Mapping the Galactic Disk with the LAMOST and Gaia Red Clump Sample. I. Precise Distances, Masses, Ages, and 3D Velocities of ~140,000 Red Clump Stars

Yang Huang1, Ralph Schönrich2, Huawei Zhang3,4, Yaqian Wu5, Bingqiu Chen1, Haifeng Wang1,9, Chun Wang3,9, Haibo Yuan7, Xinyi Li1, Weixiang Sun1, Ji Li5, and Xiaowei Liu1

1 South-Western Institute for Astronomy Research, Yunnan University, Kunming 650500, People’s Republic of China; yanghaiang@ynu.edu.cn, x.liu@ynu.edu.cn
2 Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Dorking, Surrey, RH5 6NT, UK; r.schoenrich@ucl.ac.uk
3 Department of Astronomy, Peking University, Beijing 100871, People’s Republic of China
4 Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, People’s Republic of China
5 National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, People’s Republic of China
6 Max-Planck Institute for Astronomy, Königsthul, D-69117, Heidelberg, Germany
7 Department of Astronomy, Beijing Normal University, Beijing 100875, People’s Republic of China
8 Department of Space Science and Astronomy, Hebei Normal University, Shijiazhuang 050024, People’s Republic of China

Received 2020 February 14; revised 2020 June 1; accepted 2020 June 2; published 2020 August 5

Abstract

We present a sample of ~140,000 primary red clump (RC) stars of spectral signal-to-noise ratios higher than 20 from the Large Sky Area Multi-Object Fibre Spectroscopic Telescope (LAMOST) Galactic spectroscopic surveys, selected based on their positions in the metallicity-dependent effective temperature–surface gravity and color–metallicity diagrams, supervised by high-quality Kepler asteroseismology data. The stellar masses and ages of those stars are further determined from the LAMOST spectra, using the kernel principal component analysis method and are trained with thousands of RCs in the LAMOST–Kepler fields with accurate asteroseismic mass measurements. The purity and completeness of our primary RC sample are generally higher than 80%. For the mass and age, a variety of tests show typical uncertainties of 15% and 30%, respectively. Using over 10,000 primary RCs with accurate distance measurements from the parallaxes of Gaia Data Release 2 (DR2), we recalibrate the Ks absolute magnitudes of primary RCs by, for the first time, considering both the metallicity and age dependencies. With the the new calibration, distances are derived for all the primary RCs, with a typical uncertainty of 5–10%, which is even better than the values yielded by the Gaia parallax measurements for stars beyond 3–4 kpc. The sample covers a significant volume of the Galactic disk of 4 ≤ R ≤ 16 kpc, |Z| ≤ 5 kpc, and −20 ≤ φ ≤ 50°. Stellar atmospheric parameters, line-of-sight velocities, and elemental abundances derived from the LAMOST spectra and proper motions of Gaia DR2 are also provided for the sample stars. Finally, the selection function of the sample is carefully evaluated in the color–magnitude plane for different sky areas. The sample is publicly available at https://zenodo.org/deposit/3875974.

Unified Astronomy Thesaurus concepts: Distance indicators (394); Red giant clump (1370); Galaxy abundances (574); Stellar ages (1581); Milky Way disk (1050); Stellar masses (1614); Galaxy structure (622)

1. Introduction

Primary red clump (RC) stars10 are metal-rich low-mass stars (typically smaller than 2 \( M_\odot \)) of intermediate to old age in the core helium-burning phase (ignited degenerately). They are widely used as distance indicators given their quite stable luminosities that are weakly dependent on chemical composition and age (e.g., Cannon 1970; Paczynski & Stanek 1998). By carefully calibrating the absolute magnitude of RC stars using the Hipparcos parallaxes (ESA 1997) of few hundred nearby RCs, precise distances to the Galactic center (Paczynski & Stanek 1998), to the Local Group galaxies, e.g., the Large Magellanic Cloud (LMC; Stanek et al. 1998), the Small Magellanic Cloud (SMC; Subramanian & Subramanian 2012), and Andromeda (M31; Stanek & Garnavich 1998) have been determined.

As standard candles widely distributed across the entire Galactic disk, RCs are excellent tracers to explore the three-dimensional structure (e.g., Bovy et al. 2016), study the chemical and kinematic properties (e.g., Bovy et al. 2014, 2015; Nidever et al. 2014; Huang et al. 2015, 2016), and unravel the assemblage history of the Galactic disk. However, it is a challenging task to identify the individual RCs among the numerous field stars. In principle, the separation between RCs and other types of stars, e.g., red giant branch (RGB) stars, can be carried out based on their large frequency separation \( \Delta \nu \) of the acoustic modes and period spacing \( \Delta P \) of the gravity modes measured from high-precision light curves, e.g., obtained with the Kepler and Convection, Rotation and planetary Transits (CoRoT) satellites. However, at the moment no more than a few 10,000 stars have \( \Delta \nu \) and \( \Delta P \) measurements in the Kepler and CoRoT fields. More generally, with values of effective temperature \( T_{\text{eff}} \) and surface gravity \( \log g \) available from the large-scale spectroscopic surveys, such as the RAdial Velocity Experiment and Large Sky Area Multi-Object Fibre Spectroscopic Telescope (LAMOST) for a huge number of stars, a large number of RC candidates can be collected by selecting stars located in a relatively “small box” in the \( T_{\text{eff}} – \log g \) diagram of 4800 ≤ \( T_{\text{eff}} \) ≤ 5200 K and 2.0 ≤ \( \log g \) ≤ 3.0 (e.g., Siebert et al. 2011; Williams et al. 2013). Unfortunately, RCs thus selected suffer from significant contamination, typically from 20–50% (e.g., Williams et al. 2013; Wan et al. 2015), from RGB stars. By applying further cuts in the metallicity-dependent \( T_{\text{eff}} – \log g \) plane and in the color \( (J – K)_{\text{F}11} \)–metallicity Z plane, Bovy et al. (2014) successfully selected over 10,000 RC-like stars with a purity greater than 97%.

9 LAMOST Fellow.
10 In contrast, secondary RCs are helium-burning (ignited non-degenerately) descendants of high-mass (typically greater than 2 \( M_\odot \)) stars and are nonstandard candles.
from the Apache Point Observatory Galactic Evolution Experiment (APOGEE) survey. After calibration with the stellar evolution model and the high-quality asteroseismology data from Kepler, the cut in the metallicity-dependent $T_{\text{eff}}-\log g$ plane can separate the RC-like and RGB stars very well. The second cut in the color $(J-K_s)_{\text{c}}$–metallicity $Z$ plane can further eliminate contamination from the secondary RCs. The distances of the selected RCs are then estimated with a precision better than 5–10%. By applying similar cuts to stars selected from the LAMOST Galactic spectroscopic surveys, Huang et al. (2015, hereafter H15) obtained a clean RC sample of nearly 100,000 stars, covering a significant volume of the Galactic disk. More recently, using a training set of stars selected from the Kepler asteroseismology data, Ting et al. (2018) derived spectroscopy-based estimates for $\Delta \nu$ and $\Delta P$ and then used the results to select out over 200,000 RCs, claiming a contamination of $\sim 9\%$.

This paper is an update of H15. The update is twofold. First, in H15, RCs are selected from the second data release of LAMOST Galactic spectroscopic surveys. Here the analysis is extended to the fourth release and thus yields more RCs that cover a larger Galactic disk volume. Second, we attempt to derive the masses and ages of those selected RCs. For a low-mass giant star, robust ages can be estimated with a well-measured mass. At present, age estimates of a precision of about 20% (e.g., Martig et al. 2016a; Ness et al. 2016; Wu et al. 2018) have been achieved for red giants in the Kepler fields, along with very precise asteroseismic masses (better than 8–10%; e.g., Huber et al. 2014; Yu et al. 2018) and atmospheric parameters from spectroscopy (e.g., the APOGEE and LAMOST surveys). Those masses and ages, inferred from precise asteroseismic measurements, are then taken as training sets to derive masses and ages of over 100,000 giant stars selected from the LAMOST and APOGEE surveys, using either an empirical relation based on the spectroscopic carbon to nitrogen abundance ratios (e.g., Martig et al. 2016a; Ness et al. 2016; Sanders & Das 2018) or a data-driven approach (e.g., Ness et al. 2016; Ho et al. 2017; Ting et al. 2018; Wu et al. 2019). The physics underlying the method of deriving masses and ages from the medium-/high-resolution spectra is that the carbon to nitrogen abundance ratio $[C/N]$ is tightly correlated with stellar mass as results of the convective mixing through the CNO cycle (aka the first-dredge up process). In this sense, $[C/N]$ ratios deducible from the optical/infrared spectra are good indicators of stellar masses for red giants and thus can be further used to derive their ages. In the current work, we applied a similar technique used by Wu et al. (2018, 2019) to the selected RCs and derived their masses and ages, using a training set selected from the LAMOST–Kepler common stars (see Section 4). Moreover, atmospheric parameters, line-of-sight velocities, and elemental abundances estimated from the LAMOST spectra (Xiang et al. 2017b, 2017c) and proper motions from Gaia Data Release 2 (DR2; Gaia Collaboration et al. 2018; Lindegren et al. 2018) are also provided for the RC sample.

The paper is structured as follows. In Section 2, we describe the data employed in the current analysis. We then introduce the selection of primary RCs in Section 3 and determine their masses and ages in Sections 4. In Section 5, we present a new calibration of the $K_s$ absolute magnitudes of primary RCs and derive their distances based on the new calibration. The selection function of the primary RC sample is presented in Section 6. We present the final catalog of selected primary RCs in Section 7. Finally, a summary is given in Section 8.

2. Data

2.1. LAMOST Galactic Surveys

Being a 4 m quasi-meridian reflecting Schmidt telescope equipped with 4000 fibers distributed in a field of view of about 20 sq. deg., LAMOST can simultaneously collect up to 4000 spectra per exposure covering the wavelength range of 3800–9000 Å at a resolution of about 1800 (Cui et al. 2012). For the scientific motivations and target selections of the surveys, please refer to Zhao et al. (2012), Deng et al. (2012), and Liu et al. (2014) for details. After one year Pilot Surveys between 2011 September and 2012 June, the five year Phase-I LAMOST Regular Surveys began in 2012 September and completed in the summer of 2017. By adding a new component of medium-resolution ($R \sim 7500$) surveys, the Phase-II LAMOST Regular Surveys were initiated in 2019 September, following one year Pilot Surveys between 2017 September and 2018 June.

To derive stellar atmospheric parameters and line-of-sight velocities from the LAMOST spectra, two independent stellar parameter pipelines have been developed—the official LAMOST Stellar Parameter Pipeline (LASP; Luo et al. 2015) and the LAMOST Stellar Parameter at Peking University (LSP3; Xiang et al. 2015, 2017b). Typical uncertainties of $5 \text{ km s}^{-1}$ in $v_{\text{rot}}$, 100 K in $T_{\text{eff}}$, 0.25 dex in $\log g$, and 0.10 dex in metallicity [Fe/H] have been achieved by both pipelines for “normal”-type (FGK) stars. Based on a kernel principal component analysis (KPCA), the latest version of LSP3 (Xiang et al. 2017b) is able to deliver log $g$ of red giant stars with precision of about 0.1 dex for red giant stars, using a training set from the LAMOST–Kepler common stars. With the same technique, estimates of [$\alpha$/Fe], [C/H], and [N/H] of giant stars with precisions of 0.05, 0.10, and 0.10 dex, respectively, are also achieved, using a training set from the LAMOST–APOGEE common stars. Based on the derived atmospheric parameters, values of the interstellar extinction are also derived for the stars, using the “star pair” method developed by Yuan et al. (2013). The precision of the estimated $E(B-V)$ values is about 0.04 mag. In the current work, we adopt the atmospheric parameters (i.e., $T_{\text{eff}}$, $\log g$, and [Fe/H]) derived with LSP3 from the low-resolution spectra ($R \sim 1800$) in the fourth data release of the LAMOST Galactic surveys (http://dr4.lamost.org/doc/vac; Xiang et al. 2017c).

2.2. Asteroseismic Samples from the Kepler Data

In addition to the LAMOST spectra and related catalogs of stellar parameters, different asteroseismic samples constructed from the Kepler data are also used for training, examining, and mass estimation purposes. Similar to H15, we use the asteroseismic sample with the evolutionary stages of the stars classified (based on $\Delta \nu$ and $\Delta P$), from Stello et al. (2013), to select primary RCs but exclude RGBs and secondary RCs. We also have tested the selection of primary RCs trained by more recent asteroseismic samples (e.g., the sample of Vrard et al. 2016) but no significant improvements on the selection are found. Finally, we adopt the global asteroseismic parameters $\nu_{\text{max}}$ and $\Delta \nu$ measured by Yu et al. (2018) for mass estimation. Using a modified version of the SYD pipeline (Huber et al. 2009), Yu et al. (2018) systematically characterized solar-like
oscillations and granulation for 16,094 oscillating red giants, based on the end-of-mission long-cadence data of Kepler. The derived parameters are quite precise, with typical relative uncertainties of a few percent.

3. Selection of Primary RCs

As mentioned, to select a clean sample of primary RCs, we adopt the method developed in H15. The method relies on the accurate estimation of surface gravity log g. To training sample from LAMOST–Kepler fields (De Cat et al. 2015; Ren et al. 2016; Zong et al. 2018), a high precision (of about 0.1 dex) has been achieved with LSP3 for the estimation of log g from the LAMOST spectra (Xiang et al. 2017b). Such a precision of log g, together with a typical uncertainty of 100 K of $T_{\text{eff}}$, allow one to select primary RCs with very low contamination (Bovy et al. 2014; H15; Ting et al. 2018).

The method of H15 includes two steps. The first step is to select RCs in the metallicity-dependent $T_{\text{eff}}$-$\log g$ plane, supervised by an asteroseismic sample with the evolutionary stage classified by Stello et al. (2013). In principle, one can directly adopt the relation found by H15. However, the values of $T_{\text{eff}}$ and $\log g$ used by H15 have been updated by the latest version of LSP3 (Xiang et al. 2017b). We therefore re-derive the relation using the updated $T_{\text{eff}}$ and $\log g$. To do so, we crossmatch red giant stars from the LAMOST–Kepler fields with those of Stello et al. (2013), and a total of 1547 common stars (hereafter LE sample for short) with LAMOST spectral signal-noise ratios ($S/N_s$) greater than 50 are found. With the LE sample, we derive the following cuts in the metallicity-dependent $T_{\text{eff}}$-$\log g$ plane to select RC-like stars:

$$1.8 \leq \log g (\text{dex}) \leq 0.0006 K^{-1}[T_{\text{eff}} - T_{\text{eff}}^{\text{Ref}}] + 2.5,$$

where

$$T_{\text{eff}}^{\text{Ref}} = -873.1 \text{ K dex}^{-1}[\text{Fe/H}] + 4255 \text{ K}.$$  

Similar in H15, the cuts are empirically obtained by maximizing the product of completeness and purity of the RC stars in the LE sample. As an example, the distribution of LE sample stars of solar metallicity ($-0.05 < [\text{Fe/H}] < 0.05$) in the $T_{\text{eff}}$-$\log g$ plane is presented in Figure 1. Clearly, the cuts defined by above separate RC-like and RGB stars very well. As in H15, the distribution of stars predicted by the Padova and TRieste Stellar Evolution Code (PARSEC) stellar isochrones (Bressan et al. 2012) is also shown as background in Figure 1. The prediction was generated assuming a constant star formation history for the last 10 Gyr, a lognormal initial mass function from Chabrier (2001) and a metallicity distribution similar to that of the LE sample stars for the individual metallicity bins. The distribution of the LE sample stars is generally in agreement with the predicted one, except for possibly some minor systematic offsets between the LSP3 and PARSEC values of $T_{\text{eff}}$.

In H15, a second step was developed to remove secondary RCs, using cuts in the color–metallicity plane. Since the metallicities [Fe/H] yielded by the latest version of LSP3 change little compared to earlier results, we have directly adopted the cuts obtained by H15:

$$Z < 2.58[(J - K_s)_0 - 0.400]^3 + 0.0034,$$

and

$$Z > 1.21[(J - K_s)_0 - 0.085]^9 + 0.0011.$$
Here the values of intrinsic color \((J - K_s)_{\odot}\) are from the Two Micron All-Sky Survey (2MASS; Skrutskie et al. 2006), after correcting for reddening estimated with the “star pair” method (see Section 2.1; Yuan et al. 2013). Values of metallicity \(Z\) are converted from those of [Fe/H] using the relation given by Bertelli et al. (1994) assuming a solar value \(Z_{\odot} = 0.017\). As shown by Figure 1, the cuts defined by Equations (3) and (4) can efficiently eliminate secondary RCs. The latter are descendants of high-mass \((\geq 2 M_\odot)\) in the helium-burning phase, ignited non-degenerately. They are nonstandard candles. In addition to the above cuts developed in H15, we consider only stars of \(T_{\text{eff}} \leq 5000\) K and \(\log g \leq 2.75\) to further suppress the contamination of secondary RCs and RGBs (see the black pluses in the right panel of Figure 1).

By applying above cuts to the whole set of LAMOST Data Release 4 processed with the latest version of LSP3 (LMDR4 hereafter; Xiang et al. 2017c), about 150,000 primary RC candidates of spectral \(S/N > 10\) are obtained. To characterize the purity and completeness of the selected primary RC sample, we use another asteroseismic sample with the evolutionary stage classified by Vrard et al. (2016). The sample contains more than 6100 stars. To do so, we first exclude stars in the sample of Vrard et al. (2016) that are shared in common with the LE sample and crossmatch the remaining ones with the whole LMDR4. This yields 1720 and 2836 common stars, respectively. We divide the 1720 common stars into different \(S/N\) bins and calculate the numbers of primary RCs \((N_{\text{ApRC}})\), secondary RCs \((N_{\text{SpRC}})\), and RGBs \((N_{\text{SRGB}})\) in each \(S/N\) bin. Similarly, the 2836 common stars are also divided into the same \(S/N\) bins and the numbers of primary RCs \((N_{\text{ApRC}})\), secondary RCs \((N_{\text{ArRC}})\), and RGBs \((N_{\text{ARGB}})\) in the individual bins are derived. The purity and completeness for the individual spectral \(S/N\) bins are then simply given by ratios of \(N_{\text{SpRC}}/(N_{\text{SpRC}} + N_{\text{SRRC}} + N_{\text{SRGB}})\) and \(N_{\text{SpRC}}/N_{\text{ArRC}}\), respectively. The results are presented in Figure 2. The purity is around 65% for an \(S/N\) of 10 and increases rapidly to about 85% for \(S/N\)s higher than 20. The completeness is low, about 40% at \(S/N \sim 10\), and increases rapidly to over 85% for \(S/N > 20\). Based on the results, we further exclude primary RC candidates of \(S/N < 20\). This leaves us with nearly 140,000 stars. Over 87% of the remaining stars have \(S/Ns\) higher than 30, implying a purity and a completeness exceeding 80% for most of our primary RCs in the final sample. If only concerning RCs that include both primary and secondary RCs, our technique is expected to deliver a sample of purity over 95% at \(S/N > 50\) (see the squares in Figure 2), comparable to the recent results of Ting et al. (2018). Here the above examination performance could be decreased, if any systematic differences between the training sample of Stello et al. (2013) and the test sample of Vrard et al. (2016) are found. In this degree, the reported purity and completeness in Figure 2 are actually lower bounds.

### 4. Masses and Ages of Primary RCs

#### 4.1. Training and Test Sets

In this section, we estimate the masses and ages for our selected primary RCs. To do so, we first construct a primary RC sample with precise mass and age values deduced using the asteroseismic information provided by Yu et al. (2018) and the stellar atmospheric parameters from LAMOST. We then crossmatch the sample of Yu et al. (2018) with our primary RC sample. A total of 4285 common stars are found. Masses of those common stars are determined using the standard seismic scaling relation:

\[
\frac{M}{M_{\odot}} = \left( \frac{\Delta \nu}{\Delta \nu_{\odot}} \right)^{-1} \left( \frac{\nu_{\text{max},\odot}}{\nu_{\text{max},\odot}} \right)^3 \left( \frac{T_{\text{eff},\odot}}{T_{\text{eff},\odot}} \right)^{3/2}.
\]

Here we adopt solar values of \(T_{\text{eff},\odot} = 5777\) K, \(\nu_{\text{max},\odot} = 3090\) \(\mu\)Hz, and \(\Delta \nu_{\odot} = 135.1\) \(\mu\)Hz (Huber et al. 2011). However, as pointed out by previous studies (e.g., Huber et al. 2011; Viani et al. 2017), the standard scaling relation may induce systematic errors of 10% to 15% in the derived masses. Similar to Wu et al. (2019), we adopt a modified scaling relation:
relation from Sharma et al. (2016) to reduce the systematic errors. Based on the theoretical stellar models, a correction factor, $f\Delta\nu$, is added in the standard scaling relation,

$$\frac{M}{M_\odot} = \left( \frac{\Delta\nu}{f\Delta\nu} \right)^{-4} \left( \frac{\nu_{\text{max},\odot}}{\nu_{\text{max}}} \right)^3 \left( \frac{T_{\text{eff}}}{T_{\text{eff},\odot}} \right)^{3/2},$$

where $f\Delta\nu$ can be obtained for all stars using the publicly available code Asfgrid provided by Sharma et al. (2016). With this modified scaling relation, masses of the 4285 common stars are derived using $\Delta\nu$ and $\nu_{\text{max}}$ provided in Yu et al. (2018) and $T_{\text{eff}}$ yielded by the LSP3. As Figure 3 shows, masses thus derived for those primary RCs largely fall between 0.4 and 2.0 $M_\odot$, except for a few of masses greater than 2 $M_\odot$, which are probably from contamination of secondary RCs.

With the masses estimated above and the spectroscopic stellar atmospheric parameters ($T_{\text{eff}}$, $\log g$, and $[\text{Fe}/\text{H}]$), the ages of those primary RCs can be further derived using the stellar isochrones. To do so, we have adopted a Bayesian approach similar to Xiang et al. (2017a). The input constraints include the mass $M$ and surface gravity $\log g$ inferred from asteroseismic parameters and the effective temperature $T_{\text{eff}}$ and metallicity $[\text{Fe}/\text{H}]$ estimated from the LAMOST spectra with LSP3. For the stellar isochrones, we adopted the PARSEC ones calculated with a mass-loss parameter $\eta_{\text{Reimers}} = 0.2$ (Bressan et al. 2012). Details of age estimation with Bayesian approach are described by Xiang et al. (2017a). To show the results of our age determinations, posterior probability distributions as a function of age for three stars of typical ages are shown in Figure 4. Thanks to the very precise mass estimates from the asteroseismology information, the distributions show prominent peaks, yielding well-constrained ages. As Figure 5 shows, the typical uncertainties of the estimated ages are about 15–20%, except for a few (about 3%) uncertainties larger than 40% (as a result of the large errors in the derived mass and/or in the atmospheric parameters).

With masses and ages estimated for those 4285 stars, we then divide them into two subsamples, a training and a test one. For the training sample, 2120 stars are randomly picked out with a spectral S/N greater than 50, a mass error smaller than 15%, and an age error smaller than 40%. In Figure 6, the distributions of the training sample in the $[\text{Fe}/\text{H}]-[\alpha/\text{Fe}]$ plane, color coded by mass and age, and in the age–$[\alpha/\text{Fe}]$ plane, color coded by mass and metallicity, are shown. We note that the discarded stars do not change the distributions of mass, age, $[\text{Fe}/\text{H}]$, and $[\alpha/\text{Fe}]$ of the test set. The remaining 987 stars of mass errors smaller than 15% and age errors smaller than 40% are adopted as the test set.

4.2. Mass and Age Determinations for Primary RCs

In this section, we estimate masses and ages from the LAMOST spectra, using the KPCA (Schölkopf et al. 1998) method trained with the training sample described above.
Generally, PCA is a classic and powerful method that can convert observations (e.g., spectra) into a set of linearly uncorrelated orthogonal variables or principal components (PCs). KPCA works like a PCA but can be extended to nonlinear feature extractions with kernel techniques (a Gaussian radial basis function is adopted here). The approach has been successfully used to estimate stellar atmospheric parameters, as well as masses and ages from the LAMOST spectra by H15, Xiang et al. (2017b), and Wu et al. (2019), respectively. For details of this method, please refer to those papers.

Similar to Xiang et al. (2017b), the LAMOST blue-arm spectra (3900–5500 Å) are adopted to train the relations between the spectral features and stellar mass and age. The key of the training is to find an optimal number of principal components ($N_{PC}$). Generally, a small value of $N_{PC}$ is not sufficient to construct tight relations between the spectral features and the parameters to be estimated (stellar masses and ages here). On the other hand, an excessively large value of $N_{PC}$ has the risk of over-fitting and causes problems when dealing with spectra of relatively low S/Ns.

To determine an optimal value of $N_{PC}$, we have tried different values of $N_{PC}$ and compared the dispersions of the relative residuals of the derived masses and ages for both training and test sets. For the training sample, the dispersions of the relative residuals of the deduced masses and ages decrease monotonically from 0.24 to 0.09 and from 0.56 to 0.18, respectively, as $N_{PC}$ increases from 5 to 900 (see Figure 7). However, for the test set, the dispersion of the relative mass residuals decreases from 24% to 13% as $N_{PC}$ increases from 5 to 125 and then remains unchanged as $N_{PC}$ continuously increases to 900. A similar result is also found for the relative age residuals (see Figure 7). Obviously, for $N_{PC} \geq 125$, the algorithm over-fits the training set for both stellar masses and ages. Finally, the turning point, $N_{PC} = 100$, in the dispersions of relative mass and age differences as a function of $N_{PC}$ (as shown in Figure 7) is adopted to train the relations between the spectral features and stellar masses and ages. The relative residuals of masses and ages, given by the training set for $N_{PC} = 100$, as a function of atmospheric parameters ($T_{\text{eff}}$, log $g$, and [Fe/H]) are shown in Figure 8, and no obvious trends of variations with those parameters are detected. The
typical dispersions of the relative residuals are 13% and 27%, respectively.

Using the relations as trained above, we derived masses and ages for the whole primary RC sample. The values estimated from spectra of low quality may suffer from large systematics (e.g., Xiang et al. 2017b and Wu et al. 2019), since the KPCA-based multivariate linear regression approach is quite sensitive to the quality of the spectra (e.g., S/N). As in Xiang et al. (2017b) and Wu et al. (2019), internal calibrations are performed for the derived masses and ages using duplicate observations. To do so, a parameter, \( d_g \), is defined. This parameter is given by the maximal kernel value between the target spectrum and the spectra in the training set and describes their similarities. The value of \( d_g \) can vary from 0 to 1, with unity representing exactly the same between a target spectrum and one in the training spectra. Generally, there is a good relation between \( d_g \) and the S/N, but \( d_g \) is found to be more efficient than S/Ns for internal calibration. Small values of \( d_g \) appear due to little similarities between a target spectra and those in the training set. The reason for the few similarities could be either the low spectral S/Ns or some unexpected features (e.g., residuary cosmic rays and/or sky lines) in the target spectra.

To calibrate masses and ages estimated with a small value of \( d_g \), duplicated observations with one yielding \( d_g > 0.8 \) and another yielding \( d_g < 0.8 \) are adopted. Parameters deduced from spectra of \( d_g \geq 0.8 \) are assumed to be free of the systematics and thus are adopted as reference values for the calibration. Those parameters derived from the latter observations that yield \( d_g < 0.8 \) are divided into different bins of \( d_g \) with a bin size of 0.05 for \( 0.50 \leq d_g < 0.80 \) and 0.075 for \( 0.30 \leq d_g < 0.50 \). Results from spectra with \( d_g < 0.30 \) are discarded considering the potentially large errors. A total of about 10,000 stars are excluded in this way. For each \( d_g \) bin, a first-order polynomial is applied to calibrate masses and ages estimated from spectra with \( d_g > 0.8 \) to those from spectra with \( d_g \geq 0.8 \). Stellar masses and ages of nearly 60,000 stars with \( d_g < 0.80 \) are internally calibrated in this way.

4.3. Validation of Estimated Masses and Ages

In this section, we examine the accuracies of the masses and ages of primary RCs as determined above. First, the internal uncertainties are evaluated using the duplicate observations. Second, external checks are performed by member stars of well-studied open clusters (OCs).
First, we select candidates from our primary RC sample within four cluster diameters from the center of the OC. The diameters and central positions of the OCs are adopted from the catalog of Dias et al. (2002). Second, only stars located close to the red clump region of the appropriate cluster isochrone in the color–magnitude diagram (CMD), i.e., the clump region, are selected. Finally, we require that the member stars must have proper motions similar to the cluster values given in the Gaia DR2, $|\mu_x - \mu_x^{\text{cluster}}| < 3\sigma^{\text{cluster}}$ and $|\mu_y - \mu_y^{\text{cluster}}| < 3\sigma^{\text{cluster}}$. In total, 22 member stars are selected for the seven OCs. In Figure 11, the weighted mean ages (see Table 2) given by our KPCA technique for the seven OCs are compared to the literature values, and they generally agree with each other.

5. New $M_K$, Calibration and Distances of Primary RCs

5.1. New $M_K$, Calibration of Primary RCs

To derive distances of the selected primary RCs, an accurate calibration of RC absolute magnitudes is required. Based on the Hipparcos parallaxes and the OC member stars, absolute magnitudes of RCs in different bands (e.g., $VIJHK_s$) have been calibrated previously by a number of studies (e.g., Paczyński & Stanek 1998; Stanek & Garnavich 1998; Grocholski & Sarajedini 2002; Salaris & Girardi 2002; Groenewegen 2008; Laney et al. 2012, hereafter L12). Generally, the absolute magnitudes in different bands are found to be constant and in the near-infrared bands, especially the $K_s$ band, the absolute magnitudes are believed to be most stable least affected by the population effects. The $K_s$ band is also less affected by the interstellar reddening compared to the visual bands (e.g., Salaris & Girardi 2002; L12). Nevertheless, even for the near-infrared bands, ignoring the population effects (i.e., metallicity and age) may introduce systematics on the level of 5–10% (see Figure 3 in Salaris & Girardi 2002 and also Figure 6 in Girardi et al. 2016).

To further improve the calibration, we have collected a large sample of common stars of the Gaia DR2 and our RC sample to recalibrate $K_s$ absolute magnitude of RCs as a function, for the first time, of both metallicity and age. The sample is constructed with the following cuts:

1. The stars must have a Galactic latitude $|b| > 10^\circ$ and a value of $E(B-V)$ less than 0.1 mag, either from Schlegel et al. (1998; hereafter SFD98) for high latitudes

### Table 1

| Coeff. | Mass  | Age   |
|--------|-------|-------|
| $a$    | 0.0258| 0.0541|
| $b$    | 0.9879| 1.0475|
| $c$    | 3.7579| 11.7923|

To assess the uncertainties of estimated masses and ages for primary RCs, results from duplicate observations of similar spectral S/Ns (differed by less than 20%) and collected during different nights are used. The relative mass and age estimate residuals (after divided by $\sqrt{2}$) as a function of the mean spectral S/N are shown in Figure 9. The dispersions of the residuals show significant variations with S/Ns over 15% (mass) and 45% (age) for S/Ns smaller than 30, and 5–8% (mass) and 10% (age) for S/Ns larger than 100. To assign proper random errors to the estimated masses and ages, we fit the relative residuals with

$$
\sigma_i = a + \frac{c}{(S/N)^b}
$$

where $\sigma_i$ represents the random error. The resulting fit coefficients for mass and age estimates are presented in Table 1. In addition to the random errors, method errors are also considered, and the final errors are given by $\sqrt{\sigma^2_i + \sigma^2_m}$. Here $\sigma_m$ denotes the method error and the errors are 13% and 27% (from the training residuals, see Section 4.1), respectively, for mass and age estimates. Most of the stars in our primary RC sample have spectral S/Ns higher than 30 and thus the uncertainties are dominated by $\sigma_m$.

Member stars of a given OC are usually believed to form from a single giant molecular cloud and thus have similar ages, distances, kinematics, and chemical compositions. Therefore, they serve as good test beds to check the accuracy of our age determinations for primary RCs, as well as for distance and other parameter estimates yielded by LSP3 (Xiang et al. 2017b). To do so, seven OCs of a wide age range (1–10 Gyr) observed by the LAMOST surveys are selected. Then member stars of those OCs are picked out from our primary RC sample. As an example, the selection of member stars in the RC evolutionary stage for NGC 2682 is illustrated in Figure 10.
1. The stars must have LAMOST spectra of S/Ns higher than 60 and relative uncertainties of age, as determined above, less than 40%.

2. The stars must have LAMOST spectra of S/Ns higher than 60 and relative uncertainties of age, as determined above, less than 40%.

3. The photometric errors in the $K_s$ band must be smaller than 0.03 mag.

4. The stars have distances estimated by Schönrich, et al. (2019, hereafter SME19), using the Gaia DR2 parallaxes (Lindegren et al. 2018). Note that in the treatment of SME19, the Gaia parallax uncertainties have been increased by 0.043 mas in quadrature and have been corrected for a parallax offset of 0.054 mas.

5. Following the suggestions of SME19, further cuts on the Gaia measurements have been applied to minimize the biases in the estimated distances: (1) relative parallax uncertainties smaller than 10%, (2) parallax uncertainties smaller than 0.05 mas, (3) estimated distances greater than 80 pc, (4) G-band magnitudes smaller than 14.0 mag and $0.5 < G_{BP} - G_{RP} < 1.4$ mag, (5) the number of measurements of $n > 5$ and excess noise $< 1$, and (6) an $G_{BP} - G_{RP}$ excess flux factor of $1.172 < E_{BPRP} < 1.3$.

The first and last three cuts are to ensure precise determinations of $K_s$-band absolute magnitudes. The second one is to ensure low contamination from the secondary RCs and RGBs, as well as to ensure accuracies of the estimated metallicity and age. Finally, a total of 13,764 stars have been selected. Their $K_s$-band absolute magnitudes are derived using the distances estimated by SME19 from the Gaia DR2 parallaxes (Lindegren et al. 2018) and the $K_s$ magnitudes from 2MASS (Skrutskie et al. 2006) after reddening corrections. $K_s$-band absolute magnitudes of primary RCs as a function of age are presented in Figure 12 and a significant trend is detected for primary RCs of ages greater than 5 Gyr. By further grouping the stars into different metallicity bins, we show the median values of $M_K$ for the individual metallicity and age bins. Generally the trends of variations with metallicity and age are in excellent agreement with predictions of stellar evolution models.
agreement with the theoretical predictions of Salaris & Girardi (2002). For stars with age greater than 4–5 Gyr, the $K_s$-band absolute magnitudes of primary RCs decrease with age by few hundredth in magnitude per Gyr and increase with [Fe/H] also by few hundredth in magnitude per dex.\footnote{Note that for stars of [Fe/H] $>$ −0.10, the dependence of $M_{Ks}$ on [Fe/H] tends to be minor.} We note that a similar trend of $M_{Ks}$ with age is also found by Chen et al. (2017), based on $\sim$170 seismically identified RCs of solar metallicities in the Kepler field. For stars of ages younger than 5 Gyr, $M_{Ks}$ show weak dependence on both age and metallicity, with a median comparable to that Hipparcos local sample (e.g., L12). The typical standard deviation, after subtracting contribution from the photometric and astrometric uncertainties, is around 0.15 mag (i.e., 7% in distance). To consider both the metallicity and age effects, we use a third-order polynomial to fit $M_{Ks}$ as a function of age for different metallicity bins:

$$M_{Ks} = a - b\tau + c\tau^2 + d\tau^3. \tag{8}$$

Here $\tau$ is stellar age. The coefficients for the different metallicity bins are given in Table 3.

### Table 3

| [Fe/H] | $a$ | $b$ | $c$ | $d$ |
|--------|-----|-----|-----|-----|
| $>$−0.10 | −1.643 | −5.261×10$^{-3}$ | 2.045×10$^{-3}$ | −5.387×10$^{-3}$ |
| $>$−0.25 | −1.478 | −7.120×10$^{-2}$ | 1.103×10$^{-2}$ | −4.370×10$^{-4}$ |
| $>$−0.10 | −1.406 | −1.202×10$^{-1}$ | 1.938×10$^{-2}$ | −8.098×10$^{-4}$ |
| $>$−0.25 | −1.862 | 1.282×10$^{-2}$ | 3.979×10$^{-3}$ | −6.046×10$^{-4}$ |

#### Notes.

\footnote{From Dias et al. (2002).}

\footnote{From Cantat-Gaudin et al. (2018).}

\footnote{References: (1) Sandquist et al. (2016); (2) Molenda-Zakowicz et al. 2014; (3) Salaris et al. (2004); (4) Jacobson et al. (2011); (5) Anthony-Twarog et al. (2014); (6) Lee-Brown et al. (2015); (7) Brewer et al. (2016); (8) Hole et al. (2009); (9) An et al. (2007); (10) Heiter et al. (2014); (11) Stello et al. (2016); (12) Geller et al. (2015); (13) Meibom et al. (2009); (14) Geller et al. (2008); (15) An et al. (2015); (16) Tofflemire et al. (2014); (17) Bragaglia et al. (2006); (18) Friel et al. (2005).}

#### 5.2. Distances of Primary RCs

Using the newly calibrated relation, distances have been derived for all the selected primary RCs using the LSP3 values of [Fe/H], stellar ages have estimated in Section 4, and $K_s$-band magnitudes from 2MASS (Skrutskie et al. 2006), after applying corrections for the reddening as given by SFD98 for
stars of high latitudes ($|b| > 30^\circ$) or derived with the “star pair” technique (Yuan et al. 2013) for stars close to the disk plane ($|b| < 30^\circ$).

We further validate the distances estimated from our new calibration using the statistical method of Schöning et al. (2012, hereafter SBA12). As shown in the upper panel of Figure 13, the potential distance bias estimated with the SBA12 technique as a function of age for the different metallicity bins are all within a few (no more than 5%) percent. The results show that our new calibration is very accurate for primary RCs of all populations. In the bottom panel of Figure 13, we show the same results but plotted against the estimated distances. Again, the zero offsets of distances estimated with our new calibration are within 5% for essentially all the distance bins. For the distances derived assuming a constant value of $M_K$ as calibrated by L12 using Hipparcos local sample, the resulting distances are biased less than 5% for $d < 2$ kpc but overestimated by 5%–10% for $d > 2$ kpc. The results are easily understandable. The local RCs ($d < 2$ kpc) have ages and metallicities similar to those of the Hipparcos local sample adopted by L12, leading to minor distance biases. For more distant RCs ($d > 2$ kpc), their ages and metallicities may deviate significantly from those of the Hipparcos local sample, resulting in significant distance biases using the calibration of $M_K$ by L12. In addition, similar examinations are applied against the $J - K_s$ color and metallicity [Fe/H] of RC stars. The results again show that the systematic biases of the distance estimated by our new calibration are within 5% for almost all the color and metallicity bins (see Figure A1). We also examine the distances of primary RCs selected from the APOGEE DR14 (Bovy et al. 2014), using essentially the same technique described in the current work. In their work, the distances have been derived using the PARSEC stellar isochrones, scaled to the local calibration of L12 by adding a constant offset (see Section 3 of Bovy et al. 2014 for details). For the local RCs ($d < 2$ kpc), the distance biases are minor, on the level of a few percent. But to our surprise, the biases increase almost linearly with the distance for more distant RCs of $d > 2$ kpc and reach over 40% at $d \sim 4.5$ kpc.

6. Selection Function of the Sample

To assess the selection function of our primary RC sample, we follow the procedure of Chen et al. (2018). We calculate the selection function of the different fields (each field has an area around 3.36 sq. deg., partitioned with HEALPix) by comparing the number of stars observed with LAMOST
and with robust stellar parameter estimates to that given by the completed 2MASS photometric catalog in the \(J - K_s\) versus \(K_s\) plane. The final results are shown in Figure 14. One can see significant variations of target selection effects across the \(J - K_s\) versus \(K_s\) diagram. Generally, the fainter the magnitude and the redder the color, the lower sampling rate. The reasons for the trend are that we have used only blue-arm LAMOST spectra in deriving the stellar parameters and also applied a spectral S/N (at 4650 Å) cut (≥20) in selecting the primary RCs. For more details about the target selection and the selection function of LAMOST data, we refer to Chen et al. (2018).

7. Properties of the LAMOST RC Sample

7.1. Spatial Coverage

Figure 15 shows the number density distributions of our sample primary RCs in the \(X-Y\) and \(X-Z\) planes.\(^{14}\) The sample covers a large volume of the Galactic disk, with \(R\) (the projected Galactocentric distance in a Galactocentric cylindrical system) ranging from 4 to 16 kpc, \(Z\) from −5 to 5 kpc, and \(\phi\) (in the direction of Galactic rotation) from −20 to 50°. This large volume should certainly allow one to probe the structural, chemical, and kinematic properties of both the Galactic thin and thick disks.

7.2. Age Distribution

The distribution of mean stellar ages for 2°.5 × 2°.5 projected patches on the sky in the Galactic coordinate system is presented in Figure 16. As expected, a positive age gradient with an increasing absolute Galactic latitude (|\(b|\)) is clearly seen. Generally, the mean age is younger than 3 Gyr for \(|b| < 10^\circ\) and can be as old as 8 Gyr at higher latitudes, e.g., \(|b| > 50^\circ\). Figure 17 shows the mean [\(\alpha/\text{Fe}\)] ratio and age of stars at different positions across the \(R-Z\) plane of the Galactic disk. Generally, both the mean [\(\alpha/\text{Fe}\)] ratio and age show a negative gradient in the radial direction and a positive gradient in the vertical direction. Strong flaring of the young Galactic disk is clearly detected in both distributions, with younger populations (indicated either by the [\(\alpha/\text{Fe}\)] ratio or the stellar age) extended to higher heights of the Galactic plane as \(R\) increases.

The radial and vertical age gradients of the Galactic disk have been reported by Casagrande et al. (2016) and

\(^{14}\) Here \(X\), \(Y\), and \(Z\) are axes of a Galactocentric, right-handed Cartesian system, with \(X\) pointing in the direction opposite to the Sun, \(Y\) in the direction of Galactic rotation, and \(Z\) toward the north galactic pole.
Martig et al. (2016b), respectively. Using a large sample of stars with ages derived from the LAMOST spectra, Xiang et al. (2017a) and Wu et al. (2019) present similar age maps similar to those reported here. By carefully considering the selection effects of the primary RC sample (see Section 6), the structure of the Galactic disk of different mono-abundance or mono-age populations can be studied in detail with the current sample (Z. Yu et al. 2020, in preparation).

7.3. Age, [Fe/H], and [α/Fe] Relations

Figure 18 plots the distributions of stellar number densities and mean stellar ages in the [Fe/H]–[α/Fe] plane. As expected, a bimodal distribution of [α/Fe] is clearly seen. By applying an empirical cut in the [Fe/H]–[α/Fe] plane, one finds that 15% stars belong to the high [α/Fe] sequence, i.e., associated with the so-called chemical thick disk. In Figure 18, stars of the high [α/Fe] sequence typically have ages older than 9–10 Gyr. In contrast, stars of the low [α/Fe] sequence have much younger ages. Our results are in excellent agreement with the previous finding for the solar neighborhood based on high-resolution spectroscopy (e.g., Fuhrmann 1998; Bensby et al. 2003; Haywood 2008; Haywood et al. 2013; Hayden et al. 2015).

The middle panels of Figure 18 show the distributions of stellar number densities and mean values of [α/Fe] in the age–[Fe/H] plane. Generally young (<8 Gyr) populations show a wide [Fe/H] range while there is an obvious lack of metal-rich stars in the old (≥8 Gyr) populations. Such an age–metallicity relation is consistent with the results found for stars in the solar neighborhood (e.g., Haywood et al. 2013; Bergemann et al. 2014) and also from other large samples of stars (e.g., Xiang et al. 2017a and Wu et al. 2019). The lack of a tight age–metallicity relation for young populations could be explained by the effects of stellar radial migration (e.g., Sellwood & Binney 2002; Schönrich & Binney 2009).

Finally, the bottom panels of Figure 18 show the distributions of stellar number densities and mean values of [Fe/H] in the age–[α/Fe] plane. Two sequences are clearly seen—one of the young (<8–9 Gyr) populations with low [α/Fe] ratios and another of older (≥8 Gyr) populations with high values of [α/Fe]. The underlying physics leading to the two sequences is still under hot debate (e.g., Haywood et al. 2013) and our current sample should certainly help solve the issue. In addition to the two sequences, an obvious excess of stars of age ≤6 Gyr and [α/Fe] ≥ 0.15 (the so-called young α-enhanced stars; e.g., Chiappini et al. 2015; Martig et al. 2015) is detected in the age–[α/Fe] plane. This pattern is also seen in our training sample (see the bottom left panel of Figure 6). We will discuss the origin(s) of those stars in a separate paper (W.-X. Sun et al. 2020, in preparation).

Before summarizing, it is worth mentioning that we have started a series of studies of the Galactic disk(s) using this primary RC sample, including: (1) mapping the 3D asymmetrical kinematics and detecting new substructures in the Galactic disk (Wang et al. 2019, 2020), (2) determining the structural properties of different mono-abundance disk populations (Yu et al. 2020), (3) detecting kinematic signatures of the Galactic warp (Li et al. 2020), and (4) exploring the origin(s) of the so-called young α-enhanced stars (W.-X. Sun et al. 2020, in preparation).

8. Summary

Based on stellar atmospheric parameters deduced with the latest version of LSP3 for whole LMDR4, nearly 140,000 primary RC stars of spectral S/Ns higher than 20 have been successfully singled out, based on their positions in the metallicity-dependent effective temperature–surface gravity and color–metallicity diagrams, supervised by a high-quality asteroseismology data set. Based on the various tests, a purity and a completeness of over 85% have been achieved for this current sample of primary RCs.

Using the relations trained by a KPCA method with thousands of RC stars in the LAMOST–Kepler fields that have accurate asteroseismic mass measurements, stellar masses, and ages are further determined from the LAMOST.
spectra for our primary RC sample stars. Various tests show typical uncertainties of 15% and 30%, respectively, for the estimated masses and ages.

Using over 10,000 primary RCs with accurate distance measurements from the Gaia DR2 parallaxes, the $K_s$-band absolute magnitudes of the primary RC are recalibrated by

---

Figure 18. Distributions of our RC sample stars in the (i) $\rm{[Fe/H]}-\rm{[\alpha/Fe]}$ plane, binned by 0.025 dex by 0.01 dex and color coded by the number density (panel (a)) and mean age (panel (b)); (ii) $\rm{[Fe/H]}$–age plane, binned by 0.25 Gyr by 0.025 dex and color coded by the number density (panel (c)) and mean $\rm{[\alpha/Fe]}$ (panel (d)); and (iii) $\rm{[\alpha/Fe]}$–age plane, binned by 0.25 Gyr by 0.01 dex and color coded by the number density (panel (e)) and mean $\rm{[Fe/H]}$ (panel (f)). The gray line in panels (a) and (b) represent an empirical cut to separate the chemical thin and thick disk stars. Stars with $\rm{[\alpha/Fe]} > 0.125$ and age younger than 5 Gyr (the so-called young $\alpha$-enhanced stars, see Section 7.3 for details) are discarded when drawing panel (b).
considering the effects of both metallicity and age, for the first time. With this new calibration, very accurate distances are derived for the whole sample, with typical uncertainties of 5–10%, which are even better than the Gaia measurements for stars beyond 3–4 kpc.

The sample covers a significant volume of the Galactic disk of $4 \leq R \leq 16$ kpc, $|Z| \leq 5$ kpc, and $-20 \leq \phi \leq 50^\circ$. Stellar parameters, line-of-sight velocities, and elemental abundances deduced from the LAMOST spectra and proper motions from the Gaia DR2 are also provided in the whole sample stars. The sample is of vital importance to probe the structural, chemical, and kinematic properties of the Galactic disk(s) and is available at https://zenodo.org/record/3875974.

The 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 has made use of data from the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium).

It is a pleasure to thank J. Binney for a thorough read of the manuscript and helpful comments.

This work is supported by National Natural Science Foundation of China grants 11903027, 11973001, 11833006, 11811530289, U1731108, and U1531244; and National Key R&D Program of China No. 2019YFA0405503. Y.H. is supported by the Yunnan University Grant C176220100007. R. S. is supported by a Royal Society University Research Fellowship. H.F.W. is supported by the LAMOST Fellow project, National Key Basic Research Program of China via SQ2019YFA040021 and funded by China Postdoctoral Science Foundation via grant 2019M653504 and Yunnan province postdoctoral Direct Culture Foundation. H.F.W. is also supported by the Cultivation Project for LAMOST Scientific Payoff and Research Achievement of CAMS-CAS.

Appendix

Distance Biases versus Color and Metallicity of RCs

The biases of the distances estimated by our new calibration (see Section 5.1) against the color $J - K_s$ and metallicity $[\text{Fe}/\text{H}]$ of RCs are shown in Figure A1.

Figure A1. Upper panel: the relative error vs. $J - K_s$ for the LAMOST primary RCs of $S/N > 60$. The binning scheme is the same as for Figure 13. Lower panel: the relative error vs. $[\text{Fe}/\text{H}]$ for the LAMOST primary RCs of $S/N > 60$. The binning scheme is again the same as for Figure 13.
