1}$) |
| Bensby et al. (2014) | 67  | 38  | 35  | $-$46             |
| This work   | 65  | 39  | 35  | $-$44             |

---

Figure 8. The velocity dispersions for $U_{\text{LSR}}, V_{\text{LSR}}, W_{\text{LSR}}$ and $V_{\text{tot}}$ vs. age for the selected calibration star sample. The 30 bins have approximately equal numbers of stars (~4350 in each). The solid black lines denote the respective best fit of the AVR (Equation (14)) using the coefficients in Table 5.
were set as the range of 50 ± 34.1 percentiles of these 1000 sets of best fits.

The values of the fitting parameters on the AVR are summarized in Table 5. We obtained β = 0.34±0.02, 0.43±0.02, 0.54±0.02, and 0.40±0.02, for $U_{LSR}$, $V_{LSR}$, $W_{LSR}$, and $V_{int}$, respectively, which are consistent with the values derived from previous studies (Holmberg et al. 2009; Aumer & Binney 2009; see Section 3.5 for detailed comparisons).

### 3.2. AVRs of Different Galactic Components

Our calibration sample consists of 98,486 (75.51%) thin disk stars, 23,572 (18.08%) thick disk stars, 8152 (6.25%) Hercules stars, and 193 (0.15%) halo stars. To explore the differences between AVRs of different Galactic components, for each component, we divided stars into bins of approximately equal size according to their ages. Due to sample size, the bin numbers are set as 20, 10, 10, and 5 for the thin disk, thick disk, Hercules stream, and halo, respectively. For each component, we performed the same method as in Section 3.1 to obtain the AVR.

Figure 9 displays the AVRs of different components. As can be seen, the AVRs obtained from the thin and thick disk components fit well to the power-law AVR derived from the whole calibration sample. For the Hercules stream, the dispersion of velocity component $\sigma_{p}$ generally follows the power-law AVR, but other components, i.e., $\sigma_{t}$ and $\sigma_{p}$, seem to be independent of age. For the halo, the velocity dispersions are much larger than those predicted by the power-law AVR, and there is no clear trend between velocity dispersions and ages. Therefore, we conclude that the power-law AVR can only apply to the thin/thick disk components.

### 3.3. Radial Variation of AVR

To explore how the AVR varies with Galactic radius, we divided the foregoing calibration sample (Section 2.3.1) into five subsamples according to their Galactic radii, i.e., $R = 7.5–8.0$ kpc, $R = 8.0–8.5$ kpc, $R = 8.5–9.0$ kpc, $R = 9.0–9.5$ kpc, and $R = 9.5–10.0$ kpc. For each subsample, we performed the same method as in Section 3.1 to fit the AVR. Figure 10 shows the fitted AVRs for the five subsamples. The fitting parameters of the five AVRs are summarized in Table 5. As can be seen, the AVRs generally decrease with increasing $R$, which is caused by the decrease of the velocity dispersion with $R$ (Equation (7)). Therefore, the typical uncertainties in $k$ and $\beta$ of AVRs obtained from the whole sample are generally larger than in the five subsamples due to the radial variation. However, the values of $k$ and $\beta$ differ mildly with $R$ (mean value: ~5% for $k$ and 4% for $\beta$). As can be seen in Figure 10, the AVRs for the five subsamples are generally within the 1σ range of that for the whole sample (gray regions). This is consistent with the result derived from simulations in Aumer et al. (2016), which showed that the shape of the AVR is almost independent of $R$ when $R \gtrsim 6$ kpc.

### 3.4. Vertical Selection Effect of AVR

Selecting stars from a limited vertical volume introduces phase correlations between stars that influence the values of velocity dispersions at the time of selection and before (Aumer et al. 2016). For example, when stars are selected close to $Z = 0$, they are all close to their maxima in $W_{LSR}$. Tracking them back in time, their vertical velocity dispersions therefore have to be lower than at the time of selection (Aumer & Schönrich 2015). To explore the vertical selection effect in our sample, following the method in Aumer et al. (2016), we compared the AVRs for stars with $|Z| < 100$ pc and for all stars irrespective of the $Z$ position. The result is displayed in Figure 11. As can be seen, there is no significant difference between the AVR for $|Z| < 100$ pc (red) and that for all stars (gray region). We conducted a KS test between the velocity dispersions of stars with $|Z| < 100$ pc and those of all stars. The $p$-values are 0.86, 0.97, 0.99, and 0.94 for $U_{LSR}$, $V_{LSR}$, $W_{LSR}$, and $V_{int}$, respectively. Such high $p$-values

---

**Table 5**

Fitting Parameters of the AVR (Equation (14))

| $U$   | $V$   | $W$   | $V_{int}$ |
|-------|-------|-------|-----------|
| $k$   | $\beta$ | $k$   | $\beta$  | $k$   | $\beta$  |
|       |        |       |           |
| Whole sample | 23.66±0.66 | 0.34±0.02 | 12.49±0.49 | 0.43±0.02 | 8.50±0.47 | 0.54±0.02 | 27.55±0.82 | 0.40±0.02 |
| $R: 7.5–8.0$ kpc | 24.16±0.17 | 0.33±0.01 | 14.32±0.10 | 0.39±0.01 | 8.74±0.08 | 0.57±0.01 | 30.90±0.16 | 0.36±0.01 |
| $R: 8.0–8.5$ kpc | 25.31±0.08 | 0.33±0.01 | 14.18±0.06 | 0.39±0.01 | 8.84±0.06 | 0.54±0.02 | 29.94±0.10 | 0.38±0.01 |
| $R: 8.5–9.0$ kpc | 23.40±0.13 | 0.34±0.01 | 11.84±0.03 | 0.46±0.02 | 8.30±0.05 | 0.55±0.01 | 27.02±0.11 | 0.40±0.02 |
| $R: 9.0–9.5$ kpc | 22.28±0.09 | 0.35±0.01 | 12.27±0.05 | 0.42±0.02 | 8.32±0.05 | 0.54±0.02 | 26.42±0.10 | 0.40±0.02 |
| $R: 9.5–10.0$ kpc | 22.73±0.25 | 0.32±0.01 | 12.66±0.15 | 0.44±0.04 | 8.27±0.08 | 0.53±0.04 | 27.74±0.21 | 0.38±0.02 |

**Notes.** The unit of $k$ is km s$^{-1}$.

* Here the parameters and their uncertainties were obtained by fitting the data of Holmberg et al. (2009) with our methods shown in Section 3.1.

---

Chen et al.
demonstrate the vertical selection effect has little influence on AVR.

3.5. Comparison to Previous Works

Here, we compare our fitted AVR (Section 3.1) to those AVRs fitted with different samples in previous studies. The result is plotted in Figure 12. As can be seen, our result is in good agreement with those of Holmberg et al. (2009), Yu & Liu (2018), and Mackereth et al. (2019). Nevertheless, we note that our result differs significantly from those of Binney et al. (1997), Bovy et al. (2012b), and Sharma et al. (2014) on the younger and older ends, respectively.

All these studies fit their AVRs with different stellar samples. Gomez et al. (1997) used 2812 stars from the Hipparcos data. Holmberg et al. (2009) used 2640 FG main-sequence stars from the Geneva–Copenhagen Survey (GCS) data. Bovy et al. (2012b) used 3365 stars from the APOGEE data. Sharma et al. (2014) used 5201 stars from the GCS and RAVE data. Yu & Liu (2018) used 3564 subgiant and red giant stars from the LAMOST–Gaia data. Mackereth et al. (2019) used 65,719 stars from the APOGEE and Gaia data.
Figure 12. The vertical AVRs from different studies. The dashed lines present results from previous studies. The gray region denotes the $1\sigma$ range of the best fit obtained from our calibration sample.

Our study used 130,403 stars, which were selected from the LAMOST DR4 value-added catalog with well-constrained kinematics and ages (Section 2.3.1). Due to the increase in sample size, as well as in the quality of stellar characteizations, the uncertainties of the AVR parameters have been largely reduced in this work. To better demonstrate this point, we performed a more detailed comparison to the AVR of Holmberg et al. (2009), which are widely used and very close to our best fit of the AVR (Figure 12). We adopted the same procedures (Section 3.1) to obtain the uncertainties of AVR but used the data of Holmberg et al. (2009). The results are summarized at the bottom part of Table 6. As can be seen, for AVR from Holmberg et al. (2009), see the bottom part of Table 5, the typical uncertainties of age from previous studies. The gray region denotes the $1\sigma$ range of the best fit obtained from our calibration sample.

3.6. Kinematic Age and Uncertainty

For a group of stars, the typical kinematic age can be derived by using the AVR (solving Equation (14)), which gives

$$ t = \left( \frac{\sigma}{k \text{ km s}^{-1}} \right)^2 \text{ Gyr}, $$

(15)

By means of error propagation, the relative uncertainty of kinematic age can be estimated as

$$ \frac{\Delta t}{t} = \sqrt{\left( \frac{\partial \ln t}{\partial \beta} \frac{\Delta \beta}{\beta} \right)^2 + \left( \frac{\partial \ln t}{\partial k} \frac{\Delta k}{k} \right)^2 + \left( \frac{\partial \ln t}{\partial \sigma} \frac{\Delta \sigma}{\sigma} \right)^2} $$

= \sqrt{\left( \frac{t}{\text{Gyr}} \right)^2 \left( \frac{\Delta \beta}{\beta} \right)^2 + \left( \frac{\Delta k}{k} \right)^2 + \left( \frac{\Delta \sigma}{\sigma} \right)^2},

(16)

where $\Delta$ represents the absolute uncertainty. For the sake of simplicity, here we assumed $k$, $\beta$, and $\sigma$ are independent of each other. In other words, we neglect the covariances between them. It can be seen from the formula that the first term (due to the uncertainty of $\beta$) has a positive correlation with the age ($t$), while the latter two terms are independent of age. Therefore, the relative uncertainty in age generally increases with age itself. For this reason, we set $t = 1$ Gyr and $t = 10$ Gyr to estimate the range of the relative uncertainty of age.

Based on Equation (16), we analyzed the budget of the kinematic age uncertainty derived from AVR. The results are listed in the left part of Table 6. The uncertainties in AVR fitting parameters, i.e., $\Delta k$, $\Delta \beta$, and $\Delta \sigma$, were calculated as the half width of the one-sigma interval of $k$, $\beta$, and $\sigma$ as listed in Table 5. The relative uncertainties in velocity dispersions ($\Delta \sigma/\sigma$) were adopted as the median relative uncertainties of velocity dispersions in our planet host stellar sample, which are 2.0%, 3.5%, and 2.4% for $\sigma_U$, $\sigma_V$, $\sigma_W$, and $\sigma_{tot}$, respectively.

Putting all of the above uncertainties into Equation (16), we obtained the relative age uncertainties $\Delta t/t \sim 9.7 - 14.1\%$, $\Delta t/t \sim 11.8 - 20.1\%$, $\Delta t/t \sim 12.3 - 16.9\%$, and $\Delta t/t \sim 9.2 - 16.0\%$ for $\sigma_U$, $\sigma_V$, $\sigma_W$, and $\sigma_{tot}$, respectively. The lower and upper values were calculated by assuming $t = 1$ Gyr and $t = 10$ Gyr, respectively, in Equation (16).

For comparison, we repeated the above budget calculation using the AVR of (Holmberg et al. 2009), see the bottom part of Table 5. The results are listed in the right part of Table 6. As can be seen, for AVR from Holmberg et al. (2009), the uncertainties of $k$ and $\beta$ are much larger and therefore dominant, leading to a much larger (by a factor of $\sim 3$) uncertainty in the derived kinematic age.

As discussed in Section 3.2, AVRs are not significant for the halo and Hercules stream, therefore, we suggest that this method to obtain kinematic age is only suitable for stars belonging to the Galactic disk components.

4. Application to Planet Host Stars

In this section, we apply the above revised kinematic method (Section 2) and AVR (Section 3) to a sample of planet host stars (Section 4.1), providing a catalog of their kinematic properties (Section 4.2), with a focus on the Galactic components (Section 4.3) and kinematic ages (Section 4.4).

4.1. Data Samples

This subsection describes how we constructed the planet host sample for further kinematic characterization.
4.1.1. Initializing Planet Host Sample from EA

We initialized our planet host sample with the confirmed planets table and the Kepler DR25 catalog from EA. The Kepler catalog contains 8054 Kepler Objects of Interest (KOIs) in DR25. Here, we excluded KOIs flagged by “False Positive” (FAP, Thompson et al. 2018), leaving 4034 planets (candidates) around 3069 stars. Besides Kepler, there are 1728 non-Kepler planets flagged by “Confirmed” and 1387 stars. We also removed potential binaries because additional motions caused by binary orbits could affect the results of kinematic characterization. Specifically, for Kepler planet host stars, we eliminated stars with Gaia DR2 renormalized unit-weight error (RUWE) > 1.2 (Rizzuto et al. 2018; Berger et al. 2020a). For non-Kepler planet host stars, we excluded those with pl_cflag = 0 (which indicates whether the planet orbits a binary, 0 for no) in EA. In total, we were left with 4126 stars hosting 5331 planet candidates in our initial sample.

4.1.2. Obtaining Five Astrometric Parameters from Gaia

Next, we cross-matched our initial planet host sample with Gaia to obtain astrometric parameters. The second Gaia data release (DR2, Gaia Collaboration et al. 2018a) includes five astrometric parameters: positions on the sky (α, δ), parallaxes, and proper motions (μx, μy) for more than 1.3 billion stars, with a limiting magnitude of G = 21 and a bright limit of G ≈ 3. The cross-matching was done by using the X-match service of the Centre de Données astronomiques de Strasbourg (CDS). The separation limit of the cross-matching was chosen as where the distribution of separations displayed a minimum, ∼1.5″. Besides the separation condition, we also made a magnitude cut to ensure that the matched stars are of similar brightness. The magnitude limit was set by inspecting the distribution of magnitude differences, which is 2 mag in Gaia G mag. If multiple matches satisfied these two criteria, we kept the one with the smallest angular separation. Finally, we obtained 5069 planets around 3912 stars.

4.1.3. Obtaining RV from Various Sources

We obtained radial velocities from the following five catalogs: the APOGEE DR16 catalog, the LAMOST DR4 value-added catalog, the RAVE DR5 catalog, Gaia, and EA.

1. APOGEE: The APOGEE DR16 provides a catalog of 437,485 unique stars, which contains the radial velocity (RV), effective temperature (Teff), surface gravity (log g), and chemical abundance (e.g., [Fe/H] and [α/Fe]); (Ahumada et al. 2020). We cross-matched it with the planet host sample obtained in Section 4.1.2. Here, we applied the following quality control cuts: (1) STARFLAG = 0 to only select stars with no warnings on the observation; (2) ASPCAPFLAG = 0 to ensure parameters have converged and no warnings; and (3) signal-to-noise ratio (S/N) > 80 to ensure high S/Ns (Holtzman et al. 2018), leaving 692 stars hosting 956 planets.

2. RAVE: The fifth data release (DR5) of RAVE provides radial velocities with a precision of ~1.5 km s⁻¹ and physics properties (Teff, log g, [Fe/H], etc.) from a magnitude-limited (9 < I < 12) survey for 457,588 randomly selected stars in the southern hemisphere (Kunder et al. 2017). To cross-match it with the planet host sample obtained in Section 4.1.2, we applied the following quality cuts: (1) Algo_Conv ≈ 0 to ensure that the stellar parameter pipeline has converged; (2) S/N > 40; (3) spectroscopic morphological flags (c1, c2, c3) are n; and (4) Alpha_C > −9.99 (Kunder et al. 2017). The RAVE DR5 contains stars only in a range of decl. from −88 deg to +28 deg, therefore the Kepler field that ranges from 36 deg to 53 deg in decl. is not covered. Therefore, the above cross-matching with RAVE returned only 30 stars hosting 37 non-Kepler planets.

3. Gaia: Gaia DR2 also includes radial velocities for more than 7.2 million stars with a magnitude range of G ~ 4 − 13 and a Teff range of about 3550–6900 K. The quality control cut is set such that the ratio of radial velocity and its uncertainty should be larger than 3. After cross-matching with the planet host sample obtained in Section 4.1.2, we obtained 1143 stars hosting 1479 planets (268 of them are stars with 371 planets in the Kepler field).

4. LAMOST: The LAMOST survey has several components focusing on different Galactic aspects, e.g., the Galactic halo (Deng et al. 2012), stellar clusters (Hou et al. 2013), the Galactic anticenter (LSS-GAC; Liu et al. 2014), and the Kepler fields (De Cat et al. 2015). The LAMOST DR4 value-added catalog contains parameters derived from a total of 6.5 million stellar spectra for 4.4 million unique stars (Xiang et al. 2017b). RVs, Teff, log g, and [Fe/H] have been deduced using both the official LAMOST Stellar parameter Pipeline (LASP; Wu et al. 2011) and the LAMOST Stellar Parameter Pipeline at Peking University (LSP3; Xiang et al. 2015). The typical uncertainties for RVs, Teff, log g, and [Fe/H] are 5.0 km s⁻¹, 150 K, 0.25 dex, and 0.15 dex, respectively. After applying a quality cut of S/N > 10 and cross-matching with the planet host sample obtained in Section 4.1.2, we obtained 1059 stars hosting 1421 planets. The majority (951) are stars with Kepler planets (1292).

5. EA: The NASA Exoplanet Archive (EA) also reports RVs for a portion of stars. We therefore cross-matched these stars with the planet host sample obtained in Section 4.1.2. The quality control cut is also set such that the ratio of radial velocity and its uncertainty should be larger than 3, which yields 1303 stars hosting 1737 planets. The majority (1088) are stars with non-Kepler planets (1366) from various ground-based RV and transit surveys. Note that most RV data in EA are collected from various literature works and thus are inhomogeneous.

4.1.4. Finalizing the Planet Host Sample

We finalize the planet host sample by combining various matched samples in Section 4.1.3. For stars with multiple RV measurements from different sources, we take the order of precedence as APOGEE, RAVE, Gaia, LAMOST, then EA. This generally follows the order of spectral resolution and therefore the RV uncertainty. Here, we set the EA as the lowest priority because most RVs from EA are collected from various sources, which are inhomogeneous. For the sake of reliability, we exclude stars if the differences in their RVs from different sources exceed 3 km s⁻¹. To cross-match the samples, we applied the following quality cuts: (1) S/N > 10; (2) the ratio of radial velocity and its uncertainty should be larger than 3; (3) the magnitude difference in Gaia G mag should be less than 2 mag. If multiple matches satisfied these two criteria, we kept the one with the smallest angular separation. Finally, we obtained 8054 Kepler Objects of Interest (KOIs) in total, of which 4126 stars host 5331 planet candidates in our final planet host sample.
sources are larger than three times the uncertainties. We also apply the same cut as in the calibration sample, i.e., distance $<1.54\text{kpc}$, corresponding to $7 < R < 10\text{kpc}, \theta < 10\text{deg}$, and $|Z| < 1.5\text{kpc}$. Finally, we obtain a sample of 2174 stars hosting 2872 planets. In Table 7, we summarize the composition of the sample after each step mentioned above. Figure 13 shows the location of stars in our planet host sample.

4.2. A Catalog of Planet Hosts with Kinematic Characterizations

Applying the methods described in Section 2 and Section 3 to the planet host sample (Section 4.1), we obtained a catalog (Table 8) of 2174 planet hosts with kinematic characterizations, e.g., Galactocentric velocities to the LSR ($U_{\text{LSR}}, V_{\text{LSR}},$ and $W_{\text{LSR}}$) and the relative membership probabilities between different Galactic components ($TD/D$, $TD/H$, Herc/D, and Herc/TD). For the sake of completeness, we also put in the catalog the stellar parameters that were used during the process of our kinematic characterization (e.g., parallax, proper motion, and RV) and other basic stellar parameters (e.g., $T_{\text{eff}}, \log g$, [Fe/H], and [$\alpha$/Fe]). As mentioned before, for stars with multiple sources of RV, the order of precedence follows the order of spectral resolution and therefore the uncertainty, i.e.,

Figure 13. Galactocentric radius (R) vs. height (Z, top panel) and angle ($\theta$, bottom panel) for the combined planet host sample. The diagram is color coded to represent different discovery methods and facilities. The position of the Sun is marked at $(8.34\text{kpc}, 0, 0.027\text{kpc})$, as APOGEE, RAVE, Gaia, LAMOST, then EA. While for other stellar parameters, because the estimates of stellar parameters from APOGEE are only reliable for relatively cool stars, $4000 < T_{\text{eff}} < 5500 \text{K}$ (Holtzman et al. 2018), we took APOGEE as the lowest order of precedence here instead. In what follows, we conduct some analyses on this catalog.

4.3. Galactic Components of Planet Hosts

With the derived relative membership probabilities between different Galactic components ($TD/D$, $TD/H$, Herc/D, and Herc/TD in Table 8), we then classify the 2174 planet host stars into four Galactic components, i.e., thin disk, thick disk, Hercules stream, and halo following the method as mentioned in Section 2.2. For stars not belonging to the above four components, following Bensby et al. (2014), we classify them into a category dubbed “in between”.

The results of the classification are summarized in Table 9, which lists the numbers of stars in different categories. As can be seen, about 87.1% (1894/2174) of stars in our sample are in the thin disk and about 5.2% (114/2174) of stars are in the thick disk. Forty-five stars in the planet host sample are affiliated with the Hercules stream, which has been speculated to have a dynamical origin in the inner parts of the Galaxy and then kinematically heated by the central bar (Famaey et al. 2005; Bensby et al. 2007). The fraction of halo stars is $\sim0.05\%$ (1/2174) and there are another $\sim5.5\%$ (120/2174) belonging to the “in between” category. In Table 9, we also divide planet host stars according to the method that discovered the planets. In general, we find that, first, for transiting planet hosts, those observed from space-based facilities have a higher fraction of thick disk memberships (5.8%, 77/1333) than those observed from ground-based observatories (3.9%, 12/306). Second, for ground-based planet hosts, RV planet hosts have a higher
fraction of thick disk memberships (4.8%, 25/516) than transiting planet hosts (3.9%, 12/306).

We plot the Toomre diagram of the planet host stars in Figure 14. As can be seen, the boundaries of different components are well consistent with the results of previous works (e.g., Bensby et al. 2014). Specifically, most stars with low velocities ($V_{LSR} \lesssim 50 \text{ km s}^{-1}$) are in the thin disk, while those with moderate velocities ($V_{LSR} \sim 70 - 180 \text{ km s}^{-1}$) are mainly in the thick disk. The velocity of the only halo star is larger than $220 \text{ km s}^{-1}$.

We summarize the median values of velocities and chemical abundances for different component in Table 10. As expected, the halo star has the highest Galactic velocity and the poorest [Fe/H], and the thick disk stars are kinematically hotter, more metal-poor ($\sim -0.2 \text{ dex}$) and more $\alpha$-rich ($\sim 0.1 \text{ dex}$) than are the thin disk stars. The Hercules stream stars have velocities and chemical abundances that are between those of thin and thick disk stars. We have only one halo star, which does not have an $\alpha$/Fe measurement, and therefore the median and 1σ interval of $\alpha$/Fe are not provided. In Figure 15, we plot the total

| Column | Name | Format | Units | Description |
|--------|------|--------|-------|-------------|
| 1      | Gaia_ID | Long   |       | Unique Gaia source identifier |
| 2      | LAMOST_ID | string |       | LAMOST unique spectral ID |
| 3      | APOGEE_ID | string |       | APOGEE unique spectral ID |
| 4      | RAVE_ID | string |       | RAVE unique spectral ID |
| 5      | pl_hostname | string |       | NASA Exoplanet archive unique planet host name |
| 6      | Kepler_ID | integer |       | Kepler Input Catalog (KIC) ID |
| 7      | Gaia RA | Double | deg   | Barycentric R.A. (1) |
| 8      | Gaia Dec | Double | deg   | Barycentric decl. (1) |
| 9      | Gaia parallax | Double | mas    | Absolute stellar parallax |
| 10     | Gaia e_parallax | Double | mas    | Standard error of parallax (1) |
| 11     | Gaia pmra | Double | mas yr$^{-1}$ | Proper motion in R.A. direction |
| 12     | Gaia e_pmra | Double | mas yr$^{-1}$ | Standard error of proper motion in R.A. direction |
| 13     | Gaia pmdec | Double | mas yr$^{-1}$ | Proper motion in decl. direction |
| 14     | Gaia e_pmdec | Double | mas yr$^{-1}$ | Standard error of proper motion in decl. direction |
| 15     | Gaia G mag | Double | mag | Gaia G band apparent magnitude |
| 16     | $T_{eff}$ | Float | K | Effective temperature from RAVE, LAMOST, APOGEE, Gaia, EA |
| 17     | flag_ $T_{eff}$ | integer |       | flag represents which source each value is collected from |
| 18     | log $g$ | Float |       | Surface gravity from RAVE, LAMOST, APOGEE, Gaia, EA |
| 19     | flag_ log $g$ | integer |       | flag represents which source each value is collected from |
| 20     | [Fe/H] | Float | dex | Metallicity from RAVE, LAMOST, APOGEE, Gaia, EA |
| 21     | flag_ [Fe/H] | integer |       | flag represents which source each value is collected from |
| 22     | [$\alpha$/Fe] | Float | dex | $\alpha$ element abundance from RAVE, LAMOST, APOGEE, Gaia, EA |
| 23     | flag_ [$\alpha$/Fe] | integer |       | flag represents which source each value is collected from |
| 24     | $v_r$ | Double | km s$^{-1}$ | Radial velocity from APOGEE, RAVE, Gaia, LAMOST, EA |
| 25     | $e_{r,v}$ | Double | km s$^{-1}$ | Error of radial velocity |
| 26     | flag_ $e_{r,v}$ | integer |       | flag represents which source each value is collected from |
| 27     | $R$ | Double | kpc | Galactocentric cylindrical radial distance |
| 28     | $\theta$ | Double | deg | Galactocentric cylindrical azimuth angle |
| 29     | $Z$ | Double | kpc | Galactocentric cylindrical vertical height |
| 30     | $V_R$ | Double | km s$^{-1}$ | Galactocentric cylindrical $R$ velocities |
| 31     | $V_\theta$ | Double | km s$^{-1}$ | Galactocentric cylindrical $\theta$ velocities |
| 32     | $V_Z$ | Double | km s$^{-1}$ | Galactocentric cylindrical $Z$ velocities |
| 33     | $U_{LSR}$ | Double | km s$^{-1}$ | Cartesian Galactocentric X velocity to the LSR |
| 34     | $e_{U_{LSR}}$ | Double | km s$^{-1}$ | error of Cartesian Galactocentric X velocity to the LSR |
| 35     | $V_{LSR}$ | Double | km s$^{-1}$ | Cartesian Galactocentric Y velocity to the LSR |
| 36     | $e_{V_{LSR}}$ | Double | km s$^{-1}$ | error of Cartesian Galactocentric Y velocity to the LSR |
| 37     | $W_{LSR}$ | Double | km s$^{-1}$ | Cartesian Galactocentric $Z$ velocity to the LSR |
| 38     | $e_{W_{LSR}}$ | Double | km s$^{-1}$ | error of Cartesian Galactocentric $Z$ velocity to the LSR |
| 39     | $TD/D$ | Double |       | Thick disk to thin disk membership probability ratio |
| 40     | $TD/H$ | Double |       | Thick disk to halo membership probability ratio |
| 41     | Herc/D | Double |       | Hercules stream to thin disk membership probability ratio |
| 42     | Herc/TD | Double |       | Hercules stream to thick disk membership probability ratio |

Note. The flag represents which source each value is collected from: 1 for APOGEE (Ahumada et al. 2020); 2 for RAVE (Kunder et al. 2017); 3 for Gaia (Gaia Collaboration et al. 2018a, 2018b); 4 for LAMOST (Xiang et al. 2017b); 5 for NASA exoplanet archive (https://exoplanetarchive.ipac.caltech.edu/); 0 for not available.

(This table is available in its entirety in machine-readable form.)
Table 9
The Numbers of Stars (Planets) of Our Planet Host Sample in Different Galactic Components

| Component          | Total | Thin Disk | Thick Disk | Hercules | Halo | In Between |
|--------------------|-------|-----------|------------|----------|------|------------|
| Radial Velocity    |       |           |            |          |      |            |
| Transit Keplers    | 1134  | 982       | 68         | 20       | 1    | 30         |
| K2                 | 179   | 156       | 8          | 5        | 0    | 10         |
| CoRoT              | 20    | 19        | 1          | 0        | 0    | 0          |
| Ground-based       | 306   | 278       | 12         | 5        | 0    | 11         |
| Other methods      | 19    | 19        | 0          | 0        | 0    | 0          |
| All                | 2174  | 1894      | 114        | 45       | 1    | 120        |

Note. The numbers without and with the brackets are the numbers of stars and planets.

---

4.4. Kinematic Ages of Planet Hosts

In this section, by using our catalogs of kinematic properties for planet host stars, we divided planetary host stars into various groups to study their kinematic ages. For each group, we calculated the velocity dispersion to derive the corresponding kinematic age from Equation (15). To access the uncertainties of the kinematic ages, we took a Monte Carlo method by resampling the AVR parameters (k and β) and velocity dispersion (σ) based on their uncertainties. For k and β, their values and uncertainties were adopted from Table 5. For σ, its value and uncertainty were calculated by resampling each star’s Galactic velocities from a normal distribution given its value and uncertainty. Finally, the age uncertainty was set as the 50 ± 34.1 percentiles in the resampled age distribution.

4.4.1. Kinematic Ages Derived with Total Velocity and Velocity Components

As shown in Figure 8 of Section 3, the dispersions of velocity components, i.e., $\sigma_U$, $\sigma_V$, $\sigma_W$, and total velocity $\sigma_{tot}$, all increase with age and fit well with power-law functions. Nevertheless, the fitting parameters (k and β in Table 5) are different for different velocity components, which in turn could give different kinematic ages. Here, we compare different kinematic ages calculated from the total velocity and different velocity components. To do this, we first sorted the planet host sample by the relative probabilities for $TD/D$. Next, according to $TD/D$, we divided the planet host sample into 10 bins with approximately equal sizes (~217 stars). Then, for each bin, we calculated kinematic ages using $\sigma_U$, $\sigma_V$, $\sigma_W$, and $\sigma_{tot}$. Finally, we compare these kinematic ages in Figure 16. As can be seen, kinematic ages derived using different velocities are well consistent with each other. To further see the differences quantitatively, we fit the relations between age derived from $\sigma_{tot}$ and ages from $\sigma_U$, $\sigma_V$, and $\sigma_W$. The results are:

$$t_{\sigma_U} = 0.97^{+0.07}_{-0.12} \times t_{\sigma_{tot}},$$

$$t_{\sigma_V} = 0.94^{+0.07}_{-0.11} \times t_{\sigma_{tot}},$$

$$t_{\sigma_W} = 0.99^{+0.08}_{-0.10} \times t_{\sigma_{tot}}.$$  \hspace{1cm} (17)

As can be seen, the relative differences between ages from different velocities are ~5–10%. Hereafter, unless otherwise specified, the kinematic age refers to the one derived using the dispersion of total velocity, i.e., $\sigma_{tot}$.

4.4.2. Kinematic Ages of Planet Host Stars in the Galactic Disk

Besides the element abundances and Galactic velocities, age is one of the main differences between different Galactic components. It has been known that thin disk stars are
generally younger than thick disk stars, with a dividing age of \( \sim 8 \) Gyr (Fuhrmann 1998; Hou et al. 2013). Stars in the halo are very old and the age is estimated to be \( \sim 10^{12} \) Gyr (Jofré & Weiss 2011; Kalirai 2012; Guo et al. 2016, 2019). For the Hercules stream, it has three substructures. The age distribution is peaked at 4 Gyr and extend to very old age for Hercules a and Hercules b, and Hercules c has a more uniform age distribution of \( \sim 2–10 \) Gyr (Torres et al. 2019).

As mentioned in Section 3.2, there is no clear trend between velocity dispersions and ages for stars in the Hercules stream and halo. Therefore, here we only calculate the kinematic age of planet host stars in the Galactic disks, which contain the majority (97.9%) of our sample. This will be useful in future studies of the characteristics and evolution of planetary systems related to the Galactic components and ages of host stars.

We obtained the kinematic age and uncertainty for stars in the thin disk (1894 stars) and thick disk (114 stars) with the methods described at the beginning of Section 4.3. The typical ages are \( 2.84^{+0.36}_{-0.26} \) Gyr and \( 10.42^{+1.32}_{-0.72} \) for the thin and thick disk, respectively. As shown in Figure 17, the age distribution of stars in the thick disk is generally larger than 8 Gyr, while the thin disk is populated by younger stars. This is well consistent with the division age of \( \sim 8 \) Gyr observed in previous studies.

| Galactic Component | \( V_{\text{tot}} \) (km s\(^{-1}\)) | [Fe/H] (dex) | [\( \alpha/\text{Fe} \)] (dex) |
|-------------------|-----------------|---------------|-----------------|
| Thin disk         | 34.8 (19.1, 56.0) | 0.00 (−0.16, 0.21) | 0.01 (−0.19, 0.19) |
| Thick disk        | 97.8 (80.9, 124.4) | −0.20 (−0.47, 0.11) | 0.15 (0.03, 0.26) |
| Halo              | 282.6 NA         | −0.89 NA       | NA NA           |
| Hercules a        | 73.5 (62.7, 91.1) | −0.05 (−0.40, 0.13) | 0.07 (−0.08, 0.17) |

The Astrophysical Journal, 909:115 (21pp), 2021 March 10

Chen et al.
with previous studies (Fuhrmann 1998; Bensby et al. 2003; Haywood et al. 2013; Bensby et al. 2014).

To explore the relation between $TD/D$ and kinematic age, we first sorted the plant host sample by $TD/D$. Next, according to $TD/D$, we divided the planet host sample into 10 bins of approximately equal size. Then, for each bin, we calculated their kinematic ages. As shown in Figure 18, the kinematic age generally increases with $TD/D$, demonstrating that $TD/D$ is an indicator of age for stars in the Galactic disk.

5. Discussions

5.1. Systematic Differences in Radial Velocity from Various Sources

As mentioned in Section 4.1.3, we obtained the RV from five sources: APOGEE, LAMOST, RAVE, Gaia, and EA. As the RV is one of the basic parameters used to calculate the Galactic velocity, the systematic differences in RV will induce differences in the Galactic velocity and then affect the classification of Galactic components. Therefore, it is necessary to calibrate the systematic difference in RV from various sources.

Here, we chose the APOGEE RV data as a standard reference because it has the highest resolution ($\sim R \geq 22,500$) and therefore the most accurate RV measurements. Then, we compared the difference in RV for common stars with measurements from the other sources overlapped with APOGEE: 292 for LAMOST, 2 for RAVE, 181 for Gaia, and 146 for EA. As shown in Figure 19, the systematic offsets in RV from APOGEE measurements, $\Delta RV$, are 1.21, 0.05, and 0.06 km s$^{-1}$ for LAMOST, Gaia, and EA, respectively. For RAVE, there are only two stars in common and too few to make an analysis. Here, we refer to Huang et al. (2018), which presents a new catalog of RV standard stars selected from the APOGEE data and find that the systematic offset is only 0.17 km s$^{-1}$ for RAVE. The systematic differences in RV between APOGEE and other sources are all smaller than the typical uncertainties in RV of our sample 1.58 km s$^{-1}$, and therefore they have no significant influence on the calculation of Galactic velocity and the classification of Galactic components.

5.2. Kinematic Age versus Asteroseismic Age versus Isochrone Age

In order to determine the reliability of kinematic ages derived in this work, we compare them to ages derived from asteroseismology and isochrones. The Kepler asteroseismic LEGACY sample (Silva Aguirre et al. 2017) provides a well-characterized sample of 66 Kepler planet hosts with
The Astrophysical Journal, 909:115 (21pp), 2021 March 10

Chen et al.

5.3. Future Studies

For clarity and simplicity, here we apply the revised kinematic methods (Sections 2 and 3) only to the planet host sample. Next, in our second paper of the PAST project (D.-C. Chen et al. 2021, in preparation), we will apply the revised kinematic methods to the whole Kepler star sample, enabling us to further connect stellar kinematic properties to stellar rotations and activities.

The planet host catalog provides stellar parameters, spatial position, Galactic velocity and component classification for 2174 stars that host 2872 planets. Furthermore, using AVR as we show in Section 3.6, one can obtain the kinematic age for a group of stars with these respective properties. As shown in Figure 13 and Table 9, these planet hosts are spread over different Galactic components in a wide range of distance up to 1500 pc. With such rich stellar information, future studies will be allowed to explore and answer some fundamental questions about exoplanets, such as, what are the differences in various properties of planetary systems at different positions in the Galaxy with different ages? Specifically, in a subsequent paper of the PAST project (J.-Y. Yang et al. 2021, in preparation), we will study whether/how planetary occurrence and architecture change with the Galactic environment. The answers to these questions will be crucial in constraining various models and theories of planetary formation and evolution.

6. Guidelines for Using the Methods and Catalog

In this section, we provide the guidelines, cautions, and limitations to utilize the revised kinematic methods (Sections 2 and 3) and planet host catalog (Table 8).

To classify stars into different Galactic components, the key is to calculate the relative probabilities between two different components (i.e., \(TD/D\), \(TD/H\), Herc/D, Herc/D) with Equations (4) and (5), which rely on the X factor and velocity ellipsoid (i.e., \(\sigma_U\), \(\sigma_V\), \(\sigma_W\), and \(V_{\text{sym}}\)) of different Galactic components. Here, we suggest two ways to obtain these kinematic parameters for a given Galactic position \((R, Z)\). The most easy way is to use our fitting formulae (e.g., Equation (7) with coefficients in Table 3 for \(\sigma_U\), \(\sigma_V\) and \(\sigma_W\), and Equation (8)
for $V_{\text{sym}}$). Alternatively, one can conduct interpolation based on our revised characteristics in Table 2.

To derive the kinematic age for a group of stars, one can use the AVR (Equation (15)) with the revised coefficients in Table 5. The typical age uncertainty can be estimated from Equation (16), which is $\sim 10\%$ here. For the purpose of avoiding potential spatial biases, we recommend to adopt the total velocity dispersion (Equation (13)) when using the AVR (Equation (15)).

Applying the above revised methods to our planet host sample, we provide a catalog with kinematic properties (e.g., Galactic position, velocity, and $TD/D$) and other basic parameters (e.g., $T_{\text{eff}}$ and element abundances). With this catalog, one can divide planet hosts into bins according to respective properties (e.g., planetary period, multiplicity) and calculate the kinematic age for each bin, which will be practical and useful for statistical studies of age effects on planetary systems. Although the kinematic method cannot directly measure the ages for individual stars, some of the derived kinematic properties (e.g., $TD/D$) could serve as good age tracers given their significant correlations (e.g., Figure 18).

However, there are some notable cautions and limitations, which are listed as follows:

1. The revised velocity ellipsoid and AVR are strictly applicable for stars within the region the calibration sample covers, i.e., within 1.54 kpc to the Sun (corresponding to $R = 7.5 - 10.0$ kpc, $|\theta| < 10$ deg & $|Z| = 0 - 1.5$ kpc). Taking the assumption that the Galactic disk is axis-symmetric (e.g., Yurin & Springel 2014; Aumer et al. 2016), the velocity ellipsoid and AVR will be independent of $\theta$ (e.g., Williams et al. 2013). Therefore, the criterion $|\theta| < 10$ deg is not necessary for the Galactic disk stars. Besides, with Equation (7) and Table 3, the velocity dispersion can be extrapolated to the region beyond 1.54 kpc. However, this extrapolation should be adopted with caution.

2. Since the additional motions caused by binary orbits could affect the stellar kinematic, binaries should be applied with caution.

3. Due to the small numbers of Halo and Hercules stream stars in our calibration sample, we adopt the velocity dispersion values derived from stars in the solar neighborhood as in Bensby et al. (2014). As the velocity dispersions change with the Galactic position (Williams et al. 2013), there might be some deviations in the classification of Halo and Hercules stream stars when utilizing the characteristic parameters in Table 2 and the planet host catalog in Table 8. It will be more reliable to take other parameters (e.g., velocity, element abundance, angular momentum; Bensby et al. 2007; Lee et al. 2011; Bonaca et al. 2017; Kushniruk & Bensby 2019) into consideration.

4. There is no clear trend between velocity dispersions and ages for stars belong to the Hercules stream and halo in our calibration sample. Therefore, the method to derive kinematic age is only suitable for stars in the Galactic disk.

5. The kinematic ages and uncertainties derived from Equations (15) and (16) are the typical (median/mean) values for a group of stars.

7. Summary

Since 1995, the discovered exoplanet population has expanded significantly from the solar neighborhood to a much larger area in the Galaxy (Figure 1). We are therefore entering a new era in which to study exoplanets in the context of the Galaxy. In the Galactic context, the relations between the properties of planetary systems and the kinematics as well as the ages of planet host stars have yet to be explored. To answer these questions, we perform a series of studies in a project dubbed Planets Across Space and Time (PAST). In this paper, which is Paper I and the basis of the PAST series, we revisit the kinematic methods for the classification of Galactic components (Section 2) and estimation of kinematic ages (Section 3) and apply them to planet host stars (Section 4).

For the classification of Galactic components (Section 2), we adopt the well-used kinematic approach as in Bensby et al. (2003, 2014). However, so far, the kinematic characteristics of this method have been applied only to the solar neighborhood within $\sim 100-200$ pc. For this reason, using a calibration sample based on the GAIA and LAMOST data (Section 2.3.1), we extend the kinematic characteristics to $\sim 1500$ pc (Section 2.3, Table 2) to cover the majority of planet hosts (Figures 1 and 13).

For the estimation of kinematic ages, we refit the age–velocity dispersion relation (AVR) with the calibration sample (Section 3, Figure 8). Our AVR is consistent with those in previous studies (e.g., Holmberg et al. 2009), but with much smaller internal uncertainties (Table 5) thanks to the large and high-quality calibration sample. Based on this refined AVR, we are able to derive kinematic ages with an uncertainty of 10–20% (Section 3.6), which is a factor of $\sim 3$ smaller than those from previous studies (Table 6).

Applying the above revised methods to our planet host sample, we then construct a catalog with kinematic properties and other basic parameters for 2174 stars (Section 4.2, Table 8) by combining data from Gaia, LAMOST, APOGEE, RAVE, and the NASA exoplanet archive (Table 7 and Section 4.1). The majority (1894/2174, 87.1%) of planet host stars are found to be in the Galactic thin disk, while 5.2% (114/2174) of them belong to the thick disk and only 0.05% (1/2174) reside in the halo (Table 9). As expected, we find that the total velocity, [$\alpha$/Fe], and the kinematic age generally increase with the relative probabilities for the thick disk to thin disk, i.e., $TD/D$, while [Fe/H] decreases with $TD/D$ (Figure 15, 18). The kinematic age is $2.84^{+0.32}_{-0.26}$ Gyr for the thin disk stars and $10.42^{+1.82}_{-1.71}$ Gyr for the thick disk stars in the planet host sample (Figure 17).

We also compare our derived kinematic ages with asteroseismic ages and isochrone ages (Section 5.2). Our kinematic ages match better with the asteroseismic ages, though the three kinds of ages are generally consistent with each other within their uncertainties (Figure 20).

Future studies of exoplanets in the Galactic context, e.g., the subsequent papers of our PAST series (Section 5.3), will benefit not only from the catalog of the kinematic properties but also from the revised methods that derived such a catalog in this work. The important guidelines, cautions and limitations to utilize our kinematic methods and planet host catalog are described in Section 6.

This work is supported by the National Key R&D Program of China (No. 2019YFA0405100) and the National Natural
Science Foundation of China (NSFC; grant Nos. 11933001, 11973028, 11803012, 11673011, 12003027). J.-W.X. acknowledges the support from the National Youth Talent Support Program and the Distinguished Youth Foundation of Jiangsu Scientific Committee (BK20190005). H.-F.W. is supported by the LAMOST Fellow project, National Key Basic R&D Program of China via 2019YFA0405500 and funded by China Postdoctoral Science Foundation via grants 2019M653504 and 2020T130563, Yunnan province postdoctoral directed culture Foundation, and the Cultivation Project for LAMOST Scientific Payoff and Research Achievement of CAMS-CAS. M.-S.X. and Y.H. acknowledge the National Natural Science Foundation of China (grant No.11703035). C. L. thanks National Key R&D Program of China No. 2019YFA0405500 and the National Natural Science Foundation of China (NSFC) with grant No. 11835057.

This work has included data from Guoshoujing Telescope. Guoshoujing Telescope (the Large Sky Area Multi-Object Fiber Spectroscopic Telescope LAMOST) is a National Major Scientific Project built by the Chinese Academy of Sciences. Funding for the project has been provided by the National Development and Reform Commission. LAMOST is operated and managed by the National Astronomical Observatories, Chinese Academy of Sciences. This work presents results from the European Space Agency (ESA) space mission Gaia. Gaia data are being processed by the Gaia Data Processing and Analysis Consortium (DPAC). Funding for the DPAC is provided by national institutions, in particular the institutions participating in the Gaia MultiLateral Agreement (MLA). The Gaia mission website is https://www.cosmos.esa.int/gaia. We acknowledge the NASA Exoplanet archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. Funding for RAVE has been provided by: the Australian Astronomical Observatory; the Leibniz-Institut fuer Astrophysik Potsdam (AIP); the Australian National University; the Australian Research Council; the French National Research Agency; the German Research Foundation (SPP 1177 and SFB 881); the European Research Council (ERC-StG 240271 Galactica); the Istituto Nazionale di Astrofisica at Padova; the Johns Hopkins University; the National Science Foundation of China (the AST-0908326); the W. M. Keck foundation; the Macquarie University; the Netherlands Research School for Astronomy; the Natural Sciences and Engineering Research Council of Canada; the Slovenian Research Agency; the Swiss National Science Foundation; the Science & Technology Facilities Council of the UK; Opticon: Strasbourg Observatory; and the Universities of Groningen, Heidelberg and Sydney. The RAVE website is at https://www.rave-survey.org. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. The SDSS website is www.sdss.org.
