









# The Ages of Galactic Bulge Stars with Realistic Uncertainties

Meridith Joyce<sup>1,2,3,4,9</sup> , Christian I. Johnson<sup>1</sup> , Tommaso Marchetti<sup>5</sup> , R. Michael Rich<sup>6</sup> , Iulia Simion<sup>7</sup> , and John Bourke<sup>8</sup> 

<sup>1</sup> Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA; [mjoyce@stsci.edu](mailto:mjoyce@stsci.edu)<sup>9</sup>

<sup>2</sup> Kavli Institute for Theoretical Physics, University of Santa Barbara, CA 93106, USA

<sup>3</sup> Konkoly Observatory, Research Centre for Astronomy and Earth Sciences, H-1121 Budapest, Konkoly Th. M. út 15-17, Hungary

<sup>4</sup> CSFK, MTA Centre of Excellence, Budapest, Konkoly Thege Miklós út 15-17, H-1121, Hungary

<sup>5</sup> European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748 Garching bei München, Germany

<sup>6</sup> Department of Physics and Astronomy, UCLA, 430 Portola Plaza, Box 951547, Los Angeles, CA 90095-1547, USA

<sup>7</sup> Shanghai Key Lab for Astrophysics, Shanghai Normal University, 100 Guilin Road, Shanghai 200234, People's Republic of China

<sup>8</sup> Mathematical Sciences Publishers, 798 Evans Hall No. 3840, Berkeley, CA 94720, USA

Received 2022 May 16; revised 2023 January 19; accepted 2023 January 22; published 2023 March 23

## Abstract

Using modern isochrones with customized physics and carefully considered statistical techniques, we recompute the age distribution for a sample of 91 microlensed dwarfs in the Galactic bulge presented by Bensby et al. and do not produce an age distribution consistent with their results. In particular, our analysis finds that only 15 of 91 stars have ages younger than 7 Gyr, compared to their finding of 42 young stars in the same sample. While we do not find a constituency of very young stars, our results do suggest the presence of an  $\sim 8$  Gyr population at the highest metallicities, thus contributing to the long-standing debate about the age–metallicity distribution of the Galactic bulge. We supplement this with attempts at independent age determinations from two sources of photometry, BDBS and Gaia, but find that the imprecision of photometric measurements prevents reliable age and age uncertainty determinations. Lastly, we present age uncertainties derived using a first-order consideration of global modeling uncertainties in addition to standard observational uncertainties. The theoretical uncertainties are based on the known variance of free parameters in the 1D stellar evolution models used to generate isochrones, and when included, result in age uncertainties of 2–5 Gyr for this spectroscopically well-constrained sample. These error bars, which are roughly twice as large as typical literature values, constitute realistic lower limits on the true age uncertainties.

*Unified Astronomy Thesaurus concepts:* [Galactic bulge \(2041\)](#); [Stellar ages \(1581\)](#); [Stellar evolutionary models \(2046\)](#)

## 1. Introduction

The star formation history of the Galactic bulge remains a challenge, with age constraints on significant mass components of the population still ranging from 1 to  $>10$  Gyr. The underlying cause of this controversy is that the range in stellar metallicity, the large intrinsic depth of the bulge, severe crowding, high reddening, spatial distribution, and the intervening disk population all conspire to make it extremely difficult to constrain the age and age–metallicity relationship for the bulge. Even the advent of Gaia parallaxes has not made a significant impact: while Gaia data can remove stars in the foreground, the parallaxes are not sufficiently precise to give good age constraints on stars at bulge distances of  $\sim 8$  kpc.

Unlike the case for globular clusters, the bulge also has significant physical depth, as much of the mass is in a bar with a major axis  $\sim 20^\circ$ – $30^\circ$  from the line of sight toward  $(l, b) = (0, 0)$  (e.g., Stanek et al. 1997; Wegg & Gerhard 2013; Simion et al. 2021). While a detailed discussion of the history of the age controversy for the bulge may be found in various review articles, e.g., Rich (2013) and Barbuy et al. (2018), we offer a brief overview here.

The star formation history of the bulge remains an important problem even in the era of direct, high-redshift observations of galaxy formation and evolution. No high-redshift galaxy's evolution can be traced to the present day, but the Milky Way preserves all the stars that ever formed in our Galaxy. Metallicities derived from deep photometry (Johnson et al. 2020) and spectroscopy (e.g., Zoccali et al. 2008; Johnson et al. 2013; Ness et al. 2013) show a strong tendency for metal-rich stars to be concentrated toward the plane. Hence, if there is any age–metallicity relationship, there is an implied age gradient with presumably younger, more metal-rich stars concentrated toward the plane. Further, for ages  $\gtrsim 8$  Gyr, the differences in main-sequence turnoff (MSTO) morphology are sufficiently subtle that the spatial depth, reddening, metallicity spread, and contamination by the foreground population blur color–magnitude diagrams (CMDs). Yet, this age regime likely traces the processes that led to the formation of the bar and might therefore reveal the progression of the formation of the bulge. Given the enhanced  $\alpha$ -elements measured for bulge giants (e.g., Gonzalez et al. 2011; Johnson et al. 2014; Bensby et al. 2017; B17 hereafter; Rojas-Arriagada et al. 2020) and the rapid—if not explosive—star formation history this implies, it is possible that the strong gradient of abundance with latitude lacks a corresponding age gradient. It is just as possible that there is an age gradient, if the bulge/bar formed more slowly, or if other dynamical processes contributed to its formation. A fundamental tension is that age determinations based on photometry nearly always find a uniformly old bulge, while those based on spectroscopy find evidence of a broad age

<sup>9</sup> STScI Lasker Fellow.



dispersion (Clarkson et al. 2008; B17; Renzini et al. 2018). However, current ages are simply too coarse to tell definitively.

Despite 30 years of Hubble Space Telescope (HST) observations, the age and star formation history of the bulge remain open problems. For the reasons stated above, and due also to significant trends in abundance with Galactic latitude, HST photometry of the MSTO still fails to give the kind of age precision that characterizes, e.g., the local disk, or globular clusters. Overall, there is currently a debate between those who analyze CMDs of the bulge (often foreground-subtracted or proper motion-cleaned; e.g., Clarkson et al. 2008; Renzini et al. 2018) giving an  $\sim 10$  Gyr age, and analyses of individual microlensed dwarfs, which argue for a metal-rich component of the bulge as young as 1–5 Gyr (B17). As stellar evolution models have advanced and observations improve, we now revisit the ages computed for microlensed dwarfs to explore whether this schism in the age/star formation history derived using these different methods remains.

The deep imaging of HST offered compelling evidence of a predominantly old stellar population. Ortolani et al. (1995) compared the luminosity function of stars in Baade’s window against those of the bulge globular clusters NGC 6528 and NGC 6553 and found that the bulge was compatible in age with these clusters while possessing an age spread no larger than about 10%. Further results by Zoccali et al. (2003), Valenti et al. (2013), and more recently by Renzini et al. (2018) with HST treasury data, seemed to confirm that the bulge is largely composed of old stars with little evidence supporting a young component.

As was first shown by Kuijken & Rich (2002), one may use the proper motion in  $l$  to veto members of the foreground disk. This method was elegantly applied in Clarkson et al. (2008, 2011) to argue that the bulge is a stellar population dominated by a roughly solar metallicity,  $>10$  Gyr component. This analysis of CMDs would appear to settle the matter of the age, but that is not the case.

A series of papers has advanced the case for a young/intermediate-age population in the bulge based on the analysis of dwarf spectra serendipitously experiencing microlensing (Bensby et al. 2011, 2013, 2014, 2017). Such events can brighten a source by upwards of an order of magnitude, making possible high-resolution spectroscopy of dwarfs otherwise too faint to observe. As summarized in B17, the derived  $[\text{Fe}/\text{H}]$ ,  $\log g$ , and  $T_{\text{eff}}$  result in strikingly young ages—around 1–5 Gyr—especially for metal-rich stars. Their work finds that  $\sim 70\%$  of the stars with  $[\text{Fe}/\text{H}] > 0$  are  $< 8$  Gyr. The broad implication of these age estimates is that much of the population that constitute the bar, with  $-0.5 < [\text{Fe}/\text{H}] < +0.5$ , is younger than 8 Gyr, and that non-negligible portions are younger than 5 Gyr. As these ages are in direct conflict with the results from MSTO studies, it is of the utmost importance to evaluate these conclusions carefully using the most current isochrones, the most appropriate statistical methods, and meticulous consideration of all sources of uncertainty (and, when available, actual photometry).

In this study, we perform an independent reanalysis of the spectroscopic parameters presented in B17 using modern, customized stellar models and a combination of statistical techniques including maximum likelihood methods,  $\chi^2$  frequentist approaches, and Monte Carlo (MC) simulations. For the bulk of our analysis, we use modern isochrones taken directly from the publicly available Modules for Experiments

in Stellar Astrophysics (MESA) Isochrones and Stellar Tracks (MIST) database (Choi et al. 2016). We supplement this with explorations using isochrones computed with the Yale Rotating Stellar Evolution Code (YREC; Demarque et al. 2004), to mimic B17’s method more closely, and custom isochrones built with the 1D stellar evolution program MESA. In all cases, our analysis of the spectroscopic parameters produces an age–metallicity distribution consistent with the presence of an age gradient in the bulge, but notably inconsistent with the age results of B17.

As an independent check, we also attempt to perform age derivations from photometric subsamples of the same population using the Blanco DECam Bulge Survey (BDBS; Johnson et al. 2020; Rich et al. 2020) and Gaia EDR3 (Gaia Collaboration et al. 2021). While we can produce an age–metallicity distribution visually consistent with the absence of an age gradient—and almost totally bereft of young stars—in both cases, our attempts to compute observational age uncertainties reveal that the photometric parameters are simply too imprecise to provide meaningful estimates of either the ages or their uncertainties. Nonetheless, we include a detailed photometric analysis (see Appendix A) with the hope that it serves as both a cautionary tale and useful guide to others performing photometric age determinations.

We discuss the details of our age determination algorithm at length, and while our explicitness in this regard is not typical of similar analyses, our transparency serves to underscore the dangers of “black box” modeling results. Though machine learning and similar techniques are undoubtedly powerful, overinterpretation of results is more common when the internal workings of the algorithm or method are obscured from the user.

We conclude our analysis by performing age determinations that account for observational as well as theoretical (i.e., modeling) sources of uncertainty, where first-order estimates of the theoretical uncertainty contributions are determined by quantifying the effects of physically justified variations in the free parameter choices. While studying the impact of, for example, variations in the degree of  $\alpha$ -element enhancement in stellar models is common in more careful analyses, this is the first time that variations in the convective mixing length in Mixing Length Theory (MLT),  $\alpha_{\text{MLT}}$ , have been propagated into MIST isochrones. The age uncertainties derived using this  $\alpha_{\text{MLT}}$ -adaptive grid of isochrones are state-of-the-art and constitute a significant advancement, in terms of realism and accuracy, in the arena of global age uncertainty estimates.

This paper is organized as follows: in Section 2, we discuss the sample of the microlensed dwarfs and their potential to resolve bulge age tensions. In Section 3, we discuss our choice of isochrones, the key differences between MIST and YREC, and the effects of  $\alpha$ -element enhancement. In Section 4, we detail our age determination algorithm and present the resulting age and age–metallicity distributions for the spectroscopic data, supplemented by age uncertainties discussed in Section 5. In Section 6, we discuss the spectroscopic results in context, and in Section 7, we present realistic age error estimates based on modeling uncertainties. We compare our age–metallicity distribution to HST results in Section 8. We summarize our results and discuss their implications in Sections 9 and 10. For completeness, we apply our age determination technique to BDBS and Gaia photometry in Appendix A, taking care to discuss the shortcomings of these results and the reasons why

these, and all, photometric age determinations should be approached with caution.

## 2. Ages of the Microlensed Sample

The microlensed sample described in detail by Cohen et al. (2008, 2010) and Bensby et al. (2010, 2011, 2013, 2014, 2017) is a rare and powerful tool. Microlensing produces brightness magnification of 10–1000 $\times$ , enabling the inference of spectroscopic parameters for subdwarf stars that would normally be much too faint for most instruments to detect. In microlensing conditions, it is possible to perform high-resolution spectroscopy with 8 m class telescopes on these targets, thus providing measurements of their physical parameters—effective temperature and surface gravity—directly and independently of distance. A sample of microlensed bulge stars, then, has the potential to provide robust fundamental parameters for stars in this region. Given the implications the bulge age distribution carries for its star formation history, the origin of the bulge, and the formation history of the Galaxy itself, stellar ages are of particular importance.

Recently, age determinations for the sample of 91 microlensed stars in the bulge were conducted by B17. Their analysis concluded that the Galactic bulge has complex age and abundance properties; in particular, that the peaks in the bulge metallicity distribution, star formation episodes, and the abundance trends suggest the Galactic bulge/bar may be an inner extension of the Galactic thin/thick disk. They also found that the star formation rate appears to have been faster in the bulge than in the local thick disk, which indicates a denser stellar environment closer to the Galactic center.

In sum, their results support the idea of a secular origin for the Galactic bulge, suggesting the bulge was built up slowly from existing stellar populations in the central region of the Galaxy. The primary piece of evidence for these assertions is B17’s finding of a large numbers of young stars, which implies prolonged star formation in the region.

The bulge, however, has long been understood to be old, and an overabundance of young stars in this region calls into question our understanding of the formation history of the Galaxy and of galaxy evolution mechanisms more generally. An old bulge implies that the bulge assembled first, which is consistent with several chemical enrichment model fits to the observed [Fe/H] and [alpha/Fe] distributions (e.g., Ballero et al. 2007; Grieco et al. 2012; Matteucci et al. 2019). Further, our knowledge of the chemical distribution of the Galaxy is not consistent with recent star formation episodes in this region, which is in conflict with the fact that younger stars must have been formed more recently. Alternatively, if the young stars did not form there, we must then explain how younger stars moved into the region.

There is a known tension in the literature between age determinations from photometry, which produce a uniformly old bulge (e.g., Ortolani et al. 1995; Zoccali et al. 2003; Clarkson et al. 2008; Valenti et al. 2013; Renzini et al. 2018; Bovy et al. 2019; Hasselquist et al. 2020), and age determinations from microlensing, which produce a very broad age distribution (e.g., Bensby et al. 2013, 2017). However, age determinations from stellar models are difficult and sometimes unreliable. Often, they are plagued by inappropriate statistical techniques, out-of-date physical assumptions in the models, and inadequate treatment of uncertainties in both observations

and models. The first step toward resolving this tension, then, is to examine rigorously the age determination process itself.

## 3. Choice of Isochrones

For the bulk of our analysis, we use nonrotating isochrones from the modern and publicly available MIST database (Choi et al. 2016). The stellar tracks underlying these models are computed with MESA (Paxton et al. 2011, 2013, 2015, 2018, 2019). The models adopt the solar abundance scale of Asplund et al. (2009), the OPAL opacities of Iglesias & Rogers (1996), a default value for the convective mixing length parameter of  $\alpha_{\text{MLT}} = 1.82$  with the MLT (Boehm-Vitense 1979) prescription of Henyey et al. (1964), photosphere tables based on the 1D model atmospheres of Kurucz (1970) as surface boundary conditions, and a small degree of core convective overshoot,  $f_{\text{ovs}} = 0.016$ . The models include the effects of diffusion on the main sequence, where diffusion is modeled via five representative elemental species (Thoul et al. 1994). The key physical assumptions made in MIST are summarized in Table 3 of Choi et al. (2016) and discussed in further detail throughout.

We begin by refitting B17’s sample directly. Figure 1 provides a reproduction of B17’s Figure 6, with the panels divided according to the same metallicity bins, as indicated in the upper left-hand caption. Each panel contains four MIST isochrones with ages 1, 5, 10, and 15 Gyr shown in dotted, dotted-dashed, dashed, and solid lines, respectively. The metallicities assigned to the isochrones are the median of each bin: from metal-poor to metal-rich, [Fe/H]<sub>iso</sub> =  $-1.8$ ,  $-1.4$ ,  $-1.0$ ,  $-0.6$ ,  $-0.2$ , and the most metal-rich bin shows [Fe/H] =  $+0.2$  and  $+0.3$ . Darker colored tracks indicate higher metallicities (purple for highest, yellow for lowest), as shown on the color bar.

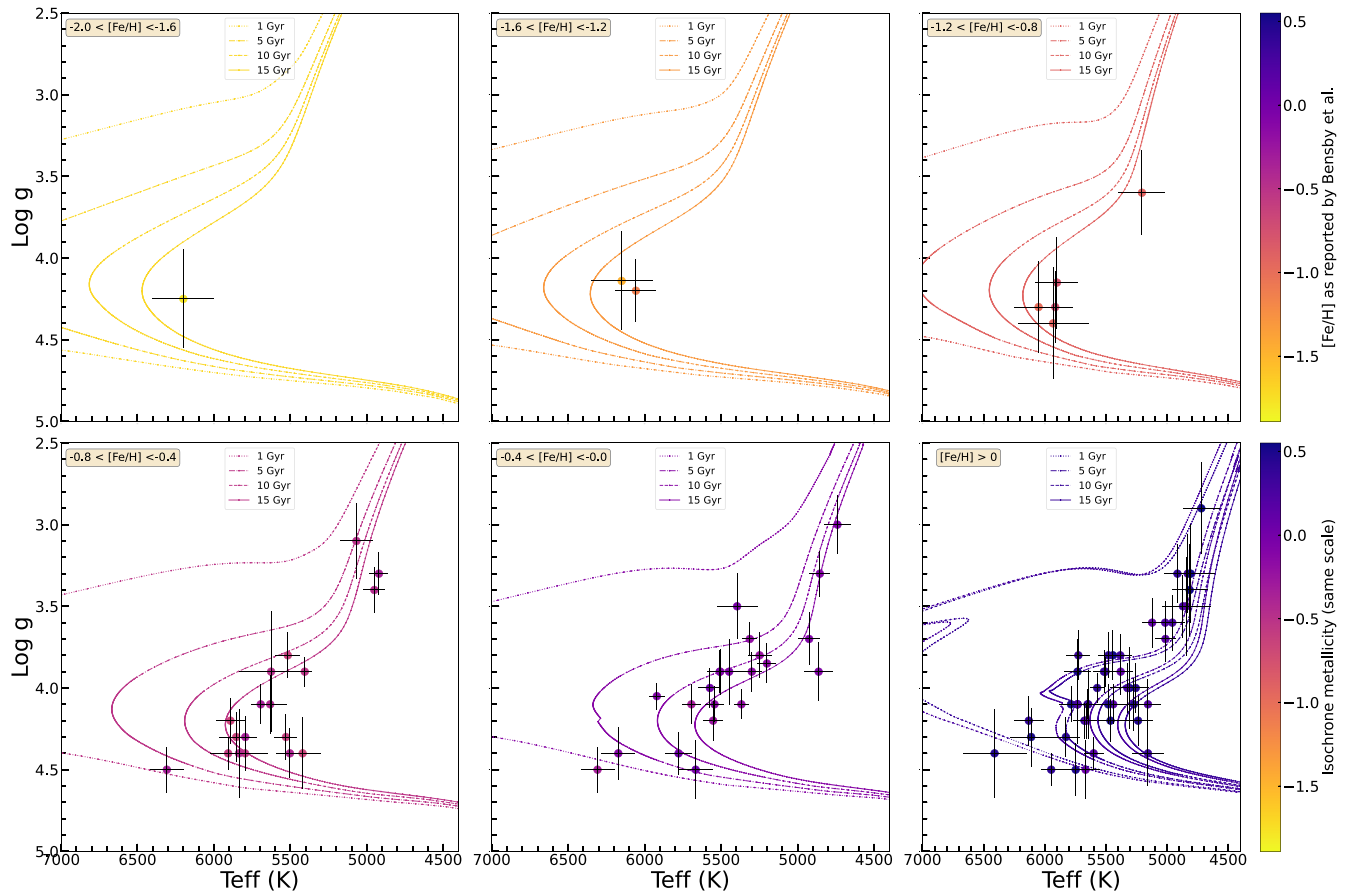
A side-by-side comparison of Figure 1 with B17’s Figure 6 already suggests some discrepancy. In general, the MIST isochrones are hotter (left-shifted) relative to the observations than are the isochrones in B17. If all physical assumptions in both sets of models were identical, this would translate to the prediction of older ages for the same stars with MIST. The choices of physics are, of course, not identical. Differences in the models’ treatment of abundances, in particular, complicates the physical picture.

We control for the impact of differences in the choice of modeling physics in two ways: first, we examine the choice of isochrone basis by comparing results computed with MIST to results computed with the Yonsei–Yale (hereafter abbreviated YY; Demarque et al. 2004) isochrones, the models used in B17. Second, we compare results from MIST isochrones with simulated  $\alpha$ -element enhancement to those without.

### 3.1. YY versus MIST

The version of the YY isochrones used in B17 date back to 2003, making them nearly 15 yr older than MIST. They also constitute a sparser theoretical data set, as the number of parameter variations covered by the publicly available YY grid is smaller, and the metallicity resolution lower, than in MIST’s case. Despite this, the YY isochrones were among the first to incorporate the effects of helium and  $\alpha$ -element enhancement carefully (Demarque et al. 2004).

Figure 2 shows a selection of the 2004 YY isochrones (dotted lines) overlaid on a selection of MIST isochrones (solid lines)



**Figure 1.** Reproduction of B17’s Figure 6 but using MIST isochrones. Metallicity bins are the same as in B17.

with similar metallicities for a range of ages. Curves with similar colors are similar metallicities, with the same color scale used in Figure 1.

For fixed age and metallicity, the Yale isochrones generally predict cooler temperatures compared to the MIST isochrones—which is what we should expect based on the comparison of Figure 1 to B17’s Figure 6—but this relationship inverts, or becomes ambiguous, near the base of the red giant branch (RGB).

There is noticeable discrepancy in the locations of the MSTO regions in the 2 and 5 Gyr panels. Since the majority of stars in the B17 sample fall exactly in this region, especially those stars identified as young, we should expect a nontrivial difference between the ages predicted by MIST and the ages predicted by Yale, all else being equal.

It is well known that the choices of modeling physics have a large effect on the MSTO and base of the RGB (e.g., Joyce & Chaboyer 2018a), as the isochrones demonstrate. The key differences in modeling physics between MIST and YY can be summarized as follows:

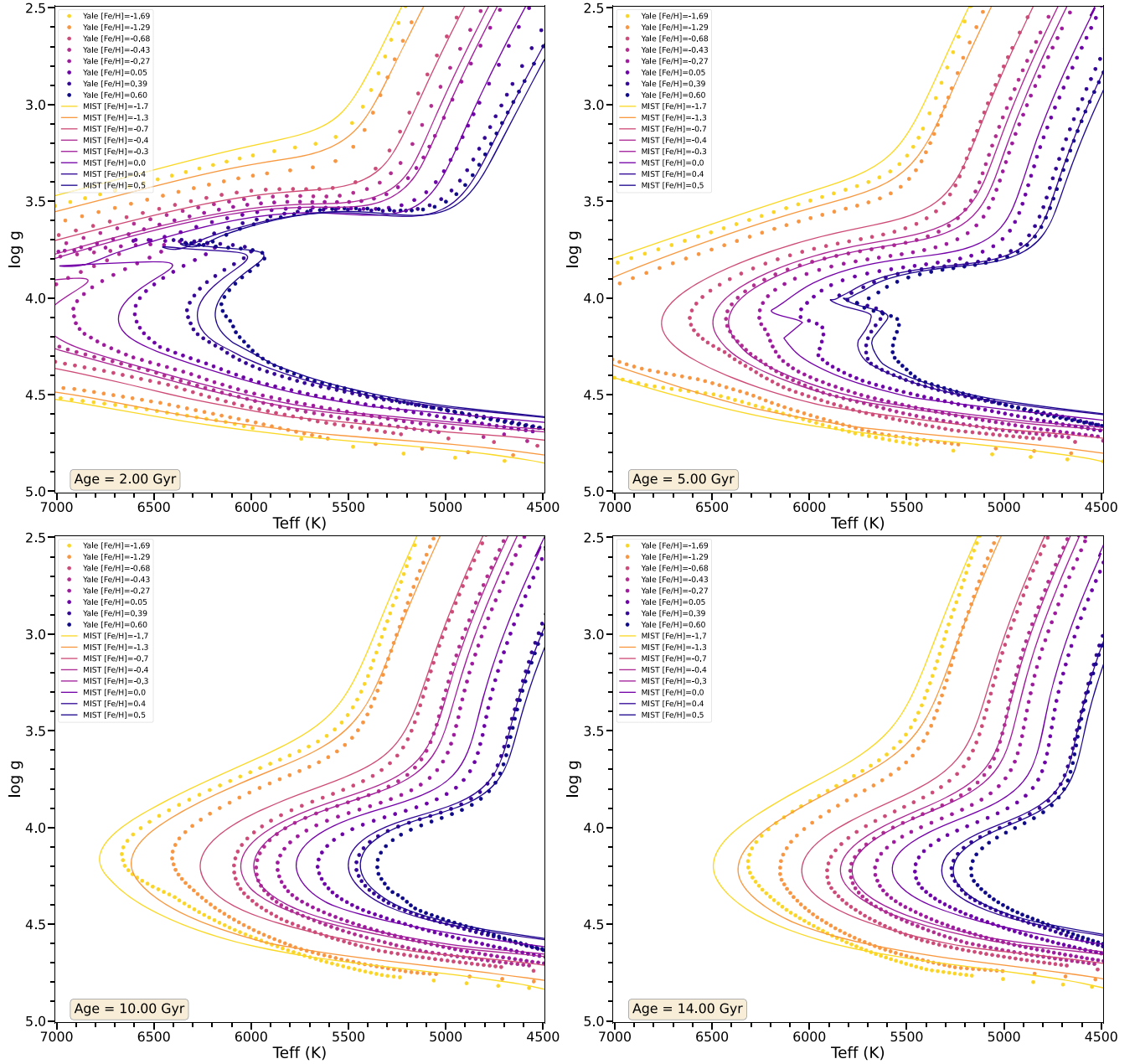
1. Choice of solar abundance scale: YY uses the abundances of Grevesse & Sauval (1998), which place the solar metallicity at  $Z = 0.020$ , whereas MIST uses the scale of Asplund et al. (2009), which places the solar metallicity at  $Z = 0.014$ ;
2. Atmospheric boundary conditions: YY uses a gray model atmosphere, whereas MIST uses photosphere tables based on 3D atmosphere simulations;

3. Treatment of diffusion: YY includes the effects of diffusion throughout, whereas MIST only uses diffusion on the main sequence; and
4. Mixing length: the YY stellar tracks adopt a value of  $\alpha_{\text{MLT}} = 1.743$  for the convective mixing length, whereas the MESA models use  $\alpha_{\text{MLT}} = 1.82$ .

Among the first three, the difference in the solar abundance scale will dominate. Regarding point 3, we note that the incorporation of diffusion will not affect the evolution beyond the main sequence, but it is worth mentioning this difference in the isochrone databases given that many of our targets are not necessarily more evolved than the upper main sequence. We defer a discussion of point four until Section 7. For a more comprehensive comparison of the assumptions made in these and other stellar evolution programs, see Table 3 of Tayar et al. (2022).

### 3.2. Effects of $\alpha$ -element Enhancement

Although there is no way to request isochrones with a particular degree of  $\alpha$ -enhancement, or  $[\alpha/\text{Fe}]$ , from the publicly available MIST database directly,  $\alpha$ -enhancement can be approximated in the models by rescaling the baseline abundance,  $[\text{Fe}/\text{H}]_0$ , to a new value  $[\text{Fe}/\text{H}]_{\text{rescaled}}$  that encapsulates the relative increase in global metallicity imparted by larger numbers of  $\alpha$ -capture elements. This is done by computing an array of global metal mass fractions,  $Z$ , associated with a densely sampled grid of ( $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ ) combinations. This dictionary can then be used to find an



**Figure 2.** A selection of MIST and Yale isochrones are shown for a selection of ages: 2, 5, 10, and 14 Gyr. MIST models are shown in dots, Yale with solid lines. Models of similar color have similar metallicities.

appropriate pair ( $[\text{Fe}/\text{H}]_{\text{rescaled}}$ ,  $[\alpha/\text{Fe}] = 0$ ) for any input combination of ( $[\text{Fe}/\text{H}]$ ,  $[\alpha/\text{Fe}]$ ). Since this relation defines parameter pairs as equivalent according to their production of the same  $Z$ , the choice of solar scale will affect this relationship.

We adopt an  $\alpha$ -element enhancement distribution for the isochrones according to the functions

$$\begin{aligned} [\alpha/\text{Fe}] &= +0.3 \text{ for } [\text{Fe}/\text{H}] < -0.3, \\ [\alpha/\text{Fe}] &= -[\text{Fe}/\text{H}] \text{ for } -0.3 < [\text{Fe}/\text{H}] < 0, \text{ and} \\ [\alpha/\text{Fe}] &= 0 \text{ for } [\text{Fe}/\text{H}] > 0, \end{aligned}$$

which are based on the consistently observed  $[\alpha/\text{Fe}]$  versus  $[\text{Fe}/\text{H}]$  trends reported by numerous authors (e.g., Gonzalez et al. 2011; Johnson et al. 2014; B17; Rojas-Arriagada et al. 2020). This produces the rescaled  $[\text{Fe}/\text{H}]$  values given in Table 1.

The difference between  $\alpha$ -enhanced and standard MIST isochrones can be seen in Figure 3. The primary effect of including  $\alpha$ -element enhancement is to shift the isochrones rightward toward cooler temperatures (redder colors) for equivalent ages and baseline metallicities. This appears similar to the difference between the MIST and YY isochrones. However, the reason for the difference in this case is that, as metallicity increases, stars maintain the same reaction rates at lower temperatures. The number of donor electrons increases in the atmosphere, which causes an increase in the continuous opacity and a subsequent decrease in the surface temperature. The isochrones constructed from such models will therefore skew cooler, but their evolution also lags behind isochrones without  $\alpha$ -enhancement. When  $\alpha$ -enhancement is handled properly, the age predictions will increase.

**Table 1**  
Rescaled Isochrone Metallicities Accounting for  $\alpha$ -element Enhancement

[Fe/H]	$[\alpha/\text{Fe}]$	Rescaled [Fe/H]
-2.0	+0.3	-1.787
-1.9	+0.3	-1.687
-1.8	+0.3	-1.587
-1.7	+0.3	-1.487
-1.6	+0.3	-1.387
-1.5	+0.3	-1.287
-1.4	+0.3	-1.187
-1.3	+0.3	-1.087
-1.2	+0.3	-0.987
-1.1	+0.3	-0.887
-1.0	+0.3	-0.787
-0.9	+0.3	-0.687
-0.8	+0.3	-0.588
-0.7	+0.3	-0.486
-0.6	+0.3	-0.384
-0.5	+0.3	-0.285
-0.4	+0.3	-0.184
-0.35	+0.3	-0.137
-0.32	+0.3	-0.108
-0.3	+0.3	-0.088
-0.26	+0.26	-0.077
-0.2	+0.2	-0.060
-0.1	+0.1	-0.037
0 to +0.5	0.0	original [Fe/H]

Figure 4 is a reproduction of Figure 1 using  $\alpha$ -enhanced MIST isochrones. They provide a slightly better visual fit to the B17 sample and are more reflective of real  $\alpha$ -abundance trends with metallicity. They are also most similar in composition to the abundances assumed in the YY isochrones used in B17’s analysis, while still being up-to-date and adopting the most modern physical conventions. We proceed using the  $\alpha$ -enhanced MIST isochrones as the basis for our quantitative age determinations.

#### 4. Age Determinations and Statistical Analysis

Though it is straightforward to fit an isochrone to an observation “by eye,” it is much more difficult to construct a mathematically rigorous definition of a best-fitting model. This is especially true in a situation where:

1. Many observations are plausibly consistent with a large number of isochrones (i.e., the isochrone falls within the star’s  $1\sigma$  uncertainties);
2. The isochrones are discretely spaced in age and metallicity and thus limited in resolution; and
3. The models contain their own intrinsic uncertainties that are both difficult to quantify and highly variant over different parameter regimes—a point to which we will return in later discussion.

To deal with these issues, we apply a combination of known statistical techniques to estimate the age and observational age uncertainty for each star according to an appropriate basis of isochrones. Our method incorporates elements of maximum likelihood analysis, MC simulations, and a frequentist approach to the goodness-of-fit/ $\chi^2$  (see, for instance, Joyce & Chaboyer 2018b; Murphy et al. 2021; Tang & Joyce 2021). The chosen isochrone basis constitutes a set of candidate hypotheses against which we test our stars. In all cases, the

MIST isochrones range in age from 0.5 to 20 Gyr, sampled at a resolution of 0.5 Gyr, and in metallicity from  $[\text{Fe}/\text{H}] = -2.0$  to  $[\text{Fe}/\text{H}] = +0.5$ , sampled at a resolution of 0.1 dex. We permit ages above the formal age of the universe (13.5 Gyr; Ade et al. 2016) for consistency with B17’s lack of upper bound on their own age determinations, and the standard MIST database permits ages up to 20 Gyr.

For each of B17’s targets, we measure the goodness-of-fit between the star and every point along each isochrone. Restricting to the appropriate evolutionary regime (we may confidently exclude the pre-main sequence and the RGB above the RGB bump, for example), the number of points along a particular isochrone is 220.<sup>10</sup>

The goodness-of-fit for each theoretical–observational pair is determined according to a  $\chi^2$  statistic, given by

$$\chi^2_{\text{for } i \text{ DoF}} = \sum_i \frac{(o_i - t_i)^2}{\sigma_i^2}, \quad (1)$$

where  $o_i$  are the values derived from observations,  $t_i$  are the values assumed in the models (theory), and  $\sigma_i$  are the observational uncertainties. The number of degrees of freedom (DoFs) is equal to the number of terms  $i$  in the summation, which determines the distribution of the  $\chi^2$  scores. This metric, which ultimately underlies our claim that an isochrone of particular age is more or less consistent with a measurement, depends strongly on the number of independent<sup>11</sup> observations, the uncertainties in the observations themselves, and the uncertainties implicit in, e.g., distance when a transformation from apparent to absolute magnitude is required. It is therefore crucial that we carefully consider the definitions of the constituent uncertainty terms in each case.

Explicitly, in the case of B17’s sample, there are three data sources and their associated uncertainties: the error in  $\log g$ , the error in the effective temperature  $T_{\text{eff}}$ , and the error in  $[\text{Fe}/\text{H}]$  (abbreviated below as  $Z$ , the convention for representing metallicity in stellar models). As these are already physical coordinates and the uncertainties are reported directly, this  $\chi^2$  definition is straightforward

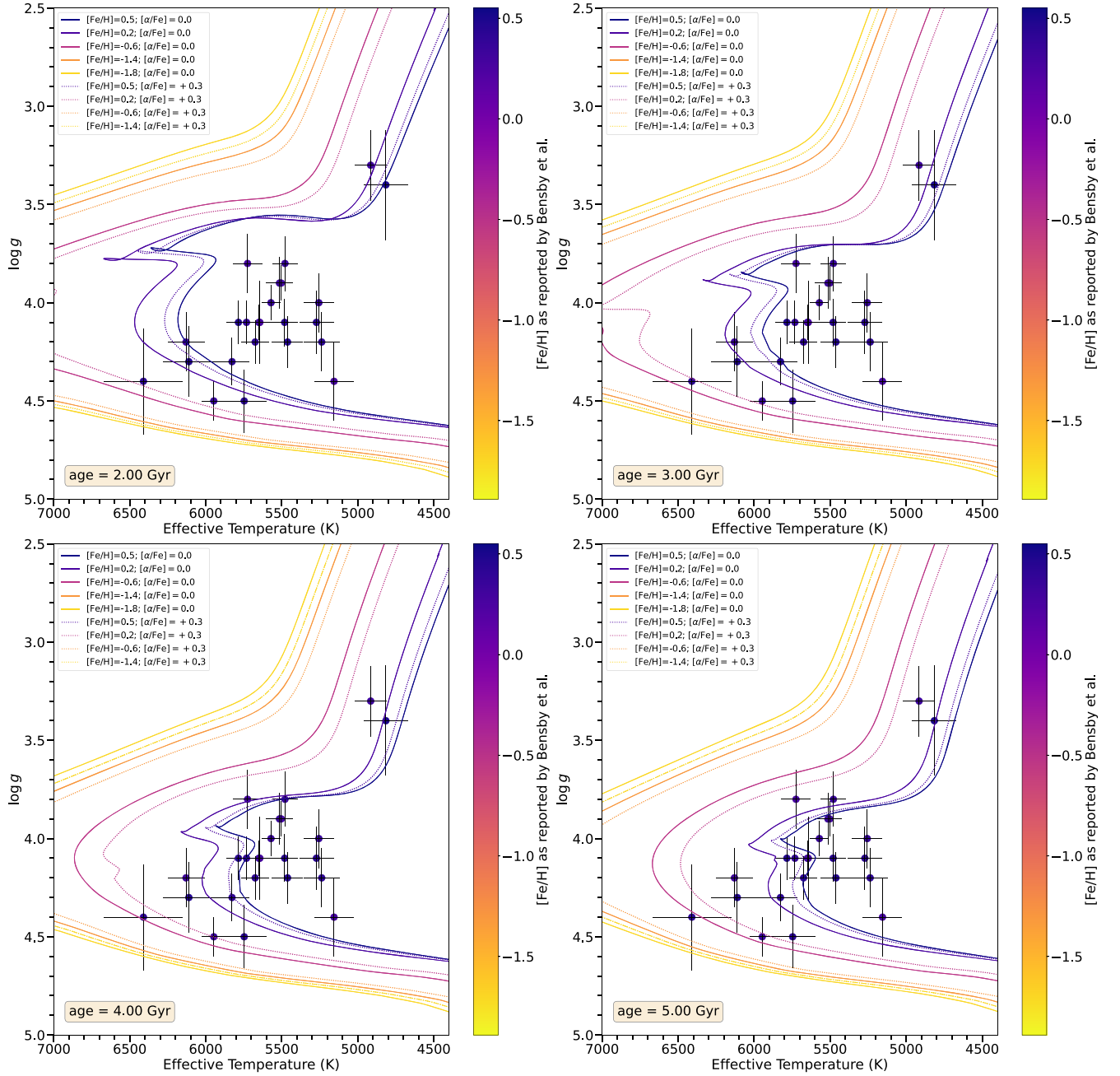
$$\chi^2_{\text{B17}} = \frac{(\log g_o - \log g_t)^2}{\sigma_{\log g,o}^2} + \frac{(T_{\text{eff},o} - T_{\text{eff},t})^2}{\sigma_{T_{\text{eff},o}}^2} + \frac{(Z_o - Z_t)^2}{\sigma_{Z,o}^2}, \quad (2)$$

as are the definitions for  $\sigma_{\log g,o}$ ,  $\sigma_{T_{\text{eff},o}}$ , and  $\sigma_{Z,o}$ . We note that this statistic applies equal weighting to uncertainties in  $T_{\text{eff}}$ ,  $\log g$ , and metallicity.

Fits to photometric data, on the other hand, are more complicated; they must take into account not only the uncertainties in the apparent magnitudes for the star in each filter used to build the CMD, but uncertainties in the distance to the star, extinction, and observational instrumental systematics. Further, the relationship among these quantities and appropriate relative weightings are not obvious. Definitions of the

<sup>10</sup> We note, however, that this restriction was removed when custom isochrones were used in the fits due to the modified evolutionary trajectories of the underlying tracks caused by, e.g., very low assumptions for the convective mixing length; see Section 7. A test comparing age determinations when using a restricted evolutionary domain on the isochrones to those using an unrestricted domain produced a shift toward older ages of less than 1%.

<sup>11</sup> Observations are not strictly independent, but we must make this assumption for the sake of practicality.



**Figure 3.** Isochrones at a small range of metallicities and their  $\alpha$ -scaled counterparts are shown for ages of 2–5 Gyr, as indicated. Isochrones that take into account  $\alpha$ -element enhancement are indicated on these four panels by dashed lines. A solid line of the same color indicates the same isochrone without  $\alpha$ -enhancement. In the case of the isochrone with  $[\text{Fe}/\text{H}] = 0.5$ , no  $\alpha$ -enhanced counterpart exists due to the global metallicity upper limit of  $+0.5$  imposed by the MIST database. The most visible effect of including  $\alpha$ -element enhancement in the isochrones is to shift the isochrones toward cooler temperatures, but the net effect of increased metallicity is that stars with equivalent temperatures and luminosities have lower masses, translating to slower evolution and an increase in the inferred age as a function of  $[\text{Fe}/\text{H}]$ .

uncertainty terms appropriate for the photometric analyses conducted here are discussed further in Appendix A.

For each star  $S$  in the B17 sample, we compute an agreement statistic, or “score,” comparing  $T_{\text{eff},S}$ ,  $\log g_S$ , and  $Z_S$  to each of 220 evolutionary points  $\times$  40 ages  $\times$  29 metallicities for a total of 255,200 instances of  $T_{\text{eff},i}$  and  $\log g_i$  in Equation (2). Like age,  $[\text{Fe}/\text{H}]$  (aka  $Z_i$ ) is fixed per isochrone, so  $Z_i$  does not vary over the set of evolutionary points. This algorithm thus returns the relative likelihoods of all candidate age–metallicity hypotheses (isochrones) for  $S$ .

We then construct a definition of the weighted average age through which each age hypothesis is penalized according to its

likelihood: Each score  $\chi_n^2$ ,  $n = 1, \dots, 255200$  has a relative likelihood,  $p_n$ , given by the probability density function (PDF) for the  $\chi^2$  distribution with three DoFs. We caution that this association implicitly assumes the three variables  $\log g$ ,  $T_{\text{eff}}$ , and  $[\text{Fe}/\text{H}]$  are independent and normally distributed. Though this is not strictly true (we note, for instance, that  $\log g$  may more closely follow a log normal distribution, and any parameters derived from the same observational source—in this case, spectroscopy—are correlated by definition), we confirm that this is a reasonable enough approximation via MC simulations, discussed in detail in Section 5 and the Appendix.

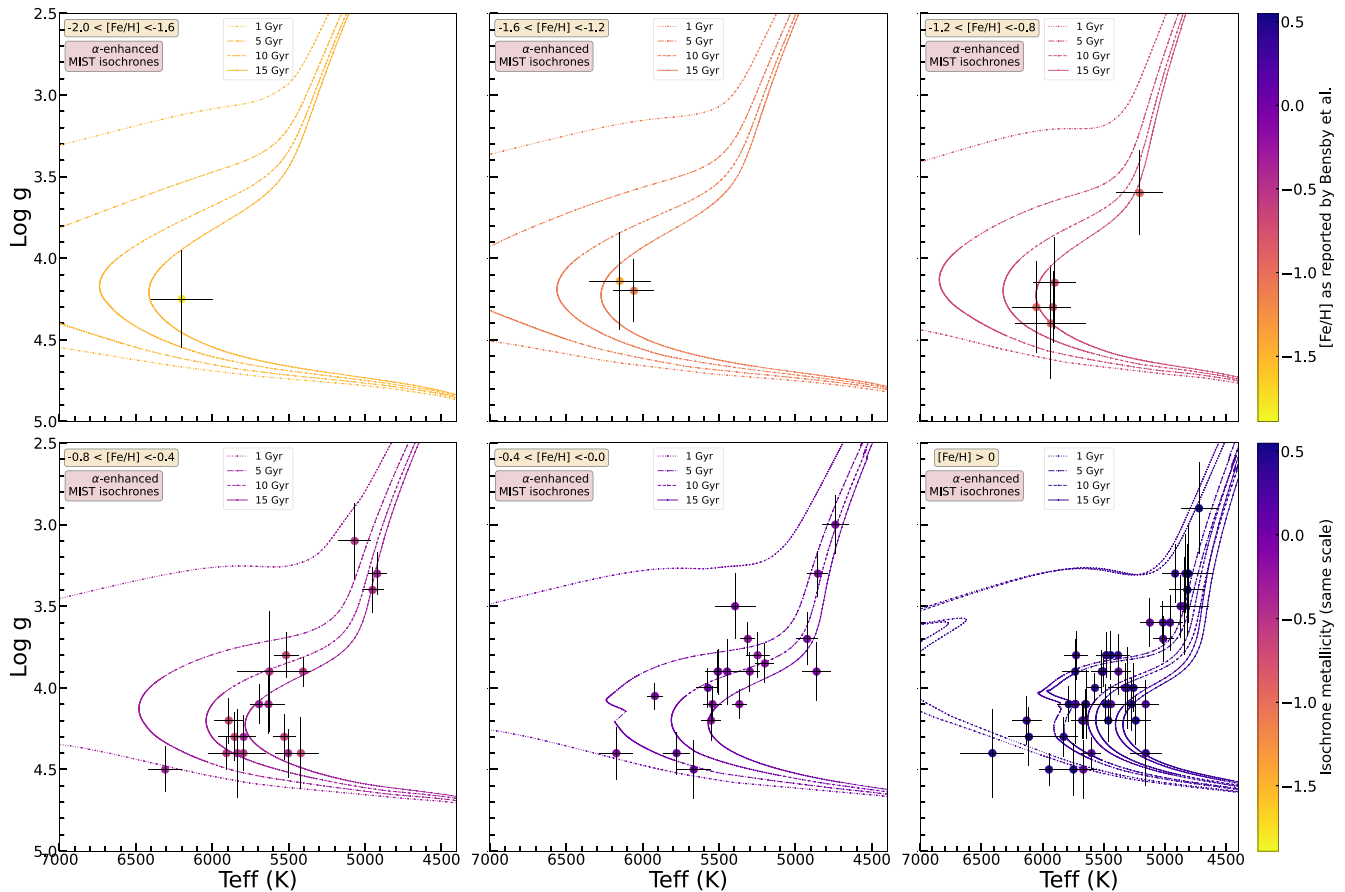


Figure 4. Same as Figure 1 but using  $\alpha$ -enhanced MIST isochrones.

The weighted average age of star  $S$  is given by

$$t_S = \frac{\sum_n t_n p_n}{\sum_n p_n}, \quad (3)$$

where  $t_n$  is the age of the isochrone hosting  $\log g_t$ ,  $T_{\text{eff},t}$  and  $Z_t$ ,  $p_n$  is the value of the  $\chi^2$  PDF at the score for that hypothesis given by Equation (2), and  $n$  is the total number of candidate hypotheses considered. The MIST-based age we report for a star therefore represents the weighted average over all age solutions, with the most consistent age fits weighted highest.

#### 4.1. Comparison to Other Methods in the Literature

Bensby et al. (2011) briefly describe how they form a probability distribution for the age estimates as a weighted average of “probabilities,” which are products of the normal density functions comparing the observed parameters for a star to those of a suite of isochrones. This is similar to our method: both assume the observed parameters come from independent normal random variables, for example, though in our case the weights are the values of the density function for the  $\chi^2$  distribution with three DoFs. However, Bensby et al. (2011) relegate a more detailed discussion of their method to J. Meléndez et al. (in preparation), which we were unable to find, and B17 refer only to their 2011 paper regarding their use of the same method. While we do find some discussion of a seemingly similar method in Meléndez et al. (2012), it is unclear how they define their  $1\sigma$  uncertainties. We speculate that, rather than fitting their derived distribution to a normal

distribution, they extend their estimates outwards to collect  $1\sigma$ 's worth of points assuming normality and do so in either direction independently, which could be the cause of their highly asymmetric error bars. However, this is not stated explicitly, so we are unable to make a more rigorous comparison at this time.

In some capacity, Bensby et al. (2011) derive their  $1\sigma$  error estimates from the shapes of the distributions they find. We instead use MC resampling for two reasons. First, this requires no additional assumptions, whereas they must make some assumptions about these shapes. Second, through random resampling, we are able to incorporate model uncertainties (see Section 7), which they did not take into account.

Various other methods have been used to estimate stellar ages using a suite of isochrones. Angus et al. (2019) used Markov Chain Monte Carlo random walks based on Foreman-Mackey et al. (2013) to find the parameter region with the best-fitting isochrones. We found this inappropriate for our data, as our larger uncertainties mean too broad a range of isochrones offering plausible fits, which would lead to random walkers ending up in local minima as an artifact of the discrete isochrones. Similarly, Creevey et al. (2022) used a combination of Gaia DR3 photometry, spectroscopy, and BaSTI isochrones to derive ages for millions of stars, but their method is more suited to targets in the solar neighborhood and has limited utility for stars with nonsolar metallicities (see also Kordopatis et al. 2023). Sanders & Das (2018) obtained age estimates for  $\sim 3$  million stars from Gaia and spectroscopic data using a Bayesian method that calculates a weighted average over isochrones due to Burnett & Binney (2010). Our method is

similar in principle, and it would be worthwhile to compare our method to theirs in future work. Their method more readily accounts for the nonuniform spread of their isochrone grid, but we do not expect our lack of consideration for this feature to make a significant difference given that we tested the sensitivity of our results to the input isochrone grid and increased its age and metallicity resolution as needed.

Another machine-learning method is The Cannon (Ness et al. 2015), used for example by Ho et al. (2017), Hasselquist et al. (2020), and Sit & Ness (2020) to estimate ages for large numbers of stars based on training data derived from APOGEE and other surveys. Our target, however, is a small number of individual stars for which we can make direct inferences, so this method is not ideal for our situation. We also prefer to avoid machine-learning techniques in general, as they can still suffer from the same sensitivities to input (i.e., training sets) as any other flawed method, but these sensitivities are obscured by “black box” usage, leading to overconfidence in the results. Any age determinations for imperfect data will inevitably be biased, whether by the training sample, the method, the priors, or data incompleteness and other problems. Though our method is not without its shortcomings and questionable-but-necessary assumptions, it is clearly defined, reproducible, and its limitations are not obscured by a black box.

#### 4.2. Ages according to the Basic MIST Isochrones

The age distribution resulting from the application of our method to 91 B17 targets is given in Figure 5. The most striking feature of the distribution shown in the upper panel of Figure 5 is the strong peak at an age of approximately 13 Gyr. While some stars with ages below 8 Gyr are identified, only four stars are found to have ages below 5 Gyr, and we find none with ages less than 3 Gyr. On the contrary, B17’s age distribution, constructed from the same data and overlaid in red in the lower panel of Figure 5, shows a nearly uniform distribution of ages between 0.5 and 14.5 Gyr.

#### 4.3. Ages according to Yale Isochrones

We test the hypothesis that the Yale isochrones should produce older ages, presented in Section 3, by applying the algorithm presented here to the publicly available set of 2003 Yale isochrones. In this case, the range of ages available is reduced to 0.5–20 Gyr in increments of 1 Gyr (half the resolution of the MIST grid), and there are eight metallicities available spanning  $[\text{Fe}/\text{H}] = -1.7$  to  $+0.6$ . The resulting age distribution is shown in Figure 6.

First, we note that the mean of the Yale-based age distribution is younger than MIST’s by a non-negligible 1.3 Gyr. Second, we observe a larger number of young stars (below 5 Gyr) detected and a significantly diminished high-age tail above 12 Gyr in the Yale distribution. There is also a more prominent peak around 6 Gyr than in Figure 5. However, per the lower panel of Figure 6, the shape of the Yale distribution still differs noticeably from the distribution presented by B17; it is much closer to the distribution found with MIST.

#### 4.4. Ages with $\alpha$ -enhanced MIST Isochrones

We quantify the effects of  $\alpha$ -element enhancement on stellar ages by recomputing the distribution presented in Figure 5 using the set of  $\alpha$ -enhanced MIST isochrones described in Section 3.2. Figure 7 shows the resulting distribution.

The impact on the median age is small, producing an overall shift toward older ages of about 1%: roughly 10.8 Gyr (no  $\alpha$ -enhancement) versus 10.9 Gyr (with  $\alpha$ -enhancement). However, the effect on the age bin near 8–9 Gyr is substantial and makes more apparent one of the key differences between the age distribution derived in B17 and the ones derived presently: where B17 find a three-peaked distribution—or, arguably, a near-uniform distribution<sup>12</sup>—with an overabundance of young stars near 3 Gyr, fits with the age determination algorithm presented here reveal a distribution with one prominent peak around 13 Gyr, a subordinate peak around 8 Gyr (particularly in the fits of MIST isochrones with  $\alpha$ -enhancement and Yale isochrones) and a clear, monotonic decrease in count below 8 Gyr.

The shift of the distribution toward older ages, particularly for the oldest and most metal-poor stars that make up the most  $\alpha$ -enhanced population, is consistent with our expectations. We proceed with our analysis by adopting the  $\alpha$ -basis MIST isochrones as the default physical configuration.

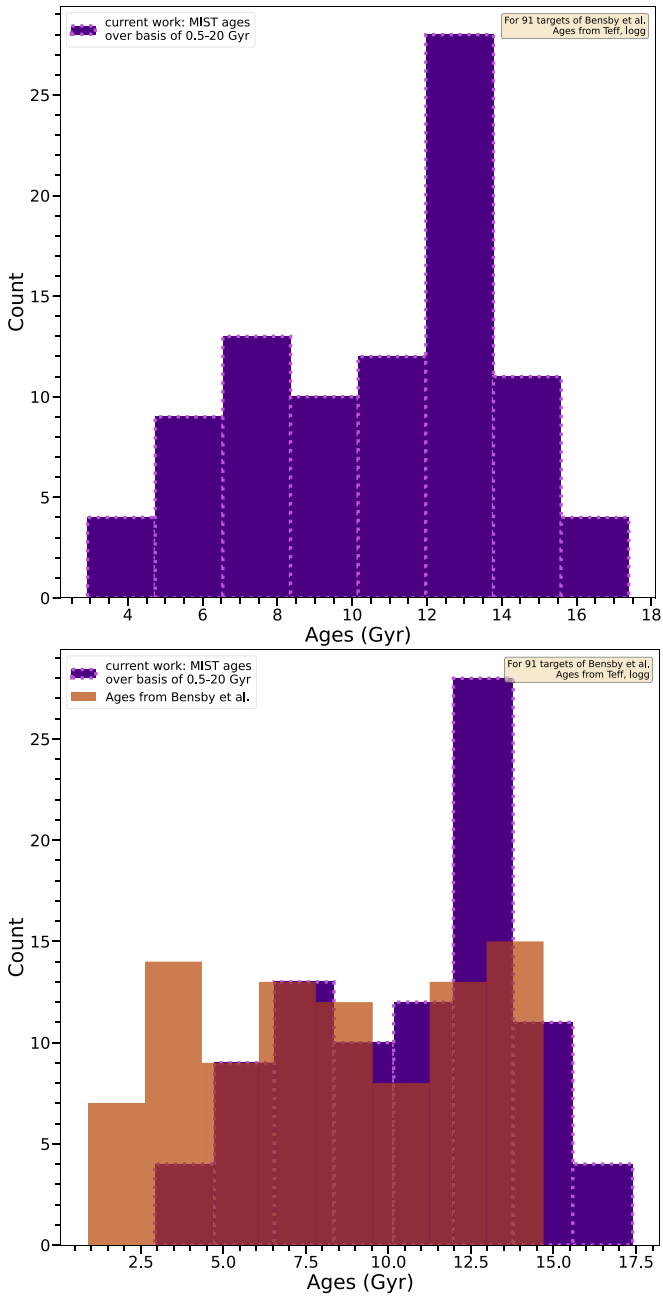
#### 4.5. The Age–Metallicity Distribution

Figure 8 compares the age–metallicity distribution found with MIST ( $\alpha$ -enhanced basis) to the distribution reported in B17 as well as to the age–metallicity relation from Haywood et al. (2016). The upper left panel shows age versus  $[\text{Fe}/\text{H}]$  for the three data sets. The purple diamonds (MIST) suggest a characteristic broken power law, with ages roughly constant below  $[\text{Fe}/\text{H}] = -0.5$  but declining with increasing metallicity above  $[\text{Fe}/\text{H}] = -0.5$ . The red circles (B17), on the other hand, show a large degree of age scatter at all but the lowest metallicities, and it would be difficult to claim that a broken power law describes the data. A comparison between our new MIST-based age–metallicity relation and the relation from Haywood et al. (2016) (green stars) shows that the present work extends the population of old ( $>10$  Gyr) bulge stars to higher  $[\text{Fe}/\text{H}]$  than is predicted by the Haywood et al. (2016) model. Our results are in better agreement with the age–metallicity relation from Bernard et al. (2018; see also their Figure 10).

The remaining three panels break the distribution into metallicity histograms: the upper right shows the youngest age bin,  $<7.0$  Gyr, the lower left shows  $7 \leq t \leq 10$ , and the lower right shows  $t > 10$  Gyr. The youngest and intermediate panels provide a clear picture of where the distributions differ most: first, there is a large discrepancy in the number of young stars identified at all—this is most apparent in the  $<7$  Gyr panel, where B17 report 42 of 91 stars with ages less than 7 Gyr (nearly 50%), as compared to our finding of 15 young stars (16%). Second, though the intermediate-age panel shows similar numbers of stars identified in this age range by both MIST and B17, these are spread across a much larger range in metallicity in B17’s case. The distributions in the oldest panel are reasonably consistent, though MIST finds that fully two-thirds (66%) of the stars have ages  $>10$  Gyr compared to B17’s 40%.

Figure 9 presents this relationship another way: binned average age is shown as a function of metallicity, with the number of stars in each bin annotated. The data show an increase in age dispersion with increasing metallicity, which is not consistent with (primarily photometry-based) claims that

<sup>12</sup> A Kolmogorov–Smirnov test confirms that B17’s distribution is in fact consistent with a uniform distribution: the null hypothesis cannot be rejected at the 5% level.

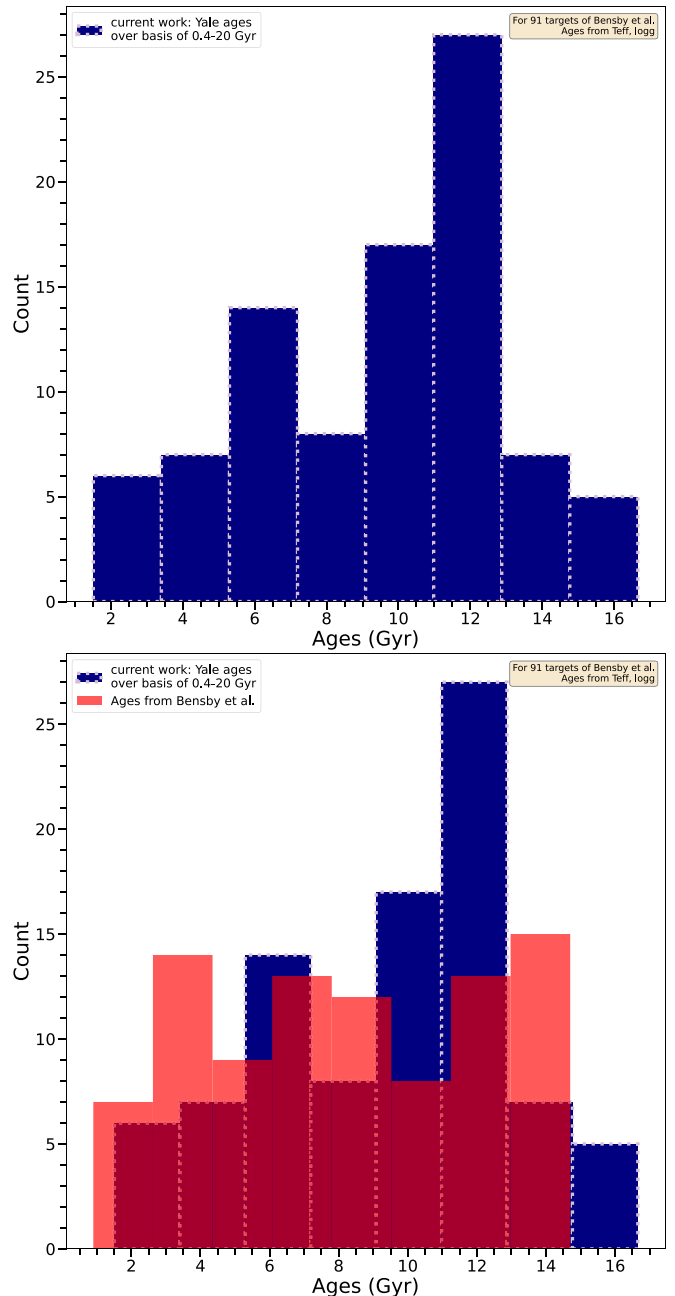


**Figure 5.** Upper: the age distribution computed using MIST isochrones and the algorithm described in Section 4. Lower: two histograms overlaid, comparing the ages presented in B17 (transparent red bars, foreground) to ages derived using MIST isochrones and the same  $\log g$  and  $T_{\text{eff}}$  measurements (purple bars, background).

stellar ages display no metallicity dependence. However, this trend also does not suggest a large constituency of very young stars. Another feature apparent in the upper left panel of Figure 8 is the size of the age uncertainties reported by B17 relative to MIST’s, with the MIST age uncertainties being significantly reduced relative to those presented in B17. We move now to a rigorous discussion of the calculation of the age uncertainties presented here.

### 5. Age Uncertainty Analysis

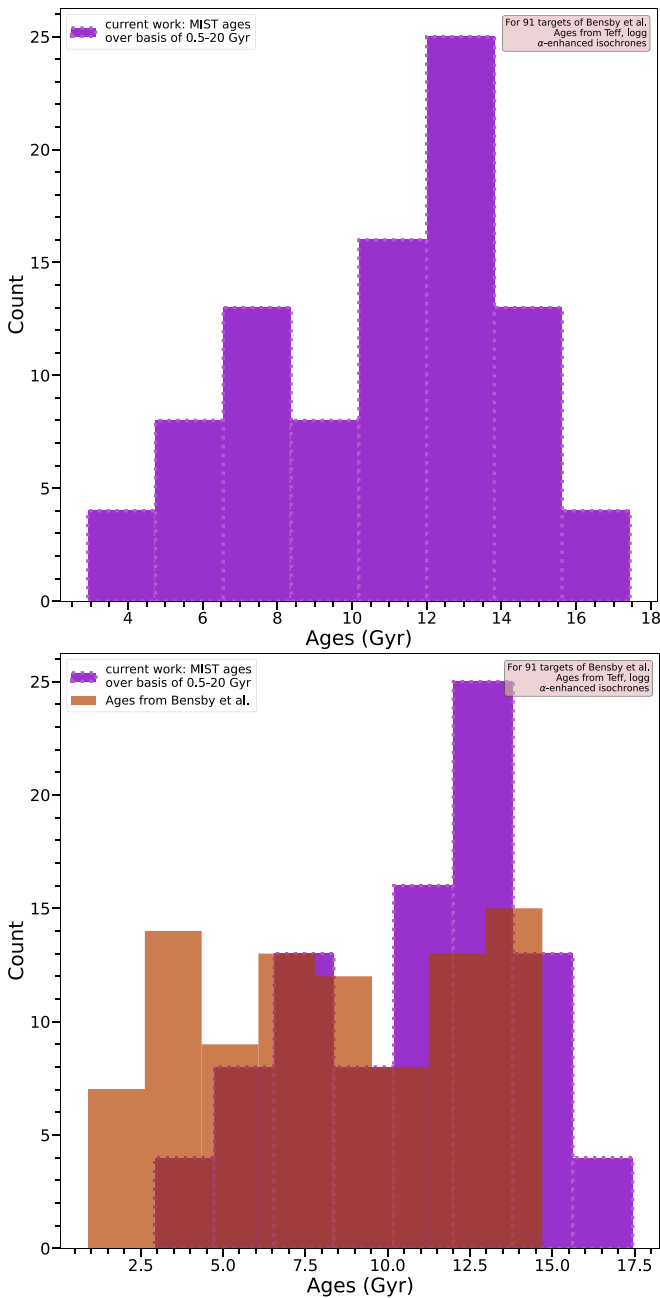
The method described in Section 4 produces an age estimate for a star according to a particular basis of isochrones.



**Figure 6.** Upper: the panel shows the ages of 91 B17 targets computed using the age determination algorithm described in this manuscript, but instead adopting a basis of 2003 Yale isochrones instead of MIST isochrones. These were the isochrones used in B17’s original age determinations and adopt different physical assumptions than MIST. Lower: same as left, but with the B17 distribution overlaid, as in previous figures. Though the presence of a larger number of young stars is clear, the shape of the Yale-based age distribution is still not consistent with the distribution presented in B17.

However, it does not provide a straightforward way of estimating the uncertainty in the  $t_S$  of Equation (3) due to limitations in our knowledge of the observational data set.<sup>13</sup>

<sup>13</sup> Though formal statistics provides a standard definition for the variance in Equation (3), the validity of that definition hinges on features that our algorithm and data set cannot satisfy: first, that  $t_0, \dots, t_n$  be normally distributed—they are not, as these are a set of uniformly spaced candidate ages. Second, all contributing uncertainties must be dimensionless and log-free. They are not, and in some cases we do not have enough information to convert them to the appropriate coordinates.



**Figure 7.** Upper: same as Figure 5, but assuming the  $\alpha$ -enhancement distribution of Table 1 in the MIST isochrones. Lower: the red foreground histogram shows the results from B17 overlaid on the histogram in the left panel.

Therefore, to estimate the variance ( $1\sigma$  error) in  $t_S$ , we use an MC technique inspired by discussion in Angus et al. (2019) and thoroughly detailed in Andrae (2010).

To find the observational age uncertainty for star  $S$ , we adopt as our starting assumption the effective temperature, gravity, metallicity,  $1\sigma$  temperature uncertainty,  $1\sigma \log g$  uncertainty, and  $1\sigma$  [Fe/H] uncertainty reported for  $S$  by B17. From these, we construct three independent normal distributions with densities

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad (4)$$

defined by

$$\begin{aligned} \mu &= T_{\text{eff},S}, & \sigma &= \sigma_{T_{\text{eff},S}}, \\ \mu &= \log g_S, & \sigma &= \sigma_{\log g_S}, \\ \mu &= Z_S, & \sigma &= \sigma_{Z_S}, \end{aligned} \quad (5)$$

for the temperature,  $\log g$ , and [Fe/H], respectively.

In MC resampling, we repeat the age determination 100 times. We calculate the weighted average for one star using 100 independent samples of the parameters  $T_{\text{eff}}$ ,  $\log g$ , and [Fe/H] from the distributions of Equation (5) in place of  $\log g_o$ ,  $T_{\text{eff},o}$ , and  $Z_o$ , respectively, in Equation (2), but always use the B17 uncertainties. This gives a spread of age estimates that we could have seen if B17’s measurements had varied according to their uncertainties; an example is given for MOA-2013-BLG-402S in Figure 10, and similar histograms for all stars can be found in the Appendix. The distribution in Figure 10 has a standard deviation (root variance) of  $\sigma = 1.83$  Gyr for MOA-2013-BLG-402S. In general, this technique gives our estimate of the standard error ( $1\sigma$  uncertainty) in our age estimate,  $t_S$ .

We note that as the number of MC trials increases, the distribution of the age estimates should approach a normal distribution with a mean equal to our age estimate,<sup>14</sup>  $t_S = 10.62$  in the case of MOA-2013-BLG-402S. We see good agreement ( $\mu_{\text{simulated}} = 10.66$  versus  $t_S = 10.62$ , within 0.5%) at 100 trials.

It is important to note that a valid age determination technique must be basis-invariant; that is, it must work independently of the color–magnitude or brightness–temperature system in which the isochrone fits are performed. Our technique satisfies this requirement, as demonstrated through its application to photometry, described in Appendix A.

## 6. Comparison of Age Determinations

Table 2 summarizes all age determinations: the MIST-based ages according to the default and  $\alpha$ -enhanced isochrone bases as well as the BDBS-based and Gaia-based photometric ages discussed in Appendix A. The age determinations from B17 are reproduced for comparison. All ages and uncertainties are in Gyr; [Fe/H] and its uncertainty (in dex) are taken from B17 and adopted in all age determinations for all data sets, physical and photometric. Where observations for a particular star are not available in a given photometric systems, the entry is marked “NA.” Membership in either or both of the BDBS and Gaia gold samples is indicated with markers at the end of the star’s name. In all four systems, the age uncertainties are estimated via the MC technique described in Section 7 using 100 trials.

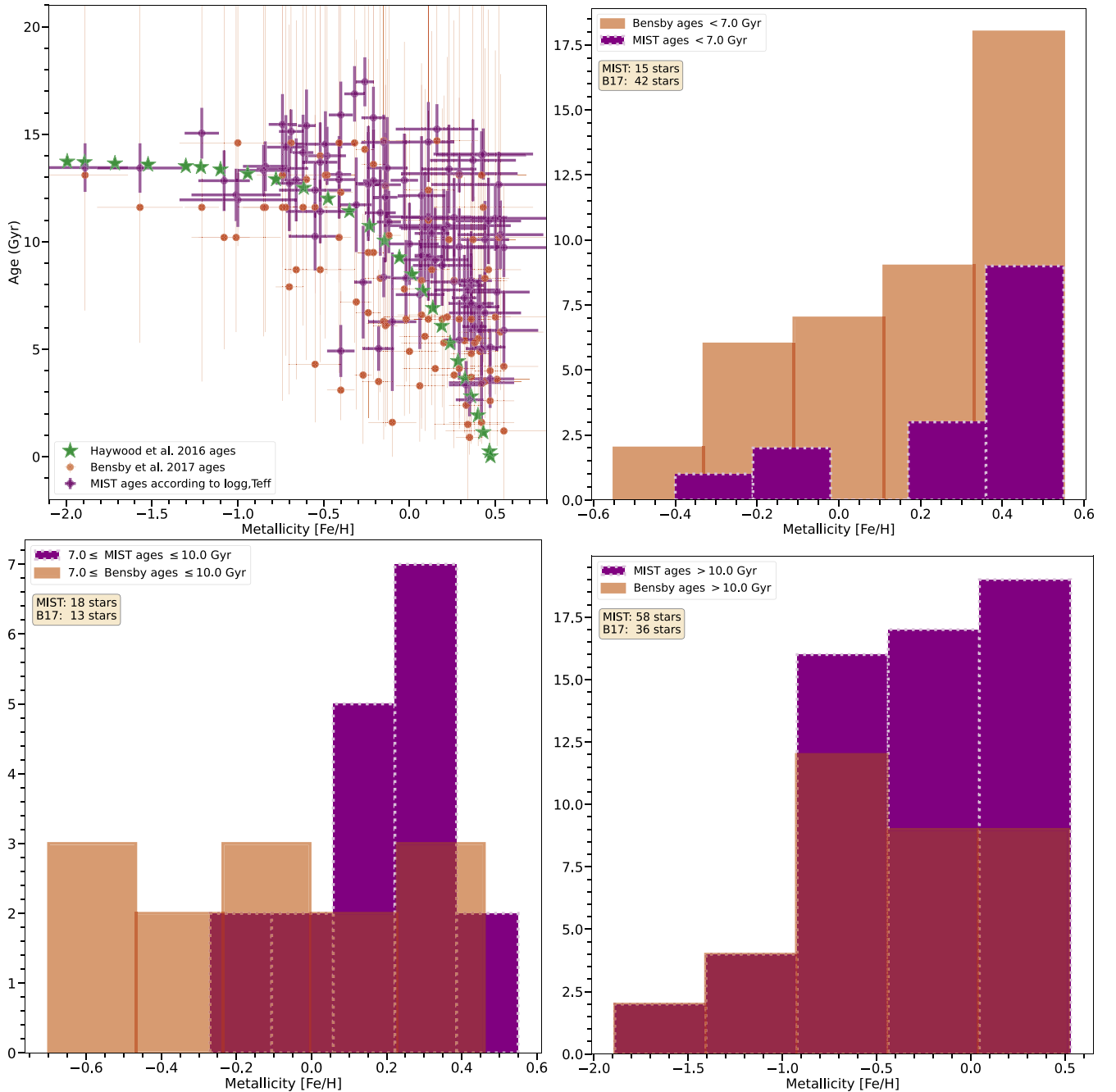
The two most important caveats on results thus far are the following:

1. In all cases discussed previously, the age uncertainties are computed using observational sources of error only; and
2. The error bars are only physically meaningful in the spectroscopic cases.

Point (2) is discussed in the Appendix; we elaborate on point (1) in the next section.

Among the spectroscopic determinations, the ages derived according to default and  $\alpha$ -enhanced isochrone bases are consistent with each other. For a given star, the differences

<sup>14</sup> If the variables in the  $\chi^2$  definition are indeed normally distributed, then the mean of the MC age distribution,  $\mu_{\text{MC}}$ , should approach the predicted age,  $t_S$ , as the number of trials increases.



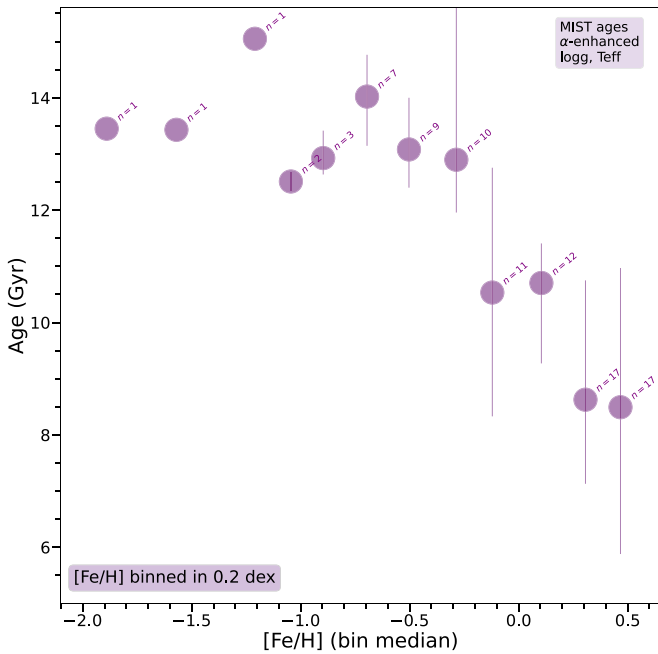
**Figure 8.** Upper left: age vs. metallicity using  $\log g$ ,  $T_{\text{eff}}$ , and the  $\alpha$ -enhanced basis of isochrones, corresponding to the distribution shown in Figure 7. Upper right: histogram showing the distribution of metallicities [Fe/H] for the population of stars with ages less than 7.0 Gyr according to the ages determined in the current work (blue, solid borders) and B17 (red, dashed borders), respectively. Lower left: same as upper right, but for an intermediate age distribution as indicated. Lower right: same as left, but for stars with ages greater than 10 Gyr according to the two age determinations, respectively. In all histogram panels, “MIST ages” refers to the ages determined according to the  $\alpha$ -element enhanced basis of isochrones.

between them range from about 5%–10%. Though this difference is not necessarily negligible, it is eclipsed by nearly a factor of two in the observational age uncertainty, which hovers around 1–2 Gyr, or roughly 20%. This finding suggests that careful consideration of abundances in the models may not be worthwhile given existing spectroscopic precision, but such considerations could become important in light of modest observational uncertainty reductions. Age uncertainties of this order are consistent with other estimates for subgiants and early RGB stars in the literature (e.g., Tayar et al. 2022), but they are not comparable to the upper and lower age limits quoted

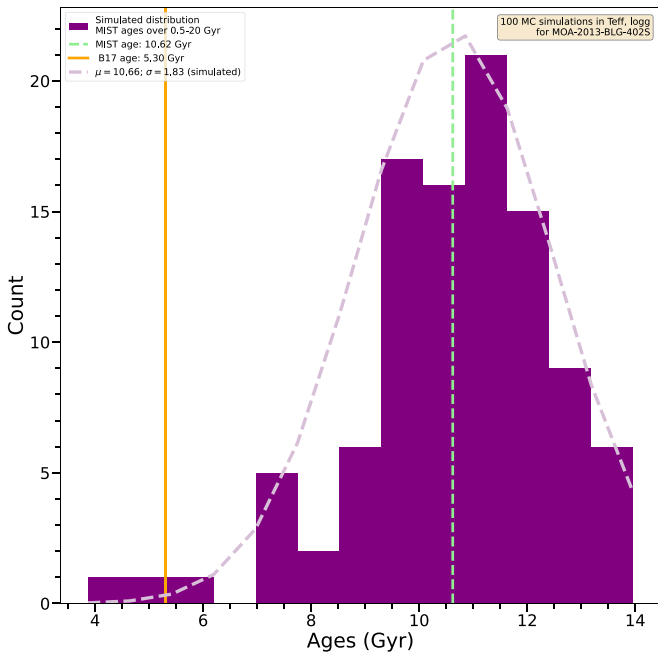
by B17. However, the relationship between the age and uncertainty estimation method presented here and the one used to derive the upper and lower age bounds in B17 is unclear (per the discussion in Section 4.1).

## 7. Modeling Uncertainties

While a comprehensive analysis of modeling systematics in MESA and MIST is beyond the scope of the current study, we must make an effort to avoid the most prominent flaw in all stellar age determinations: incomplete, or completely absent, consideration of theoretical uncertainties.



**Figure 9.** Ages derived with MIST  $\alpha$ -enhanced isochrones are shown as a function of metallicity. Ages represent the average age within the corresponding metallicity bin, where each metallicity bin spans 0.2 dex. The asymmetric vertical bars represent the age dispersion in a given bin: the first and fourth quartiles are the upper and lower segments, respectively. Points without dispersions correspond to bins in which there is only one age measurement. The number of age measurements in each bin is given by  $n$  next to the corresponding marker.



**Figure 10.** A MIST-based age distribution simulated from 100 MC trials in physical coordinate ( $\log g$ ,  $T_{\text{eff}}$ ) for the target MOA-2013-BLG-402S is shown. The variance (second moment,  $\sigma^2$ ) of this distribution provides a reasonable estimate for the age uncertainty.

As mentioned in Section 4, the isochrones themselves adopt a complex network of physical assumptions, each of which has its own associated uncertainty. This patchwork of parameters is highly covariant, and the sum of their effects

is difficult to quantify. While some sources of modeling uncertainty have been studied extensively, such as metallicity and abundance assumptions, atmospheric boundary conditions, convective overshoot, and treatment of diffusion (e.g., Feiden 2015; Choi et al. 2018; Joyce & Chaboyer 2018a; Tayar et al. 2022), these considerations are not typically applied to more than one idealized observational target at a time. Further, there are many additional sources of modeling uncertainty that are neglected entirely; in many cases, this is due to poor or nonexistent constraints on the values of free parameters. A prime example of such a case is the convective mixing length,  $\alpha_{\text{MLT}}$ . More than a decade of research shows that (1) inappropriate choices for  $\alpha_{\text{MLT}}$  can impart huge inaccuracies to model-derived fundamental stellar parameters, and (2) that the appropriate choice mixing length is dependent on metallicity, evolutionary stage, and probably mass (Trampedach et al. 2014; Tayar et al. 2017; Joyce & Chaboyer 2018a, 2018b; Viani et al. 2018; Tang & Joyce 2021; Cinquegrana & Joyce 2022).

Yet, the de facto approach in stellar modeling and isochrone construction is to adopt a fixed, solar value with no further exploration. As Figure 11 makes clear, this is not acceptable.

Figure 11 shows a suite of isochrones computed with the same MESA version and physical settings used in MIST, but with variations in  $\alpha_{\text{MLT}}$  permitted. In the upper left panel, the metallicity ( $Z = 0.0142$ ; roughly solar) and age (10 Gyr) are identical for every isochrone; the only difference is in the value of the mixing length adopted in its constituent stellar tracks, indicated by color. The mixing length values range from 0.4 to 2.3. These choices are physically justified, as some magnetically active M dwarfs have been found to be fit best by mixing lengths as low as  $\alpha_{\text{MLT}} = 0.4 H_P$  (e.g., Mann et al. 2015; Ireland & Browning 2018), and values above  $2.0 H_P$  are common in fits to solar-like stars.

The effects are striking, particularly in terms of effective temperature: in the evolutionary region most pertinent for these observations—the MSTO and subgiant branch—the range in  $T_{\text{eff}}$  induced by varying  $\alpha_{\text{MLT}}$  alone is more than 1500 K. Just above this, on the early RGB, it balloons to 2000 K (the danger of using the RGB for age determinations for precisely this reason has, at least, been noted in Joyce & Chaboyer 2018a). In terms of  $\log g$ , the variation is at least  $\Delta \log g = 0.5$ . In the lowest-valued cases, below about  $\alpha_{\text{MLT}} = 1.2$ , the models show the development of a convective core, indicated by the hooked feature at the MSTO.

In the remaining three panels, the span in mixing length is truncated at a lower value of  $\alpha_{\text{MLT}} = 1.4$  to capture more realistically the range of values appropriate for subgiant and early RGB stars (which are not fully convective). The B17 data set is shown overlaid. Even with a more conservative estimate of the variance in this parameter, the spread in the H-R diagram cannot be ignored. The two right-hand panels show the same age assumption but different metallicities:  $[\text{Fe}/\text{H}] = 0$  (solar) on top versus  $[\text{Fe}/\text{H}] \approx -0.5$  underneath. Here we can see clearly that adjustments to the mixing length (that are well within reason) will comparably or more significantly impact age determinations than metallicity variations of 0.5 dex. Comparing instead the lower two panels, which have the same metallicity assumption but different ages, we find that a difference of  $\alpha_{\text{MLT}} = 1.4$  versus  $\alpha_{\text{MLT}} = 2.3$  is nearly equivalent, in terms of shifts in  $T_{\text{eff}}$ , to an age difference of 5 Gyr.

**Table 2**  
Age Summary

Name	[Fe/H]	B17 Age	MIST Age	... $\alpha$ -enhanced	...Model Err.	BDBS Age	Gaia Age	
1	MOA-2009-BLG-174S	0.11 ± 0.08	6.4 <sup>+9.7</sup> <sub>-3.1</sub>	8.9 ± 2.0	9.3 ± 2.0	9.3 ± 4.0	9.3 ± 1.4	10.0 ± 1.2
2	MOA-2009-BLG-259S	0.34 ± 0.19	1.5 <sup>+8.9</sup> <sub>-99.9</sub>	7.2 ± 2.4	7.8 ± 2.4	7.8 ± 3.8	NA	NA
3	MOA-2010-BLG-167S ★	-0.60 ± 0.05	12.9 <sup>+13.5</sup> <sub>-8.4</sub>	15.3 ± 1.7	15.4 ± 1.7	15.4 ± 3.8	9.5 ± 1.5	NA
4	MOA-2010-BLG-311S	0.51 ± 0.19	3.6 <sup>+7.5</sup> <sub>-3.1</sub>	7.6 ± 2.7	7.7 ± 2.7	7.7 ± 3.3	10.5 ± 1.0	NA
5	MOA-2010-BLG-446S	-0.40 ± 0.08	3.1 <sup>+4.6</sup> <sub>-1.4</sub>	5.1 ± 1.2	4.9 ± 1.2	4.9 ± 2.0	NA	NA
6	MOA-2010-BLG-523S ★	0.06 ± 0.14	3.3 <sup>+8.2</sup> <sub>-2.6</sub>	7.4 ± 2.5	7.5 ± 2.5	7.5 ± 3.8	8.9 ± 1.8	NA
7	OGLE-2011-BLG-0950S ★†	0.33 ± 0.10	2.4 <sup>+3.6</sup> <sub>-1.6</sub>	3.4 ± 1.1	3.3 ± 1.1	3.3 ± 1.6	7.1 ± 1.6	8.3 ± 2.2
8	OGLE-2011-BLG-0969S ★	-1.57 ± 0.25	11.6 <sup>+13.2</sup> <sub>-6.3</sub>	13.6 ± 1.1	13.4 ± 1.1	13.4 ± 1.5	10.1 ± 1.4	NA
9	MOA-2011-BLG-034S	0.11 ± 0.15	12.4 <sup>+13.4</sup> <sub>-8.1</sub>	13.8 ± 1.9	14.6 ± 1.9	14.6 ± 2.8	8.0 ± 1.8	NA
10	MOA-2011-BLG-058S	0.37 ± 0.25	10.1 <sup>+12.9</sup> <sub>-6.1</sub>	13.6 ± 1.9	13.8 ± 1.9	13.8 ± 2.8	11.5 ± 1.0	NA
11	OGLE-2011-BLG-1072S	0.36 ± 0.12	4.8 <sup>+8.4</sup> <sub>-4.1</sub>	8.2 ± 2.3	8.2 ± 2.3	8.2 ± 3.4	NA	NA
12	MOA-2011-BLG-090S	-0.26 ± 0.05	14.3 <sup>+99.9</sup> <sub>-10.7</sub>	17.4 ± 1.1	17.4 ± 1.1	17.4 ± 2.7	12.3 ± 1.1	NA
13	MOA-2011-BLG-104S	-0.85 ± 0.15	11.6 <sup>+13.0</sup> <sub>-6.8</sub>	13.4 ± 1.2	13.3 ± 1.2	13.3 ± 1.8	NA	NA
14	OGLE-2011-BLG-1105S	0.00 ± 0.17	4.9 <sup>+10.3</sup> <sub>-2.9</sub>	9.7 ± 1.9	9.9 ± 1.9	9.9 ± 2.9	NA	NA
15	MOA-2011-BLG-174S	-0.18 ± 0.09	3.5 <sup>+5.2</sup> <sub>-1.8</sub>	5.5 ± 1.0	5.0 ± 1.0	5.0 ± 2.5	9.5 ± 1.7	NA
16	MOA-2011-BLG-191S	0.26 ± 0.14	3.8 <sup>+8.6</sup> <sub>-3.4</sub>	8.0 ± 3.1	8.1 ± 3.1	8.1 ± 4.1	NA	NA
17	MOA-2011-BLG-234S †	-0.02 ± 0.08	6.4 <sup>+8.6</sup> <sub>-2.9</sub>	8.1 ± 1.8	8.3 ± 1.8	8.3 ± 3.0	6.0 ± 2.4	6.8 ± 2.2
18	MOA-2011-BLG-278S	0.52 ± 0.39	10.2 <sup>+12.6</sup> <sub>-5.0</sub>	12.3 ± 1.5	12.7 ± 1.5	12.7 ± 1.8	NA	NA
19	OGLE-2011-BLG-1410S	0.22 ± 0.37	6.5 <sup>+12.1</sup> <sub>-3.1</sub>	10.7 ± 1.6	10.8 ± 1.6	10.8 ± 2.4	NA	NA
20	MOA-2011-BLG-445S	0.26 ± 0.31	8.2 <sup>+12.1</sup> <sub>-4.0</sub>	10.7 ± 1.8	11.1 ± 1.8	11.1 ± 2.8	10.5 ± 1.5	NA
21	MOA-2012-BLG-022S	0.42 ± 0.09	3.4 <sup>+5.1</sup> <sub>-1.9</sub>	5.1 ± 1.3	5.0 ± 1.3	5.0 ± 2.2	NA	NA
22	OGLE-2012-BLG-0026S	0.50 ± 0.44	6.5 <sup>+12.4</sup> <sub>-3.6</sub>	11.0 ± 1.5	11.1 ± 1.5	11.1 ± 2.5	NA	NA
23	MOA-2012-BLG-187S ★	-0.74 ± 0.08	11.6 <sup>+13.4</sup> <sub>-9.0</sub>	15.5 ± 1.4	15.5 ± 1.4	15.5 ± 2.2	9.8 ± 1.8	NA
24	MOA-2012-BLG-202S ★†	-0.15 ± 0.13	14.7 <sup>+99.9</sup> <sub>-8.5</sub>	14.7 ± 1.3	14.6 ± 1.3	14.6 ± 2.1	9.0 ± 1.5	7.4 ± 2.0
25	OGLE-2012-BLG-0211S	-0.03 ± 0.08	7.8 <sup>+11.5</sup> <sub>-5.9</sub>	12.5 ± 2.2	12.9 ± 2.2	12.9 ± 3.5	8.7 ± 1.6	NA
26	OGLE-2012-BLG-0270S	-0.84 ± 0.13	11.6 <sup>+13.0</sup> <sub>-5.8</sub>	13.6 ± 1.2	13.5 ± 1.2	13.5 ± 2.0	NA	NA
27	MOA-2012-BLG-291S	0.16 ± 0.24	14.7 <sup>+99.9</sup> <sub>-7.9</sub>	15.1 ± 1.2	15.2 ± 1.2	15.2 ± 2.1	8.6 ± 1.8	NA
28	MOA-2012-BLG-391S	-0.24 ± 0.07	6.7 <sup>+11.8</sup> <sub>-4.9</sub>	13.0 ± 2.7	12.7 ± 2.7	12.7 ± 3.5	NA	NA
29	MOA-2012-BLG-410S	-0.14 ± 0.07	6.1 <sup>+11.2</sup> <sub>-4.7</sub>	12.2 ± 3.0	12.1 ± 3.0	12.1 ± 3.9	NA	NA
30	OGLE-2012-BLG-0521S	0.09 ± 0.15	5.6 <sup>+11.5</sup> <sub>-4.4</sub>	10.2 ± 2.6	10.7 ± 2.6	10.7 ± 3.7	NA	NA
31	MOA-2012-BLG-532S ★	-0.55 ± 0.21	11.6 <sup>+12.6</sup> <sub>-4.0</sub>	12.7 ± 1.3	12.4 ± 1.3	12.4 ± 1.6	9.6 ± 1.6	NA
32	OGLE-2012-BLG-0563S	-0.66 ± 0.07	8.7 <sup>+11.6</sup> <sub>-5.1</sub>	12.9 ± 1.9	12.9 ± 1.9	12.9 ± 3.6	9.0 ± 1.8	NA
33	OGLE-2012-BLG-0617S ★	-0.14 ± 0.09	12.6 <sup>+13.3</sup> <sub>-6.8</sub>	13.0 ± 1.9	12.7 ± 1.9	12.7 ± 3.2	9.3 ± 1.5	NA
34	OGLE-2012-BLG-0816S	-0.10 ± 0.14	1.6 <sup>+6.4</sup> <sub>-1.6</sub>	6.3 ± 3.2	6.3 ± 3.2	6.3 ± 3.4	12.1 ± 1.1	NA
35	OGLE-2012-BLG-1156S	-1.89 ± 0.25	13.1 <sup>+13.4</sup> <sub>-6.3</sub>	13.5 ± 1.1	13.4 ± 1.1	13.4 ± 1.6	NA	NA
36	OGLE-2012-BLG-1217S	-0.41 ± 0.07	10.2 <sup>+12.4</sup> <sub>-6.3</sub>	13.2 ± 1.8	13.1 ± 1.8	13.1 ± 3.1	10.7 ± 1.2	NA
37	OGLE-2012-BLG-1274S	0.07 ± 0.04	8.2 <sup>+9.1</sup> <sub>-6.9</sub>	9.3 ± 1.1	9.4 ± 1.1	9.4 ± 3.2	9.1 ± 1.7	NA
38	OGLE-2012-BLG-1279S ★	-0.62 ± 0.06	11.6 <sup>+12.8</sup> <sub>-5.5</sub>	14.2 ± 1.8	14.2 ± 1.8	14.2 ± 3.5	10.0 ± 1.3	NA
39	OGLE-2012-BLG-1526S ★	-0.24 ± 0.06	9.5 <sup>+12.6</sup> <sub>-6.0</sub>	12.9 ± 2.4	12.7 ± 2.4	12.7 ± 3.6	9.5 ± 1.7	NA
40	OGLE-2012-BLG-1534S	-0.15 ± 0.04	6.3 <sup>+7.5</sup> <sub>-5.2</sub>	8.6 ± 0.9	8.3 ± 0.9	8.3 ± 3.2	NA	NA
41	MOA-2013-BLG-063S	0.47 ± 0.14	2.6 <sup>+3.5</sup> <sub>-1.2</sub>	3.7 ± 1.4	3.6 ± 1.4	3.6 ± 1.8	NA	NA
42	MOA-2013-BLG-068S	-0.27 ± 0.05	3.8 <sup>+7.5</sup> <sub>-3.2</sub>	8.1 ± 2.6	8.1 ± 2.6	8.1 ± 4.5	NA	NA
43	MOA-2013-BLG-299S	-0.48 ± 0.11	13.1 <sup>+13.1</sup> <sub>-4.6</sub>	14.2 ± 1.4	14.0 ± 1.4	14.0 ± 2.1	9.2 ± 1.4	9.4 ± 1.3
44	MOA-2013-BLG-402S	0.20 ± 0.25	5.3 <sup>+11.5</sup> <sub>-3.8</sub>	10.2 ± 1.8	10.6 ± 1.8	10.6 ± 2.8	7.9 ± 2.0	NA
45	OGLE-2013-BLG-0446S	0.40 ± 0.08	5.5 <sup>+7.5</sup> <sub>-5.0</sub>	7.0 ± 1.2	7.0 ± 1.2	7.0 ± 2.5	4.3 ± 2.5	NA
46	MOA-2013-BLG-517S	-1.08 ± 0.15	10.2 <sup>+12.7</sup> <sub>-5.2</sub>	12.9 ± 1.4	12.8 ± 1.4	12.8 ± 1.6	10.1 ± 1.6	NA
47	MOA-2013-BLG-524S	-1.01 ± 0.26	10.2 <sup>+12.7</sup> <sub>-4.4</sub>	12.3 ± 1.2	12.2 ± 1.2	12.2 ± 1.5	NA	10.3 ± 1.0
48	MOA-2013-BLG-605S	-0.17 ± 0.09	8.3 <sup>+12.8</sup> <sub>-5.3</sub>	11.9 ± 2.0	11.3 ± 2.0	11.3 ± 5.0	9.6 ± 1.5	8.6 ± 1.7
49	OGLE-2013-BLG-0692S ★†	0.15 ± 0.19	4.1 <sup>+9.8</sup> <sub>-3.3</sub>	8.6 ± 2.4	9.2 ± 2.4	9.2 ± 3.8	9.6 ± 1.4	9.2 ± 2.0
50	OGLE-2013-BLG-0835S	0.53 ± 0.25	5.8 <sup>+12.0</sup> <sub>-2.7</sub>	10.5 ± 1.6	10.3 ± 1.6	10.3 ± 2.0	NA	NA
51	OGLE-2013-BLG-0911S	0.47 ± 0.09	4.0 <sup>+5.7</sup> <sub>-3.6</sub>	5.1 ± 0.9	5.1 ± 0.9	5.1 ± 2.2	NA	NA
52	OGLE-2013-BLG-1015S	0.36 ± 0.07	6.4 <sup>+8.2</sup> <sub>-5.3</sub>	8.0 ± 1.4	8.0 ± 1.4	8.0 ± 3.5	NA	NA
53	OGLE-2013-BLG-1114S	0.42 ± 0.19	1.6 <sup>+3.3</sup> <sub>-1.0</sub>	3.3 ± 1.8	3.5 ± 1.8	3.5 ± 1.7	NA	NA
54	OGLE-2013-BLG-1125S	-0.12 ± 0.07	10.3 <sup>+12.6</sup> <sub>-4.2</sub>	11.3 ± 1.4	10.9 ± 1.4	10.9 ± 3.8	NA	NA
55	OGLE-2013-BLG-1147S	0.29 ± 0.09	4.1 <sup>+5.7</sup> <sub>-2.8</sub>	5.4 ± 1.9	5.4 ± 1.9	5.4 ± 2.6	10.7 ± 1.4	9.9 ± 1.4
56	OGLE-2013-BLG-1259S	-0.21 ± 0.06	9.5 <sup>+12.6</sup> <sub>-5.9</sub>	13.0 ± 2.4	12.8 ± 2.4	12.8 ± 4.1	7.1 ± 2.2	8.1 ± 1.8
57	OGLE-2013-BLG-1768S	0.11 ± 0.25	11.0 <sup>+12.7</sup> <sub>-4.0</sub>	11.2 ± 1.3	11.2 ± 1.3	11.2 ± 1.6	NA	NA
58	OGLE-2013-BLG-1793S	0.32 ± 0.07	5.4 <sup>+7.6</sup> <sub>-4.3</sub>	7.4 ± 2.0	7.4 ± 2.0	7.4 ± 3.3	NA	NA

**Table 2**  
(Continued)

	Name	[Fe/H]	B17 Age	MIST Age	... $\alpha$ -enhanced	...Model Err.	BDBS Age	Gaia Age
59	OGLE-2013-BLG-1868S $\star$	$0.38 \pm 0.07$	$5.3^{+6.8}_{-4.3}$	$6.0 \pm 1.1$	$6.0 \pm 1.1$	$6.0 \pm 2.5$	$9.6 \pm 1.4$	$9.5 \pm 1.6$
60	OGLE-2013-BLG-1938S $\dagger$	$-0.41 \pm 0.09$	$14.6^{+99.9}_{-6.5}$	$12.9 \pm 1.8$	$12.9 \pm 1.8$	$12.9 \pm 4.6$	$9.0 \pm 1.8$	$8.7 \pm 1.3$
61	MOA-2014-BLG-131S	$0.29 \pm 0.34$	$13.1^{+99.9}_{-3.8}$	$13.2 \pm 0.9$	$13.2 \pm 0.9$	$13.2 \pm 1.3$	$9.1 \pm 1.9$	NA
62	OGLE-2014-BLG-0157S $\star\dagger$	$-0.52 \pm 0.10$	$14.0^{+99.9}_{-7.4}$	$13.7 \pm 1.9$	$13.7 \pm 1.9$	$13.7 \pm 3.8$	$10.0 \pm 1.1$	$10.3 \pm 1.2$
63	OGLE-2014-BLG-0801S	$0.55 \pm 0.34$	$1.2^{+11.5}_{-99.9}$	$10.1 \pm 1.4$	$9.7 \pm 1.4$	$9.7 \pm 2.6$	$9.4 \pm 1.6$	NA
64	OGLE-2014-BLG-0953S	$-0.13 \pm 0.12$	$6.2^{+12.3}_{-4.6}$	$13.3 \pm 2.4$	$13.4 \pm 2.4$	$13.4 \pm 3.0$	NA	NA
65	OGLE-2014-BLG-0987S	$0.23 \pm 0.16$	$10.1^{+12.8}_{-6.0}$	$13.1 \pm 2.1$	$13.4 \pm 2.1$	$13.4 \pm 2.6$	$2.4 \pm 1.6$	NA
66	OGLE-2014-BLG-1122S $\star$	$-0.32 \pm 0.06$	$14.6^{+99.9}_{-9.3}$	$16.9 \pm 1.3$	$16.9 \pm 1.3$	$16.9 \pm 2.8$	$9.3 \pm 1.6$	NA
67	OGLE-2014-BLG-1370S	$0.43 \pm 0.29$	$11.6^{+13.3}_{-7.2}$	$13.8 \pm 1.2$	$14.1 \pm 1.2$	$14.1 \pm 2.6$	NA	NA
68	OGLE-2014-BLG-1418S	$0.44 \pm 0.14$	$8.3^{+10.7}_{-6.2}$	$10.1 \pm 1.8$	$10.1 \pm 1.8$	$10.1 \pm 3.5$	$8.9 \pm 1.9$	$9.8 \pm 0.9$
69	OGLE-2014-BLG-1469S	$-0.70 \pm 0.08$	$7.9^{+12.2}_{-5.0}$	$12.7 \pm 2.2$	$12.7 \pm 2.2$	$12.7 \pm 3.5$	NA	NA
70	OGLE-2014-BLG-2040S	$0.19 \pm 0.17$	$6.4^{+9.8}_{-4.1}$	$8.2 \pm 1.9$	$8.9 \pm 1.9$	$8.9 \pm 2.8$	NA	NA
71	OGLE-2015-BLG-0078S	$-0.55 \pm 0.15$	$4.3^{+11.6}_{-2.7}$	$10.5 \pm 1.6$	$10.2 \pm 1.6$	$10.2 \pm 2.9$	NA	NA
72	MOA-2015-BLG-111S	$0.07 \pm 0.14$	$6.6^{+11.7}_{-4.8}$	$11.9 \pm 2.8$	$12.2 \pm 2.8$	$12.2 \pm 3.6$	NA	NA
73	OGLE-2015-BLG-0159S $\star$	$0.36 \pm 0.13$	$3.7^{+7.3}_{-3.1}$	$7.1 \pm 2.4$	$7.1 \pm 2.4$	$7.1 \pm 3.6$	$7.9 \pm 2.1$	NA
74	MOA-2010-BLG-037S	$0.55 \pm 0.20$	$4.2^{+6.2}_{-3.1}$	$5.8 \pm 1.8$	$5.9 \pm 1.8$	$5.9 \pm 2.1$	NA	NA
75	MOA-2010-BLG-049S	$-0.40 \pm 0.07$	$12.3^{+13.4}_{-8.8}$	$15.9 \pm 1.6$	$15.9 \pm 1.6$	$15.9 \pm 3.2$	$9.6 \pm 1.3$	NA
76	OGLE-2009-BLG-076S	$-0.72 \pm 0.11$	$11.6^{+13.3}_{-6.9}$	$14.4 \pm 1.5$	$14.4 \pm 1.5$	$14.4 \pm 2.3$	$10.2 \pm 1.5$	NA
77	MOA-2010-BLG-078S	$-1.00 \pm 0.34$	$14.6^{+99.9}_{-4.8}$	$12.0 \pm 1.3$	$11.9 \pm 1.3$	$11.9 \pm 1.5$	NA	NA
78	MOA-2006-BLG-099S	$0.44 \pm 0.21$	$3.5^{+7.0}_{-1.7}$	$6.5 \pm 1.9$	$6.7 \pm 1.9$	$6.7 \pm 2.7$	NA	NA
79	MOA-2009-BLG-133S $\star$	$-0.69 \pm 0.07$	$14.6^{+99.9}_{-6.1}$	$15.2 \pm 1.0$	$15.1 \pm 1.0$	$15.1 \pm 1.9$	$11.3 \pm 1.4$	NA
80	OGLE-2008-BLG-209S $\star\dagger$	$-0.31 \pm 0.09$	$7.2^{+12.2}_{-5.0}$	$12.0 \pm 2.2$	$11.7 \pm 2.2$	$11.7 \pm 3.4$	$9.1 \pm 1.6$	$11.0 \pm 1.3$
81	OGLE-2006-BLG-265S	$0.46 \pm 0.19$	$8.7^{+12.1}_{-5.8}$	$10.7 \pm 2.0$	$11.0 \pm 2.0$	$11.0 \pm 3.1$	$7.9 \pm 1.9$	NA
82	MOA-2010-BLG-285S $\star\dagger$	$-1.21 \pm 0.10$	$11.6^{+13.4}_{-8.1}$	$15.1 \pm 1.2$	$15.1 \pm 1.2$	$15.1 \pm 1.7$	$11.2 \pm 1.4$	$9.3 \pm 1.1$
83	MOA-2008-BLG-310S	$0.41 \pm 0.11$	$4.9^{+8.2}_{-4.2}$	$7.0 \pm 1.5$	$7.0 \pm 1.5$	$7.0 \pm 2.8$	$1.4 \pm 1.2$	NA
84	MOA-2008-BLG-311S	$0.35 \pm 0.08$	$0.9^{+2.9}_{-0.8}$	$2.9 \pm 0.8$	$2.6 \pm 0.8$	$2.6 \pm 1.7$	$12.1 \pm 1.1$	NA
85	OGLE-2007-BLG-349S	$0.42 \pm 0.26$	$13.1^{+13.5}_{-6.3}$	$13.8 \pm 1.1$	$14.0 \pm 1.1$	$14.0 \pm 2.2$	NA	NA
86	MOA-2009-BLG-456S	$0.13 \pm 0.10$	$8.7^{+11.1}_{-6.3}$	$9.7 \pm 2.1$	$10.4 \pm 2.1$	$10.4 \pm 3.4$	$5.3 \pm 2.8$	NA
87	MOA-2009-BLG-475S	$-0.52 \pm 0.20$	$8.7^{+12.3}_{-4.3}$	$11.9 \pm 1.5$	$11.4 \pm 1.5$	$11.4 \pm 2.1$	NA	NA
88	MOA-2009-BLG-489S	$-0.21 \pm 0.07$	$13.6^{+13.8}_{-9.1}$	$16.3 \pm 1.4$	$15.8 \pm 1.4$	$15.8 \pm 3.8$	NA	NA
89	MOA-2009-BLG-493S	$-0.74 \pm 0.15$	$13.1^{+99.9}_{-4.3}$	$13.7 \pm 0.7$	$13.4 \pm 0.7$	$13.4 \pm 1.2$	NA	NA
90	OGLE-2007-BLG-514S	$0.29 \pm 0.23$	$6.4^{+10.3}_{-4.6}$	$8.9 \pm 1.8$	$9.8 \pm 1.8$	$9.8 \pm 2.5$	NA	NA
91	MACHO-1999-BLG-022S	$-0.49 \pm 0.17$	$13.1^{+13.5}_{-7.6}$	$14.9 \pm 1.5$	$14.5 \pm 1.5$	$14.5 \pm 2.1$	$11.0 \pm 1.5$	NA

**Note.** All ages are shown in Gyr. “MIST Age” refers to the refitting of B17’s own physical coordinates ( $T_{\text{eff}}$ ,  $\log g$ ) in the present analysis. The  $\star$  symbol indicates membership in the BDBS gold sample;  $\dagger$  indicates membership in the Gaia gold sample. We strictly emphasize that the uncertainties reported for the photometric results (BDBS, Gaia) are not physical (see the discussion in the Appendix); they are simply the results of simulations performed with insufficiently precise observational constraints. Only the uncertainties associated with spectroscopic age determinations should be interpreted as physical, and only those that further account for modeling uncertainties should be taken as meaningful lower limits on real age uncertainties.

Based on these order-of-impact arguments, we perform a simplified estimate of the total age uncertainty through the inclusion of theoretical terms motivated by Figure 11. In a new set of MC simulations, we adopt the same definitions for  $\chi^2$  and  $t_S$  given in Equations (2) and (3), respectively, and sample from the same distributions in observational parameters given in Equation (5), but we introduce an additional term accounting for variation in the isochrones. In each MC simulation, we sample normal distributions given by

$$\mu = 0, \quad \sigma = \sigma_{T_{\text{eff,th}}}, \quad (5a)$$

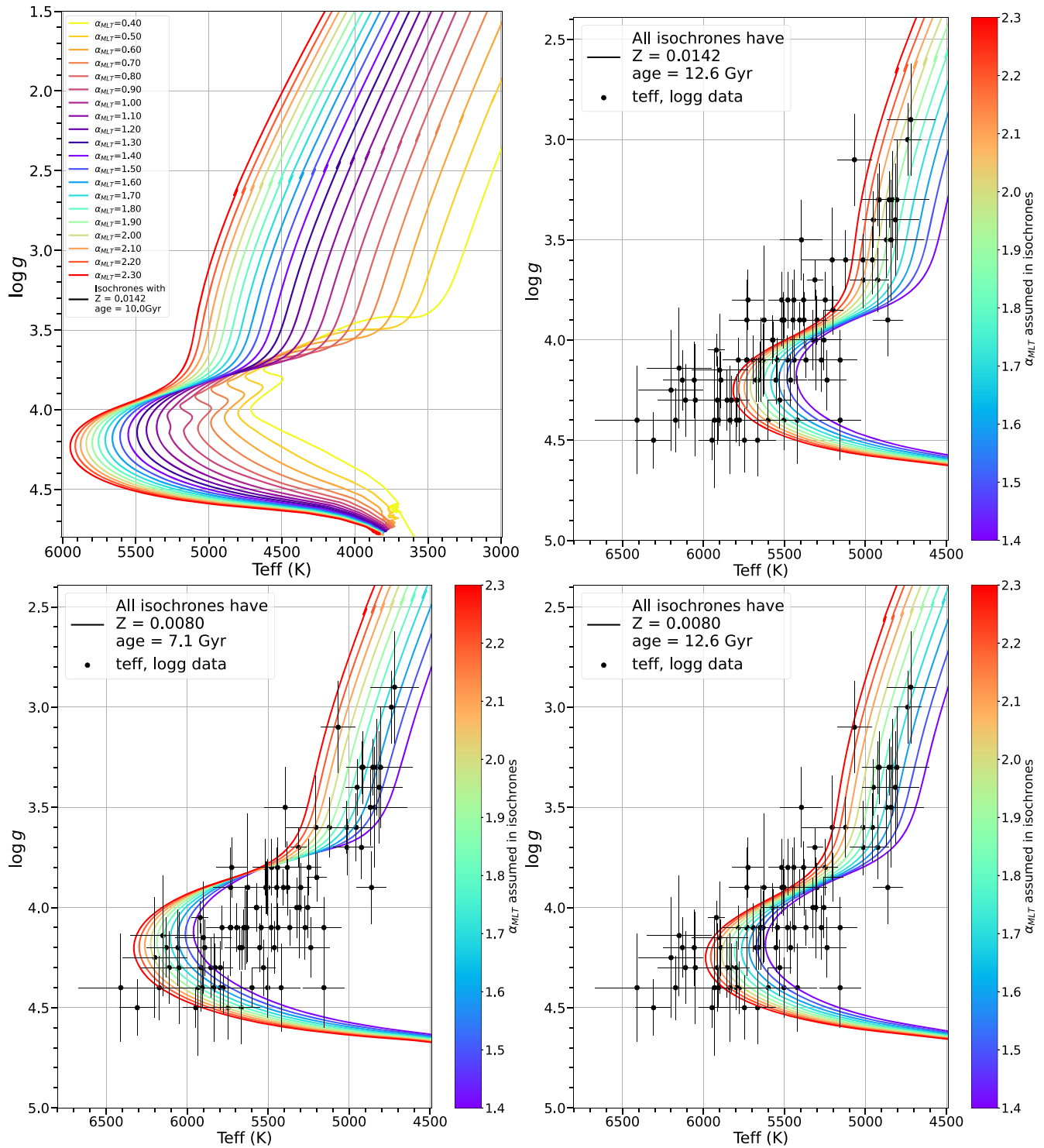
$$\mu = 0, \quad \sigma = \sigma_{\log g_{\text{th}}}, \quad (5b)$$

uniformly shifting the entire basis of isochrones horizontally by a random sample of Equation (5a) and vertically by a random sample from Equation (5b). The rest of the algorithm proceeds as normal. The values adopted for the order-of-magnitude

theoretical uncertainties in each direction are  $\sigma_{T_{\text{eff,th}}} = 200$  K and  $\sigma_{\log g_{\text{th}}} = 0.17$  dex, representing approximately one-third of the spans<sup>15</sup> shown in the H-R diagrams of Figure 11. As these choices are based on mixing length arguments alone, these numbers should be taken to represent lower limits on the true degree of model variation. While a rigorous estimate of the total model uncertainty is beyond the scope of this paper, uncertainty in  $\alpha_{\text{MLT}}$  is likely to dominate, so this is not an unreasonable first-order approximation.

The results of these simulations are given in the “model err” column of Table 2. For a given star, the typical increase from the observational age uncertainty to the global age uncertainty is 1–2 Gyr, thus putting the true age uncertainties anywhere between 2 and 5 Gyr. These are not only more realistic, they are realistic lower limits.

<sup>15</sup> The span is taken to represent a  $3\sigma$  spread, as a crude approximation.



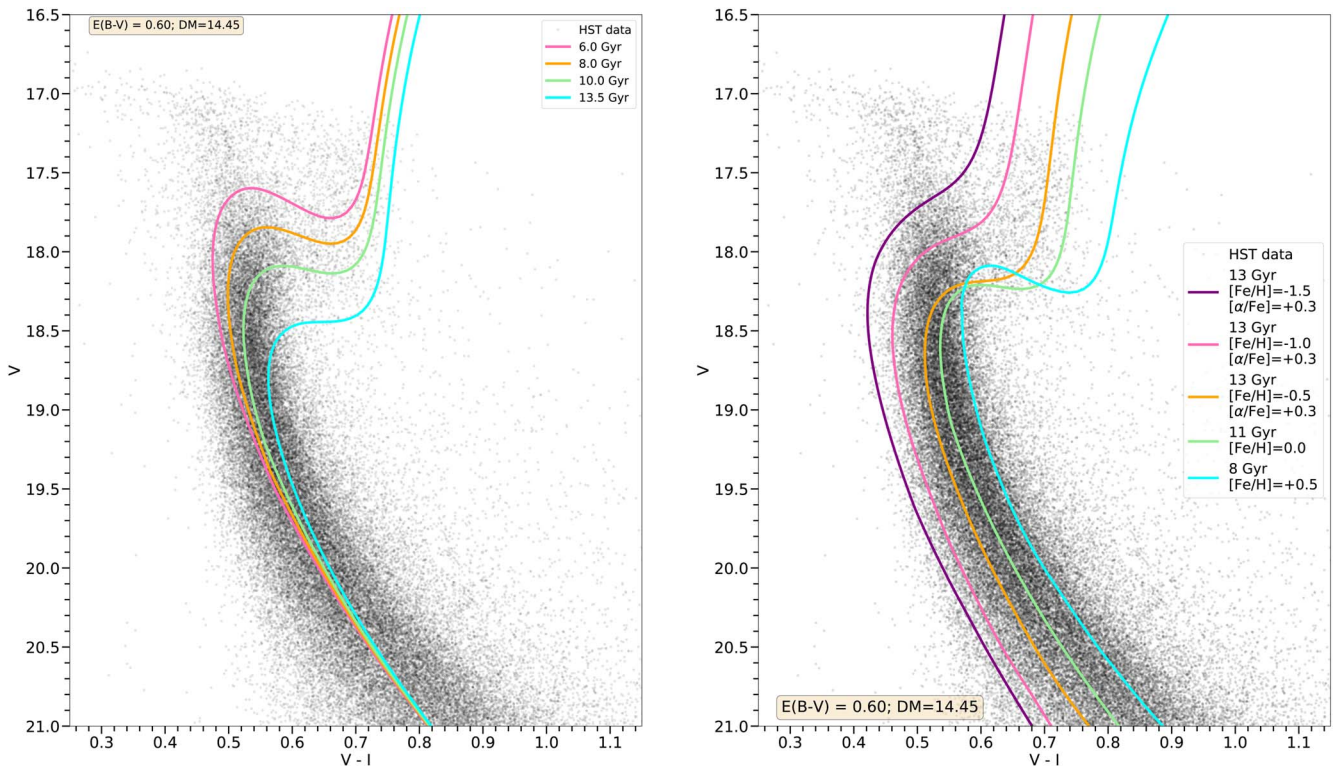
**Figure 11.** As an example, the intrinsic temperature scatter among custom isochrones (computed with MESA models) due to ad hoc assumptions about convection (choice of mixing length,  $\alpha_{MLT}$ ) is at least half of the nearly 2000 K temperature span shown here near the MSTO region and base of the subgiant branch.

### 8. Comparison with HST Data

As discussed in Section 1, most HST analyses have concluded that the bulge is uniformly old with a small age spread, typically finding a mean age of  $\sim 10\text{--}11$  Gyr and a spread of  $\sim 1\text{--}2$  Gyr (e.g., Kuijken & Rich 2002; Clarkson et al. 2008; Brown et al. 2010; Clarkson et al. 2011; Gennaro et al. 2015; Renzini et al. 2018). In contrast, Haywood et al. (2016) note that the MSTOs of HST CMDs are too narrow in color to

be compatible with a population that simultaneously possesses a large metallicity spread and a small age spread. Haywood et al. (2016) instead claim that incorporating an age–metallicity relation that forces metal-rich stars to be progressively younger reduces the MSTO color spread and resolves the tension between photometric and spectroscopic studies finding discrepant age distributions.

However, the age–metallicity relation adopted by Haywood et al. (2016), which follows Bensby et al. (2013), includes a



**Figure 12.** Left: a proper motion–cleaned HST CMD of the low-latitude SWEEPS field using data from Calamida et al. (2015) is shown with a set of MIST isochrones spanning ages of 2–13 Gyr and with  $[\text{Fe}/\text{H}] = 0$  over plotted. We find that the bulge MSTO is bracketed by ages of  $\sim 6$ –13 Gyr, and that stars younger than 4 Gyr are too blue and bright to be compatible with the data. Note that most of the stars brighter and bluer than the nominal MSTO are known to be blue stragglers (Clarkson et al. 2011). Right: similar isochrones are shown that represent our derived age–metallicity relation (Figure 9). Since our age–metallicity relation restricts the youngest stars to those with  $[\text{Fe}/\text{H}] > 0$ , the younger ages, which push stars bluer, are offset by the increased metallicity, which drives stars redder, such that the breadth of the expected MSTO is reduced without producing the large contingent of bright, blue main-sequence stars seen in Haywood et al. (2016).

significant number of stars younger than 5 Gyr. As a result, Barbuy et al. (2018) pointed out that the Haywood et al. (2016) age–metallicity relation produces too many bright, blue MSTO stars to be compatible with the observed CMDs. Their age–metallicity relation also contrasts with ours presented in Figure 9, which is much flatter, contains only a small number of stars  $< 6$  Gyr in age, and has a median age of  $\sim 8$  Gyr for the most metal-rich bin. Both Haywood et al. (2016) and the present study agree that the bulge is “predominantly old,” but our work strongly removes the youngest age bins. As a result, Figure 12, which shows a proper motion–cleaned HST CMD of the low-latitude SWEEPS field described in Calamida et al. (2015), illustrates that our derived age–metallicity relation is more compatible with the observations. The mild slope of the age–metallicity relation found here retains the benefit of reducing the MSTO color spread but does not include such young ages as to generate a large contingent of unobserved bright, blue stars. Therefore, Figure 12 reinforces our earlier conclusion that the bulge is bracketed by ages of  $\sim 6$ –13 Gyr, and that no appreciable fraction of stars have ages  $< 4$ –5 Gyr.

## 9. Implications of Our Derived Ages

The application of multiple approaches to age determination has not relieved the tension between isochrone-derived ages and ages derived from analysis of HST CMDs with proper-motion cleaning (Clarkson et al. 2008; Renzini et al. 2018) or foreground subtraction (Zoccali et al. 2003), though our new age–metallicity relation may help improve compatibility between some of these methods (see Section 8). Figure 8

illustrates a sharp break in the derived age dispersion at  $[\text{Fe}/\text{H}] \sim -0.5$ , with more metal-rich stars spanning a range from 12 to 13 Gyr in age, and being in the mean  $\sim 3$  Gyr younger than their metal-poor counterparts. In agreement with Bernard et al. (2018), we find that most bulge stars are old ( $> 10$  Gyr) or of intermediate ( $> 5$ –8 Gyr) age; however, we do not find any stars  $< 3$  Gyr old and the prevalence of intermediate-age stars rises sharply for  $[\text{Fe}/\text{H}] > 0.3$ . We cannot find any aspect of our analysis or the isochrone physics that would lead to a systematic derivation of younger ages at high metallicity. In contrast to B17, our analysis gives an increase in intermediate-age stars that is clearly pronounced at high metallicity; we conclude this is an actual physical aspect of our sample.

We may find some support for the age range at high metallicity from other observations. Catchpole et al. (2016) find that long-period Mira variables (attributed to intermediate age populations) tend to follow the bar structure, while short-period (older) Miras do not. Hasselquist et al. (2020) use a large sample of APOGEE data to argue that a large fraction of the bulge—at least half their sample—is old ( $> 8$  Gyr), and that the population of young (2–5 Gyr) stars they do find is strongly constrained to the Galactic plane. The present work, along with Marchetti et al. (2022), do not confirm the presence of such young stars. However, we note that Marchetti et al. (2022) does not sample inside  $b \sim -3.5^\circ$ , where young stars may be more prevalent. On the other hand, the B17 sample analyzed here shows no significant correlation between age and Galactic latitude.

It is difficult to age date populations of late-type stars such as SiO masers from their bolometric luminosities (uncertain also

due to distance), but there are indications of a widespread population of luminous SiO masers that follow the bulge/bar population kinematics (Trapp et al. 2018). However, the most luminous late-type stars are extremely concentrated to the Galactic center (Blommaert & Groenewegen 2007). This work and our foreground cleaning of disk stars from the Galactic bulge (Marchetti et al. 2022) argue strongly against a significant population of main-sequence progenitors younger than 3 Gyr. Considering the lack of main-sequence progenitor candidates, we conclude that no substantial population of massive, evolved AGB stars is present in large numbers in the bulge.

Studies of galaxy morphology (Sheth et al. 2008) find that the fraction of barred galaxies decreases strongly with redshift, dropping from 0.6 at  $z = 0.1$  to 0.3 at  $z = 0.8$ . The evolution of bar fraction is striking in that it occurs at relatively low redshift. This suggests that the formation of bars in spiral galaxies may be a relatively recent phenomenon that took place within the last 5–10 Gyr.

Dynamical models (e.g., Shen et al. 2010) have long found that the formation of bars from a massive stellar disk occurs rapidly in just a few rotation times. Such a scenario would appear at first to be in stark disagreement with an age range for stars in the bar, and would especially challenge a scenario in which stars of higher metallicity are more concentrated to the plane. It is possible that once the bar forms, gas flows support additional star formation from the enriched gas, with the younger, more metal-rich population observed closer to the plane. Martinez-Valpuesta & Gerhard (2013) advance a scenario in which the initial disk has a radial metallicity gradient that, after the bar forms, is mapped into a comparable vertical metallicity gradient. Their model preserves the basic premise of a rapid buckling of the bar while allowing for an observed metallicity gradient. The present-day region of the bar closest to the Galactic center, however, shows negligible evidence of a metallicity gradient (Johnson et al. 2020).

The chemistry of the bulge offers some additional support for this age–metallicity relationship. The trend of  $[\alpha/\text{Fe}]$  versus  $[\text{Fe}/\text{H}]$  is clearly toward scaled solar with increasing metallicity (e.g., Gonzalez et al. 2011; Johnson et al. 2014), consistent with Type Ia supernovae having the time to contribute enough iron to force this trend. However, similar trends are seen in the thick disk, for which the evidence for an extended formation history is less compelling, and the long timescales needed for Type I supernovae to pull  $[\alpha/\text{Fe}]$  toward solar-scaled composition (Tinsley 1979; Greggio et al. 2008; Maoz et al. 2010) may not be available.

In conclusion, our new ages do not find support for a population of young, massive stars in the bulge (age  $< 3$  Gyr). However, the tension between the ages of microlensed dwarfs derived from spectroscopic analysis and isochrone fitting (and also via photometric isochrone fits) and proper motion–cleaned HST-based CMDs remains unresolved. The new, more carefully computed ages point toward a growing likelihood that the most metal-rich stars in the bar are younger than their more metal-poor counterparts. This conclusion looks increasingly attractive as new evidence accumulates for a progressive concentration of metal-rich stars toward the plane. However, we emphasize that while a number of disparate approaches point toward support for modestly younger ages or a wider age dispersion in the bar, all age determination methods face challenges in precise measurements for stars older than  $\sim 5$

Gyr; additional robust measurements are needed. While challenging, it would be desirable to construct proper-motion-separated HST-based CMDs that are also segregated by metallicity, to confirm whether the metal-rich population shows evidence of a greater age dispersion. The work by Renzini et al. (2018) is a good first step in this direction, and we envision building on their methods to provide more granular examinations of the age–metallicity relation in the bulge. It will be important to find new approaches to test the age dispersion of the bulge/bar population, as there are important implications for the formation and evolution of the bulge/bar.

## 10. Summary

We present new age determinations of the Milky Way bulge microlensed dwarf sample presented in B17 by combining their spectroscopic stellar parameters with out-of-the-box and custom modern MIST isochrones and carefully considered statistical analysis techniques. We also include an analysis of the impact that various error sources have on the derived ages, such as: the choice of isochrone grid (MIST versus YY), including (or not) metallicity dependent  $\alpha$ -element enhancement, and reasonable variations in the mixing length parameter. The newly derived ages are then combined with the previously published  $[\text{Fe}/\text{H}]$  values to construct an updated age–metallicity relation for the bulge.

The full microlensed dwarf sample analyzed here has a mean age of about 11 Gyr and a spread of  $\sim 3.5$  Gyr. We find a dominant old peak near 13 Gyr and a secondary peak near 8 Gyr, but we do not find any younger local maxima in the bulge age distribution. Our newly derived mean age is nearly 3 Gyr older than the mean age presented in B17, and unlike the B17 results, we do not find any appreciable population of stars  $< 5$  Gyr in age. The strong reduction in the number of young bulge stars is most clearly illustrated when comparing our results against those in B17 for stars  $< 7$  Gyr in age. The latter finds that nearly 50% of their sample is younger than 7 Gyr while we find only 16%. Similarly, we find 66% of the microlensed sample have ages  $> 10$  Gyr while B17 find only 40%. Therefore, our derived age distribution is both more strongly weighted toward older stars and has a sharper cut-off on the young end compared to the B17 distribution. These results are robust against variations in the assumed  $[\alpha/\text{Fe}]$  ratios and choice of stellar isochrone grid as we find nearly the same distributions, modulo some small-scale zero-point shifts, when purely solar-scaled compositions are used and when adopting the YY grid used by B17.

In addition to deriving ages using the spectroscopic parameters, we crossmatched the B17 sample against stars observed in the  $g$  and  $i$  bands with the BDBS and with the Gaia EDR3 database. In the former, we found 51 stars in common, while in the latter, we found 17 matches. In both photometric cases, we included all dominant error sources in the MC fitting procedure and found mean ages of  $\sim 9$ –10 Gyr and spreads of  $\sim 1$ –2 Gyr, which is about 1–2 Gyr younger than the spectroscopic sample but with a much more narrow distribution. Interestingly, our photometric age distribution for individual stars closely matches those found in the literature when fitting cleaned MSTOs in bulge CMDs. However, further analysis showed that the combined effects of photon noise, distance, bulge membership, and small-scale reddening uncertainties were too large to provide any meaningful results from isochrone fitting. Instead, when using only photometry, the

fitting procedure simply trended toward the median of the isochrone grid, thus providing no useful information on the age distribution (see Appendix A).

We investigated the age sensitivity to the MIST isochrone’s adopted mixing length, which is one of the most significant sources of uncertainty in stellar models, especially near the MSTO. Simply changing  $\alpha_{\text{MLT}}$  between 1.4 and 2.3, which are all plausible values for low-mass stars in the parameter space analyzed here, added at least 1–2 Gyr to the global age uncertainty. Modifying this parameter alone increases the lower limit of the age uncertainty of an individual star to between 2 and 5 Gyr, thus showing that the age uncertainty values from many previous works are likely severely underestimated.

In order to investigate further the discrepancies between photometric and spectroscopic age distributions of bulge MSTO stars, we applied the age–metallicity relation derived here from spectroscopy and overlaid the resulting isochrones onto a proper-motion-cleaned HST CMD of the bulge SWEEPS field. A previous analysis by Haywood et al. (2016) suggested that the bulge must have a significant age dispersion, rather than being uniformly old, because a purely old bulge that also has a large metallicity spread would yield too broad of an MSTO to match the observations. However, the age–metallicity relation adopted by Haywood et al. (2016), which follows the results of Bensby et al. (2013), includes a substantial number of stars younger than 5 Gyr in age. As a result, their simulated CMD produces a prominent plume of bright, blue stars that are not observed in the cleaned CMDs. However, we found that applying our new age–metallicity relation, which has very few stars <5 Gyr in age and is on average older, is able to reduce the color spread at the MSTO to the expected levels while avoiding the production of much younger, brighter stars that are clearly not present in the data. Therefore, we conclude that the bulge is bracketed by ages of about 6–13 Gyr with little room for a significant contingent of younger stars.

M.J. was supported by the Lasker Data Science Fellowship, awarded by the Space Telescope Science Institute. M.J. acknowledges the Kavli Institute for Theoretical Physics at the University of California, Santa Barbara, whose collaborative residency program TRANSTAR21 was supported by the National Science Foundation under grant No. NSF PHY-1748958. M.J. wishes to thank László Molnár, Jamie Tayar, Aaron Dotter, and Brian Chaboyer for helpful discussion. M.J. further acknowledges the European Commission’s Widening Fellowship: this project has received funding from the European Union’s Horizon 2020 research and innovation programme. C.I.J. gratefully acknowledges Annalisa Calamida and Will Clarkson for sharing data and details of their HST reductions for the HST SWEEPS field. T.M. acknowledges an ESO Fellowship.

*Facilities:* HST(STIS), Swift(XRT and UVOT), AAVSO, CTIO:1.3 m, CTIO:1.5 m, CXO.

## Appendix A

### Age Fits Using Photometric Systems

As an independent check on our age determinations from B17’s own measurements, we attempt to derive ages from colors and magnitudes for those B17 sources which also have photometry from the BDBS (Rich et al. 2020; Johnson et al. 2020) and/or Gaia EDR3 (Gaia Collaboration et al.

2021). Among B17’s 91 stars, we find 51 that have reliable BDBS  $g$ - and  $i$ -band photometry and 16 that have reliable  $G$ ,  $B_P$ , and  $R_P$  Gaia photometry. Within these, especially well-determined targets are selected for “gold samples,” described below.

#### A.1. BDBS Photometry

The BDBS survey is a deep, UV–optical Dark Energy Camera (DECam) imaging campaign spanning >200 contiguous square degrees of the southern Milky Way bulge. The survey reaches below the MSTO along most sight lines, and includes photometry in the DECam  $ugrizY$  bands that are calibrated to the Sloan Digital Sky Survey (SDSS) ( $u$ ) and PanSTARRS ( $grizy$ ) systems (Tonry et al. 2012). Nearly 250 million unique point sources are included in the BDBS catalog, along with a  $1' \times 1'$  reddening map spanning  $|l| < 10^\circ$  and  $b > -10^\circ$ . Details regarding the project goals, performance, and early science results are presented in Rich et al. (2020) and Johnson et al. (2020).

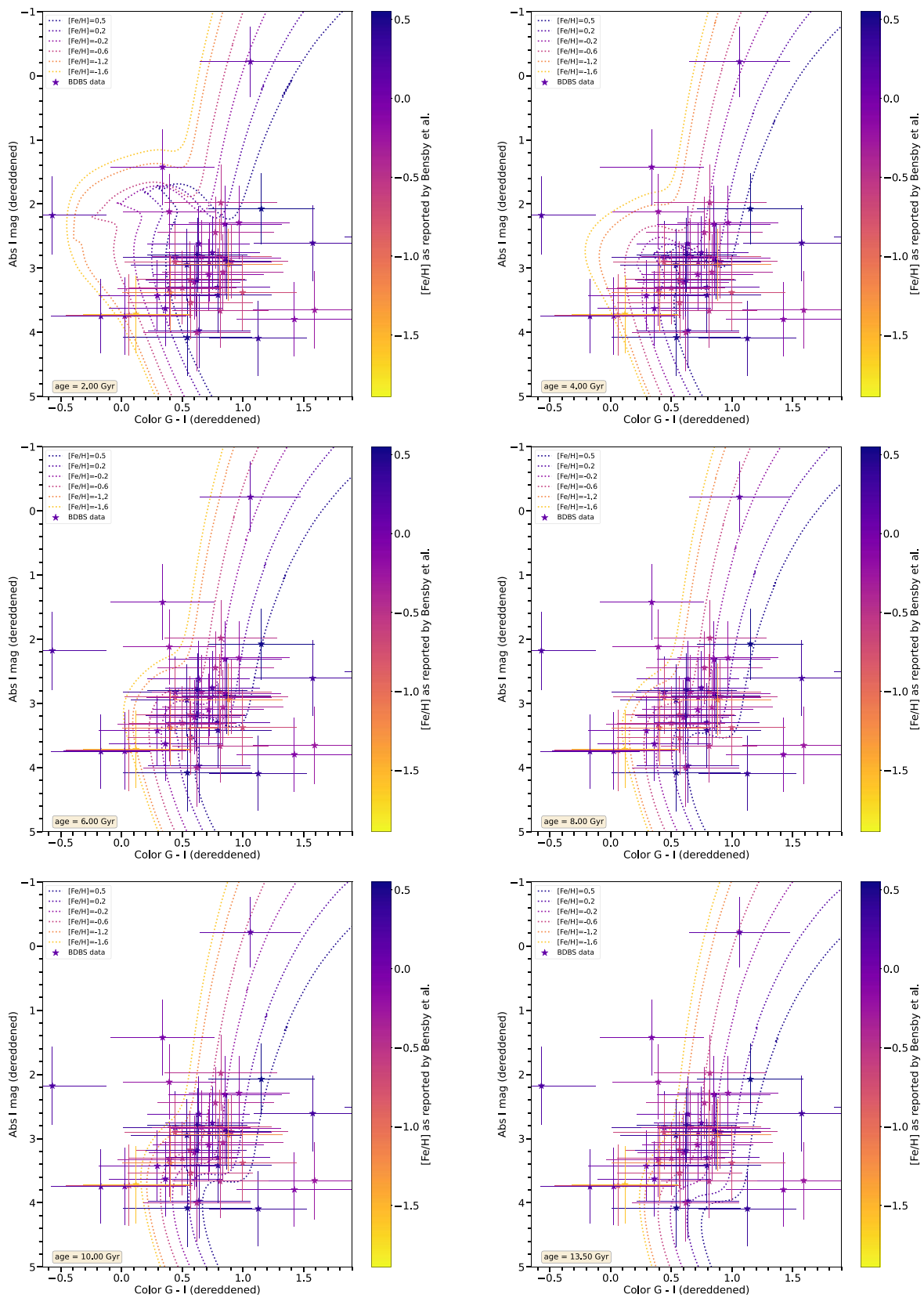
Crossmatching the BDBS catalog with the microlensed sample presented in B17, we find 60 stars in common. Stars that were not crossmatched either reside outside the BDBS footprint or were in a very crowded region where we could not be certain that the correct star was identified. Among the 60 crossmatched targets, nine were removed due to poor photometry or unavailability of distances. Among the remaining 51 candidates, we further identified a set of 18 stars as belonging to a “gold sample.” These stars were selected as having higher quality photometric measurements. Specifically, the gold sample BDBS stars were observed a minimum of three times, were not located in high background regions, and did not include any measurements with peculiar image shape parameters.

CMDs showing the absolute, dereddened magnitudes and colors for 51 BDBS targets against MIST isochrones are presented in Figure 13. Panels show ages ranging from 2 to 13.5 Gyr. Metallicities are assigned to the stars from B17 directly and indicated for both data and isochrones according to the color bar.

To fit models to BDBS photometry, the MIST isochrones must be converted to the PanSTARRS color–magnitude system for consistency with the BDBS  $i$ - and  $g$ -band filters, and the observations must be shifted into an absolute magnitude system, which requires knowledge of the distance and extinction. Where these transformations are possible (i.e., where sufficiently reliable distance and reddening information is available), Equation (1) still holds as the appropriate definition of  $\chi^2$ . The analog of Equation (2) for BDBS, Gaia, or any photometric fit is then

$$\begin{aligned} \chi_{\text{BDBS}}^2 &= \frac{(\text{color}_o - \text{color}_t)^2}{\sigma_{\text{color},o}^2} + \frac{(\text{mag}_o - \text{mag}_t)^2}{\sigma_{\text{mag},o}^2} + \frac{(Z_o - Z_t)^2}{\sigma_{Z,o}^2}. \end{aligned} \quad (\text{A1})$$

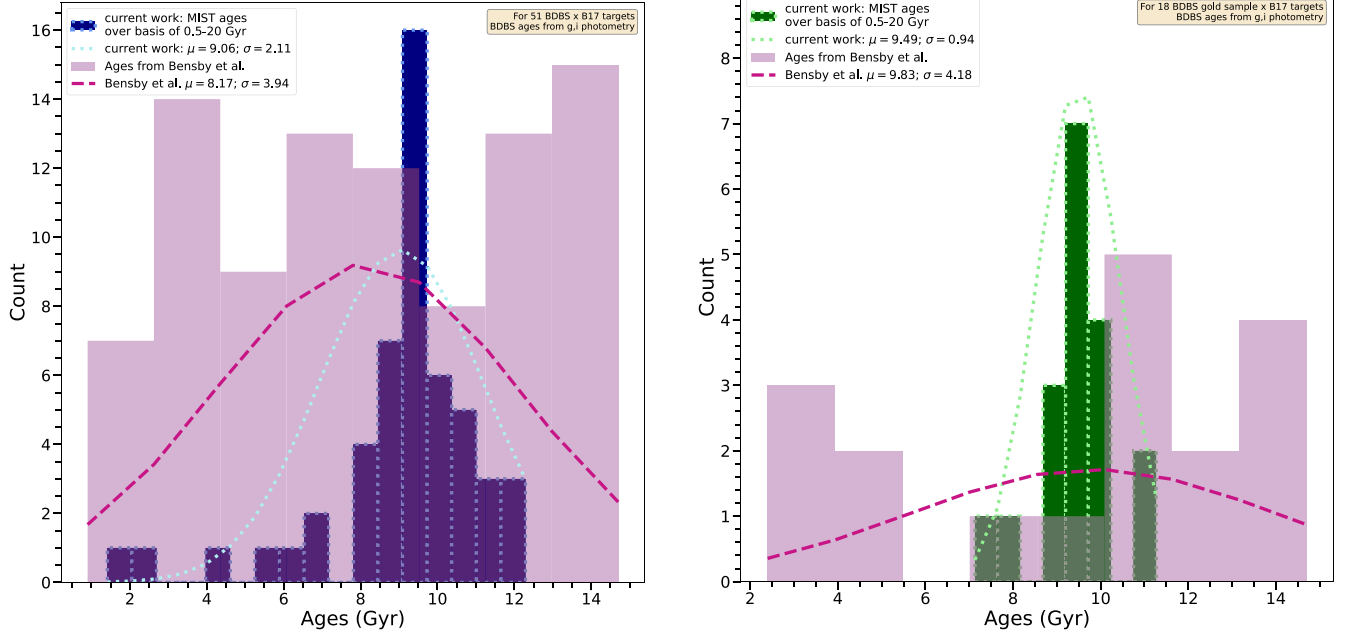
However, unlike in the fits to B17’s physical coordinates, the terms  $\sigma_{\text{color},o}$  and  $\sigma_{\text{mag},o}$  are composite uncertainties. The full definition of color,  $G - I$ , is actually  $(g - A_g) - (i - A_i)$ , where  $A_g$  and  $A_i$  are the extinctions in those bands, respectively. However, we also note that the extinction estimates are based on a reddening map that aimed to minimize the infrared color/



**Figure 13.** MIST isochrones in the PanSTARRS color basis are shown against BDBS photometry in  $I$  vs.  $G - I$ .

magnitude spread of red clump stars inside  $1' \times 1'$  bins on the sky; this is not the same as measuring or fitting the extinction for each star individually in the B17 sample. We also do not

account for possible variations or errors in the reddening laws, such as changes in the total-to-selective extinction ratio. Further, there is an uncertainty in both  $g$  and  $i$  due to



**Figure 14.** Age histograms in the style of Figures 5 and 7 but with ages derived from BDBS photometry. Left: the full intersection of the B17 and BDBS target lists for which complete photometric and distance information was available, totaling 51. The age distribution according to B17 for the same subset of stars is overlaid in pink. Right: same as the left panel, but for the gold sample containing 18 members. The ages for the same 18 stars according to B17 is overlaid in pink.

instrumental systematics. Therefore, the uncertainty in color as a whole is given by

$$\sigma_{\text{color},o}^2 = \sqrt{\sigma_{i\text{-band}}^2 + \sigma_{g\text{-band}}^2 + \sigma_{A_g}^2 + \sigma_{A_i}^2 + \sigma_{i,\text{sys}}^2 + \sigma_{g,\text{sys}}^2}, \quad (\text{A2})$$

with  $\sigma_{i,\text{sys}} = 0.03$  mag and  $\sigma_{g,\text{sys}} = 0.02$  mag (Johnson et al. 2020). Similarly,  $\sigma_{\text{mag}}$  must incorporate uncertainty in the  $i$ -band magnitude,  $A_i$ , instrumental systematics in  $i$ , and uncertainty in the distance modulus, DM

$$\sigma_{\text{mag},o} = \sqrt{\sigma_{i\text{-band}}^2 + \sigma_{A_i}^2 + \sigma_{i,\text{sys}}^2 + \sigma_{\text{DM}}^2}. \quad (\text{A3})$$

For each star’s line of sight, we calculated the distance at which the bulge density distribution reaches a maximum. We adopt the bulge parametric model described in Simion et al. (2017) (Model E) which provided the best fit to the infrared Vista Variables in the Via Lactea (VVV; Minniti et al. 2010) survey. The dispersion encompasses the 0.16–0.84 density fractions and is typically  $\sim 1.5$  kpc.

Since the Gaia EDR3 parallaxes are insufficient for deriving accurate distances for faint bulge stars, the second set of distances is estimated from a sample of 2.6 million red clump stars described in Johnson et al. (2022). The red clump distance map covers sight lines between  $|l| = \pm 10^\circ$  and  $b < -3^\circ 25'$  and encompasses a large fraction of the microlensed dwarf sample. For microlensed stars residing inside the red clump map, a mean distance and dispersion were calculated from all red clump stars that had a similar ( $|\Delta[\text{Fe}/\text{H}] < 0.1$  dex) metallicity and were within a  $0.5$  radius. Microlensed stars that reside outside the red clump map were excluded from further analysis due to insufficient distance information.

The presence of both data sets allows us to fit as many BDBS targets as possible. Due to the effects on the uncertainty terms, and therefore on the  $\chi^2$  and  $p$  scores, stars that do not

have distances according to either map cannot be fit and are removed from the sample.

One of the largest error sources for a given star is the uncertainty in the extinction-correction factors for the BDBS  $g$  band ( $A_g$ ) and  $i$  band ( $A_i$ ). We deredden the BDBS photometry following the approach described in Johnson et al. (2020). We employ a high-resolution  $E(J - K_s)$  dust map computed using the VVV survey. The  $(J - K_s)$  color dispersion  $\sigma_{E(J-K_s)}$  caused by the extinction correction was propagated to the Pan-STARRS photometry using the photometric transformations outlined in Johnson et al. (2020).

Using definitions (A1), (A2), and (A3) in the calculation of  $\chi^2$  and  $p$ , we compute ages for 51 BDBS targets. The resulting distribution is shown in Figure 14.

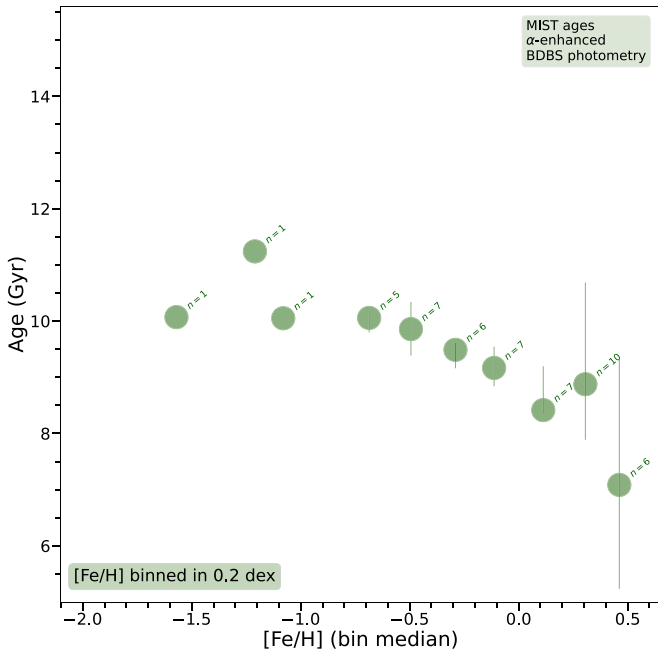
To understand this distribution—in particular, its narrowness—let us return briefly to the CMDs in Figure 13, where we see that the observational error bars from the BDBS photometry are much larger relative to the spacing of the isochrones than in analogous Figures 1 and 3. Given the larger number of constituent uncertainty terms embedded in the photometric  $\sigma_{\text{color}}$  and  $\sigma_{\text{mag}}$  definitions (and the well-established difficulty of making photometric age determinations for this reason), it is not surprising that this is the case. It follows, then, that a star adopting these  $1\sigma$  color and magnitude uncertainty definitions will intersect with a much larger number of isochrones, therefore achieving good (or over-fit)  $\chi^2$  scores for several hundreds of evolutionary points spanning a large range of ages and metallicities. The resulting age distribution should therefore reflect the weakly constrained star’s plausible consistency with nearly any model, thus yielding a sharply peaked distribution near the median of the trial ages (10.25 Gyr for ages 0.5–20 Gyr) and low variance, as shown.

What the sensitivity of the photometric ages to the set of input age candidates (i.e., grid of isochrones) really tells us is that the stars are not well enough constrained through photometry alone to be discerning among models. Much more

**Table 3**  
Sources with Reliable Gaia Photometry

	Name	Gaia Source ID	[Fe/H]	[Fe/H] Error	B17 Age	B17 Lower Age Limit	B17 Upper Age Limit
1	OGLE-2013-BLG-0692S	4052447161630182400	0.15	0.19	4.10	3.30	9.80
2	MOA-2012-BLG-202S	4064990283016815744	-0.15	0.13	14.70	8.50	99.90
3	MOA-2011-BLG-234S	4064581887363003520	-0.02	0.08	6.40	2.90	8.60
4	OGLE-2011-BLG-0950S	4043354643038544256	0.33	0.10	2.40	1.60	3.60
5	OGLE-2014-BLG-0157S	4042474694255127168	-0.52	0.10	14.00	7.40	99.90
6	OGLE-2013-BLG-1938S	4041757847027545984	-0.41	0.09	14.60	6.50	99.90
7	OGLE-2008-BLG-209S	4050303698056121728	-0.31	0.09	7.20	5.00	12.20
8	MOA-2010-BLG-285S	4056277031255698304	-1.21	0.10	11.60	8.10	13.40

**Note.** The Gaia source IDs and fundamental parameters as reported by B17 are shown for eight crossmatched Gaia targets determined to have reliable photometry as described in Appendix A.2.



**Figure 15.** Same as Figure 9, but using the ages derived from BDBS photometry. The axes are on the same scale in both panels for direct comparison. In both cases, there is a trend toward younger ages at higher metallicities, though the age spread is significantly smaller in the photometric case. We emphasize once again that the age spread and age dispersions are smaller in this case not because the photometric ages are more reliable, but because the photometric uncertainties are not precise enough to discern between age hypotheses.

precise brightness, distance, and reddening measurements would be necessary to produce truly valid, input-agnostic age determinations using the method presented here.<sup>16</sup> A further consequence of insufficient observational precision is unreliable age uncertainties. The MC calculations redetermine a star’s age across variations in its observed parameters, but if the observational uncertainties are so forgiving that good fits are found across all isochrones regardless of changes to the central values of color and magnitude (the values substituted for color, and mag<sub>0</sub> in Equation (A1)), the age determination will not change much from trial to trial. This produces an output distribution with low variance because the limiting factor on age variation is the maximum and minimum of the isochrone grid rather than the observational error bars. As such, age

<sup>16</sup> This statement is likely true for any similar statistical method as well.

determinations made using this photometry should be approached with caution.

The distribution clearly bears no resemblance to the near-uniform distribution of the B17 ages for the same subsample (reduced from 91 to the 51 stars for which we have BDBS measurements), shown overlaid in pink, but neither does it resemble the MIST age determinations for the microlensed coordinates shown in Figure 7. However, this distribution is typical of other photometrically derived age distributions for this region in the literature (e.g., Gennaro et al. 2015).

When the age determination algorithm is applied to the BDBS gold sample only (lower panel of Figure 14), we find an even sharper distribution, with lower variance, and starker disagreement with the B17 age determinations for the equivalent subsample. Though the problems associated with small number statistics should not be discounted when the sample is reduced to 18 candidates, it is worth noting that the mean of the distribution does increase when the only most reliable BDBS targets are fit. The lower panel of Figure 14 also makes clear that the stars being classified as young or intermediate-age by the MIST and BDBS photometry are not the same as those being classified as young by B17.

However, we wish to make clear that a direct comparison between the spectroscopic and photometric age distributions is not strictly valid given the caveats regarding the photometric uncertainties discussed above. It is not surprising, then, that there is tension between the age distributions derived from photometry and those derived from microlensing and spectroscopy; the impact of uncertainties cannot be ignored.

A comparison between Figures 15 and 9 underscores this point. The average binned age is shown as a function of metallicity, with number of stars in each bin annotated, for the BDBS photometry. While both data sets show an increase in age dispersion with increasing metallicity, the photometric data are consistent with the claim that the stellar ages display no metallicity dependence. This is clearly not the same story told by the spectroscopic data. A naïve interpretation of this discrepancy is that there is some fundamental tension between the data collection methods; but, in reality, this is once again explained by photometric imprecision in absolute magnitudes, driven particularly by distance and small-scale reddening variations along with the extreme crowding of low-latitude bulge fields. The age spread and age dispersions are smaller in the photometric case because the photometric uncertainties are of the same order as the entire ranges in color and magnitude covered by the isochrones.

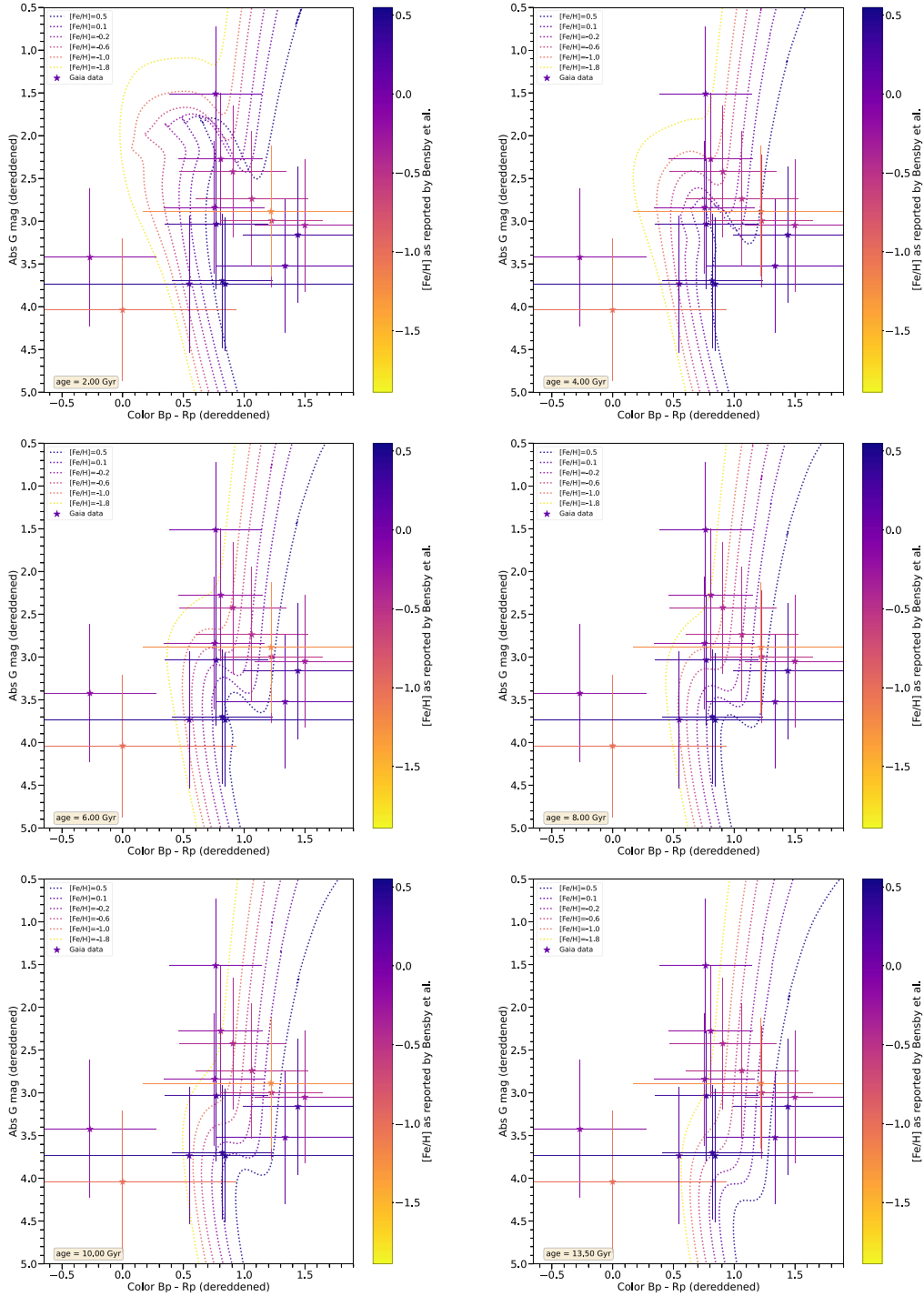
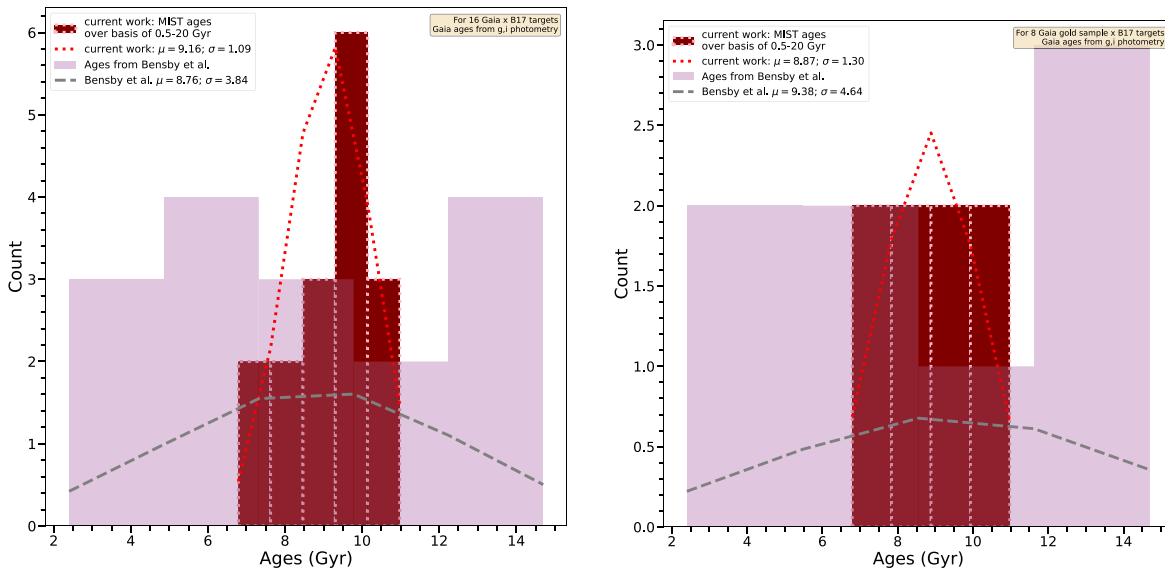


Figure 16. Same as Figure 13, but for Gaia photometry.

### A.2. Gaia Photometry

Nonetheless, it is worthwhile to repeat this exercise with an orthogonal photometric data set. We crossmatch the B17 sample with Gaia EDR3, finding a total of 17 stars with Gaia photometry and colors. We correct the apparent  $G$ -band magnitudes using the formula provided by Riello et al. (2021). To select a sample of stars with reliable Gaia photometry, we also correct the photometric flux excess factor following the approach described in Riello et al. (2021). The corrected excess factor  $C^*$  is computed comparing the integrated flux in the  $G_{BP}$  and  $G_{RP}$  bands to the one in the

$G$  band, and it is therefore a measure of the consistency of the photometry in the different bands. We then estimate the corresponding uncertainty  $\sigma_{C^*}$  as a function of (corrected)  $G$  magnitude using Equation (18) in Riello et al. (2021). We find a total of eight stars that satisfy the condition  $C^* \leq 5\sigma_{C^*}$ , and their identifiers are reported in Table 3. Gaia EDR3 parallaxes and proper motions are available for a subset of five stars. Even though Gaia EDR3 parallaxes for these faint stars are not precise enough to determine reliable individual distances, for one of the targets, OGLE-2013-BLG-0692S, the reported parallax  $\varpi = (2.75 \pm 0.84)$  mas might indicate that the star is



**Figure 17.** Left: the same as Figure 14, but for the intersection of the B17 and Gaia target lists examining only those stars selected according to the description in Section A.2. This is the “Gaia gold sample.” Right: same as the left panel, but for the entire intersection of the B17 and Gaia target lists. This totals 16 stars after the removal of Gaia stars with either (1) bad photometry or (2) for which two independent distance determinations with uncertainties were not available.

a foreground star and therefore does not belong to the bulge. This is further confirmed by Gaia EDR3 proper motions. For each of the five stars, we compute its probability to belong kinematically to the bulge based on its Galactic proper motions ( $\mu_{l,*}$ ,  $\mu_b$ ), which we determine by fitting the distribution in proper motion for all the stars within  $0.5^\circ$  from the center of the fields presented in Marchetti et al. (2022). By this metric, we find OGLE-2013-BLG-0692S to be the only star with a probability marginally lower than 50% of belonging to the bulge,  $P_{\text{bulge}} = 0.48$ . We include it our gold sample and fits regardless.

Distances are once again provided by Johnson et al. (2022) and Simion et al. (2017), with extinction uncertainties in the BP and RP filters calculated as described in Appendix A.1. A lack of reliable distance reduces the full crossmatched sample from 17 to 16 targets, while the gold sample includes eight targets. The color and magnitude uncertainty definitions for the Gaia targets are given by

$$\sigma_{\text{color},o} = \sqrt{\sigma_{\text{BP}}^2 + \sigma_{\text{RP}}^2 + \sigma_{\text{ABP}}^2 + \sigma_{\text{ARP}}^2} \quad (\text{A4})$$

and

$$\sigma_{\text{mag},o} = \sqrt{\sigma_{G-\text{band}}^2 + \sigma_{A_G}^2 + \sigma_{\text{DM}}^2}, \quad (\text{A5})$$

respectively. In this case, the analogs of the instrumental systematics terms (e.g.,  $\sigma_{i,\text{sys}}$ ) used in Equations (A2) and (A3) are the Gaia zero-point offsets per filter ( $G$ ,  $B_p$ , and  $R_p$ ), which are already folded into the definitions of  $\sigma_{G-\text{band}}$ ,  $\sigma_{\text{BP}}$ , and  $\sigma_{\text{RP}}$ .

Figure 16 shows the CMDs for all 16 available Gaia targets with uncertainties defined according to Equations (A4) and (A5). The MIST isochrones are transformed into the  $UBV(RI)c/2\text{MASS}/\text{Kepler}/\text{Hipparcos}/\text{Gaia}$  (DR2/EDR3)/Tess photometric system. As in the BDBS CMDs of Figure 13, we see that the  $1\sigma$  observational uncertainties for most stars overlap with nearly the entire set of isochrones (horizontal error bars) and cover a similarly wide expanse of evolutionary time and theoretical brightness (vertical error bars).

We perform the age determinations as in previous sections. The resulting distributions for the entire sample (left panel) and gold sample (right panel) are shown in Figure 17, with the Gaia/MIST ages shown in dark red and the histograms built from B17 overlaid. Qualitatively, we observe once again the narrowness of the MIST-based distributions compared to the equivalent subsamples in B17, though we reiterate the point made in Appendix A regarding the impropriety of claiming robust photometric ages for stars with such large uncertainties—and whose age determinations display such clear model-dependence—in the first place. Further, the small number of targets in both cases makes meaningful statistical inferences difficult: there is no apparent relationship between the Gaia/MIST ages and equivalent B17 ages across the two samples.

## Appendix B

### Comparison of Photometric Age Determinations

While the age uncertainties quoted in Table 2 for the photometric fits are the true, numerical results of MC simulations, they should not be interpreted as meaning anything beyond this. Neither BDBS nor Gaia photometric data are precise enough that the age determination algorithm can distinguish between candidate ages.

The two photometric ages are typically inconsistent with their spectroscopic counterparts but consistent with each other. Both of these features speak once again to the unreliability of these photometric age determinations and, in particular, their sensitivity to the input age grid. The fact that the photometric age “uncertainties” are on par with, or lower than, the spectroscopic age uncertainties for the same star should not be misinterpreted either: the photometric age variance over 100 MC simulations is low because the trial age determinations are always near the median of the input grid, which is always the same, regardless of parameter variations.

### Appendix C Results of MC Error-bar Simulations

For each star fit according to each of the three sets of observational parameters, we provide the age histograms generated from the MC simulations of their uncertainties.

Figures 18, 19, and 20 show histograms computed from 100 MC uncertainty simulations for MIST-based age determinations according to spectroscopic coordinates, BDBS photometry, and Gaia photometry, respectively.

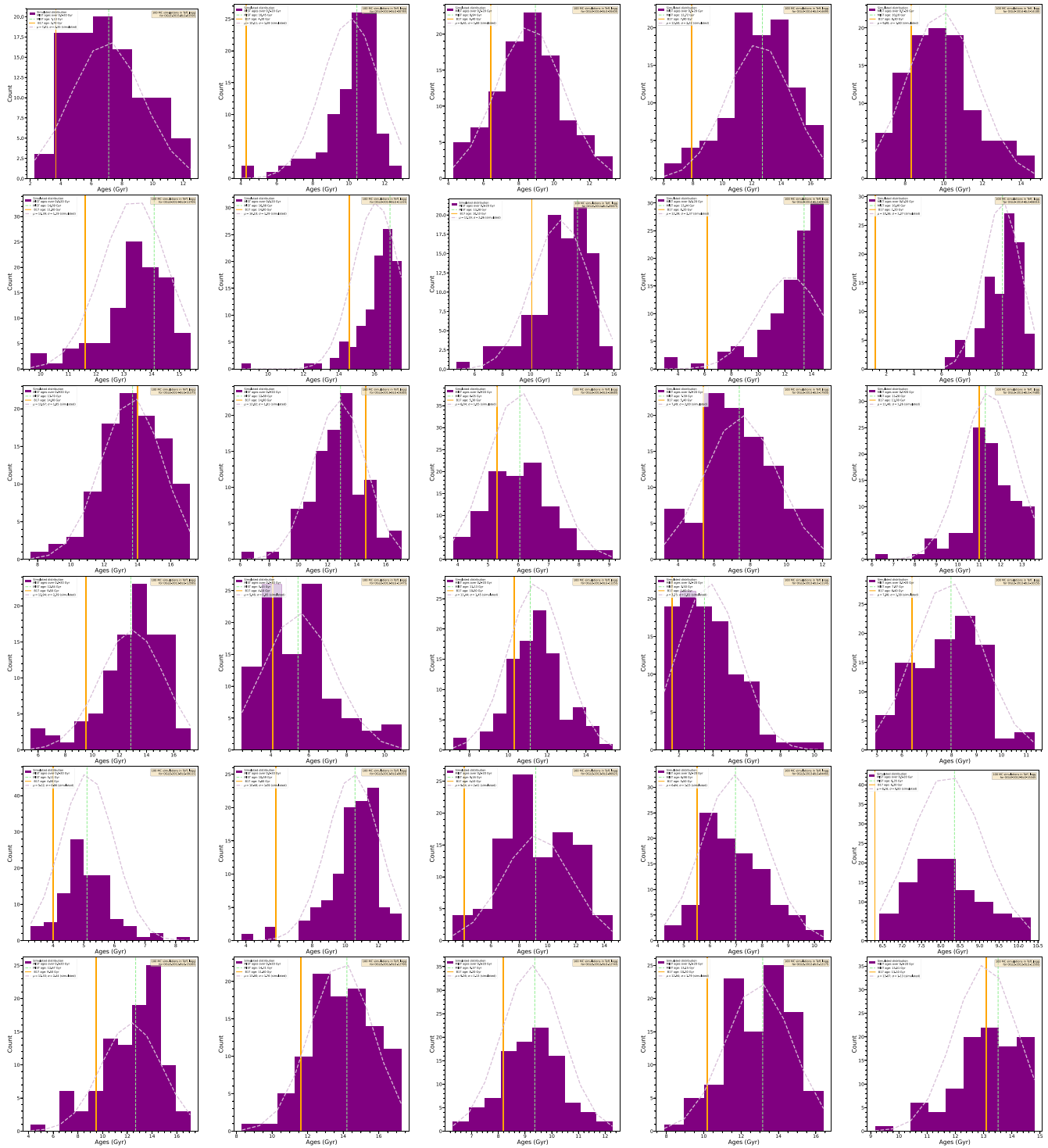


Figure 18. Error-bar simulations for the MIST-based refit to physical coordinates:  $T_{\text{eff}}$  and  $\log g$ .

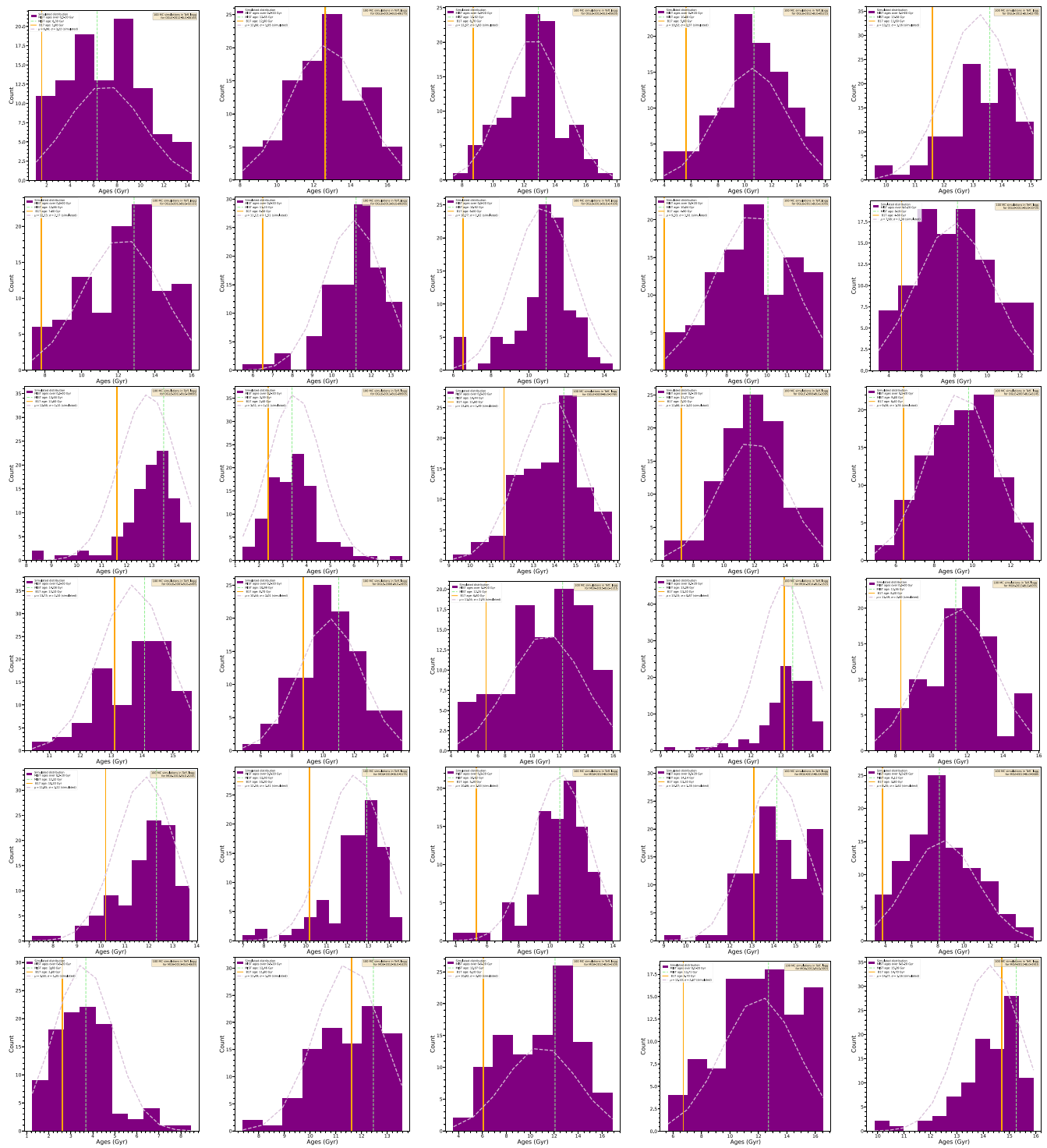


Figure 18. (Continued.)

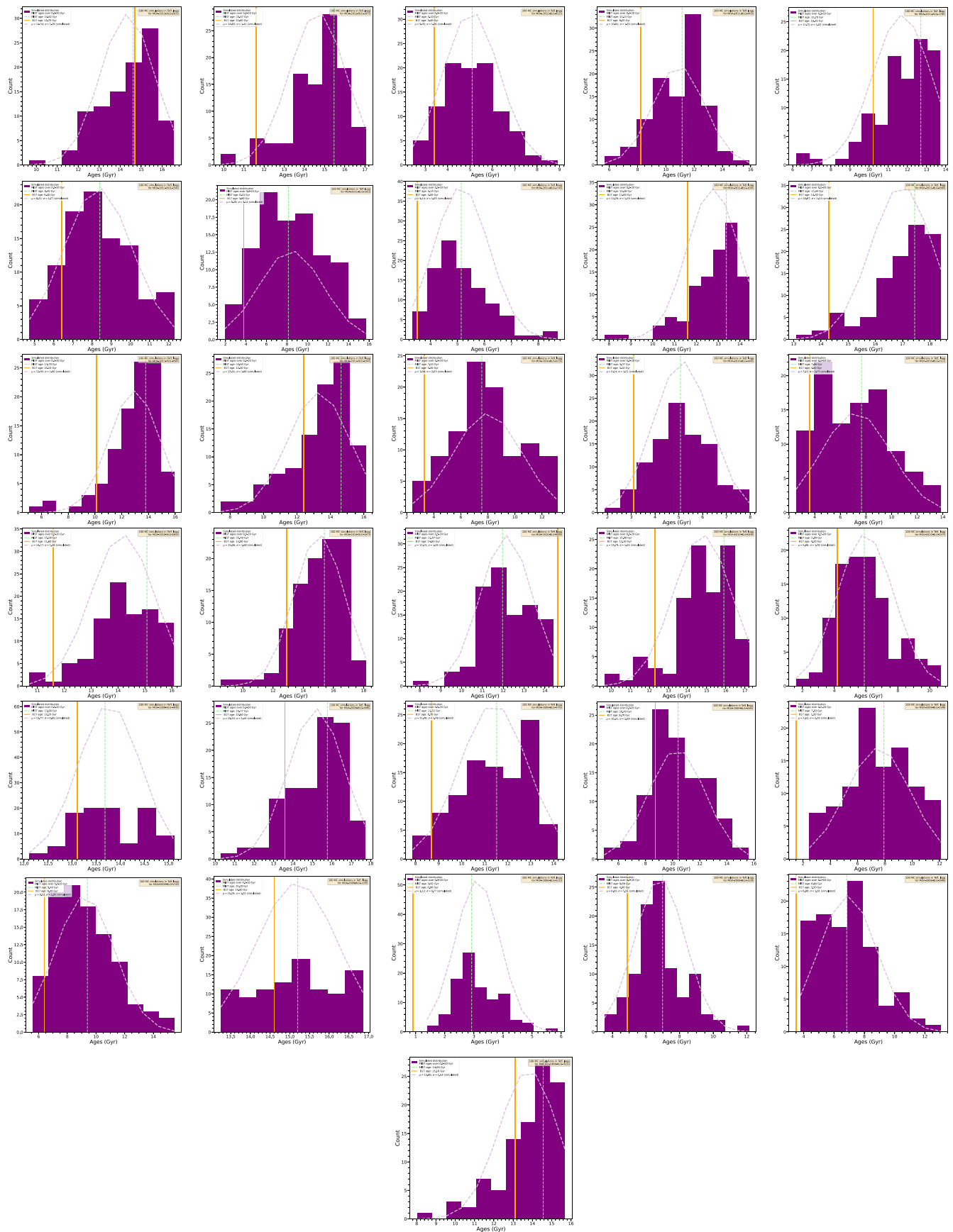


Figure 18. (Continued.)

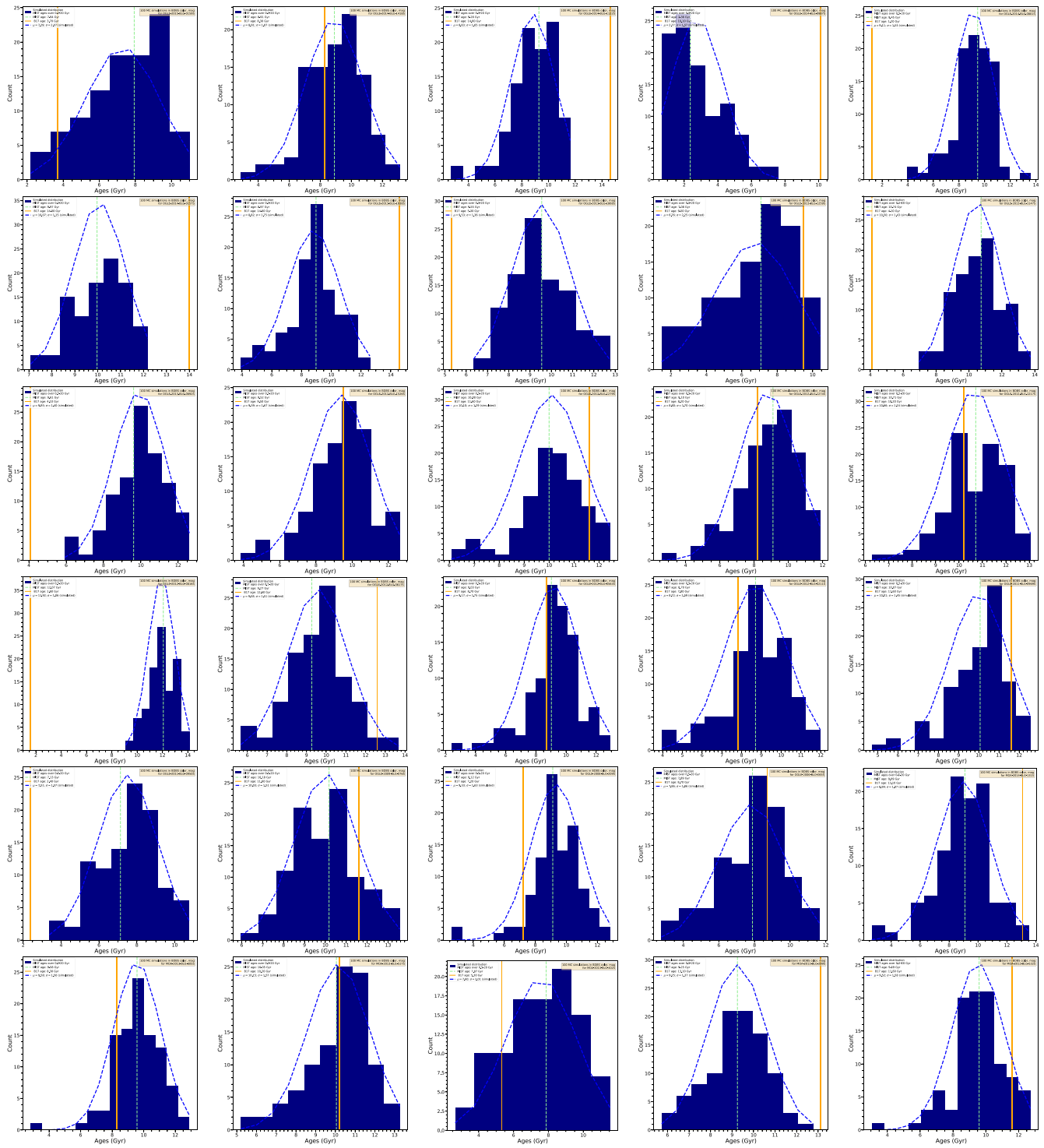


Figure 19. Error-bar simulation for BDBS targets.



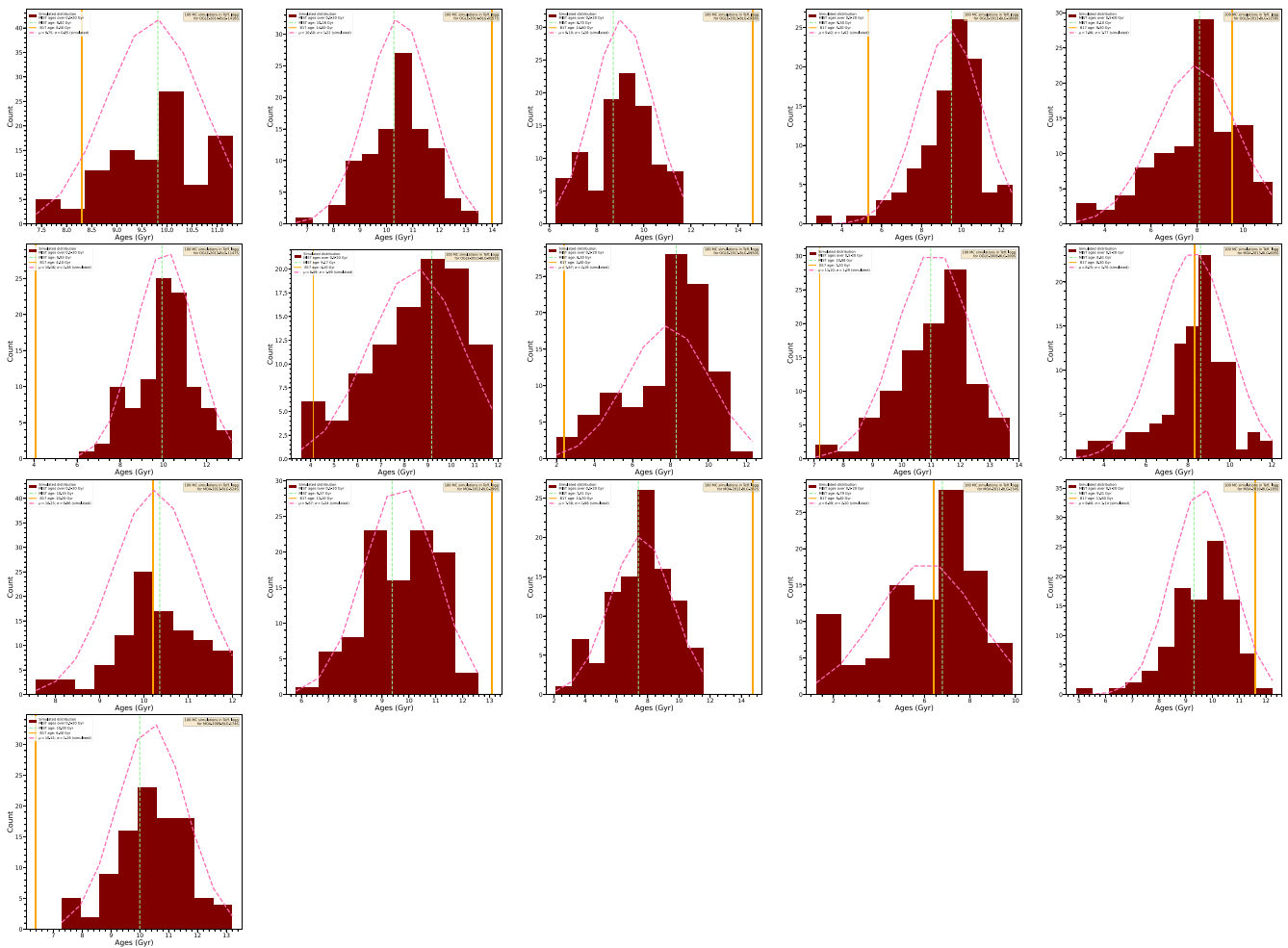



Figure 20. Error-bar simulation for Gaia targets.

## ORCID iDs

Meridith Joyce  <https://orcid.org/0000-0002-8717-127X>  
 Christian I. Johnson  <https://orcid.org/0000-0002-8878-3315>  
 Tommaso Marchetti  <https://orcid.org/0000-0003-0064-0692>  
 R. Michael Rich  <https://orcid.org/0000-0003-0427-8387>  
 Iulia Simion  <https://orcid.org/0000-0001-8889-0762>  
 John Bourke  <https://orcid.org/0000-0002-4958-5404>

## References

- Ade, P. A. R., Aghanim, N., Arnaud, M., et al. 2016, *A&A*, 594, A13  
 Andrae, R. 2010, arXiv:1009.2755  
 Angus, R., Morton, T. D., Foreman-Mackey, D., et al. 2019, *AJ*, 158, 173  
 Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, 47, 481  
 Ballero, S. K., Matteucci, F., Origlia, L., & Rich, R. M. 2007, *A&A*, 467, 123  
 Barbuy, B., Chiappini, C., & Gerhard, O. 2018, *ARA&A*, 56, 223  
 Bensby, T., Feltzing, S., & Oey, M. S. 2014, *A&A*, 562, A71  
 Bensby, T., Feltzing, S., Johnson, J. A., et al. 2010, *A&A*, 512, A41  
 Bensby, T., Adén, D., Meléndez, J., et al. 2011, *A&A*, 533, A134  
 Bensby, T., Yee, J. C., Feltzing, S., et al. 2013, *A&A*, 549, A147  
 Bensby, T., Feltzing, S., Gould, A., et al. 2017, *A&A*, 605, A89  
 Bernard, E. J., Schultheis, M., Di Matteo, P., et al. 2018, *MNRAS*, 477, 3507  
 Blommaert, J. A. D. L., & Groenewegen, M. A. T. 2007, in ASP Conf. Ser. 374, From Stars to Galaxies: Building the Pieces to Build up the Universe, ed. A. Vallenari et al. (San Francisco, CA: ASP), 193  
 Boehm-Vitense, E. 1979, *ApJ*, 234, 521  
 Bovy, J., Leung, H. W., Hunt, J. A. S., et al. 2019, *MNRAS*, 490, 4740  
 Brown, T. M., Sahu, K., Anderson, J., et al. 2010, *ApJL*, 725, L19  
 Burnett, B., & Binney, J. 2010, *MNRAS*, 407, 339  
 Calamida, A., Sahu, K. C., Casertano, S., et al. 2015, *ApJ*, 810, 8  
 Catchpole, R. M., Whitelock, P. A., Feast, M. W., et al. 2016, *MNRAS*, 462, 3765  
 Choi, J., Dotter, A., Conroy, C., et al. 2016, *ApJ*, 823, 102  
 Choi, J., Dotter, A., Conroy, C., & Ting, Y.-S. 2018, *ApJ*, 860, 131  
 Cinquegrana, G. C., & Joyce, M. 2022, *RNAAS*, 6, 77  
 Clarkson, W., Sahu, K., Anderson, J., et al. 2008, *ApJ*, 684, 1110  
 Clarkson, W. I., Sahu, K. C., Anderson, J., et al. 2011, *ApJ*, 735, 37  
 Cohen, J. G., Huang, W., Udalski, A., Gould, A., & Johnson, J. A. 2008, *ApJ*, 682, 1029  
 Cohen, J. G., Gould, A., Thompson, I. B., et al. 2010, *ApJL*, 711, L48  
 Creevey, O. L., Sordo, R., Pailler, F., et al. 2022, arXiv:2206.05864  
 Demarque, P., Woo, J.-H., Kim, Y.-C., & Yi, S. K. 2004, *ApJS*, 155, 667  
 Feiden, G. A. 2015, in ASP Conf. Ser. 496, Living Together: Planets, Host Stars and Binaries, ed. S. M. Rucinski, G. Torres, & M. Zejda (San Francisco, CA: ASP), 137  
 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  
 Gennaro, M., Tchernyshyov, K., Brown, T. M., & Gordon, K. D. 2015, *ApJ*, 808, 45  
 Gonzalez, O. A., Rejkuba, M., Zoccali, M., et al. 2011, *A&A*, 530, A54  
 Greggio, L., Renzini, A., & Daddi, E. 2008, *MNRAS*, 388, 829  
 Grevesse, N., & Sauval, A. J. 1998, *SSRv*, 85, 161  
 Grieco, V., Matteucci, F., Pipino, A., & Cescutti, G. 2012, *A&A*, 548, A60  
 Hasselquist, S., Zasowski, G., Feuillet, D. K., et al. 2020, *ApJ*, 901, 109  
 Haywood, M., Di Matteo, P., Snaith, O., & Calamida, A. 2016, *A&A*, 593, A82  
 Henry, L. G., Forbes, J. E., & Gould, N. L. 1964, *ApJ*, 139, 306  
 Ho, A. Y. Q., Ness, M. K., Hogg, D. W., et al. 2017, *ApJ*, 836, 5  
 Iglesias, C. A., & Rogers, F. J. 1996, *ApJ*, 464, 943  
 Ireland, L. G., & Browning, M. K. 2018, *ApJ*, 856, 132

- Johnson, C. I., Rich, R. M., Kobayashi, C., Kunder, A., & Koch, A. 2014, *AJ*, **148**, 67
- Johnson, C. I., Rich, R. M., Kobayashi, C., et al. 2013, *ApJ*, **765**, 157
- Johnson, C. I., Rich, R. M., Young, M. D., et al. 2020, *MNRAS*, **499**, 2357
- Johnson, C. I., Rich, R. M., Simion, I. T., et al. 2022, *MNRAS*, **515**, 1469
- Joyce, M., & Chaboyer, B. 2018a, *ApJ*, **856**, 10
- Joyce, M., & Chaboyer, B. 2018b, *ApJ*, **864**, 99
- Kordopatis, G., Schultheis, M., McMillan, P. J., et al. 2023, *A&A*, **669**, A104
- Kuijken, K., & Rich, R. M. 2002, *AJ*, **124**, 2054
- Kurucz, R. L. 1970, SAOSR, 309
- Mann, A. W., Feiden, G. A., Gaidos, E., Boyajian, T., & von Braun, K. 2015, *ApJ*, **804**, 64
- Maoz, D., Sharon, K., & Gal-Yam, A. 2010, *ApJ*, **722**, 1879
- Marchetti, T., Johnson, C. I., Joyce, M., et al. 2022, *A&A*, **664**, A124
- Martinez-Valpuesta, I., & Gerhard, O. 2013, *ApJL*, **766**, L3
- Matteucci, F., Grisoni, V., Spitoni, E., et al. 2019, *MNRAS*, **487**, 5363
- Meléndez, J., Bergemann, M., Cohen, J. G., et al. 2012, *A&A*, **543**, A29
- Minniti, D., Lucas, P. W., Emerson, J. P., et al. 2010, *NewA*, **15**, 433
- Murphy, S. J., Joyce, M., Bedding, T. R., White, T. R., & Kama, M. 2021, *MNRAS*, **502**, 1633
- Ness, M., Hogg, D. W., Rix, H.-W., Ho, A. Q., & Zasowski, G. 2015, *ApJ*, **808**, 16
- Ness, M., Freeman, K., Athanassoula, E., et al. 2013, *MNRAS*, **430**, 836
- Ortolani, S., Renzini, A., Gilmozzi, R., et al. 1995, *Natur*, **377**, 701
- Paxton, B., Bildsten, L., Dotter, A., et al. 2011, *ApJS*, **192**, 3
- Paxton, B., Cantiello, M., Arras, P., et al. 2013, *ApJS*, **208**, 4
- Paxton, B., Marchant, P., Schwab, J., et al. 2015, *ApJS*, **220**, 15
- Paxton, B., Schwab, J., Bauer, E. B., et al. 2018, *ApJS*, **234**, 34
- Paxton, B., Smolec, R., Schwab, J., et al. 2019, *ApJS*, **243**, 10
- Renzini, A., Gennaro, M., Zoccali, M., et al. 2018, *ApJ*, **863**, 16
- Rich, R. M. 2013, *Planets, Stars and Stellar Systems* (Dordrecht: Springer), 271
- Rich, R. M., Johnson, C. I., Young, M., et al. 2020, *MNRAS*, **499**, 2340
- Riello, M., De Angeli, F., Evans, D. W., et al. 2021, *A&A*, **649**, A3
- Rojas-Arriagada, A., Zasowski, G., Schultheis, M., et al. 2020, *MNRAS*, **499**, 1037
- Sanders, J. L., & Das, P. 2018, *MNRAS*, **481**, 4093
- Shen, J., Rich, R. M., Kormendy, J., et al. 2010, *ApJL*, **720**, L72
- Sheth, K., Elmegreen, D. M., Elmegreen, B. G., et al. 2008, *ApJ*, **675**, 1141
- Simion, I. T., Belokurov, V., Irwin, M., et al. 2017, *MNRAS*, **471**, 4323
- Simion, I. T., Shen, J., Koposov, S. E., et al. 2021, *MNRAS*, **502**, 1740
- Sit, T., & Ness, M. K. 2020, *ApJ*, **900**, 4
- Stanek, K. Z., Udalski, A., Szymański, M., et al. 1997, *ApJ*, **477**, 163
- Tang, J., & Joyce, M. 2021, *RNAAS*, **5**, 117
- Tayar, J., Claytor, Z. R., Huber, D., & van Saders, J. 2022, *ApJ*, **927**, 31
- Tayar, J., Somers, G., Pinsonneault, M. H., et al. 2017, *ApJ*, **840**, 17
- Thoul, A. A., Bahcall, J. N., & Loeb, A. 1994, *ApJ*, **421**, 828
- Tinsley, B. M. 1979, *ApJ*, **229**, 1046
- Tonry, J. L., Stubbs, C. W., Lykke, K. R., et al. 2012, *ApJ*, **750**, 99
- Trampedach, R., Stein, R. F., Christensen-Dalsgaard, J., Nordlund, Å., & Asplund, M. 2014, *MNRAS*, **445**, 4366
- Trapp, A. C., Rich, R. M., Morris, M. R., et al. 2018, *ApJ*, **861**, 75
- Valenti, E., Zoccali, M., Renzini, A., et al. 2013, *A&A*, **559**, A98
- Viani, L. S., Basu, S., Joel Ong, J. M., Bonaca, A., & Chaplin, W. J. 2018, *ApJ*, **858**, 28
- Wegg, C., & Gerhard, O. 2013, *MNRAS*, **435**, 1874
- Zoccali, M., Hill, V., Lecureur, A., et al. 2008, *A&A*, **486**, 177
- Zoccali, M., Renzini, A., Ortolani, S., et al. 2003, *A&A*, **399**, 931