



SPar: Estimating Stellar Parameters from Multiband Photometries with Empirical Stellar Libraries

Mingxu Sun¹ , Bingqiu Chen² , Helong Guo² , He Zhao³ , Ming Yang⁴ , and Wenyuan Cui¹ ¹ Department of Physics, Hebei Key Laboratory of Photophysics Research and Application, Hebei Normal University, Shijiazhuang 050024, People's Republic of China; bchen@ynu.edu.cn² South-Western Institute for Astronomy Research, Yunnan University, Kunming 650500, People's Republic of China³ Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023, People's Republic of China⁴ Key Laboratory of Space Astronomy and Technology, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, People's Republic of China

Received 2023 May 23; revised 2023 July 6; accepted 2023 July 8; published 2023 August 23

Abstract

Modern large-scale photometric surveys have provided us with multiband photometries of billions of stars. Determining the stellar atmospheric parameters, such as the effective temperature (T_{eff}) and metallicities ($[\text{Fe}/\text{H}]$), absolute magnitudes (M_G), distances (d), and reddening values ($E(G_{\text{BP}} - G_{\text{RP}})$) is fundamental to study the stellar populations, structure, kinematics, and chemistry of the Galaxy. This work constructed an empirical stellar library that maps the stellar parameters to multiband photometries from a data set with Gaia parallaxes, LAMOST atmospheric parameters, and optical to near-infrared photometry from several photometric surveys. Based on the stellar library, we developed a new algorithm, SPar (Stellar Parameters from multiband photometry), which fits the multiband stellar photometries to derive the stellar parameters (T_{eff} , $[\text{Fe}/\text{H}]$, M_G , d , and $E(G_{\text{BP}} - G_{\text{RP}})$) of the individual stars. The algorithm is applied to the multiband photometric measurements of a sample of stars selected from the SMSS survey, which have stellar parameters derived from the spectroscopic surveys. The stellar parameters derived from multiband photometries by our algorithm are in good agreement with those from the spectroscopic surveys. The typical differences between our results and the literature values are 170 K for T_{eff} , 0.23 dex for $[\text{Fe}/\text{H}]$, 0.13 mag for M_G , and 0.05 mag for $E(G_{\text{BP}} - G_{\text{RP}})$. The algorithm proved to be robust and effective and will be applied to the data of future large-scale photometric surveys such as the Mephisto and CSST surveys.

Unified Astronomy Thesaurus concepts: [Interstellar dust extinction \(837\)](#); [Low mass stars \(2050\)](#)

1. Introduction

Modern large-scale photometric surveys, such as the Sloan Digital Sky Survey (SDSS; York et al. 2000), the Pan-STARRS 1 Survey (Chambers et al. 2016), the Two Micron All Sky Survey (2MASS; Skrutskie et al. 2006), and the Wide-Field Infrared Survey Explorer (WISE; Wright et al. 2010), have provided us broadband photometry covering the wavelength from the optical to the infrared (IR) bands of billions of stars. The multiband photometric measurements of stars contain the spectral energy distribution (SED) information of these stars, from which we can obtain the stellar atmosphere parameters (i.e., the effective temperature T_{eff} , the metallicity $[\text{Fe}/\text{H}]$, and the surface gravity $\log g$), as well as the distance and extinction values of stars (Berry et al. 2012; Chen et al. 2014; Green et al. 2014, 2015).

Stellar atmospheric parameters are typically determined from their spectra. While large-scale multifiber spectroscopic surveys like the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST; Luo et al. 2015), the Apache Point Observatory Galactic Evolution Experiment (APOGEE; Abdurro'uf et al. 2022), the Galactic Archaeology with HERMES (GALAH; Buder et al. 2021), and the Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration et al. 2022) have provided spectra for tens of millions of stars, we

can now determine the parameters of billions of stars with comparable precision using photometric data of exceptionally high accuracy or data obtained through specially designed filters. This approach allows us to obtain accurate stellar parameters without relying solely on spectroscopic data. For example, Xu et al. (2022) have measured metallicities of ~ 27 million stars from the Gaia Early Data Release 3 (Gaia EDR3; Gaia Collaboration et al. 2021) photometric data of unprecedented millimagnitude precision. The typical metallicity precision is about $\delta[\text{Fe}/\text{H}] = 0.2$ dex. Based on the narrowband photometry from the SkyMapper Southern Survey (SMSS; Wolf et al. 2018; Onken et al. 2019), Huang et al. (2022) have determined stellar atmospheric parameters for ~ 24 million stars. The precision of their metallicity estimates has typical values around 0.05–0.15 dex. Yang et al. (2022) obtained stellar parameters (effective temperature, T_{eff} , surface gravity, $\log g$, and metallicity, $[\text{Fe}/\text{H}]$) from the narrowband photometries of the J-PLUS survey (Cenarro et al. 2019). They have achieved precisions of $\delta T_{\text{eff}} \sim 55$ K, $\delta \log g \sim 0.15$ dex, and $\delta [\text{Fe}/\text{H}] \sim 0.07$ dex, respectively. We note that these works are only for stars located at high Galactic latitudes, where the interstellar extinction is small. For stars at low Galactic latitudes, where the extinction effects are large, the stellar atmospheric parameters need to be estimated along with the stellar extinction values.

The future Multi-channel Photometric Survey Telescope (Mephisto; Yuan et al. 2020) photometric survey and the Chinese Space Station Telescope (CSST; Zhan 2011, 2018; Cao et al. 2018) optical survey will have both high precision

and specially designed filters, which will provide great opportunities for us to obtain accurate stellar parameters of billions of stars. Mephisto is a wide-field survey telescope with a 1.6 m primary mirror. It is equipped with three CCD cameras and is capable to image the same patch of sky in three bands simultaneously, which will provide us with real-time colors of stars with unprecedented accuracy. The filters of Mephisto (*uvgriz*) are similar to those of SkyMapper (Bessell 2001). The Mephisto-W Survey (Lei et al. 2021, 2022; B. Chen et al. 2023, in preparation) will target the northern sky of decl. between -21° and 75° with a coverage of over $27,000 \text{ deg}^2$. CSST is a 2 m space survey telescope, which shares the same orbit as the Chinese Space Station. The CSST optical survey (CSST-OS) will observe a large sky area of $\sim 18,000 \text{ deg}^2$ in seven photometric filters (*NUV*, *u*, *g*, *r*, *i*, *z*, and *y*) covering the wavelengths from the near-ultraviolet (NUV) to near-infrared (NIR).

Deriving the stellar atmospheric parameters as well as the distance and extinction values is a fundamental task for the Mephisto and CSST surveys. In this work, we present a new algorithm, SPar (Stellar Parameters from multiband photometry), to estimate stellar atmospheric parameters, such as the effective temperature (T_{eff}) and metallicities ($[\text{Fe}/\text{H}]$), absolute magnitudes (M_G), distances (d), and reddening values ($E(G_{\text{BP}} - G_{\text{RP}})$) of a large sample of stars from their multiband photometries. Previous algorithms such as the Star-Horse (Anders et al. 2019) and the General stellar Parameterized from photometry (GSP-PHOT; Andrae et al. 2023) rely on the theoretical stellar models, which may suffer systematic effects (Green et al. 2021). One method of dealing with these inaccuracies in theoretical models is to apply empirical corrections based on the observed photometry of stars of known type. Berry et al. (2012), Chen et al. (2014), and Green et al. (2015) measure stellar parameters and extinction based on the empirical stellar locus in the color-color space. Sun et al. (2018); Chen et al. (2019a), and Sun et al. (2021) calculated the intrinsic colors and reddening values of the individual stars based on a spectroscopic sample selected from the spectroscopic surveys. In this work, we will construct an empirical stellar library that maps the stellar parameters to multiband photometries from a data set with Gaia parallaxes, LAMOST atmospheric parameters, and optical to NIR photometry from several photometric surveys.

The paper is structured as follows: Section 2 presents the relevant data set. Section 3 describes in detail the empirical stellar library we constructed. Section 4 describes our algorithm and Section 5 tests it. Finally, Section 6 summarizes the algorithm.

2. Data

As the Mephisto and CSST surveys have not yet started, in the current work we have used photometric data from the SMSS survey for the experiment. This work is based on the Gaia Data Release 3, the broadband photometry from SMSS, 2MASS (Skrutskie et al. 2006), and WISE (Wright et al. 2010), and the spectroscopic data from LAMOST, APOGEE, and GALAH.

The SMSS is an ongoing photometric survey of the Southern sky (Wolf et al. 2018; Onken et al. 2019). The survey depth is between 19.7 and 21.7 mag in six optical bands: *u*, *v*, *g*, *r*, *i* and *z*. In this work, we adopt the data from its second data release (SMSS DR2; Onken et al. 2019). The photometry has a

internal precision of 1% in the *u* and *v* bands, and 0.7% in the other four bands (*g*, *r*, *i*, and *z*).

To break the degeneracy of effective temperature (or intrinsic colors) and extinction for the individual stars, we combine the SMSS photometry with the IR photometry of 2MASS and WISE. The 2MASS survey is a full-sky survey undertaken in three filters: *J*, *H*, and *K_s*. The systematic errors of the 2MASS photometry are estimated to be less than 0.03 mag (Skrutskie et al. 1997). The WISE survey is a full-sky survey undertaken in four bands: *W1*, *W2*, *W3*, and *W4*. In the current work, we adopt the AllWISE Source Catalog (Kirkpatrick et al. 2014). We use only the data in the *W1* and *W2* bands, as the *W3* and *W4* measurements have lower sensitivities and poorer angular resolutions.

The Gaia DR3 (Gaia Collaboration et al. 2023) photometric data and parallax measurements are also adopted to derive the stellar parameters when available. Gaia DR3 was released by the Gaia mission (Gaia Collaboration et al. 2016). It contains more than a billion sources with five astrometric parameters (position, parallaxes, and proper motions) and three-band photometry (*G*, *G_{BP}*, and *G_{RP}*). The median uncertainty of the parallax is $\sim 0.02\text{--}0.03 \text{ mas}$ for $G < 15 \text{ mag}$ sources, 0.07 mas at $G = 17 \text{ mag}$, and 0.5 mas at $G = 20 \text{ mag}$ (Gaia Collaboration et al. 2021). At $G = 20 \text{ mag}$, the typical uncertainties for the Gaia DR3 photometries are 6, 108, and 52 mmag for the Gaia *G*, *G_{BP}*, and *G_{RP}* bands, respectively (Gaia Collaboration et al. 2021).

The LAMOST, APOGEE, and GALAH spectroscopic data are adopted in the current work for two aims: to construct the empirical stellar library and to validate the resulting stellar properties. In this work, we use the ‘‘LAMOST LRS Stellar Parameter Catalog of A, F, G, and K Stars’’ from the LAMOST data release 8 (LAMOST DR8; Luo et al. 2015). The catalog contains stellar atmospheric parameters (T_{eff} , $\log g$ and $[\text{Fe}/\text{H}]$) derived from over 6 million low-resolution spectra. For the APOGEE data, we use the APOGEE stellar parameters catalog from the SDSS data release 17 (SDSS DR17; Abdurro’uf et al. 2022). The catalog contains stellar atmospheric parameters (T_{eff} , $\log g$, and $[\text{M}/\text{H}]$) derived from over 0.6 million NIR spectra. For the GALAH data, we adopt its DR3 catalog (Buder et al. 2021), which contains stellar parameters (T_{eff} , $\log g$ and $[\text{Fe}/\text{H}]$) of over 0.5 million nearby stars.

3. Empirical Stellar Library

An empirical stellar library is first created based on a sample of stars selected from the LAMOST spectroscopic data and Gaia DR3, for which atmospheric parameters, distances, and extinction values can be well measured. We crossmatch the LAMOST and Gaia stars with the optical and NIR photometric data, i.e., the SMSS, 2MASS, and WISE, using a radius of $1''$. To exclude the stars with bad observations, a sample of stars are selected by the criteria as below:

1. LAMOST spectral signal-to-noise ratio (S/N) larger than 20 and effective temperature T_{eff} between 4000 and 8000 K;
2. All Gaia $GG_{\text{BP}}G_{\text{RP}}$ photometric errors smaller than 0.1 mag and $\text{phot_bp_rp_excess_factor} > 1.3 + 0.06(G_{\text{BP}} - G_{\text{RP}})^2$;
3. Photometric errors in any SMSS *uvgriz*, 2MASS *JHK_s*, and WISE *W1W2* bands less than 0.1 mag.

Table 1
Temperature- and Reddening-dependent Reddening Coefficients of the Individual Passbands

Band	C_0	C_1	C_2	C_3	C_4	C_5
SkyMapper u	$-8.847E-01$	$1.274E-05$	$1.175E-07$	$-4.431E-01$	$-3.718E-04$	$-4.343E-01$
SkyMapper v	$-7.932E-01$	$1.136E-04$	$8.852E-09$	$-3.385E-01$	$-1.186E-04$	$-2.533E-01$
SkyMapper g	$-1.103E-01$	$3.665E-05$	$-1.261E-08$	$-5.843E-03$	$2.783E-05$	$5.486E-03$
SkyMapper r	$4.524E-01$	$3.872E-05$	$-1.091E-08$	$-1.588E-02$	$2.014E-05$	$-1.428E-02$
SkyMapper i	$1.008E+00$	$1.128E-05$	$-1.602E-08$	$-1.277E-02$	$1.162E-05$	$-1.969E-03$
SkyMapper z	$1.352E+00$	$-8.916E-06$	$-8.834E-09$	$-2.468E-02$	$1.387E-05$	$2.657E-02$
Gaia G	$5.051E-01$	$-5.638E-05$	$8.313E-09$	$6.334E-02$	$-4.587E-06$	$-2.592E-03$
2MASS J	$1.787E+00$	$-3.056E-05$	$7.046E-09$	$-4.065E-03$	$9.801E-06$	$-2.193E-03$
2MASS H	$2.029E+00$	$-5.226E-05$	$8.943E-09$	$-1.573E-02$	$2.244E-05$	$1.496E-02$
2MASS K_s	$2.149E+00$	$-5.576E-05$	$1.018E-08$	$-7.487E-03$	$1.909E-05$	$9.137E-03$
WISE $W1$	$2.235E+00$	$-6.046E-05$	$1.438E-08$	$1.043E-02$	$1.891E-05$	$-3.675E-03$
WISE $W2$	$2.276E+00$	$-7.125E-05$	$1.153E-08$	$-2.593E-03$	$3.199E-05$	$9.640E-03$

Note. The function form is $R = E(G_{BP} - x)/E(G_{BP} - G_{RP}) = C_0 + C_1x + C_2x^2 + C_3y + C_4xy + C_5y^2$, where $x = T_{\text{eff}} - 6000$ and $y = E(G_{BP} - G_{RP}) - 0.5$.

To obtain the absolute magnitude of our selected stars in each filter, the extinction values in the individual filters and the distance of each star need to be calculated.

3.1. Correct the Extinction Effect of the Individual Stars

In this work, the reddening values of our sample stars are calculated with a star-pair method (Yuan et al. 2015). In selecting the control sample for extinction correction, this study impose more rigorous constraints on the S/N of the LAMOST spectrum and photometric accuracies compared to the empirical stellar library, and only low-extinction stars are chosen. The control stars are selected via the following criteria:

1. LAMOST spectral signal-to-noise ratio (S/N) larger than 50;
2. Gaia $GG_{BP}G_{RP}$ photometric errors smaller than 0.01 mag, all SMSS $uvgriz$, 2MASS JHK_s , and WISE $W1W2$ bands photometric errors less than 0.08 mag;
3. $E(B - V)$ values from the extinction map of Planck Collaboration et al. (2014) smaller than 0.025 mag.

For a given target star in our sample, its intrinsic colors, $(G_{BP} - x)_0$ (where x denotes to the magnitude in another band, i.e., G , G_{RP} , u , v , g , r , i , z , J , H , K_s , $W1$, or $W2$), are estimated simultaneously from the corresponding values of the pair stars in the control sample. The reddening values of $E(G_{BP} - x)$ of the target star are then obtained from the differences between its observed and intrinsic colors, i.e., $E(G_{BP} - x) = (G_{BP} - x) - (G_{BP} - x)_0$.

Stars with resulted $E(G_{BP} - x)$ errors larger than 0.1 mag are excluded from our catalog. Based on the resultant reddening values $E(G_{BP} - x)$, a T_{eff} - and reddening-dependent extinction law similar to the work of Zhang & Yuan (2023) has been built. Zhang & Yuan (2023) obtained the empirical reddening coefficients for the individual colors as a function of T_{eff} and $E(B - V)$. In the current work, we derive the empirical reddening coefficients, defined as $R(G_{BP} - x) = \frac{E(G_{BP} - x)}{E(G_{BP} - G_{RP})}$, as a function of T_{eff} and $E(G_{BP} - G_{RP})$. In the current work, we adopt a binary function:

$$R(G_{BP} - x) = C_0 + C_1x + C_2x^2 + C_3y + C_4xy + C_5y^2, \quad (1)$$

where $x = T_{\text{eff}} - 6000$ and $y = E(G_{BP} - G_{RP}) - 0.5$. The resulting reddening values we derived from the star-pair

method are used to fit Equation (1) to obtain the individual coefficients C_0 to C_5 . The fitting results are listed in Table 1.

Finally, we assume that $A_{W2}/E(G_{BP} - G_{RP}) = 0.063$ (Wang & Chen 2019) for the $W2$ band has the longest wavelength and experiences the least amount of extinction of all the bands. Combining with the extinction coefficient relations we derived, the extinction value in each filter for all our sample stars can be calculated from their $E(G_{BP} - G_{RP})$ reddening values. The extinction values are then subtracted to obtain the intrinsic magnitude of the stars. In the left and middle panels of Figure 1, we show the observed and intrinsic color-magnitude diagrams of our sample stars, respectively.

3.2. The Empirical HR Diagrams

We then obtain the absolute magnitudes of the sample stars. The distances of the sample stars are calculated from the Gaia DR3 parallaxes via a simple Bayesian approach (Bailer-Jones et al. 2018; Chen et al. 2019b; Shen et al. 2022). A simple posterior probability is adopted:

$$p(d|\varpi) = d^2 \exp\left(-\frac{1}{2\sigma_\varpi^2}(\varpi - \varpi_{zp} - \frac{1}{d})\right) p(d), \quad (2)$$

where σ_ϖ and ϖ_{zp} are respectively the errors and globular zero points of the Gaia parallaxes, and $p(d)$ the space density distribution prior for the sample stars. In the current work, a zero point of $\varpi_{zp} = -0.026$ mas from Huang et al. (2021) is adopted. The Galactic structure model of Chen et al. (2017) is adopted as the spatial density distribution prior. With the resultant distances, the absolute magnitude in each band is then calculated by $M_x = x - 5 \log d + 5 - A_x$. Stars with parallax errors larger than 20% are excluded. This leads to a final sample of 3,842,671 stars. In the sample, more than 3.5 million stars have all Gaia absolute magnitude estimates, over 3.2 million stars have all IR bands (2MASS JHK_s , and WISE $W1W2$) absolute magnitudes, and over 0.3 million stars have all SMSS absolute magnitude estimates. In the right panel of Figure 1, we show the resulting Hertzsprung-Russell diagram (HRD) of our final sample stars.

The final sample of stars we obtained above is too large and not evenly distributed across the parameter spaces. Therefore, we use this sample to create a gridded stellar library. To map the stellar parameters, namely effective temperature (T_{eff}),

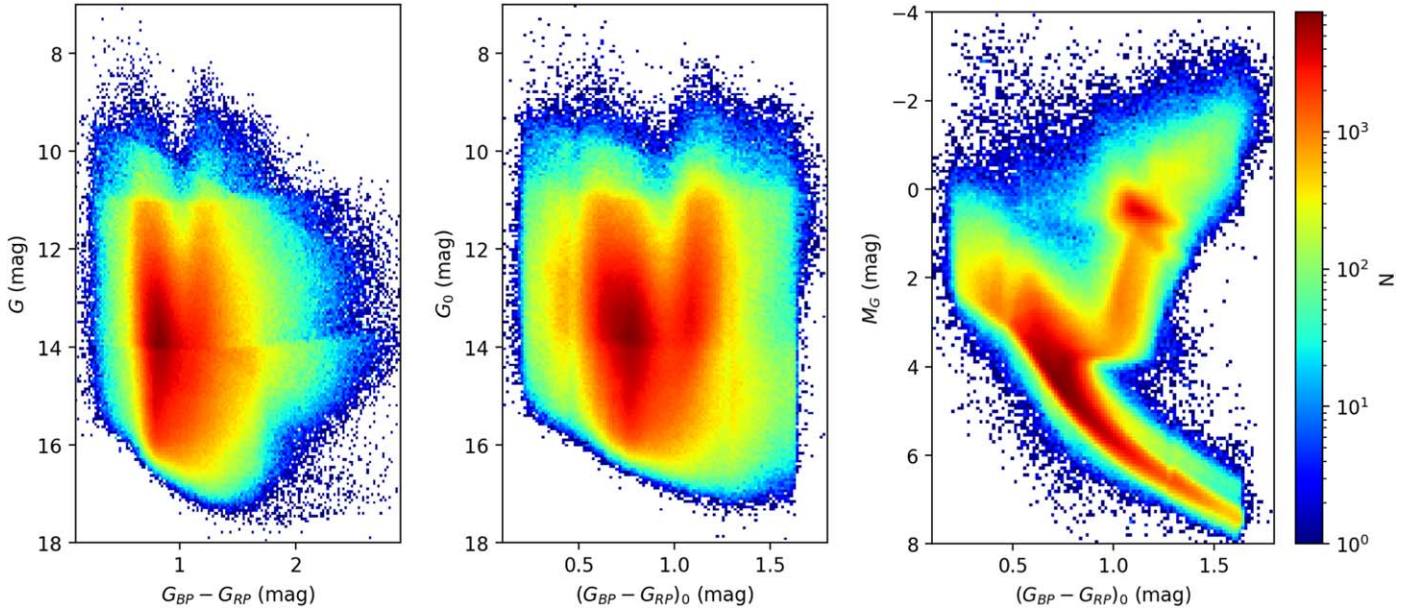


Figure 1. Left panel: Gaia observed color–magnitude diagram. Middle panel: Gaia dereddened color–magnitude diagram. Right panel: Gaia HR diagram.

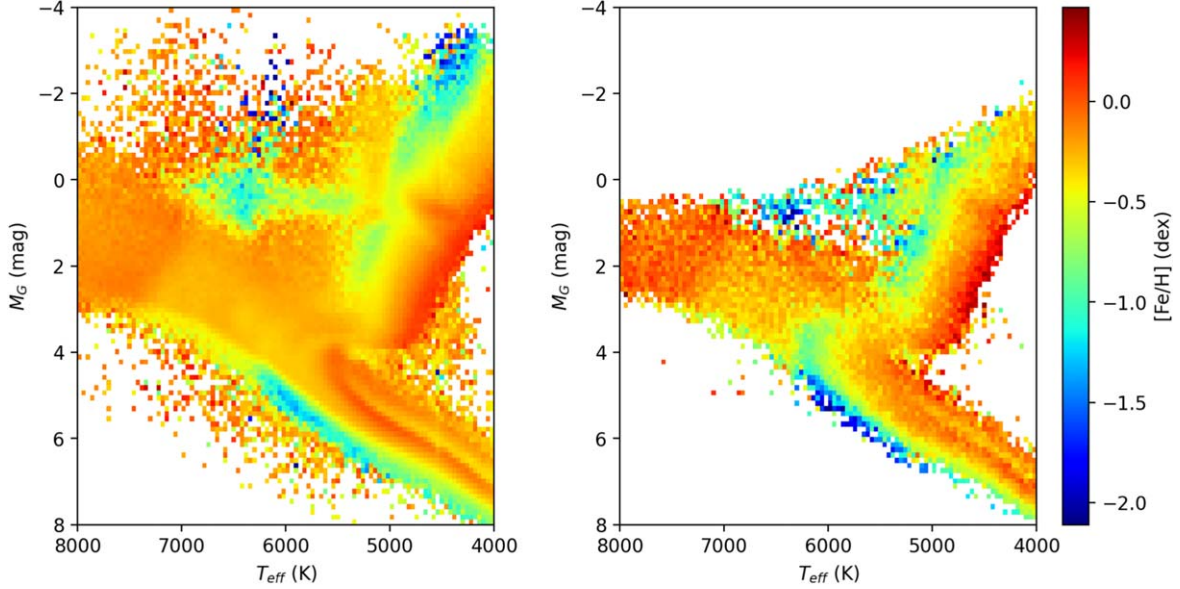


Figure 2. The HRDs plotted with our final sample stars (left) and the resultant gridded stellar library (right). The color scales represent the mean metallicities of each $T_{\text{eff}} - M_G$ bin.

metallicity ($[\text{Fe}/\text{H}]$), and Gaia G -band absolute magnitude (M_G), to absolute magnitude in each filter, we divide the parameter spaces into 100 bins for T_{eff} (ranging from 4000 to 8000 K), 50 bins for $[\text{Fe}/\text{H}]$ (ranging from -2.5 to 0.5 dex), and 100 bins for M_G (ranging from 8 to -4 mag). We use a machine-learning algorithm called random forest regression to obtain the absolute magnitudes in each passband for each bin, using the final sample stars as the training data set. We exclude any grids with fewer than five stars, resulting in an empirical stellar library with 39,905 grids in the parameter space. In Figure 2, we present the HRDs of both the final sample stars and the resultant gridded stellar library.

We performed a comparative analysis between our empirical stellar library and the PARSEC theoretical stellar isochrones (Bressan et al. 2012) as a means of verifying our results. To achieve this, we linearly interpolated the PARSEC absolute

magnitudes into our T_{eff} , $[\text{Fe}/\text{H}]$ and M_G grids, and subsequently compared them with our corresponding results. As illustrated in Figure 3, the results of this comparison indicate that our stellar library is in good agreement with the PARSEC theoretical models. Specifically, the mean values of the differences between our empirical library and the PARSEC models ranged between -0.04 and 0.04 mag, with dispersions of 0.04 and 0.03 mag observed for most of the filters. We note a slight increase in the dispersions for the WISE $W1W2$ bands, which ranged between 0.04 and 0.05 mag, and the SMSS uv bands, which show dispersions of 0.08 – 0.11 mag. We attribute the increase in dispersion to the relatively large calibration errors and uncertainties associated with the filter response curves of these particular filters. For the u band, we find a slightly higher mean difference of -0.065 mag than other filters. This difference could be due to a variety of factors,

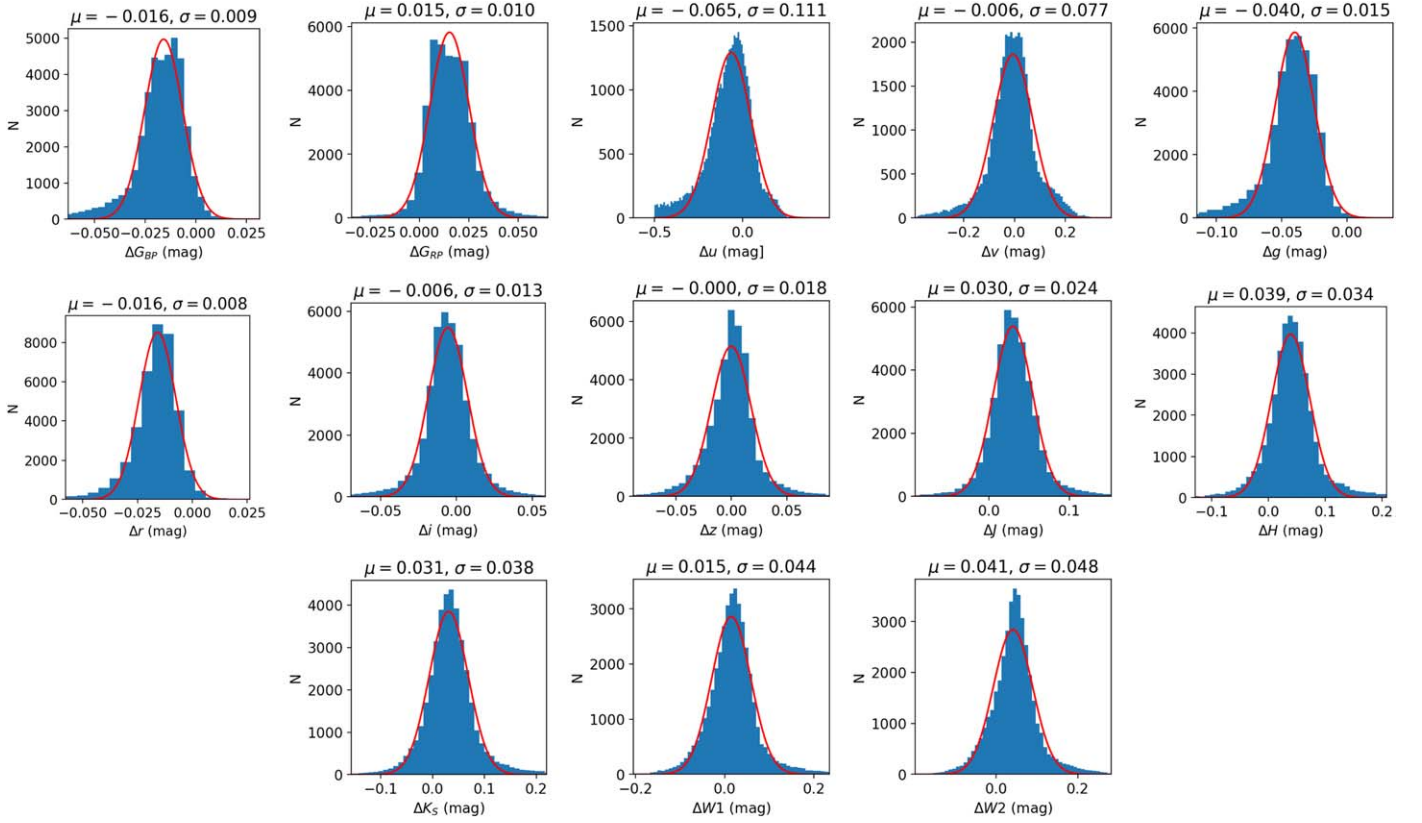


Figure 3. Histograms of the differences between our predicted absolute magnitudes of stars and those from the PARSEC isochrones (Bressan et al. 2012). The mean and dispersion values of the differences are labeled in each panel.

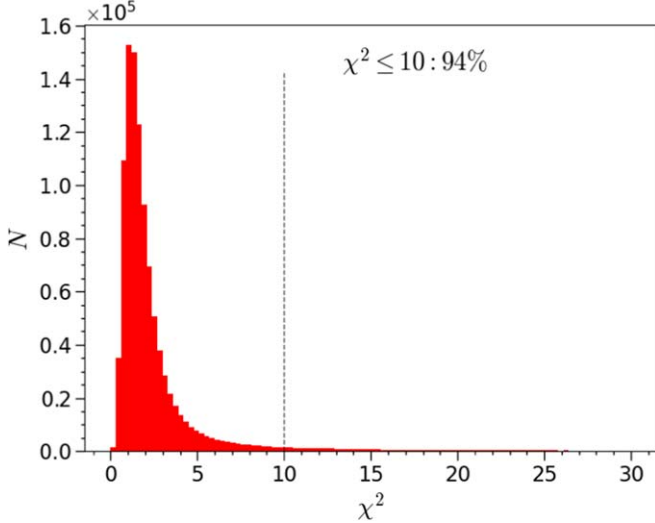


Figure 4. The distribution of χ^2 . The black dashed line is $\chi^2 = 10$ line.

including notable photometric error, extinction, and model uncertainty in the passband.

4. Estimating Stellar Parameters from Multiband Photometries

This section introduces how we derive the stellar parameters from the multiband photometries. To allow our algorithm SPAR to be applied to large samples of billions of stars, we need to minimize the computational cost. Our algorithm fits only four stellar parameters, T_{eff} , $[\text{Fe}/\text{H}]$, M_G , and $E(G_{\text{BP}} - G_{\text{RP}})$. The

distances d of stars can be further derived from the fitting results. SPAR uses an ensemble Markov Chain Monte Carlo (MCMC) method to obtain the best parameters of the individual stars by adopting a set of initial values derived from a minimum χ^2 method.

4.1. Initial Parameters from the Minimal χ^2 Method

Based on the reference stellar library and extinction law derived in Section 3, assuming a reddening value, we can predict the “distance module corrected” magnitudes in the individual filters M'_x of stars, i.e., $M'_x = M_x + A_x$. Distance module μ can then be derived by subtracting M'_x from the observed magnitude m_x of stars: $\mu = m_x - M'_x$. By substituting the resulting distance modulus into the standard magnitude equation, $m_x = M_x + A_x + \mu$, we can simulate the magnitude of the stars in the individual passbands. With T_{eff} , $[\text{Fe}/\text{H}]$, M_G , and $E(G_{\text{BP}} - G_{\text{RP}})$ as free parameters, we can model the stellar observed magnitude in each filter. We define

$$\chi^2 = \frac{1}{N - K} \sum_{x=1}^N \left(\frac{m_x^{\text{obs}} - m_x^{\text{mod}}}{\sigma_x} \right)^2, \quad (3)$$

where m_x^{obs} and m_x^{mod} are respectively the observed and simulated magnitudes of the filter x , σ_x are the photometric errors, N and K is the number of adopted filters and free

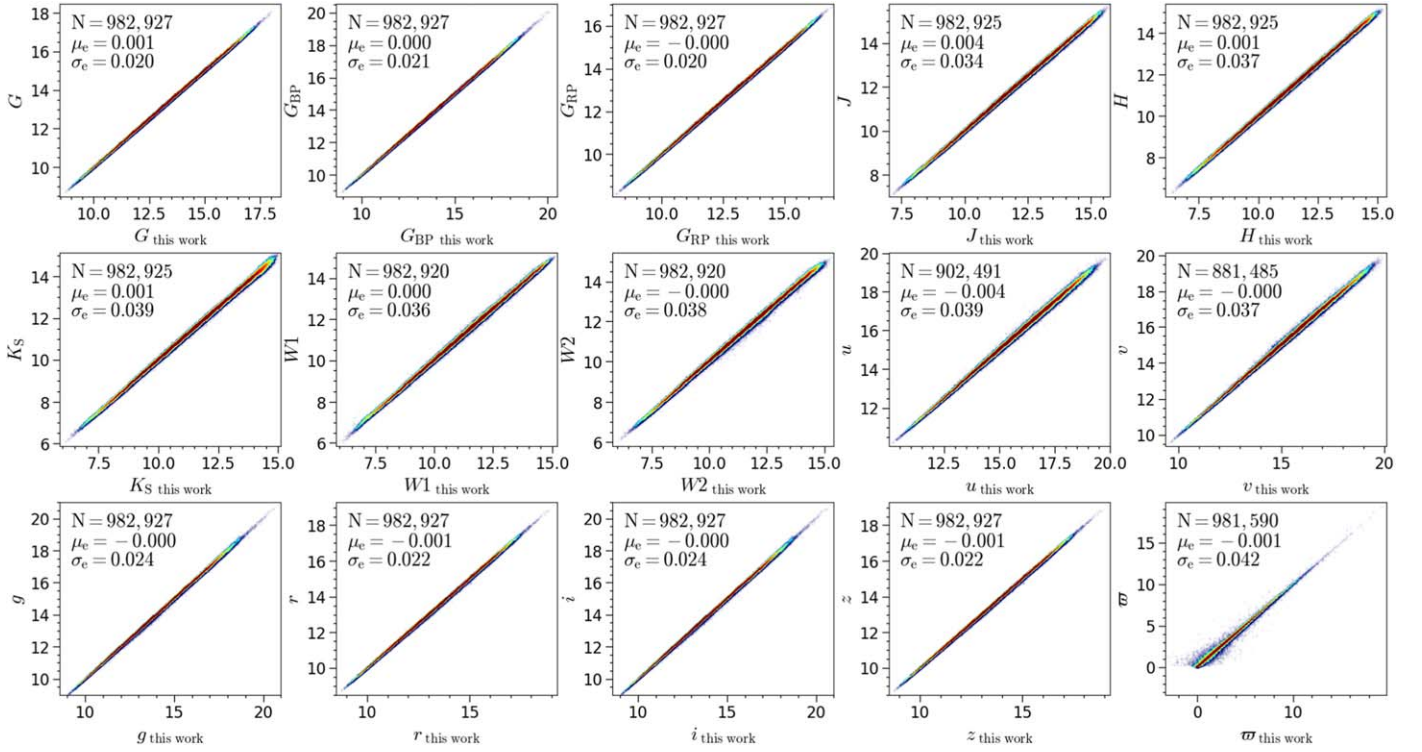


Figure 5. Comparison of simulated magnitude and parallax based on the MCMC results with the observed data. N is the total source number, and μ_e and σ_e are the median and standard deviation of differences after 3σ elimination.

parameters, respectively. If Gaia parallax exists,

$$\chi^2 = \frac{1}{N + 1 - K} \left(\sum_{x=1}^N \left(\frac{(m_x^{\text{obs}} - m_x^{\text{mod}})^2}{\sigma_x} \right) + \left(\frac{(\varpi_{\text{obs}} + \varpi_{\text{zp}} - \varpi_{\text{mod}})^2}{\sigma_{\varpi}} \right) \right), \quad (4)$$

where ϖ_{obs} and ϖ_{mod} are respectively the observed and simulated parallax, σ_{ϖ} is the parallax error, ϖ_{zp} ($\varpi_{\text{zp}} \equiv -0.026$ mas) is the zero point of the observed parallax.

We search for the minimal χ^2 parameters by running a series of $E(G_{\text{BP}} - G_{\text{RP}})$ values ranging from -0.1 to 6.0 in step of 0.02 mag and all grids in the reference stellar library. We use only the optical filters, i.e., Gaia $GG_{\text{BP}}G_{\text{RP}}$ and SMSS gri , to derive the distances of the stars. This procedure yields best-fit values of T_{eff} , $[\text{Fe}/\text{H}]$, M_G , and $E(G_{\text{BP}} - G_{\text{RP}})$ for the individual stars, which will be adopted as the initial parameters of the following MCMC analysis. If the resulted χ^2 are too large ($\chi^2 > 10$), this means that the stellar parameters in our template library do not fit the observed values well. It is possible that the star of concern is not a normal AFGK star, and then this star will not be used further in the subsequent MCMC analysis.

4.2. Final Parameters from the MCMC Analysis

In order to determine the final parameters and their uncertainties for the individual stars, we adopt the MCMC procedure described in Foreman-Mackey et al. (2013). The initial values for effective temperature T_{eff} , metallicity $[\text{Fe}/\text{H}]$, absolute magnitude M_G , and reddening $E(G_{\text{BP}} - G_{\text{RP}})$ are set according to the values derived from Section 4.1. The

maximum likelihood is defined as follows:

$$L = \prod_{x=1}^N \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left(-\frac{(X_x^{\text{obs}} - X_x^{\text{mod}})^2}{2\sigma_x^2}\right) \quad (5)$$

where X_x^{obs} and X_x^{mod} are respectively the observed and simulated magnitudes of the filter x , or parallax if available.

To run the MCMC analysis, we created 10 walkers and 20 steps chains, discarding the first five steps for burn-in purposes. The posterior distributions of the final parameters are determined by the 50th percentile values, and their uncertainties are obtained by computing the 16th and 84th percentile values.

5. Tests of Our Algorithm

In this section, we will test the SPar algorithm. We crossmatch the SMSS data with the LAMOST, APOGEE, and GALAH spectroscopic data, and have obtained a test sample of stars with parameter measurements from the spectroscopic data. The SPar algorithm is then applied to the sample stars to obtain their parameters, which are then compared with those obtained from the spectra. The test sample contains 1,046,722 stars, of which 408,671, 200,902, and 437,149 stars are from the LAMOST, APOGEE, and GALAH surveys, respectively.

Figure 4 displays the distribution of χ^2 values for all sources included in the test sample. A typical χ^2 distribution is observed, with a prominent peak at 2 and a long tail. Some stars in the sample exhibit a large χ^2 value, indicating that our empirical models are unable to effectively match their observed data. This may be due to the atmospheric parameters of these stars lying outside the range of our models. To optimize computational efficiency, we have excluded these stars from

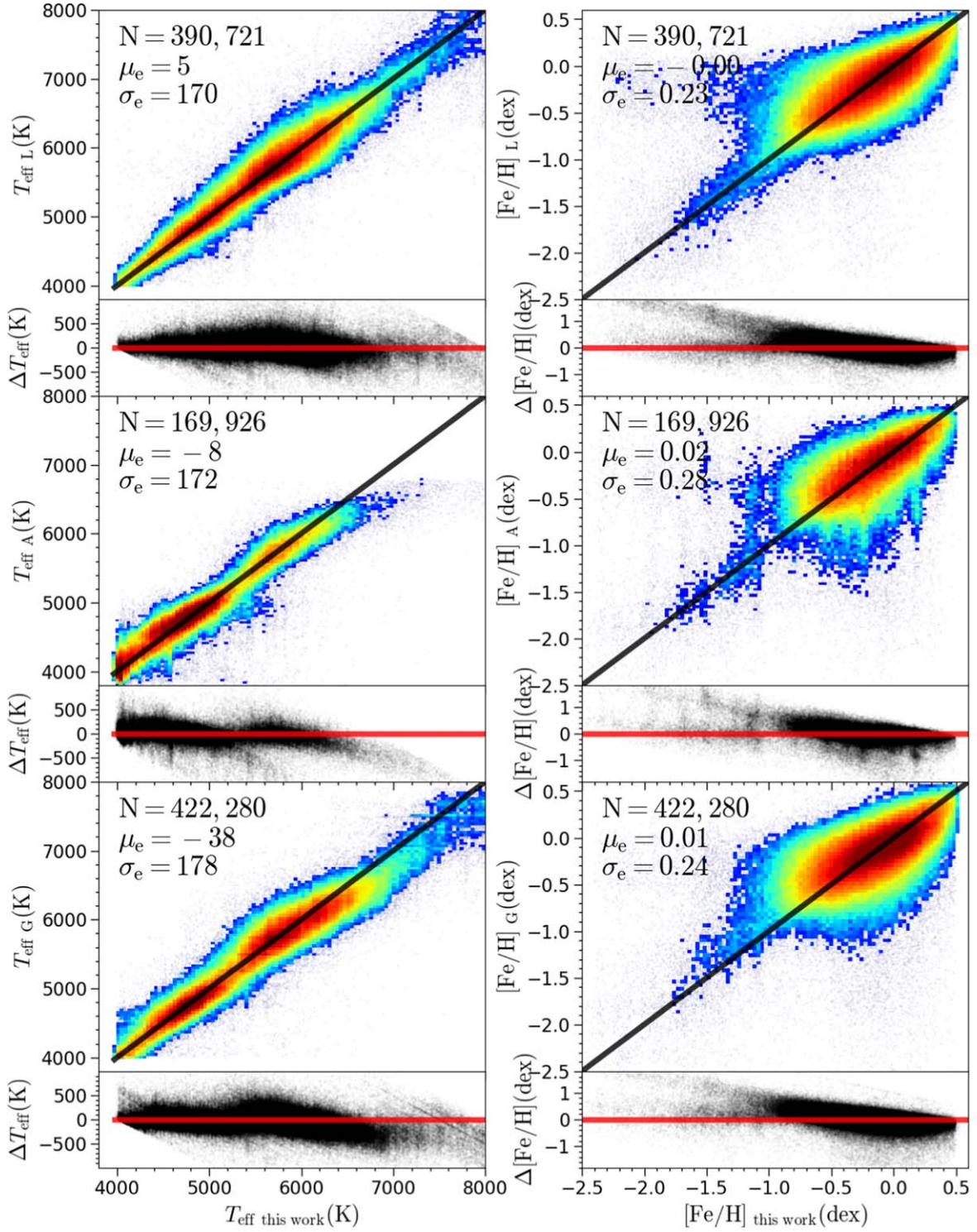


Figure 6. Comparison of T_{eff} (left) and [Fe/H] (right) of this work with LAMOST (top), APOGEE (middle), and GALAH (bottom). For every panel, the top half shows the consistency of T_{eff} or [Fe/H], and the bottom half shows the differences change with T_{eff} or [Fe/H] of this work. The black and red line in the top and bottom half panel are the equal and zero lines.

subsequent MCMC analysis. Consequently, by adopting a conservative χ^2 threshold of less than 10 in our study, we have retained 96% of LAMOST sources, 85% of APOGEE sources, and 97% of GALAH sources.

After conducting the MCMC analysis for the remaining 982,927 stars, we obtained their final parameters. To evaluate the accuracy of our algorithm, we first compared the

differences between our predictions and the observations. We simulated the observed magnitudes of the individual stars at various wavelengths and their parallaxes based on the MCMC results, which were compared with the observed data presented in Figure 5. Our simulations show good agreement with the observational results. The mean values of the differences are negligible. The dispersions of the differences are about 0.020-

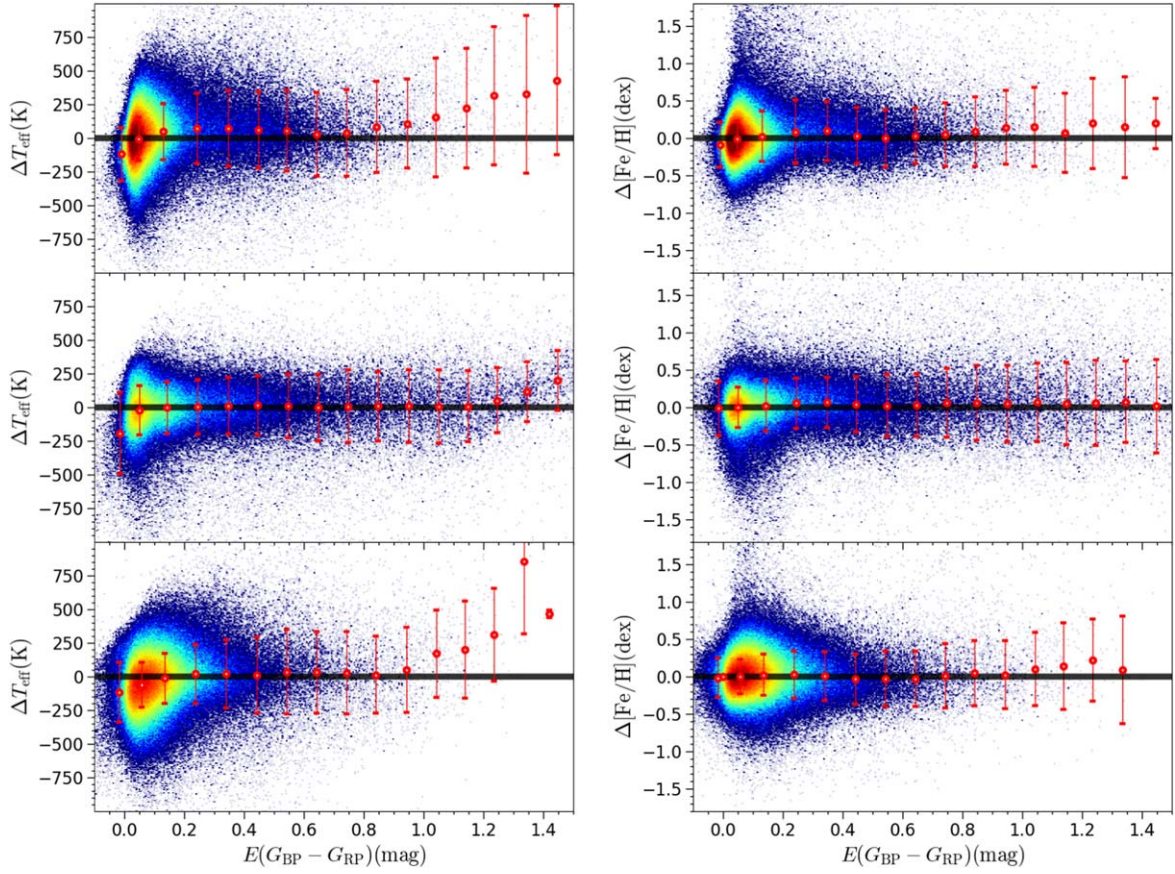


Figure 7. The changes of ΔT_{eff} (left) and $\Delta[\text{Fe}/\text{H}]$ (right) with $E(G_{\text{BP}} - G_{\text{RP}})$ of LAMOST (top), APOGEE (middle), and GALAH (bottom) stars. The red point and error bar is the median and standard deviation in every 0.1 mag bin. The black line is the zero line.

0.024 mag for the optical magnitudes (G , G_{BP} , G_{RP} , $griz$), 0.034–0.039 mag for the UV and NIR magnitudes ($uvJHK_S W1W2$), and 0.042 mas for parallax, respectively.

5.1. Comparison of the Resulted Parameters

In this section, we then compare the stellar parameters obtained by SPAR to those obtained from the spectroscopic surveys to test the accuracy of SPAR.

Figure 6 shows the comparisons of effective temperature T_{eff} and metallicities $[\text{Fe}/\text{H}]$ between the current work and the spectroscopic surveys, including LAMOST, APOGEE, and GALAH. The effective temperature from SPAR and LAMOST are in good agreement without any obvious trend of change with temperature. The dispersion of the difference between the effective temperatures is only 170 K. Compared to LAMOST, APOGEE has more low-temperature stars and fewer high-temperature stars. Some of the stars in APOGEE have T_{eff} below 4000 K, which is outside the temperature range of our empirical stellar templates. For those low-temperature stars, we would overestimate their temperatures. Regarding the GALAH data, our effective temperatures are systematically higher than the GALAH values for high-temperature stars. The possible reason for this is that the effective temperatures measured by GALAH are systematically higher than those measured by LAMOST for stars of high temperatures, as discussed in Buder et al. (2019).

The sensitivity of the SkyMapper uv filters to stellar metallicities enables accurate measurements of $[\text{Fe}/\text{H}]$ based on the SMSS multiband photometric data, as illustrated in

Figure 6. Our resulting metallicities show good agreement with those obtained from the LAMOST, APOGEE, and GALAH surveys, even for extremely metal-poor stars with $[\text{Fe}/\text{H}] \sim -2.5$ dex. However, the dispersion of the differences is relatively high, with values of 0.23, 0.28, and 0.24 dex, respectively, for LAMOST, APOGEE, and GALAH measurements. These values are larger than the dispersion values reported by Huang et al. (2022; 0.05–0.15 dex). This difference may be attributed to the fact that Huang et al. (2022) restricted their analysis to stars at high Galactic latitudes, where dust extinction values are small and readily derived from two-dimensional extinction maps. In contrast, many of the stars in our sample are located in the Galactic disk, which is subject to high extinction effects. Therefore, errors arising from dust extinction hinder accurate determination of the intrinsic colors of stars, leading to relatively large uncertainties in the derived metallicities.

We plotted the differences in effective temperature and metallicities between our work and the spectroscopic surveys against the reddening values of our sample stars in Figure 7. As the reddening values increase, the dispersions of the differences in the two parameters also increase. On average, the mean values of the differences do not vary with the reddening. However, for highly reddened stars, the mean value of the differences in the effective temperature deviates from zero. This may be due to the small number of stars in such regions and the relatively larger errors associated with them.

In addition to the stellar parameters, reddening $E(G_{\text{BP}} - G_{\text{RP}})$ and distance are also results from SPAR. We

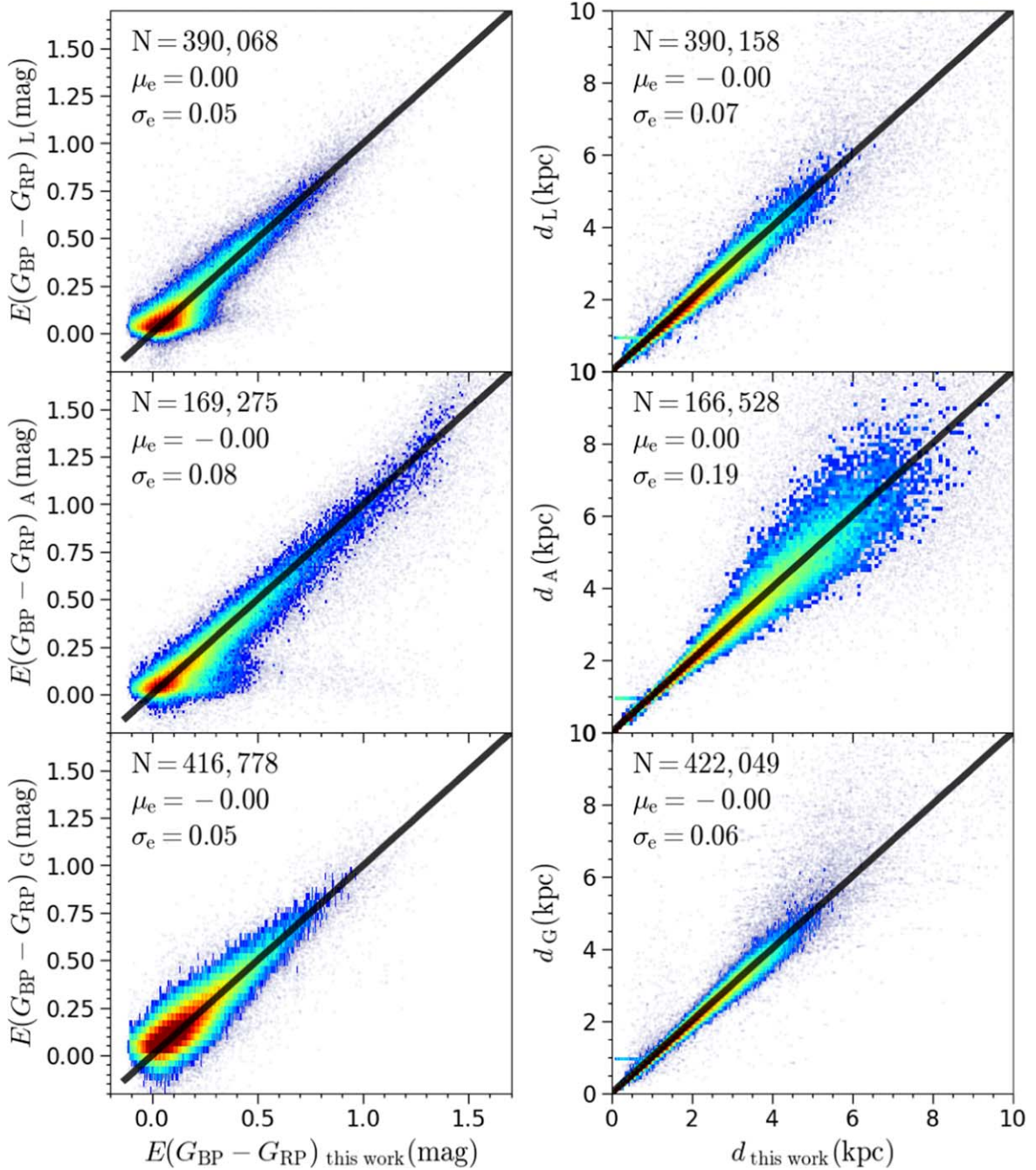


Figure 8. The same as Figure 5, but for $E(G_{\text{BP}} - G_{\text{RP}})$ and distance. The black line is the equal line.

compare our resultant $E(G_{\text{BP}} - G_{\text{RP}})$ values with those derived from the star-pair method, our derived distances, and those from Gaia DR3 parallaxes for all stars in the test sample in Figure 8. For $E(G_{\text{BP}} - G_{\text{RP}})$, the consistencies are good. There are no offsets between our results and those from the spectroscopic results. The dispersion values of the differences for the LAMOST and GALAH stars are about 0.05 mag. While for the APOGEE stars, the dispersion is larger, of about 0.08 mag. This is partly due to the high proportion of stars with large reddening values in APOGEE, and partly due to the fact that there are some low-temperature stars in APOGEE with temperatures below 4000 K. As mentioned above, we may overestimate their effective temperatures, which would lead us to overestimate their reddening values at the same time.

Regarding the distance measurements, our results are consistent with the Gaia measurements, with no significant offsets observed. The dispersions of the relative differences for both LAMOST and GALAH stars are only around 7%, while for APOGEE stars, the value is about 19%. This is because LAMOST and GALAH stars, being mainly dwarfs that are relatively close to us, have accurate parallax measurements from Gaia, resulting in smaller relative distance dispersions. Conversely, the APOGEE catalog contains many giant stars that are further away, leading to larger parallax measurement errors and, therefore, a larger dispersion in relative distance measurements.

Finally, we compared the absolute magnitudes in the Gaia G band, M_G values obtained from SPAR, with those from the three

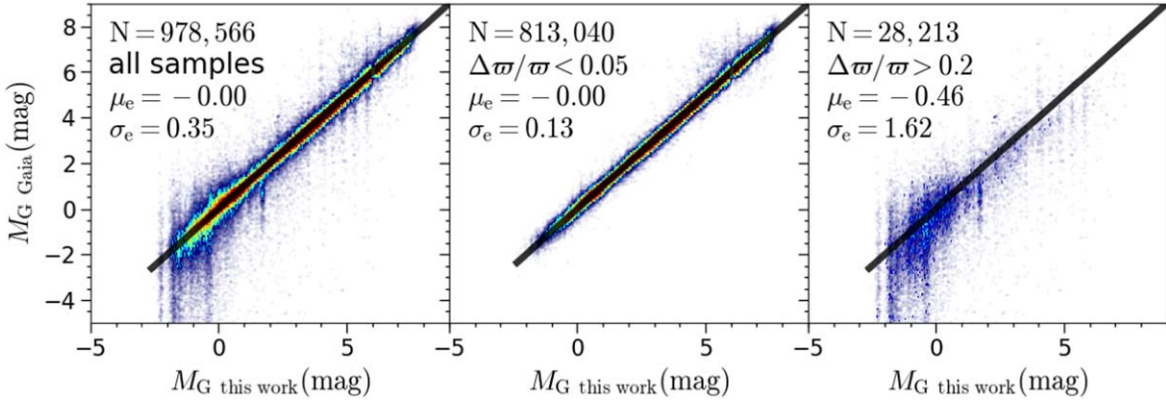


Figure 9. The same as Figure 5, but for M_G of all samples (left), high-precision parallax samples (middle), and low-precision parallax samples (right). The black line is the equal line.

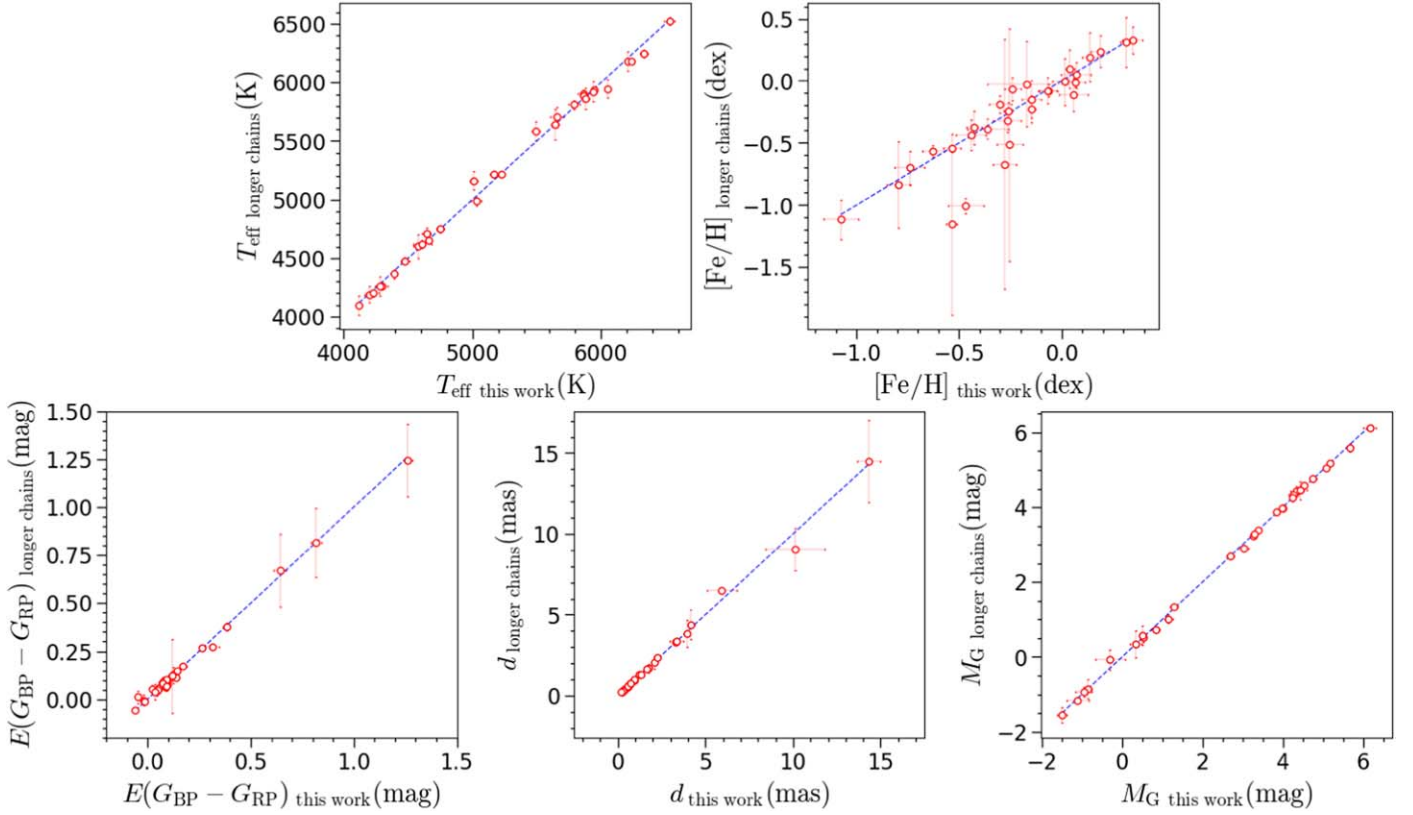


Figure 10. Comparison of final parameters (T_{eff} (top-left), $[\text{Fe}/\text{H}]$ (top-right), M_G (bottom-left), $E(G_{\text{BP}} - G_{\text{RP}})$ (bottom-middle), and distance (bottom-right)) of this work with parameters from 100 walkers and 1000 steps chains. The blue dashed line is the equal line.

surveys of all the test stars and found that, in general, the agreement is good (Figure 9). The dispersion value of the differences is about 0.35 mag, but the parallax error has a significant impact on the accuracy of the distances, and therefore, it has a great impact on the accuracy of the absolute magnitude we obtain. For stars with small relative parallax errors (less than 5%), the dispersion value of the M_G difference is only 0.13 mag. However, for stars with relative parallax errors greater than 20%, these sources exhibit significant dispersion, and the dispersion value of the M_G difference is 1.62 mag.

5.2. Comparison of Results with Longer MCMC Chains

To enable SPAr to run on large samples of stars, we employed a smaller number of chains and steps in the final MCMC analysis phase. To evaluate the effect of chain and step numbers on the results, we randomly selected 30 sources (10 sources each from LAMOST, APOGEE, and GALAH) and performed 100 walkers and 1000 step chains for each of these sources. The results obtained from SPAr and longer chains and steps are presented in Figure 10. Overall, there are good agreements between the obtained parameters. However, four sources showed relatively large deviations in $[\text{Fe}/\text{H}]$, mainly due to the large uncertainties. Despite this, we believe it is reasonable for SPAr to use relatively short chains and steps. Using longer chains and steps can significantly increase the

computational time without bringing any significant improvement in the results.

6. Summary

In this work, we have developed a new algorithm called SPAr to derive stellar parameters from multiband photometries, which can be applied to large samples of stars. The algorithm takes advantage of empirical stellar libraries constructed from Gaia, LAMOST, and other photometric surveys. It leverages the minimum χ^2 fit of the stellar SEDs to obtain the initial values for the MCMC analysis, which results in the stellar parameters, including T_{eff} , $[\text{Fe}/\text{H}]$, $E(G_{\text{BP}} - G_{\text{RP}})$, and M_G , of the individual stars. Our algorithm is tested on the LAMOST, APOGEE, and GALAH stars. The typical dispersion values of the differences between our results and literature values were 170 K for T_{eff} , 0.23 dex for $[\text{Fe}/\text{H}]$, 0.13 mag for M_G , and 0.05 mag for $E(G_{\text{BP}} - G_{\text{RP}})$.

In the future, our new SPAr algorithm will be implemented on large samples of stars obtained from the Mephisto and CSST surveys to derive atmospheric parameters, distances, and extinction values for billions of stars. This will give us crucial insights into the structure, chemistry, and other properties of the Milky Way.

Acknowledgments

We would like to thank the referee for providing us with detailed and constructive feedback that has significantly enhanced the quality of the manuscript. We are grateful to Professor Biwei Jiang for help and discussion. This work is partially supported by the National Key R&D Program of China No. 2019YFA0405500, National Natural Science Foundation of China 12173034, 11833006, 12203016, and 12173013, Natural Science Foundation of Hebei Province No. A2022205018, A2021205006, 226Z7604G, and Yunnan University grant No. C619300A034, and Science Foundation of Hebei Normal University No. L2022B33. We acknowledge the science research grants from the China Manned Space Project with NO. CMS-CSST-2021-A09, CMS-CSST-2021-A08, and CMS-CSST-2021-B03. We are grateful for the support of the Postdoctoral Research Station in Physics at Hebei Normal University.

This research made use of the crossmatch service provided by C.D.S., Strasbourg.

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>. The Gaia archive website is <https://archives.esac.esa.int/gaia>.

The national facility capability for SkyMapper has been funded through ARC LIEF grant LE130100104 from the Australian Research Council, awarded to the University of

Sydney, the Australian National University, Swinburne University of Technology, the University of Queensland, the University of Western Australia, the University of Melbourne, Curtin University of Technology, Monash University, and the Australian Astronomical Observatory. SkyMapper is owned and operated by the Australian National University's Research School of Astronomy and Astrophysics. The survey data were processed and provided by the SkyMapper Team at ANU. The SkyMapper node of the All-Sky Virtual Observatory (ASVO) is hosted at the National Computational Infrastructure (NCI). Development and support of the SkyMapper node of the ASVO has been funded in part by Astronomy Australia Limited (AAL) and the Australian Government through the Commonwealth's Education Investment Fund (EIF) and National Collaborative Research Infrastructure Strategy (NCRIS), particularly the National eResearch Collaboration Tools and Resources (NeCTAR) and the Australian National Data Service Projects (ANDS).

This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

This publication makes use of data products from the Widefield Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration.

ORCID iDs

Mingxu Sun  <https://orcid.org/0000-0002-2473-9948>
 Bingqiu Chen  <https://orcid.org/0000-0003-2472-4903>
 Helong Guo  <https://orcid.org/0000-0001-5737-6445>
 He Zhao  <https://orcid.org/0000-0003-2645-6869>
 Ming Yang  <https://orcid.org/0000-0001-8247-4936>
 Wenyuan Cui  <https://orcid.org/0000-0003-1359-9908>

References

- Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, *ApJS*, 259, 35
 Anders, F., Khalatyan, A., Chiappini, C., et al. 2019, *A&A*, 628, A94
 Andrae, R., Fouesneau, M., Sordo, R., et al. 2023, *A&A*, 674, A27
 Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, *AJ*, 156, 58
 Berry, M., Ivezić, Ž., Sesar, B., et al. 2012, *ApJ*, 757, 166
 Bessell, M. S. 2001, *PASP*, 113, 66
 Bressan, A., Marigo, P., Girardi, L., et al. 2012, *MNRAS*, 427, 127
 Buder, S., Lind, K., Ness, M. K., et al. 2019, *A&A*, 624, A19
 Buder, S., Sharma, S., Kos, J., et al. 2021, *MNRAS*, 506, 150
 Cao, Y., Gong, Y., Meng, X.-M., et al. 2018, *MNRAS*, 480, 2178
 Cenarro, A. J., Moles, M., Cristóbal-Hornillos, D., et al. 2019, *A&A*, 622, A176
 Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560
 Chen, B. Q., Huang, Y., Hou, L. G., et al. 2019b, *MNRAS*, 487, 1400
 Chen, B. Q., Huang, Y., Yuan, H. B., et al. 2019a, *MNRAS*, 483, 4277
 Chen, B.-Q., Liu, X.-W., Yuan, H.-B., et al. 2014, *MNRAS*, 443, 1192
 Chen, B. Q., Liu, X. W., Yuan, H. B., et al. 2017, *MNRAS*, 464, 2545
 DESI Collaboration, Abareshi, B., Aguilar, J., et al. 2022, *AJ*, 164, 207
 Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306
 Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, *A&A*, 649, A1
 Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, *A&A*, 595, A1
 Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2023, *A&A*, 674, A1
 Green, G. M., Rix, H.-W., Tschesche, L., et al. 2021, *ApJ*, 907, 57
 Green, G. M., Schlafly, E. F., Finkbeiner, D. P., et al. 2014, *ApJ*, 783, 114
 Green, G. M., Schlafly, E. F., Finkbeiner, D. P., et al. 2015, *ApJ*, 810, 25

- Huang, Y., Beers, T. C., Wolf, C., et al. 2022, *ApJ*, 925, 164
- Huang, Y., Yuan, H., Beers, T. C., & Zhang, H. 2021, *ApJL*, 910, L5
- Kirkpatrick, J. D., Schneider, A., Fajardo-Acosta, S., et al. 2014, *ApJ*, 783, 122
- Lei, L., Chen, B.-Q., Li, J.-D., et al. 2022, *RAA*, 22, 025004
- Lei, L., Li, J. D., Wu, J. T., Jiang, S. Y., & Chen, B. Q. 2021, *Astron. Res. Tech.*, 18, 115
- Luo, A. L., Zhao, Y.-H., Zhao, G., et al. 2015, *RAA*, 15, 1095
- Onken, C. A., Wolf, C., Bessell, M. S., et al. 2019, *PASA*, 36, e033
- Planck Collaboration, Abergel, A., Ade, P. A. R., et al. 2014, *A&A*, 571, A11
- Shen, H., Chen, B. Q., Guo, H. L., et al. 2022, *MNRAS*, 514, 4398
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, 131, 1163
- Skrutskie, M. F., Schneider, S. E., Stiening, R., et al. 1997, in *Astrophysics and Space Science Library*, Vol. 210, *The Impact of Large Scale Near-IR Sky Surveys*, ed. F. Garzon et al. (Dordrecht: Springer), 25
- Sun, M., Jiang, B., Yuan, H., & Li, J. 2021, *ApJS*, 254, 38
- Sun, M., Jiang, B. W., Zhao, H., et al. 2018, *ApJ*, 861, 153
- Wang, S., & Chen, X. 2019, *ApJ*, 877, 116
- Wolf, C., Onken, C. A., Luvaul, L. C., et al. 2018, *PASA*, 35, e010
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, 140, 1868
- Xu, S., Yuan, H., Niu, Z., et al. 2022, *ApJS*, 258, 44
- Yang, L., Yuan, H., Xiang, M., et al. 2022, *A&A*, 659, A181
- York, D. G., Adelman, J., Anderson, J. E. J., et al. 2000, *AJ*, 120, 1579
- Yuan, H. B., Liu, X. W., Huo, Z. Y., et al. 2015, *MNRAS*, 448, 855
- Yuan, X., Li, Z., Liu, X., et al. 2020, *Proc. SPIE*, 11445, 114457M
- Zhan, H. 2011, *Scientia Sinica Physica, Mechanica Astronomica*, 41, 1441
- Zhan, H. 2018, in 42nd COSPAR Scientific Assembly (Washington, DC: NASA), 16
- Zhang, R., & Yuan, H. 2023, *ApJS*, 264, 14