



Star Cluster Ages in the *Gaia* Era

Jieun Choi¹, Charlie Conroy¹, Yuan-Sen Ting^{2,3,4}, Phillip A. Cargile¹, Aaron Dotter¹, and Benjamin D. Johnson¹¹Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA²Institute for Advanced Study, Princeton, NJ 08540, USA³Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA⁴Observatories of the Carnegie Institution of Washington, Pasadena, CA 91101, USA

Received 2018 April 23; revised 2018 June 21; accepted 2018 July 3; published 2018 August 10

Abstract

We use the framework developed as part of the MESA Isochrones and Stellar Tracks (MIST) project to assess the utility of several types of observables in jointly measuring the age and 1D stellar model parameters in star clusters. We begin with a pedagogical overview summarizing the effects of stellar model parameters, such as the helium abundance, mass-loss efficiency, and mixing-length parameter, on observational diagnostics such as the color–magnitude diagram, mass–radius relation, and surface abundances, among others. We find that these parameters and the stellar age influence observables in qualitatively distinctive, degeneracy-breaking ways. To assess the current state of affairs, we use the recent *Gaia* Data Release 2 (DR2) along with data from the literature to investigate three well-studied old open clusters—NGC 6819, M67, NGC 6791—as case studies. Although there is no obvious tension between the existing observations and the MIST models for NGC 6819, there are interesting discrepancies in the cases of M67 and NGC 6791. At this time, parallax zero-point uncertainties in *Gaia* DR2 remain one of the limiting factors in the analysis of these clusters. With a combination of exquisite photometry, parallax distances, and cluster memberships from *Gaia* at the end of its mission, we anticipate precise and accurate ages for these and other star clusters in the Galaxy.

Key words: binaries: eclipsing – Hertzsprung–Russell and C–M diagrams – stars: abundances – stars: evolution – stars: interiors

1. Introduction

Over the past decade, improving the state of stellar models has become a critical and necessary step in the quest to understand the properties of thousands of exoplanets that have been discovered (e.g., Torres et al. 2012; Huber et al. 2014; Mathur et al. 2017); probe the formation and evolution histories of galaxies both near and far, including our own Milky Way (e.g., Freeman & Bland-Hawthorn 2002; Bovy et al. 2012; Martig et al. 2015); link the diverse set of transient events to their progenitors (e.g., Kochanek et al. 2008; Smartt 2009; Georgy et al. 2012); and interpret the troves of asteroseismology data that have been obtained by the *CoRoT* (Baglin et al. 2006) and *Kepler/K2* missions (Bedding et al. 2010; Gilliland et al. 2010; Huber et al. 2011). Moreover, it has become increasingly clear that the analysis and interpretation of an even larger wealth of data expected from future missions and surveys will require more complete and accurate stellar models.

Many of the essential ingredients in standard 1D stellar evolution models cannot be modeled from first principles and instead rely on physically motivated prescriptions. For example, turbulent, superadiabatic convection is usually implemented according to the mixing-length formalism, in which the mixing efficiency and stellar structure depend sensitively on α_{MLT} , a free parameter of order unity (Böhm-Vitense 1958). There are ongoing complementary efforts to address this using sophisticated 3D hydrodynamic simulations (e.g., Trampedach et al. 2014; Magic et al. 2015), as well as detailed constraints and calibrations from a variety of observations (e.g., Bonaca et al. 2012; Wu et al. 2015; Tayar et al. 2017). This work adopts the latter approach, in particular using well-studied benchmark star clusters with a comprehensive set of observations to investigate the type of information—both cluster and stellar parameters, as well as the input physics

parameters—that we can recover and the precision with which we can measure them.

This paper is the first in a series that attempts to measure stellar parameters (e.g., age) and constrain uncertain input physics (e.g., mixing-length parameter). The insights gained from this work should guide our intuition to both shape the direction of future observations and forecast what we will be able to learn from future surveys and large data sets. We will explore this more quantitatively in subsequent work. In this first paper, we lay the groundwork for our approach by qualitatively examining the effects of various uncertainties on the observable quantities. We explore uncertainties of both observational (e.g., metallicity of the cluster) and theoretical (e.g., efficiency of mass loss) origins (e.g., Magic et al. 2010; Reese et al. 2016; Angelou et al. 2017; Lagarde et al. 2017). A key aspect of this particular work is that *we consider a diverse set of observables simultaneously*. One of the goals is to explore the separation of the key parameters in the various observed planes and identify a set of suitable observables for each parameter.

The rest of the paper is organized as follows. In Section 2, we review the different types of data sets that can be employed to study the properties of star clusters and to improve stellar evolution models. In Section 3, we first provide a brief overview of the MESA Isochrones and Stellar Tracks (MIST) project, which serves as the framework for the evolutionary models explored in this work. Then, we explore the information content in these observables using theoretical models, paying particular attention to the observational feasibility, as well as degeneracies. Next, in Section 4, we present case studies of three well-studied open clusters, NGC 6819, M67, and NGC 6791. In Section 5, we discuss what we can expect to accomplish with the future *Gaia* data, and in Section 6, we

present the summary of this work. In the [Appendix](#), we present a series of figures illustrating the effect of uncertain model parameters on the color–magnitude diagram (CMD) morphologies. For this work, we adopt a Kroupa (2001) initial mass function (IMF) where necessary.

2. Observations: What Can They Tell Us?

Here we provide a broad overview of the different types of data sets and surveys, including those that are ongoing and imminent, that can be used to improve both the characterization of star clusters and the quality of the stellar models. We also discuss what type of information can be leveraged from different types of observations. We conclude each section with a discussion of the “typical” uncertainties.

2.1. Photometry

High-precision photometry in multiple filters covering a long wavelength baseline is tremendously useful for measuring the age, metallicity, extinction, and distance. Stellar evolutionary tracks must be paired with bolometric correction tables to transform the theoretical outputs, e.g., T_{eff} and $\log L$, to observed magnitudes. Under the assumption of perfect observational data, any mismatch between the models and observations can be attributed to one or both of the components: interior models and atmosphere models.

In addition to traditional CMD fitting, photometry can be used for other observational diagnostics such as number counts of different types of stars. Although taking inventory of stars can be a difficult task owing to completeness issues, as well as low number statistics in some cases, number ratios are still powerful diagnostics because they are sensitive to relative phase lifetimes. We expect to be able to reliably catalog stars in different parts of the CMD with clean membership identification from future surveys (see Section 5 for a more in-depth discussion on the improvements due to *Gaia*). A related observable is the luminosity function, e.g., along the red giant branch (RGB), which has been widely adopted in studies of globular clusters (e.g., Renzini & Fusi Pecci 1988).

We also note that multiband photometry can be used to obtain photometric metallicities (narrow- and medium-band imaging; see, e.g., Ross et al. 2014) and temperatures. Temperatures derived from color–temperature relations (e.g., Alonso et al. 1996; Ramírez & Meléndez 2005; González Hernández & Bonifacio 2009; Casagrande et al. 2010) are widely used because they are considered to be reliable and easily measurable en masse. Finally, we note that direct measurements of the stellar angular diameter (and physical diameters if the parallax distance is known) for a sample of nearby stars are available through interferometry (e.g., Boyajian et al. 2012a, 2012b). In particular, when combined with bolometric flux and multiband photometry, they provide direct constraints on the empirical color–temperature relations with a few percent accuracy (Boyajian et al. 2013).

Ground-based photometry, which is generally limited by seeing owing to Earth’s atmosphere, produces typical uncertainties of order ≈ 0.01 mag, while *Hubble Space Telescope* (*HST*) photometry can routinely yield approximately millimagnitude photometry (relative, not absolute, uncertainty). One source of uncertainty that impacts both ground- and space-based observations is the photometric zero-point, which is

necessary to convert a flux to a magnitude on some standard scale. Although high-quality photometry can produce high relative photometric precision, absolute photometric precision is tied to $\approx 1\%$ absolute flux uncertainty for flux standards such as Vega (see Bohlin 2014; see also the discussion in Carrasco et al. 2016). Due to the uncertainties associated with the detailed bandpass shape, absolute flux calibration uncertainties for broadband photometry may be as large as $\approx 2\%$, and even larger for medium- and narrow-band photometry, for stars with spectral types much different from that of the photometric standards (see Section 3 of Bohlin 2012; Evans et al. 2018).

2.2. Spectroscopy

2.2.1. Basic Stellar Parameters

There are several recent, ongoing, and planned large-scale surveys designed to obtain medium-resolution ($R \simeq 10,000$ – $25,000$) spectra of stars in different parts of the Milky Way (e.g., RAVE, Steinmetz et al. 2006; *Gaia*-ESO, Gilmore et al. 2012; APOGEE, Holtzman et al. 2015; GALAH, de Silva et al. 2015; WEAVE, Dalton et al. 2012; *Gaia*-RVS, Recio-Blanco et al. 2016; *4MOST*, de Jong et al. 2016). Their principle scientific objective is to shed light on the formation and evolution history of our Galaxy. These spectroscopic surveys yield, at minimum, radial velocity, $\log g$, T_{eff} , and metallicity, and in many cases the surface abundances of multiple elements for each star. From the stellar evolution and stellar astrophysics perspective, accurate and precise measurements of these parameters are extremely useful for testing the integrity of the stellar evolution models. With the exception of asteroseismology, surface abundances are some of the only probes of the stellar structure and interior conditions. Since the creation and destruction of different species deep within the star can be imprinted on the surface through various mixing processes, the surface abundances of different elements carry immense diagnostic power. Finally, a Hertzsprung-Russell (H-R) diagram constructed from $\log g$ and T_{eff} , also known as the “Kiel diagram,” is a useful, distance-independent diagnostic that can be compared directly with theoretical isochrones.

There is immense diversity in the analysis techniques and pipelines that are employed to measure stellar parameters for large spectroscopic samples (e.g., Smiljanic et al. 2014; Holtzman et al. 2015)—adopted line lists, optimization for the analysis of different stellar types, equivalent width versus full spectral fitting—and they produce systematically discrepant results. There is an ongoing effort to mitigate some of these concerns by carrying out detailed comparisons between different state-of-the-art methods (e.g., *Gaia*-ESO benchmark stars; Smiljanic et al. 2014). There is typically a range of values for the quoted uncertainties (combined systematic and statistical), depending on the adopted methodology and the stellar spectral types; $\log g$, T_{eff} and [Fe/H] uncertainties are generally 0.1–0.2 dex, 50–100 K, and 0.05–0.1 dex, respectively (e.g., Smiljanic et al. 2014; Holtzman et al. 2015). Systematic uncertainties are generally higher (0.1–0.2) for the other elements, though in some cases the range of reported [Fe/H] values may be comparable as well (e.g., ~ 0.2 dex for NGC 6791; Carraro et al. 2006; Gratton et al. 2006; Origlia et al. 2006; Carretta et al. 2007; Brogaard et al. 2011; Boesgaard et al. 2015; Netopil et al. 2016).

2.2.2. Carbon and Nitrogen Surface Abundances on the RGB

The surface stellar abundances generally do not reflect the initial or even the bulk interior abundances at any given time. Over the lifetime of a star, physical processes such as diffusion and dredge-up can dramatically modify the surface abundances. The magnitudes of these effects vary with the mass and metallicity of the star and the elemental species in question. For this reason, the evolution of surface abundances can be used to trace stellar mass and, thus, stellar ages.

For this work, we focus on the surface abundances of RGB stars because they constitute a significant fraction of the sample in these spectroscopic surveys that require bright beacons in distant parts of the Galaxy (see Dotter et al. 2017, for a discussion of surface vs. bulk abundances and their implications on derived stellar ages). One of the parameters that are crucial to galactic archeology is the stellar age. The classic method of inferring stellar ages using the spectroscopic $\log g - T_{\text{eff}}$ diagram is notoriously challenging owing to small T_{eff} separations between the nearly vertical RGB tracks of stars with different initial masses. As a result, small uncertainties in T_{eff} yield large uncertainties in the initial mass and therefore in age.

Recently, an alternative method using the ratio of carbon to nitrogen as an age indicator has gained traction (e.g., Masseron & Gilmore 2015; Salaris & Cassisi 2015; Martig et al. 2016; Ness et al. 2016). When the star leaves the main sequence (MS), its deepening convective envelope introduces several changes to the surface elemental abundances during what is known as the first dredge-up (FDU). Whereas some species such as iron that were depleted during the MS owing to gravitational settling are nearly restored to their initial values, other species such as nitrogen and carbon show a marked change relative to their initial abundances. The latter phenomenon occurs because the convective envelope engulfs the products of hydrogen burning in the deep interior, diluting its original bulk abundances with the CN-processed material. FDU yields an increase in surface ^{14}N and a concordant decrease in surface ^{12}C as dictated by the CNO cycle equilibrium; the $^{14}\text{N}(p, \gamma)^{15}\text{O}$ reaction is the bottleneck in the CNO cycle, resulting in the accumulation of nitrogen. Since the maximum fractional depth reached by the convective envelope increases with the initial mass, a larger decrease in the surface $[\text{C}/\text{N}]$ abundance corresponds to a larger stellar mass and therefore a younger age (Salaris & Cassisi 2015). This FDU efficiency has also been demonstrated to increase with increasing metallicity (Charbonnel 1994). A caveat of this age inference method is that the initial abundances must be known (see, e.g., the discussion in Martig et al. 2016) by disentangling the effects of stellar evolution and galactic chemical evolution. A significant advantage of studying stars in clusters is that we can obtain the abundances for a sample of MS or subgiant branch (SGB) stars in addition to the RGB stars to get a handle on their initial C and N abundances.

The end of the FDU is marked by the convective envelope receding back toward the surface ahead of the hydrogen-burning shell, which is also moving outward. Although canonical models do not show additional mixing beyond the FDU, there is solid observational evidence that extra mixing occurs beyond the RGB bump through the tip of the RGB, and possibly during the core helium burning (CHeB) phase (Gratton et al. 2000; Angelou et al. 2015). Several explanations have been put forth, including thermohaline (Charbonnel &

Zahn 2007; Charbonnel & Lagarde 2010) and rotational mixing (Sweigart & Mengel 1979; Palacios et al. 2006), though we focus on the former here.

Thermohaline mixing is a double-diffusive instability that occurs in the presence of a destabilizing composition gradient. Although positive mean molecular weight (μ) gradients are rare in the stellar interior (nuclear fusion occurs inside out, and it transforms light elements into heavy elements), they do appear in some cases, for example, during the $^3\text{He}(^3\text{He}, 2p)^4\text{He}$ reaction taking place in the external wing of the hydrogen-burning shell (Ulrich 1972; Eggleton et al. 2006). Note that this unusual reaction produces a net increase in the number of particles and thus a decrease in μ . Thermohaline mixing is established only beyond the RGB bump, a brief adjustment period the star undergoes when the hydrogen-burning shell encounters the μ -discontinuity at the base of the chemically homogeneous zone, i.e., maximum depth previously reached by the convective envelope. This instability cannot be triggered at earlier times because the magnitude of the μ gradient inversion is negligible in the presence of the stabilizing composition gradient. Once thermohaline mixing is established in the radiative layer between the hydrogen-burning shell and the convective envelope, surface abundances of numerous species, including ^3He , ^{12}C , ^{13}C , and ^{14}N , can become modified.

2.3. Asteroseismology

Asteroseismology relies on the high-precision monitoring of brightness fluctuations in the light curves originating from stellar oscillations. To date, the *CoRoT* (Baglin et al. 2006) and *Kepler* (Gilliland et al. 2010) missions have detected solar-like acoustic oscillations in well over 15,000 red giants (Kallinger et al. 2010b; Stello et al. 2013; Huber et al. 2014). The ongoing repurposed *K2* mission, the upcoming *TESS* mission, and next-generation surveys such as *WFIRST*, *Euclid*, and *Plato* are expected to increase the sample size dramatically.

The detection of oscillations requires taking the Fourier transforms of the time-series photometry. There are two main techniques for the subsequent analysis and physical interpretation of the data. The first method is called “peakbagging,” or “boutique-modeling,” which involves the detailed modeling of individual peaks in the frequency spectrum. This is a challenging and time-consuming problem owing to the sheer number of detected modes, as well as the presence of mixed dipole modes (Corsaro et al. 2015; Handberg et al. 2017).

The second method, which is more widespread given its simplicity, involves reducing the information in the frequency spectrum to two global parameters: the frequency of maximum power, ν_{max} , and the average large frequency separation, $\Delta\nu$. These parameters can be related to stellar parameters via simple scaling relations:

$$\Delta\nu \simeq \sqrt{\frac{M/M_{\odot}}{(R/R_{\odot})^3}} \Delta\nu_{\odot} \quad (1)$$

$$\nu_{\text{max}} \simeq \frac{M/M_{\odot}}{(R/R_{\odot})^2 \sqrt{T_{\text{eff}}/T_{\text{eff},\odot}}} \nu_{\text{max},\odot} \quad (2)$$

where $\Delta\nu_{\odot} = 135.1 \mu\text{Hz}$, $\nu_{\text{max},\odot} = 3100 \mu\text{Hz}$, and $T_{\text{eff},\odot} = 5777 \text{ K}$ correspond to the solar values. As can be gleaned from the equations above, $\Delta\nu$ and ν_{max} are each sensitive to the average density and surface gravity of the star, respectively.

Once they are measured from the observed frequency spectrum, $\Delta\nu$ and ν_{\max} can be combined with an external estimate of T_{eff} —from either spectroscopy or a color– T_{eff} relation—to yield masses and radii. Alternative forms of the scaling relations exist for when independent estimates of, e.g., bolometric luminosity are available. These scaling relations are used to derive masses and radii of RGB stars in the field en masse (Stello et al. 2008; Kallinger et al. 2010a; Miglio et al. 2012). Though extremely useful, these simple scaling relations have been demonstrated to harbor systematics, and thus various corrections have been proposed to improve their accuracy (e.g., White et al. 2011; Miglio et al. 2012; Sharma et al. 2016; Viani et al. 2017). There is an ongoing effort to test and validate the accuracy of scaling relations using other independent techniques such as eclipsing binaries (e.g., Gaulme et al. 2016). The general consensus is that the scaling relations provide RGB and red clump (RC) mass and radius estimates to within $\sim 10\%$ and $\sim 5\%$, respectively (see Viani et al. 2017, and references therein). Estimates of the $\log g$ from the scaling relation are better determined; systematic uncertainties and biases are estimated to be around 0.01 dex (Hekker et al. 2013; Viani et al. 2017).⁵

2.4. Detached Eclipsing Binaries

Eclipsing binary systems yield stellar masses and radii with exquisite precision, routinely below $\sim 3\%$ (Torres et al. 2010). In particular, systems that are well detached such that both members are effectively undergoing single-star evolution—also known as detached eclipsing binaries (DEBs)—provide a unique opportunity for rigorous tests of stellar evolution models; the masses, radii, and/or luminosities of both binary components must agree within the observational uncertainties at a single age (Andersen 1991). Moreover, DEBs can be used to measure stellar ages without the knowledge of distance and interstellar reddening if they are near the main-sequence turnoff (MSTO). DEBs in star clusters are especially useful because they can be combined with CMDs to provide complementary constraints on the models (e.g., Stetson et al. 2003; Brogaard et al. 2012; Gökay et al. 2013; Yakut et al. 2015; Brewer et al. 2016).

Ground-based discoveries and analyses of DEBs trace back well over a century (Russell 1912) and have yielded parameters of many stellar systems (e.g., Popper 1967; Harmanec 1988; Andersen 1991; Torres et al. 2010). The unprecedented, precise monitoring by *Kepler* has observed close to 2000 eclipsing binaries, approximately 1400 of which are classified as DEBs (Kirk et al. 2016). Ongoing and future missions such as *Gaia*, *TESS*, and *PLATO* are expected to find many more eclipsing binaries.

Accurate and precise masses and orbital parameters require high-quality radial velocities measured from double-lined spectra with high spectral resolution and signal-to-noise ratio. In a single-lined system, where only the primary component is

detected, a full orbital solution generally cannot be obtained. In these systems, the total mass must be combined with the mass ratio inferred from the light curves to obtain estimates of individual masses. Perhaps unsurprisingly, masses obtained using this method are generally less reliable owing to correlations and degeneracies among the orbital parameters and their resulting solutions (Kirk et al. 2016). Light curves provide stellar radii and orbital parameters, which can be compared with the spectroscopic determinations as a consistency check. As noted earlier, DEBs generally provide masses and radius measurements to within $\sim 3\%$ (Torres et al. 2010), and even $< 1\%$ in some cases (e.g., Brewer et al. 2016).

3. Models: Where Is the Information Content?

In this section, we first provide an overview of the MIST models. Then we present a summary of the effects of various uncertainties on the observable quantities. Next, we evaluate the sensitivity of the observables to each key parameter in order to identify the ideal set of observations, with the goal of measuring the cluster parameters (e.g., age) and constraining the uncertain free parameters (e.g., α_{MLT}). We will revisit the latter within a more quantitative and rigorous framework in subsequent work.

3.1. MIST Models

The theoretical isochrones for this work are computed within the MIST framework (Choi et al. 2016; Dotter 2016). The main objective of the MIST project⁶ is to build comprehensive grids of well-calibrated stellar evolutionary tracks and isochrones that encompass a wide range of masses, ages, metallicities, and evolutionary phases. The first set of models with solar-scaled abundances both including and excluding the effects of stellar rotation are already available (Choi et al. 2016). The second set of models consisting of non-solar-scaled abundances are currently under development (A. Dotter et al. 2018, in preparation). Stellar evolutionary tracks are computed using Modules for Experiments in Stellar Astrophysics (MESA; Paxton et al. 2011, 2013, 2015, 2018), an open-source 1D stellar evolution package. Isochrones are constructed from grids of stellar evolutionary tracks following Dotter (2016). For an in-depth overview of the MIST models, including the descriptions of the input physics, we refer the reader to Section 3 and Table 1 of Choi et al. (2016).

3.2. Overview: Model Parameters and Observations

In the following sections, we show a series of multipanel plots (Figures 1–6) illustrating the effect of an individual parameter on the various observables at a given stellar age. Here we describe each panel and the relevant observations in detail.

In each panel, we show a total of nine MIST isochrones projected onto several observed planes. The peach, pink, and purple colors correspond to different parameters (e.g., metallicities), and the solid, dot-dashed, and dotted lines correspond to three ages (2, 4, and 10 Gyr, respectively). The parameter of interest always increases from peach to purple, and the pink curve corresponds to the fiducial model, unless noted otherwise.

⁵ Currently there are interesting, unresolved discrepancies between spectroscopic and asteroseismic $\log g$ for red giants with the SDSS APOGEE spectroscopy and *Kepler* asteroseismology (<http://www.sdss.org/dr14/irspec/aspac/#calibration>). Not only does there appear to be a mild metallicity dependence on this discrepancy, but the size of the offset appears to be different for CHeB and RGB stars for reasons that are not well understood. See also Ting et al. (2018).

⁶ <http://waps.cfa.harvard.edu/MIST/>

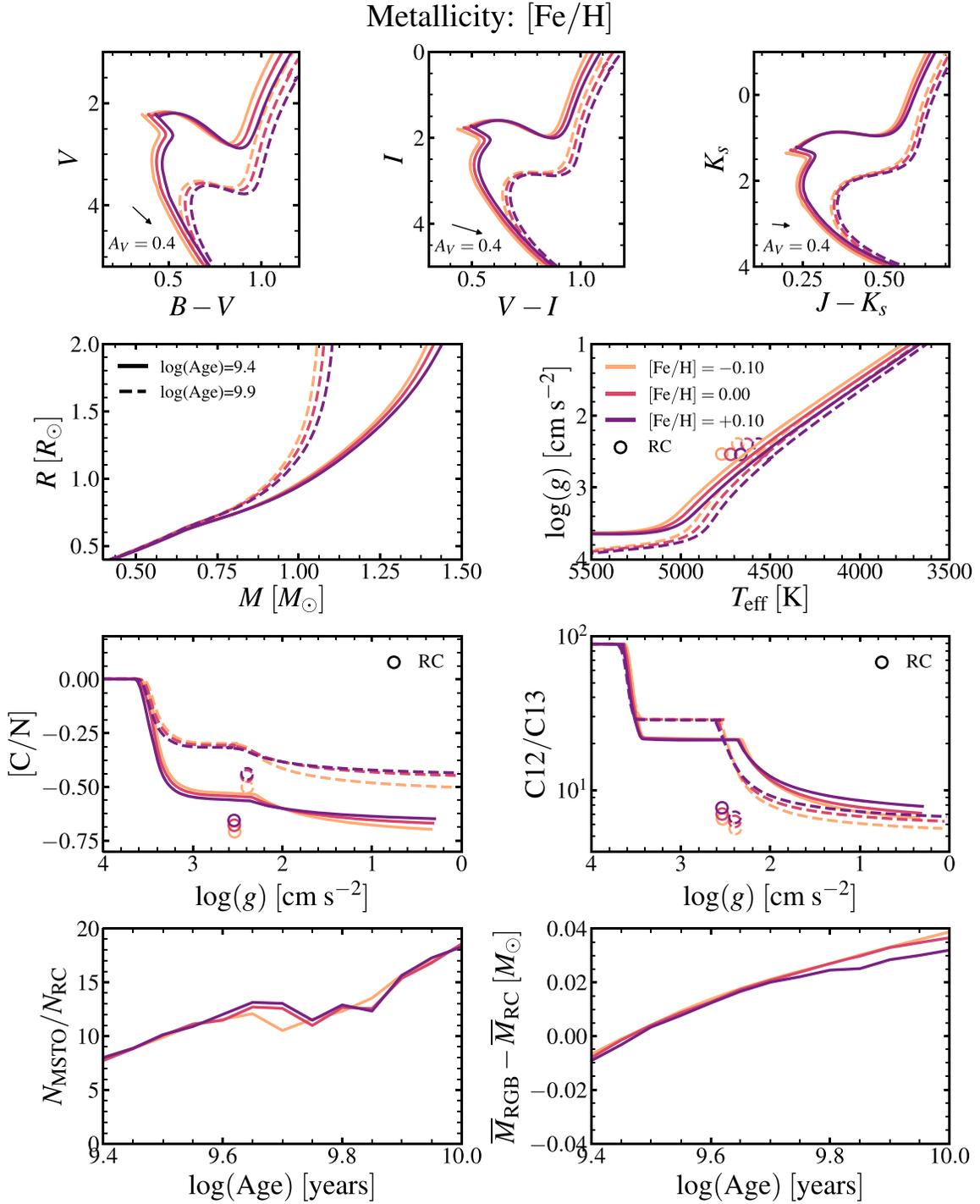


Figure 1. Effect of metallicity and age on a variety of observable parameters. There are a total of six MIST models shown, where the peach, pink, and purple colors correspond to different metallicities and the solid and dashed lines correspond to two different ages (2.5 and 8 Gyr, respectively). First row: CMDs ranging from optical to near-infrared wavelengths. There are no extinction and distance modulus applied to these CMDs, but we include reddening vectors to illustrate how the positions of the CMDs can shift. We adopt the standard $R_V = A_V/E(B-V) = 3.1$ reddening law from Cardelli et al. (1989) and assume $A_V = 0.4$ evaluated at $[\text{Fe}/\text{H}] = 0$, $\log g = 4$, and $T_{\text{eff}} = 5500$ K. Second row: mass–radius and Kiel diagrams, which can be compared with DEB data and asteroseismic $\log g$, respectively. For display purposes, we omit the transition from the core He flash (RGB tip) to the RC and mark the start of quiescent core helium burning using an open circle. Third row: surface abundance evolution as the star ascends the RGB and undergoes helium flash before settling into a quiescent, helium-burning RC phase. The first large decrease at $\log g \sim 3.5$ marks the onset of FDU, and the second large decrease at $2.5 \gtrsim \log g \gtrsim 2.0$ after the RGB bump is due to thermohaline mixing. Fourth row: number ratio of MSTO to RC stars and the difference in the average mass of the RGB and RC stars, both shown as a function of the cluster age. MSTO stars are defined to be those that fall within 0.5 mag below the hottest point of the MS in the V band. Asteroseismology can be used to both distinguish RC from RGBs and provide average masses for both types of stars. The first two rows clearly demonstrate that higher metallicity corresponds to cooler T_{eff} (redder colors) and longer MS lifetimes. The third row shows that FDU and thermohaline operate more and less efficiently, respectively, as the metallicity is increased. Finally, the last row shows that metallicity does not appear to have a strong effect on the relative numbers of MSTO to RC stars and the difference in the average mass of RGB and RC stars.

The top row features three CMDs— $B - V$, $V - I$, and $J - K_s$ —zoomed in near the MSTO, SGB, RGB, and RC. As in the standard MIST models, we use the C3K bolometric correction tables (C. Conroy et al. 2018, in preparation) constructed from the ATLAS12/SYNTHÉ atmosphere models. The ATLAS12/SYNTHÉ models include the latest atomic line list from R. Kurucz (including both laboratory and predicted lines) and many molecules whose contributions are important especially at longer wavelengths. We use the latest *Gaia* Data Release 2 (DR2) passbands and zero-points (Evans et al. 2018). There are no extinction and distance modulus applied to these CMDs.

In the first panel of the middle row, we show the theoretical mass–radius relations, which can be compared to high-precision measurements from DEBs. The second panel shows a slight variation on the classic H-R diagram with $\log g$ instead of $\log L$ on the y -axis (Kiel diagram), zoomed in on the RGB and the RC, where most of the asteroseismic targets are located. For these evolutionary phases, the predicted $\log g$ can be compared with the asteroseismic surface gravity, $\log g_{\text{astero}}$, inferred from the ν_{max} asteroseismic scaling relation (Equation (1)). Note that the ν_{max} measurement must be combined with a spectroscopic or photometric T_{eff} to infer $\log g$. For display purposes, we omit the transition from the core helium flash (the tip of the RGB) to the RC and mark the start of quiescent CHeB using an open circle.

The next row shows surface abundance ratios of carbon to nitrogen (left) and ^{12}C to ^{13}C (right) along the RGB and RC (shown as an open circle for clarity) as a function of surface gravity. Surface abundances are powerful indirect probes of the stellar interior because various mixing processes lead to changes in the surface abundances of different species at different stages of the evolution.

Finally, the bottom row features two panels that each show an integrated or averaged quantity as a function of the age of the cluster. The first panel shows the ratio of MSTO to RC stars. In this context, the MSTO stars are defined to be those that fall within 0.5 mag below the hottest point of the MS in the V band. For an isochrone with a Henyey hook, instead of the tip of the hook (the actual hottest point), we use the inflection point at which the blueward excursion begins. This is because the actual hook corresponds to a fast phase of expansion and contraction, and thus it is observationally unlikely to find many stars there. The RC stars are selected based on the MIST phase label (CHeB phase). The predicted number ratio of MSTO and RC stars is simply the ratio of the sum of the IMF weights of stars of each type. This quantity represents the relative phase lifetimes averaged over the IMF. In the right panel, we show the average mass difference between the RGB and the RC stars. Note that this is currently observable stellar mass instead of the initial mass, and that this is a simple average without the IMF weights. We can safely ignore the IMF weights in this case owing to the negligible dynamic range in mass. RGB stars are first identified using the phase label in the MIST isochrones, and then we apply further selection cuts using the criteria adopted by Miglio et al. (2012); we discard stars that are brighter than the RC magnitude to reduce possible confusion with asymptotic giant branch (AGB) stars, as well as stars fainter than 1.5 mag below the RC in the V band to avoid possible blending and low signal-to-noise ratio issues. RC stars are simply selected by their MIST phase label. The resulting

prediction can be compared to the “observed” mass difference inferred from asteroseismic masses (e.g., Miglio et al. 2012).

In each panel we also include reddening vectors to illustrate how the positions of the CMDs can be shifted owing to dust. We adopt the standard $R_V = A_V/E(B - V) = 3.1$ reddening law from Cardelli et al. (1989) and assume $A_V = 0.4$ evaluated at $[\text{Fe}/\text{H}] = 0$, $\log g = 4$, and $T_{\text{eff}} = 5500$. Reddening may appear to be degenerate with metallicity, especially at the MSTO, SGB, and RGB. But there are subtle differences in the CMD morphology, e.g., the lower MS, especially when multiple CMDs covering a wide wavelength baseline are considered simultaneously (e.g., the “kink” in the lower MS in infrared CMDs; see Pulone et al. 1998; Milone et al. 2014; Correnti et al. 2016). Given exquisite photometry (e.g., *Hubble Space Telescope* or *Gaia*) and sophisticated fitting techniques, we should be able to distinguish the two effects.

3.3. Effect of Metallicity

Figure 1 shows the effect of varying $[\text{Fe}/\text{H}]$, more specifically Z/X , the ratio of metal to hydrogen mass fractions. Note that we assume solar-scaled abundances for the models considered here, i.e., initial $[\text{C}/\text{N}]$ is held constant. The initial helium mass fraction Y is computed assuming a linear enrichment law, a common approach also adopted in MIST. More specifically, the helium abundance is tied to the metallicity assuming a linear relationship, i.e., $Y = Y_p + (Y_{\odot, \text{protosolar}} - Y_p)Z/Z_{\odot, \text{protosolar}}$, where Y_p is the primordial, big bang nucleosynthesis value. The enrichment slope in MIST is $\Delta Y/\Delta Z = (Y_{\odot, \text{protosolar}} - Y_p)/Z_{\odot, \text{protosolar}} = 1.5$. As expected, the CMDs show that increasing metallicity leads to redder colors. The mass–radius panel clearly shows that there is little separation in radius until the stars evolve away from the MS, which suggests that the sensitivity of the models to variations in the age and metallicity is concentrated in the post-MS stars. At a fixed initial mass, metal-rich stars are cooler and have longer MS lifetimes—the Kiel diagram shows a clear sequence in temperature with metallicity.

Both of the surface abundance panels show a large dip between the first two plateaus ($\log g \sim 3.5$ to $\log g \sim 3$), corresponding to the FDU (see Section 2.2.2), and the subsequent decrease is due to thermohaline mixing that is established beyond the RGB bump. Interestingly, the net changes in $[\text{C}/\text{N}]$ and $^{12}\text{C}/^{13}\text{C}$ during the two stages of mixing (FDU and thermohaline) show opposite trends with metallicity: FDU and thermohaline mixing appear to operate more efficiently in high- and low-metallicity systems, respectively. FDU is more efficient at higher metallicities and higher stellar masses because the convective envelope reaches deeper into the CN-processed region and is thus able to dredge up more material to the surface (e.g., Charbonnel 1994; Salaris & Cassisi 2015; Lagarde et al. 2017). Thermohaline is more efficient at lower metallicities and lower stellar masses (e.g., Charbonnel & Zahn 2007; Eggleton et al. 2008; Charbonnel & Lagarde 2010; Lagarde et al. 2017) owing to the compactness of the thermohaline-mixing region and the resulting steeper temperature gradient.

The bottom panels show relatively large variation with stellar age but little variation with metallicity, suggesting that these integrated quantities are more sensitive diagnostics of the stellar age than the metallicity. The number ratio is sensitive to the age because the MS lifetime increases dramatically with decreasing initial mass, but the CHeB lifetime is roughly constant for stars $\lesssim 2 M_{\odot}$. The CHeB lifetime is relatively

insensitive to the initial mass for these stars because they have roughly equal-sized degenerate helium cores that ignite once a critical temperature (corresponding to a critical mass of $\sim 0.45 M_{\odot}$) is reached.

3.4. Effect of Helium

Figure 2 shows the effects of varying the initial bulk helium mass fraction, Y , on the various observables. We hold Z fixed at the protosolar value $Z_{\odot, \text{protosolar}} = 0.0142$ and vary Y and the hydrogen abundance X . Helium abundance cannot be inferred spectroscopically because there are no photospheric helium lines owing to their high excitation potential (Asplund et al. 2009; but see Dupree & Avrett 2013). However, helioseismology provides an indirect probe of the helium abundance in the Sun, relying on the fact that the adiabatic index changes in the He II ionization zone. This technique has yielded a highly precise estimate of the solar helium abundance (0.2485 ± 0.0034 ; Basu & Antia 2004), which is at an apparent tension with the helium abundance required to reproduce the correct solar luminosity and temperature at the solar age for an interior model (Asplund et al. 2009).

The CMDs in the first row demonstrate a systematic trend with Y as they did with metallicity in Figure 1. Note that these panels illustrate the effect of helium abundance on the interior structure and evolution only, largely via changes to the opacity. Helium abundance also influences the stellar spectra and therefore the bolometric corrections, but this effect is unexplored in this work. The mass–radius relation and the $\log g - T_{\text{eff}}$ panels also show a sequence in Y ; at a fixed age, higher helium content leads to hotter stars and lower MSTO masses. The hotter temperature is due to helium’s low opacity, and the decreased MSTO mass (equivalent to a shorter lifetime) is due to the reduced hydrogen abundance. It is interesting to note that the CMDs do not show a clearly separated sequence in Y ; the effect of helium variation moves “along” the isochrone, such that the effect on the CMD is not drastic. On the other hand, the change in the surface abundances, in particular [C/N], during the FDU is less pronounced at high Y . This is because a smaller fraction of the star becomes engulfed, or dredged up, as the stellar mass is decreased. Interestingly, the efficiency of thermohaline mixing appears to be largely unaffected by initial Y . Finally, the number ratio of MSTO to RC stars does not show a clear sequence in Y , and the mass difference between RGB and RC shows marginal separation in Y . However, the difference is much smaller than the current observational uncertainties ($\sim 0.01 M_{\odot}$) and is unlikely to be detectable in the near future.

3.5. Effect of Mixing-length Parameter

In Figure 3, we show the effects of varying the mixing-length parameter α_{MLT} , a free parameter of order unity that is frequently calibrated to match the observations of the Sun (in MIST, $\alpha_{\text{MLT}} = 1.82$; Choi et al. 2016). The physical interpretation of α_{MLT} is the distance, in units of the pressure scale height, that a fluid parcel travels before depositing its energy and disintegrating into the background. Thus, it parameterizes the inefficiency of convective mixing and affects the stellar structure: a small α_{MLT} is associated with cooler T_{eff} and inflated radius. This is clearly illustrated in the CMDs, the mass–radius relation, and the Kiel diagram, particularly for the RGB stars, which have large convective envelopes.

Interestingly, the separation between the models with different values of α_{MLT} is larger in the colors than in T_{eff} , especially on the upper RGB. Given the typical observational uncertainties of $\sim 50\text{--}100$ K in T_{eff} (e.g., Holtzman et al. 2015) and tens of millimagnitude in color,⁷ this suggests that CMDs may be preferable to H-R diagrams for constraining α_{MLT} . However, the CMD approach requires a reddening correction, though a joint fitting of multiple CMDs may help alleviate the issues with degeneracies. Finally, α_{MLT} appears to have a negligible effect on the RGB surface abundances, the number ratio of MSTO to RC stars, and the average mass difference between the RGB and RC phases.

3.6. Effect of Convective Overshoot Mixing

The MLT framework on its own offers an incomplete description of convection and requires a separate prescription for extra mixing that occurs at the convective boundaries. This process, also known as convective overshoot, is meant to capture the nonzero momentum of the fluid parcel approaching the boundary of the convection zone and its subsequent penetration into the radiative region (Unno 1957; Böhm 1963; Shaviv & Salpeter 1973). Overshoot implies enhanced mixing and thus has several observable consequences, including the properties of AGB and post-AGB stars (e.g., Herwig 2000; Herwig et al. 2011), the MS width (Schaller et al. 1992), and the MSTO morphology in open clusters (e.g., Magic et al. 2010).

We adopt the “exponential diffusive overshoot” framework introduced by Freytag et al. (1996) and implemented in MESA following Herwig (2000). This prescription is meant to capture both the exponential decay of the convective velocity field and the dissolution of the fluid parcel as a diffusive process. There are two sets of free parameters available for every convective boundary in MESA: f_{ov} and $f_{0,\text{ov}}$. The first parameter, f_{ov} , determines the efficiency of overshoot mixing and describes the velocity scale height in terms of the local pressure scale height. The second parameter, $f_{0,\text{ov}}$, determines the location inside the convection zone at which the diffusion coefficient is calculated. For simplicity, we fix the latter to $f_{0,\text{ov}} = 0.008$ (half of the fiducial value for f_{ov}) as we vary f_{ov} to investigate the role of the efficiency of overshoot mixing.

Figure 4 shows the effect of varying f_{ov} in the hydrogen-burning core. Increasing the efficiency of overshoot mixing in the hydrogen core leads to a more prominent MSTO morphology and a brighter SGB owing to an enhanced MS lifetime and a larger core. Note that this parameter has no effect on old populations because their MSTO stars are sufficiently low in mass that they do not harbor convective cores during the MS.

We also tested the effect of varying $f_{0,\text{ov}}$ in the helium-burning core. Somewhat surprisingly, changing this parameter seems to have little to no effect on any of the observables; therefore, the corresponding figure is not shown. Montalbán et al. (2013) computed a series of stellar models and adiabatic frequencies and found a correlation between the average value of the asymptotic period spacing (ΔP) and the size of the helium-burning core. In a more recent work, Arentoft et al. (2017) analyzed a sample of red giants in the open cluster NGC 6811 and found that overshoot mixing in

⁷ The *Hubble Space Telescope*, which nominally represents the best-case scenario today, routinely achieves $\lesssim 1$ mmag relative uncertainty.

Helium Abundance: Y

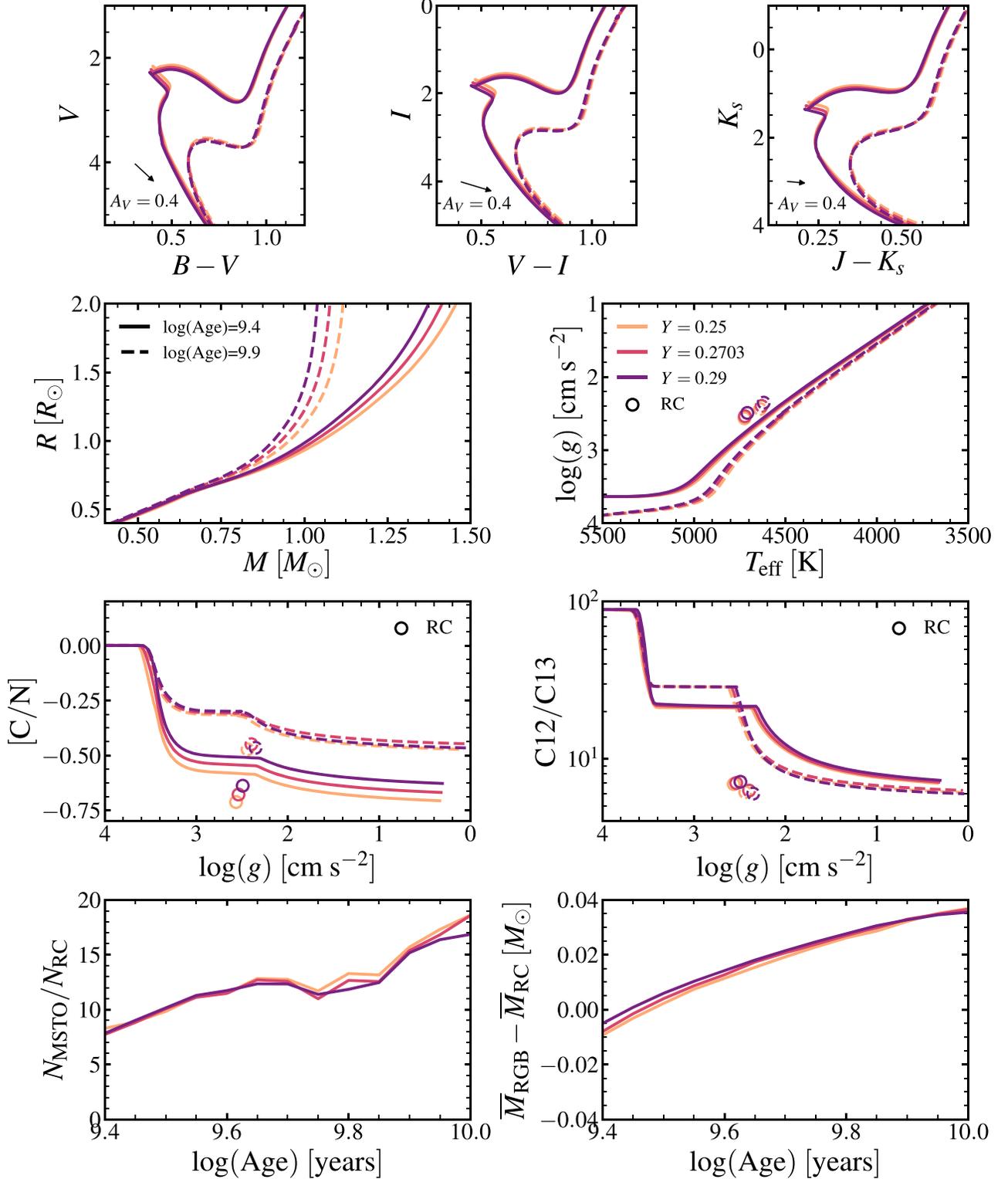


Figure 2. Same as Figure 1, except now varying the helium abundance at a fixed metal mass fraction Z . The fiducial value of $Y = 0.2703$ comes from the protosolar helium abundance in Asplund et al. (2009). At a fixed stellar age, higher helium content leads to hotter stars and shorter MS lifetimes.

the helium-burning core does not appear to have a noticeable effect on the resulting $\Delta\nu$ and ν_{\max} as long as overshoot mixing is included during the MS. However, the authors also found that $\Delta\nu$ and ΔP together have the potential to constrain the efficiency of overshoot mixing in the helium

core and shed light on the still-debated presence of breathing pulses (Castellani et al. 1985). Asteroseismic modeling that probes the detailed interior stellar structure may be required to constrain the efficiency of overshoot in the helium-burning core.

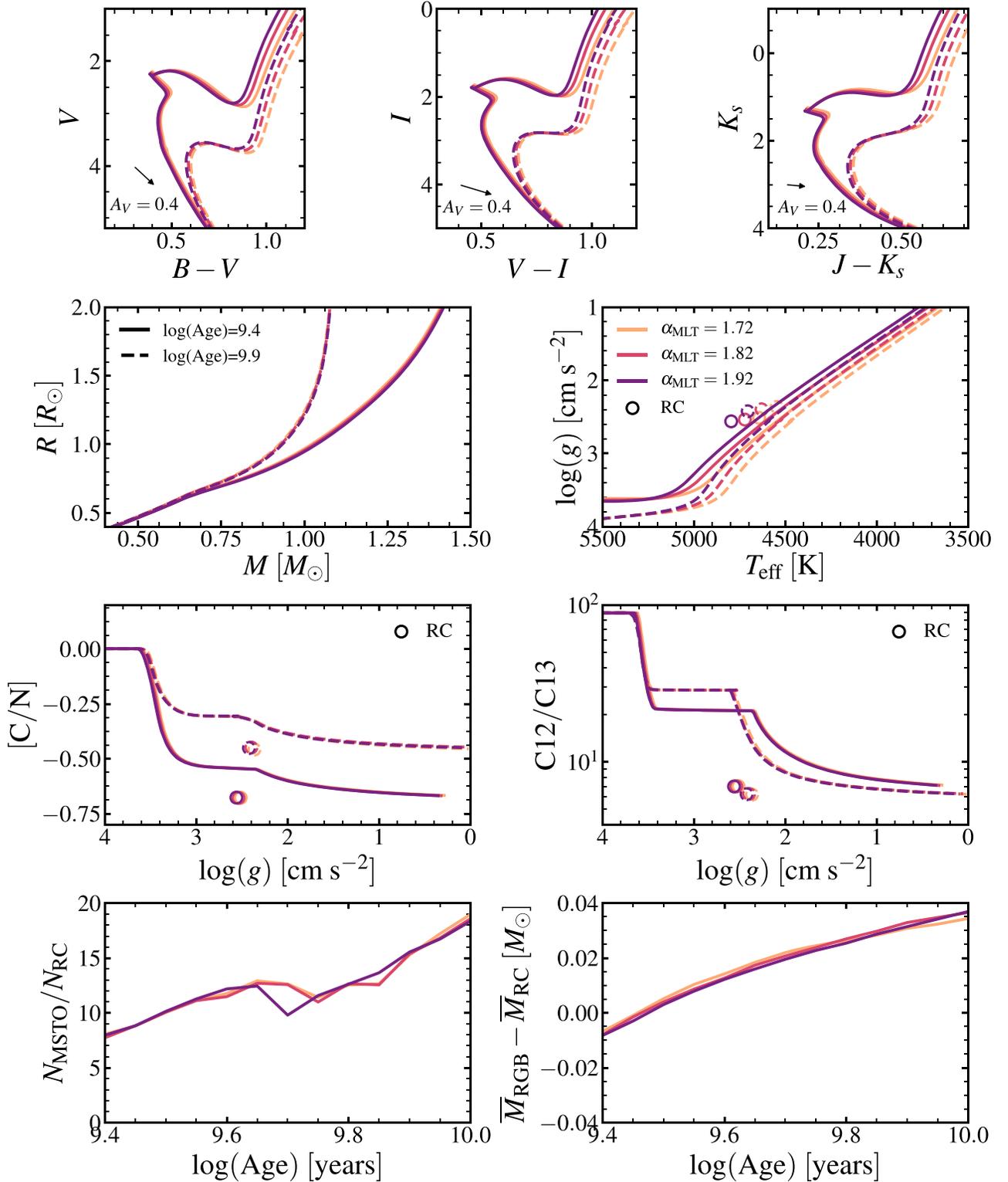
Mixing Length Parameter: α_{MLT}


Figure 3. Same as Figure 1, except now varying the mixing length α_{MLT} , which parameterizes the efficiency of convective mixing. The fiducial value of $\alpha_{\text{MLT}} = 1.82$ was chosen by calibrating to the observations of the Sun (see Section 4.1 in Choi et al. 2016, for more details). The choice of α_{MLT} has the largest effect on the temperature and therefore the color of the RGB.

3.7. Effect of Thermohaline Mixing

In Figure 5, we show the effects of varying α_{th} , the efficiency of thermohaline mixing. As described in

Section 2.2.2, thermohaline is a type of mixing that occurs in a thermally stable medium that has a destabilizing composition gradient. Standard models do not predict any

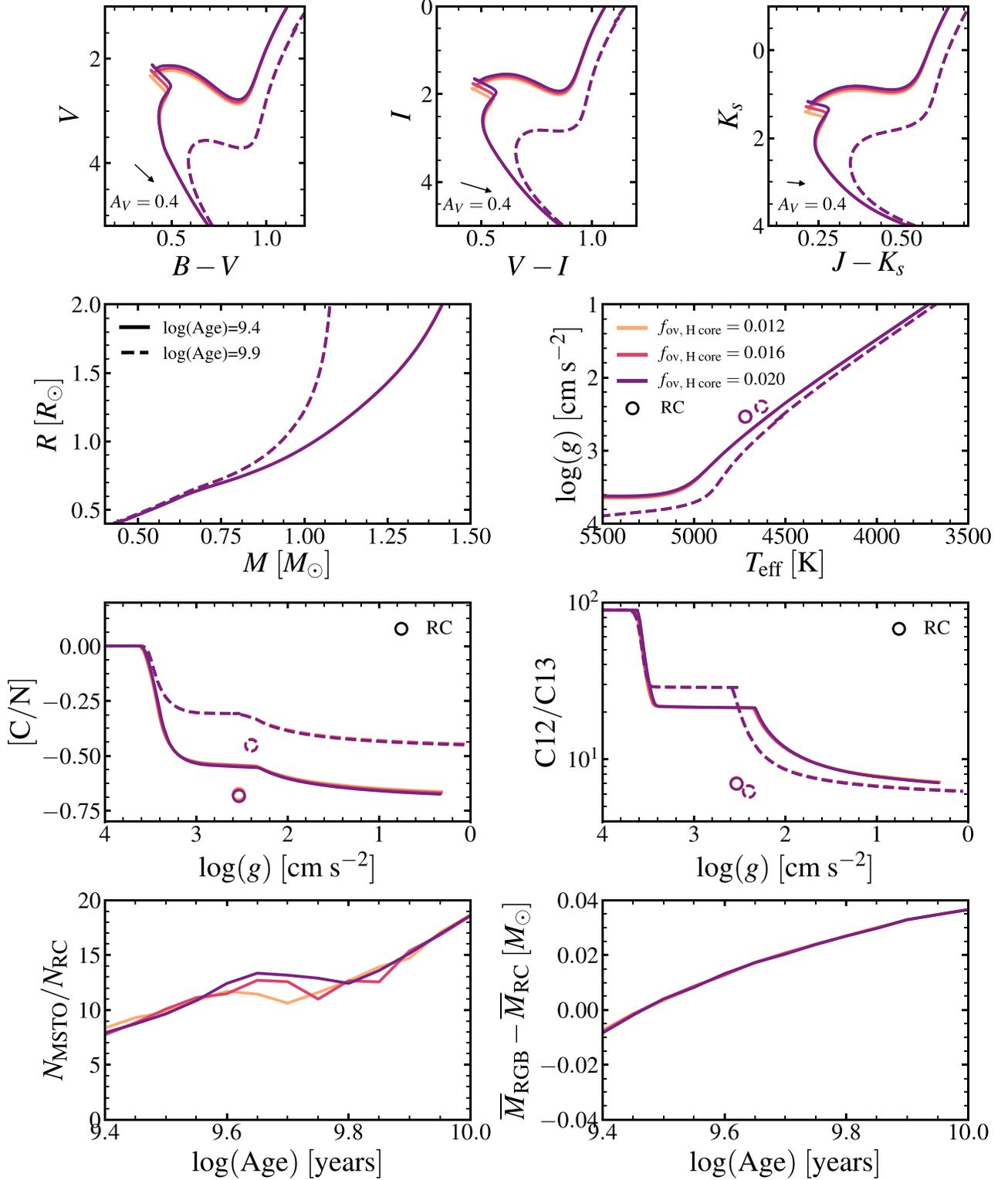
Overshoot Mixing in the Hydrogen-burning Core: $f_{\text{ov}, \text{H core}}$


Figure 4. Same as Figure 1, except now varying the efficiency of convective overshoot mixing in the hydrogen-burning core, $f_{\text{ov}, \text{H core}}$. The fiducial value of $f_{\text{ov}, \text{H core}} = 0.016$ was chosen to reproduce the MSTO morphology of the open cluster M67. Independent constraints from DEBs (Claret & Torres 2016) also lend support for this calibrated value. The choice of $f_{\text{ov}, \text{H core}}$ most noticeably affects the MSTO morphology and the luminosity of the SGB because the enhanced mixing during the MS leads to longer MS lifetimes (thus a larger MSTO mass at a fixed age) and a larger core. However, note that this has no effect on an old population because the MSTO stars are low in mass and thus harbor radiative cores.

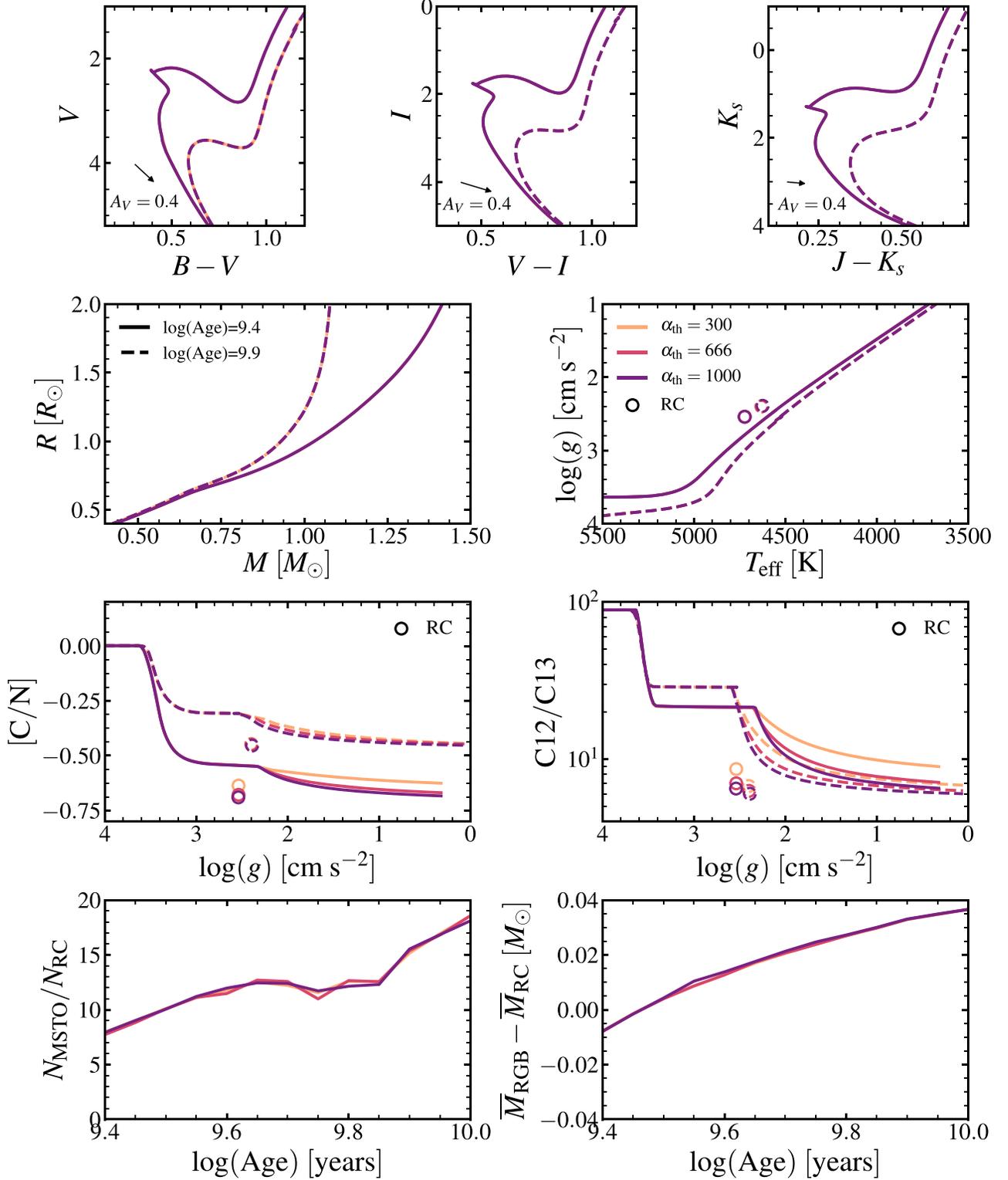
Thermohaline Mixing: α_{th}


Figure 5. Same as Figure 1, except now varying the efficiency of thermohaline mixing α_{th} . The fiducial value of α_{th} was recommended by Charbonnel & Zahn (2007) to reproduce the observed surface abundances of stars brighter than the RGB bump. The choice of α_{th} has essentially no distinguishable effect on any of the observables except for the surface abundances of RGB stars brighter than the bump.

further changes to the surface abundances along the RGB at the conclusion of the FDU, but abundance evolution beyond the RGB bump is indeed observed (e.g., Gratton

et al. 2000). Thermohaline mixing is a viable mechanism for explaining this phenomenon (but see also, e.g., Denissenkov 2010; Traxler et al. 2011; Wachlin et al. 2014), wherein an

unstable composition gradient is established by the ${}^3\text{He}$ reaction taking place in the external wing of the hydrogen-burning shell. In the framework of Ulrich (1972) and Kippenhahn et al. (1980), α_{th} has a geometric interpretation—a large value corresponds to a slender fluid element—which is also directly linked to the mixing timescale and thus the mixing efficiency. The fiducial value of α_{th} in the MIST models was recommended by Charbonnel & Zahn (2007) to reproduce the observed surface abundances of stars brighter than the RGB bump.

Perhaps unsurprisingly, changing the efficiency of thermohaline mixing has almost no observable influence except in the surface abundances beyond the RGB bump (see also Lagarde et al. 2017). Note that this effect saturates beyond some critical value of α_{th} (pink and purple curves), suggesting that there is a maximum efficiency with which thermohaline mixing operates. Given the very minor influence of thermohaline mixing on the overall evolution, we do not expect noticeable differences among the models in the bottom left panel showing the number ratio of MSTO and RC stars. The small variations are largely due to the presence of very weak breathing pulses occurring at the end of the CHeB phase, and thus the size of these variations nominally represents the minimum theoretical uncertainty on this quantity.

3.8. Effect of Mass-loss Efficiency

Figure 6 illustrates the effects of varying the wind efficiency, η . In particular, since we are focusing on the evolutionary phases preceding the AGB, the relevant mass-loss scheme is the Reimers (1975) prescription, where $\dot{M} \propto LR/M$ with a prefactor η of order unity. For the fiducial MIST models, we adopt $\eta = 0.1$, which is smaller than the value traditionally adopted in stellar models (e.g., Girardi et al. 2000; Pietrinferni et al. 2004; Ekström et al. 2012). This choice was motivated by the results from Miglio et al. (2012) and more recently Handberg et al. (2017), who demonstrated that the asteroseismic masses prefer only a modest amount of mass loss on the RGB. The mass-loss rate rises steadily as the star ascends the RGB, eventually reaching values as high as $10^{-8} M_{\odot} \text{ yr}^{-1}$ at the tip of the RGB. Variations in the mass-loss efficiency parameter have almost no effects on the observables considered here except in the average mass difference between the RGB and the RC.⁸ If there is no mass loss on the RGB, this mass difference is always less than zero because the current stellar mass is exactly equal to the initial mass and MS lifetime decreases with increasing stellar mass. On the other hand, if there is significant mass loss between the RGB and the RC, this quantity will be positive. The mass difference is negative valued at almost all stellar ages for the model with very little mass loss (peach curve), which suggests that evolutionary timescale is the dominant effect in this case.

3.9. Other Parameters to Consider

In addition to the “internal” sources of uncertainties in stellar models, there are also “external” sources of uncertainties due to difficulties associated with measuring abundances, distances,

⁸ Although the effect of mass loss on the horizontal branch morphology is not considered in this work, this subject has been studied extensively in the context of globular clusters. See, e.g., Rood (1973), Lee et al. (1990), Vink & Cassisi (2002), Dotter (2008), Catelan (2009), Gratton et al. (2010), and Salaris et al. (2016).

and reddening. Besides metallicity and helium abundance, oxygen abundance (either on its own or grouped with the other α -capture elements) is another key parameter in stellar models owing to its strong influence on the overall stellar structure and evolution and thus the inferred stellar age (e.g., VandenBerg & Bell 2001; Dotter et al. 2007; VandenBerg et al. 2012; Bond et al. 2013). Oxygen contributes significantly to the overall opacity in the stellar interior and alters the relative importance of the CNO cycle compared to the pp-chain. We generally assume metallicity⁹ to be a well-determined quantity from spectroscopy, but as noted in Section 2.2.1, different methods can yield systematically different measurements at the ~ 0.1 dex level (Smiljanic et al. 2014). For some species such as oxygen, the spectroscopic abundance may be even more uncertain owing to difficulties associated with weak, blended lines, non-LTE effects, and/or 3D effects (see the discussion in Asplund et al. 2009; but see also Ting et al. 2018).

Over the next several years, the *Gaia* mission will effectively eliminate distance and membership uncertainties for open clusters provided that they are sufficiently nearby. The predicted end-of-mission parallax error for an individual G2V star at $V = 15$ is $24 \mu\text{as}$,¹⁰ which corresponds to 2% and 12% precision at 1 and 5 kpc, respectively. Even though many clusters, including NGC 6791, have MSTOs that are significantly fainter than $V = 15$, we should still be able to obtain robust distance estimates by relying on the bright red giants and combining the constraints from many more fainter stars. For reference, the three open clusters considered for this work in Section 4 are at distances of ~ 800 pc to ~ 4 kpc and have MSTO magnitudes of $G = 13\text{--}17.5$.

Extinction (particularly differential extinction) remains a challenging problem. For this work, we assume the commonly adopted CCM reddening law (Cardelli et al. 1989), but there are several alternatives, including the Fitzpatrick (1999) and O’Donnell (1994) reddening laws. In practice, R_V , which parameterizes the slope of the optical extinction curve, is almost always assumed to be the galactic average $R_V = 3.1$ even though there is evidence for variations along different sight lines (Draine 2003). The uncertainties associated with the treatment of extinction may well be a dominant source of uncertainty in our interpretation of CMDs in the *Gaia* era. One path forward may be the use of panchromatic CMDs to infer the bolometric magnitude from modeling the spectral energy distribution, which would remove extinction from at least the y-axis of the CMD.

3.10. Separation of Information

Here we provide a succinct, visual summary of the information presented in the previous sections. Each panel in Figure 7 illustrates the sensitivity of a pair of observables to a parameter of interest and the stellar age for the purposes of disentangling their effects. The blue error bar represents a typical observational uncertainty. For some parameters such as η_{R} and α_{MLT} , their effects are nearly orthogonal to that of stellar age on their respective pairs of observables. For other parameters such as $[\text{Fe}/\text{H}]$, Y , and α_{th} , their effects are separable but covariant with the effect of stellar age. Finally, in the case of $f_{\text{ov,H core}}$ and $f_{\text{ov,He core}}$ (not shown), these

⁹ But not helium!

¹⁰ <https://www.cosmos.esa.int/web/gaia/science-performance>

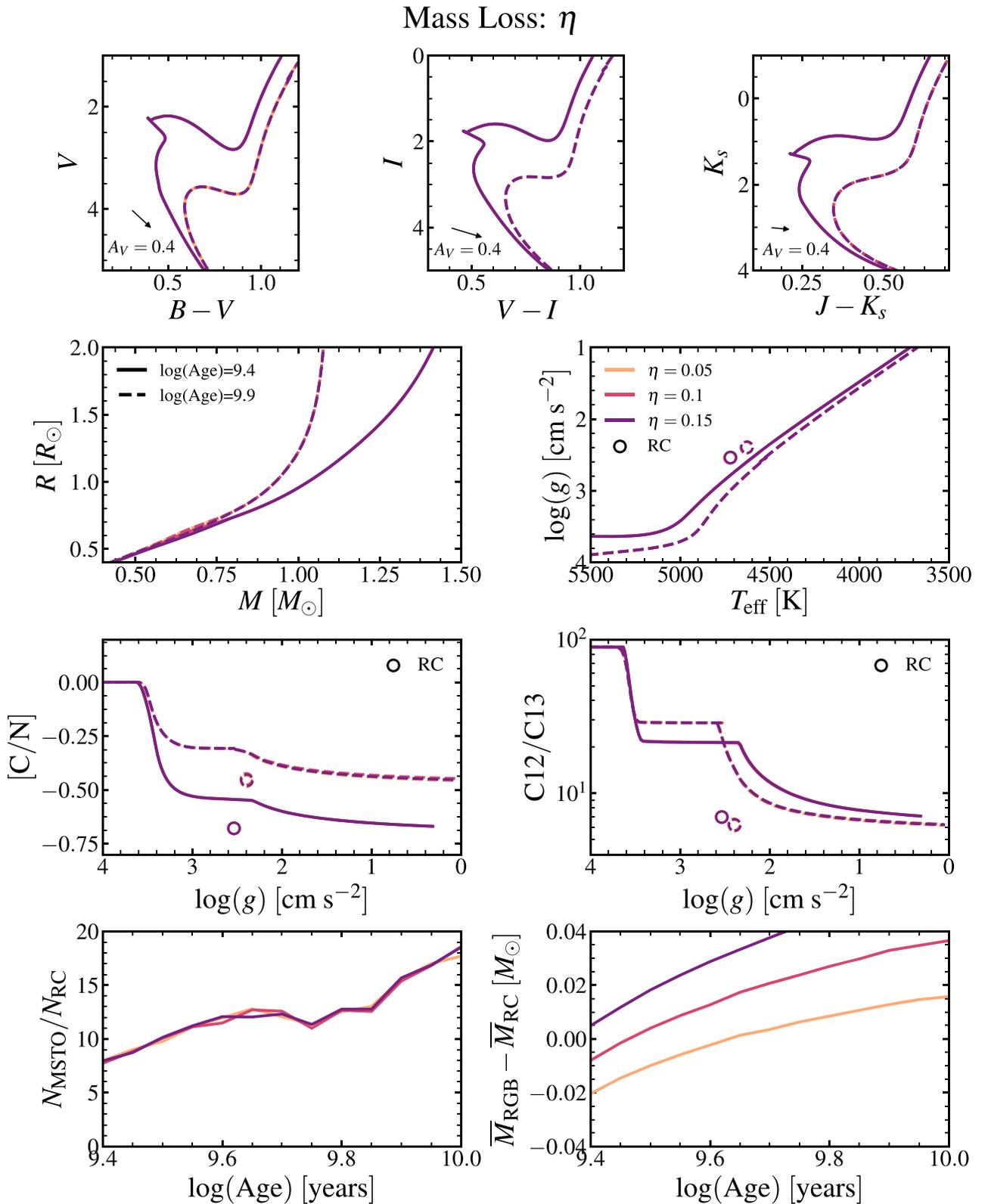


Figure 6. Same as Figure 1, except now varying the Reimers mass-loss parameter η . The fiducial value of $\eta = 0.1$ was recommended by the *Kepler* asteroseismic constraints from open clusters (Miglio et al. 2012; Handberg et al. 2017). It also reproduces the initial-final mass relation (see Section 8.2 in Choi et al. 2016, for more details). The choice of η has almost no discernible effect on any of the observables except for the masses of RGB and RC stars.

observables do not cleanly separate the parameters from stellar age.

However, we wish to emphasize that Figure 7, though it is useful for illustrative purposes, does not fully encapsulate the

sheer amount of information that is present in Figures 1–6. What these panels do not capture are the subtle yet qualitatively distinct morphologies in the panchromatic CMDs and the changes to the relative number densities of stars along the

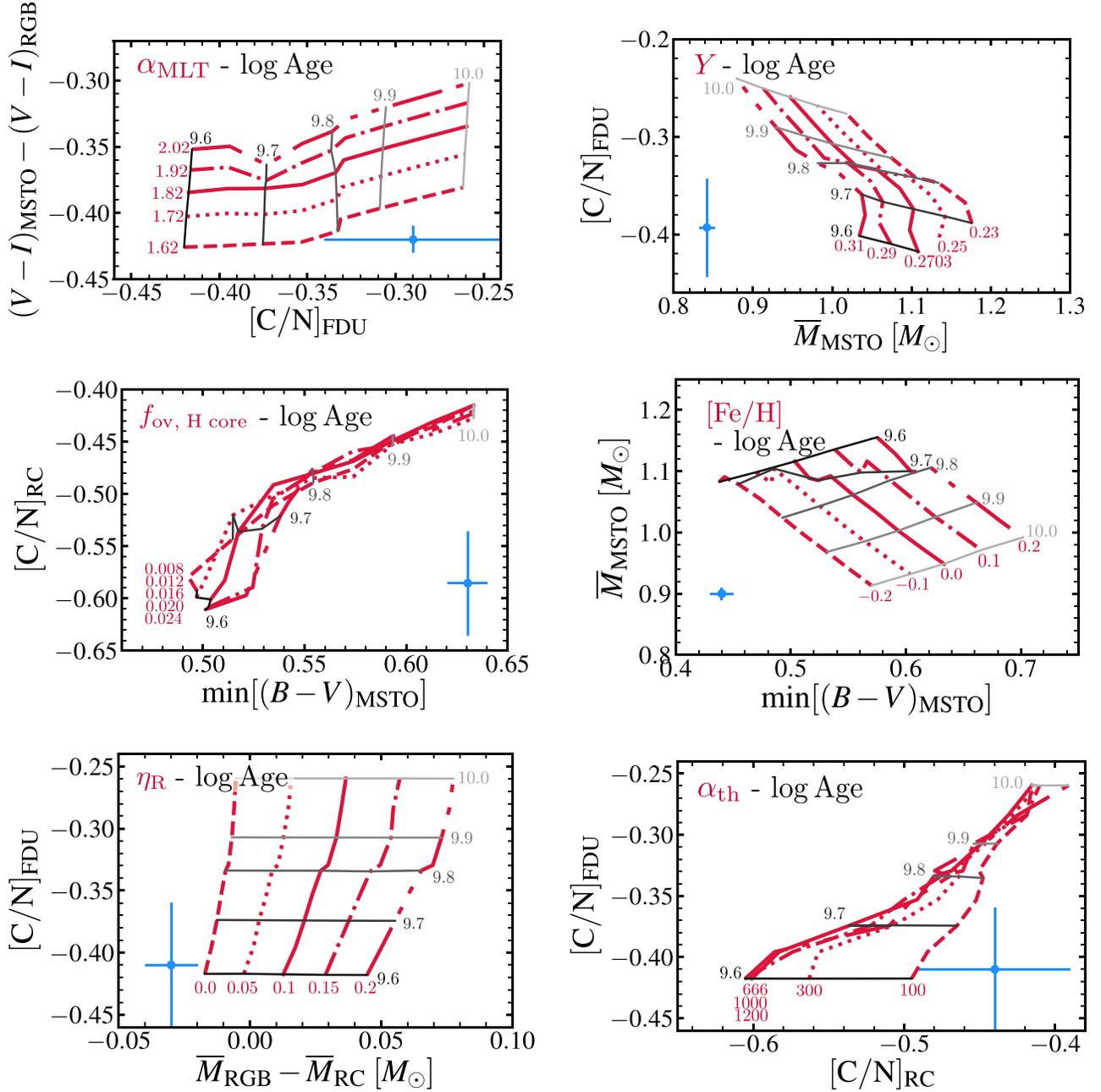


Figure 7. Diagnostic sensitivity of a pair of observables to a parameter of interest and the stellar age. For each parameter of interest, we identify a set of observables that most cleanly separate in $\log(\text{Age})$ (gray lines) and the parameter in question (red lines), though this was not always possible in every case. Top left: mixing-length parameter; $V - I$ color difference between the MSTO and the RGB (measured at $V = 1.5$) vs. surface $[\text{C}/\text{N}]_{\text{FDU}}$ abundance of post-FDU stars below the RGB bump vs. average mass of the MSTO stars. For low values of Y , there is no RGB bump at $\log(\text{Age}) = 9.6$ [yr]. Middle left: convective overshoot mixing efficiency in the hydrogen core; surface $[\text{C}/\text{N}]_{\text{RC}}$ abundance of RC stars vs. $B - V$ color at the MSTO. Middle right: metallicity; average mass of the MSTO stars vs. $B - V$ color at the MSTO. Bottom left: mass loss; surface $[\text{C}/\text{N}]_{\text{FDU}}$ abundance of post-FDU stars below the RGB bump vs. average mass difference between the RGB and RC stars. Bottom right: thermohaline mixing; surface $[\text{C}/\text{N}]_{\text{FDU}}$ abundance of post-FDU stars below the bump vs. surface $[\text{C}/\text{N}]_{\text{RC}}$ abundance of RC stars. The blue error bar represents a typical observational uncertainty (see Section 2 for an in-depth overview of different observational data sets).

CMD—in the sense of a Hess diagram—due to the model parameters. For instance, the complexities of the CMDs are boiled down to two scalar quantities, the $V - I$ difference between the MSTO and the RGB or the bluest extent of the MSTO in $B - V$, in this diagram. As such, although Figure 7 appears to suggest that it is virtually impossible to infer $f_{\text{ov}, \text{H core}}$ from observations, Figure 4 demonstrates that the detailed shape of the Henyey hook at a fixed stellar age can be used to constrain $f_{\text{ov}, \text{H core}}$. This underscores the potential of a

full CMD-fitting approach (e.g., MATCH, Dolphin 2002; BASE9, von Hippel et al. 2006) in the era of high-precision data.

4. Open Clusters: Case Studies

Now that we have qualitatively explored the effects of key uncertain parameters on several observables, we evaluate the current state of the available data and assess whether they can be used to disentangle and constrain the parameters under

consideration. For this exercise, we investigate three open clusters, NGC 6819, M67, and NGC 6791. We chose these systems for several reasons. First, all three clusters have been studied extensively and therefore represent some of the best-case scenarios. All three clusters have been observed by the *Kepler* mission as part of either the original campaign (NGC 6819 and NGC 6791) or the repurposed *K2* mission (M67). They have all been observed in several photometric bands, targeted by the APOGEE spectroscopic survey (Holtzman et al. 2015), and they are known to harbor one or more DEBs. Second, non-solar-scaled abundances and multiple stellar populations are less of a concern in open clusters compared to globular clusters (Bedin et al. 2004; Piotto et al. 2007).¹¹ While one of the major distinctions between globular clusters and open clusters is believed to be the presence or the absence of multiple stellar populations, this simple dichotomy is becoming increasingly challenged (see Geisler et al. 2012; Gratton et al. 2012; Bragaglia et al. 2014; Bastian & Lardo 2017). Nevertheless, it is possible to model globular clusters and/or multiple populations in this context as well (see Dotter et al. 2015, where the authors created tailored stellar interior and atmosphere models for NGC 6752 taking into account the individual abundances of two stellar populations). The nearby globular cluster M4 may be an interesting candidate for analysis when the α -enhanced MIST models become available, although its strong total and differential reddening (e.g., Hendricks et al. 2012) may pose a challenge; it has at least three known double-lined DEBs (Kaluzny et al. 2013), *K2* asteroseismology (Miglio et al. 2016), and APOGEE spectra (Zasowski et al. 2017). Third, these three clusters form a sequence in age and thus allow for model comparison in different stellar mass regimes. We note that NGC 6791 is noticeably more metal-rich compared to the other two clusters.

4.1. Cluster Membership and Distances with Gaia DR2 Data

We make use of the recent *Gaia* DR2 (Gaia Collaboration et al. 2018b) to identify likely cluster members and obtain clean CMDs. First, we construct a parent catalog using the Large Survey Database (Juric 2012) framework to combine the DR2 catalog with data from other surveys, such as Pan-STARRS (Chambers et al. 2016; Flewelling et al. 2016). We use a 3'' threshold to cross-match the objects, selecting the nearest candidate in the event of duplicate matches. Next, we utilize the sky positions and estimates of the cluster size from the Kharchenko et al. (2013) Milky Way Star Clusters catalog to perform a cone search centered on each cluster. In particular, we choose a search radius of twice the total apparent radius (r_2) to include as many of the potential cluster members as possible.

Next, for each cluster, we identify the likely cluster members by running the HDBSCAN clustering algorithm (Campello et al. 2013) on proper motion (μ_α , μ_δ) and parallax ϖ . Following Lindegren et al. (2018) and Gaia Collaboration et al. (2018a), we first remove objects with failed astrometric solutions ($\text{astrometric_chi2_al}/(\text{astrometric_n_good_obs_al}-5) < 1.44\text{MAX}(1, \exp(-0.4(\text{phot_g_mean_mag}-19.5)))$). HDBSCAN identifies clusters based on the density of points and, importantly, does not force all data points to belong to a detected cluster. Unlike the DBSCAN

algorithm on which it is based, HDBSCAN is more flexible in that it allows the density of the clusters to vary. Its other advantages are that there is only one important and relatively intuitive free parameter, the minimum cluster size (30), and that the algorithm returns a membership probability for every data point. We select likely cluster members with HDBSCAN membership probabilities greater than 50% for NGC 6819 and M67. For NGC 6791, we relax this threshold to 30% to retain fainter stars with lower-quality astrometric data and thus increase sampling below the MSTO.

As discussed in Luri et al. (2018), a full Bayesian inference is the preferred method for obtaining parallax-based distances given the nonlinearity between the desired (distance) and measured (parallax) quantities and the constraint that the former be necessarily positive while the latter is allowed to be zero or negative. However, as discussed in Gaia Collaboration et al. (2018a), estimating the distance through a simple inversion is acceptable as long as the relative precision in parallax is lower than $\sim 20\%$. Following Gaia Collaboration et al. (2018a), we adopt an even more strict 10% precision criterion to obtain a subsample from the likely members as identified by HDBSCAN. Although this biases against fainter members in the cluster, this is not a concern for this work because completeness is not a priority.

Before we invert the parallax measurements, we must first consider the effect of zero-point offsets in *Gaia* DR2, which unfortunately are not well characterized below the level of ~ 0.05 mas at this time (Arenou et al. 2018; Lindegren et al. 2018). Independent comparisons with quasars (Lindegren et al. 2018), external catalogs and Milky Way satellites (Arenou et al. 2018), eclipsing binaries (Stassun & Torres 2018), and RR Lyrae (Muraveva et al. 2018) have revealed both large- and small-scale ($< 1^\circ$) spatial variations in the parallax zero-points, with an average global offset of ~ -0.03 to -0.08 mas (*Gaia* parallaxes are systematically too small). We cannot reliably correct for this by adding a constant zero-point offset given that the small-scale variations are comparable in size to the average global offset. As we describe in detail in the following sections, this can lead to differences in the absolute distance modulus by a value as large as ~ 0.4 mag in the case of NGC 6819. For this work, we first compute the fiducial cluster distance using the median parallax from the high-precision subsample, apply the resulting distance modulus to the model CMDs, and assess the quality of the fits to all of the available cluster data. If the resulting fit is entirely inconsistent with the observations (i.e., impossible to reconcile by changing the age or the reddening), then we allow for a small increase in the parallax by an amount between 0 and 0.08 mas. In a future work, we will present a detailed description of the cluster membership selection process, in addition to a more rigorous determination of the cluster parameters using a hierarchical Bayesian analysis where we also model the parallax zero-points.

4.2. NGC 6819

NGC 6819 is a solar-metallicity, intermediate-age (2 Gyr), richly populated open cluster (Yang et al. 2013; Anthony-Twarog et al. 2014; Lee-Brown et al. 2015). As the youngest system in our sample, its MSTO stars are massive enough to have convective cores, giving rise to a distinctive MSTO morphology called the Henyey hook.

Figure 8 shows the multipanel plot summary of NGC 6819. The panels are analogous to those presented in Figures 1–6.

¹¹ There is evidence that NGC 6791 may be moderately α -enhanced (Linden et al. 2017), but see also Carretta et al. (2007), Boesgaard et al. (2015), and Ting et al. (2018).

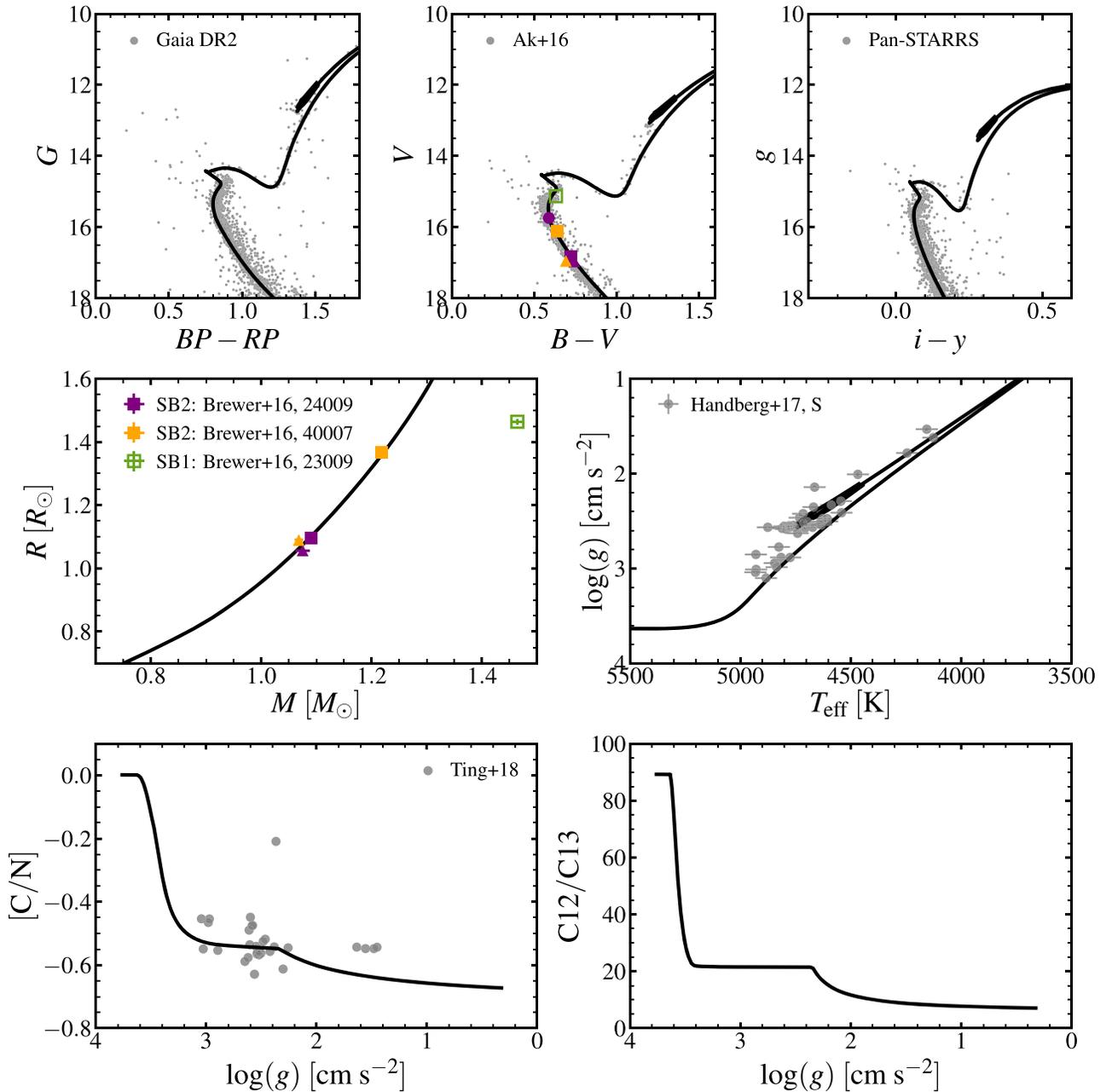


Figure 8. NGC 6819, a solar-metallicity, intermediate-age (2 Gyr) open cluster. The top three panels show photometry of likely cluster members selected using the *Gaia* DR2 data, along with the MIST models (see text for more details). The MIST CMDs adopt $m - M = 11.8$, $[\text{Fe}/\text{H}] = -0.01$, $\log(\text{Age})[\text{yr}] = 9.4$, and $A_V = 0.5$. The left panel in the middle row shows stellar parameters measured from DEB (Brewer et al. 2016) with the MIST mass–radius relation. The open and filled symbols indicate that the system is a single- (SB1) and double-lined (SB2) spectroscopic binary system, respectively. When available, the individual EB components are also shown in the CMDs, where the squares, triangles, and circles correspond to the primary, secondary, and tertiary components, respectively. The right panel shows the asteroseismic $\log g$ of single stars inferred from the scaling relations (Handberg et al. 2017), with the effective temperatures from the Casagrande & VandenBerg (2014) color–temperature relations. The bottom left panel shows the comparison between predicted surface $[C/N]$ abundances and the measured abundances, shown as gray circles. The surface abundances are obtained by reanalyzing the publicly available APOGEE DR14 spectra (Holtzman et al. 2015; SDSS Collaboration et al. 2017) with *the Payne* (Ting et al. 2018). The measurement of $^{12}\text{C}/^{13}\text{C}$ in the right panel is currently being explored.

The photometry comes from several sources: *Gaia* DR2 (Evans et al. 2018), *BVI* (Ak et al. 2016), and Pan-STARRS (Chambers et al. 2016; Flewelling et al. 2016). We utilize the *Gaia* DR2 membership (see Section 4.1) to select the likely cluster members from the Ak et al. (2016) catalog. We apply additional cuts in the *Gaia* photometry using `phot_bp_rp_excess_factor` following Gaia Collaboration et al. (2018a). The reported photometric uncertainties are ~ 2 , 3, and 5 mmag in G , V , and g_{PS} near the MSTO. Note that the

Pan-STARRS CMD is truncated above $g_{PS} \approx 15$ owing to saturation, but nevertheless we include it for its well-sampled lower MS and availability of the near-IR photometry. The black isochrone corresponds to an example “fit” to the data using the MIST isochrone. The metallicity is assumed to be the median $[\text{Fe}/\text{H}]$ value for the sample of red giants, measured from the publicly available APOGEE DR14 spectra (Holtzman et al. 2015; SDSS Collaboration et al. 2017) with *the Payne* (Ting et al. 2018; see below for more details), and the fiducial

value of the distance modulus is estimated from the *Gaia* DR2 parallax measurements as described in Section 4.1. Finally, we start with literature values of reddening and age and choose the “best-fit” values by eye for purely illustrative purposes. However, we found that $m - M = 12.22$, the distance modulus inferred from the *Gaia* DR2 parallax—and the subsequent adjustments to the age and reddening in order to fit the MS in the CMDs—is strongly ruled out by both the RC magnitude and the rest of the observations. This is consistent with the conclusion from the literature that the *Gaia* DR2 parallaxes are systematically too small (see the discussion in Section 4.1). As such, we are afforded some flexibility to apply a zero-point offset to the fiducial cluster parallax, and we find that 0.075 mas is a suitable value, which, though large, is still within the range of estimates from the literature. In summary, the MIST models assume $m - M = 11.8$, $[\text{Fe}/\text{H}] = -0.01$, $\log(\text{Age})[\text{yr}] = 9.4$, and $A_V = 0.5$.

The mass and radius measurements of the two EB systems, WOCS 24009 (Auner 665; KIC 5023948) and WOCS 40007 (KIC 5113053), are derived from a combination of *Kepler* and ground-based photometry and spectroscopy (Brewer et al. 2016). In fact, each one belongs to its own triple system: WOCS 24009 is a triple-lined system where the brightest, noneclipsing component is orbiting a short-period binary system and WOCS 40007 is a double-lined system. There is a third EB system, WOCS 23009 (Hole et al. 2009; Sandquist et al. 2013), but it is a single-lined EB, and thus the inferred parameters are less certain. Nevertheless, we include the parameters of the primary in our comparison, but in open symbols to reflect its lower fidelity. When available, the individual EB components are also shown in the CMDs, where the squares, triangles, and circles correspond to the primary, secondary, and tertiary components, respectively.

Moving on to the right panel, we show the asteroseismic $\log g$ and T_{eff} for single stars from Handberg et al. (2017). The authors used a variation of the classic scaling relations that are recast to include bolometric luminosities. The bolometric corrections and color–temperature relations that are required to estimate L and T_{eff} come from Casagrande & Vandenberg (2014). They utilize V (Milliman et al. 2014) and K_s (Cutri et al. 2003) photometry and assume a nominal reddening value of $E(B - V) = 0.15$ and $[\text{Fe}/\text{H}] = 0.02 \pm 0.10$. We adopt a T_{eff} uncertainty of 50 K following the authors’ estimates. Overall, the asteroseismic $\log g$ and T_{eff} are in good agreement with the MIST model predictions, in particular the RC magnitude and the RGB T_{eff} .

Finally, the last row shows the comparison between predicted surface [C/N] abundances and the measured observed abundances, shown as gray circles. The surface abundances are provided by *the Payne* (Ting et al. 2018) reanalysis of the APOGEE DR14 spectra (Holtzman et al. 2015; SDSS Collaboration et al. 2017). Here we only briefly describe *the Payne* since the details of the methodology are presented in Ting et al. (2018). *The Payne* utilizes the idea of generative models; it fits the variations in normalized flux with respect to stellar labels (stellar parameters and elemental abundances) with a flexible functional form approximated with neural networks. The neural networks are trained on the ATLAS12/SYNTHETIC model spectra (Kurucz 1970; Kurucz & Avrett 1981; Kurucz 1993), and the observed spectra are fit via full spectral fitting. The formal uncertainties for *the Payne* are very small (<0.01 dex), but the true uncertainties are usually dominated by model systematics. The abundance spreads measured in open

clusters, which are presumed to be chemically homogeneous, imply a precision of ≈ 0.03 dex (Ting et al. 2018). *The Payne* [C/N] abundances are in good agreement with the MIST prediction. The full spectral fitting approach in principle allows for the measurement of $^{12}\text{C}/^{13}\text{C}$, which is currently being explored.

4.3. M67

M67 is a nearby solar-metallicity, intermediate-age (4 Gyr) open cluster (Taylor 2007; Sarajedini et al. 2009; Önehag et al. 2014). One of the reasons it is so well studied is that its MSTO mass is very close to the transition mass above and below which stars burn hydrogen convectively and radiatively in their cores. Its Henyey hook is frequently used to calibrate the efficiency of convective overshoot mixing in low-mass stars (e.g., Vandenberg et al. 2006; Magic et al. 2010; Bressan et al. 2012; Choi et al. 2016).

Figure 9 shows the multipanel plot summary of M67. The optical and near-infrared photometry comes from *Gaia* DR2 (Evans et al. 2018), Yadav et al. (2008), and Sarajedini et al. (2009). For the latter two catalogs, we only show a subset of the stars with membership probabilities greater than 50%, as evaluated by Yadav et al. (2008) using their own proper-motion data. An additional proper-motion-selected sample of likely single-star members from Sandquist (2004) is overplotted for comparison. The reported photometric uncertainties are ~ 4 , 3, and 5 mmag in G , I , and K_s , respectively, near the MSTO. Similar to the approach used to model NGC 6819, we adopt the $[\text{Fe}/\text{H}]$ value from *the Payne* and estimate the distance modulus from the *Gaia* DR2 parallax. We do not apply a parallax zero-point offset, though an increase of 0.029 mas (the global zero-point as estimated in Lindegren et al. 2018), resulting in a decrease in the distance modulus of ~ 0.05 mag, is also permissible with a small decrease in age. In summary, the MIST models assume $m - M = 9.73$, $[\text{Fe}/\text{H}] = -0.01$, $\log(\text{Age})[\text{yr}] = 9.58$, and $A_V = 0.18$. In $BP - RP$, the model diverges from the data near $G \lesssim 15$ for reasons that are unknown at this time. Interestingly, $B - V$ and $B - I$ (not shown) show excellent agreement on the lower MS down to $G \approx 16.5$, below which the model diverges blueward of the data owing to the well-known M-dwarf radius inflation problem (e.g., Kraus et al. 2011; Feiden & Chaboyer 2013; Torres 2013). Moreover, there is moderate tension between the predicted and observed RGB colors at the ~ 0.05 mag level in $BP - RP$ and $V - I$, which may point to interesting model deficiencies (e.g., α_{MLT} variation; Bonaca et al. 2012; Tayar et al. 2017; but see also Choi et al. 2018; Salaris et al. 2018) and/or poorly “fit” cluster parameters, but we do not conclusively attribute the discrepancy to any one source at this time.

The left panel in the middle row compares the MIST mass–radius relation with two independent mass and radius determinations of HV Cnc. HV Cnc was initially reported to be a single-lined binary (Mathieu et al. 1990), but detections of a weak secondary and a possible tertiary component were reported in subsequent works (Melo et al. 2001; Sandquist & Shetrone 2003). Sandquist & Shetrone (2003) analyzed VI photometry and radial velocity data and found a third nonbinary component in the spectra, though its association with the HV Cnc system is still uncertain. They deconvolved the photometry of the three stars to yield the parameters of the two binary components, shown in yellow. The primary is hotter than the majority of the cluster MSTO stars, which suggests that it is either a blue straggler or undergoing the overall

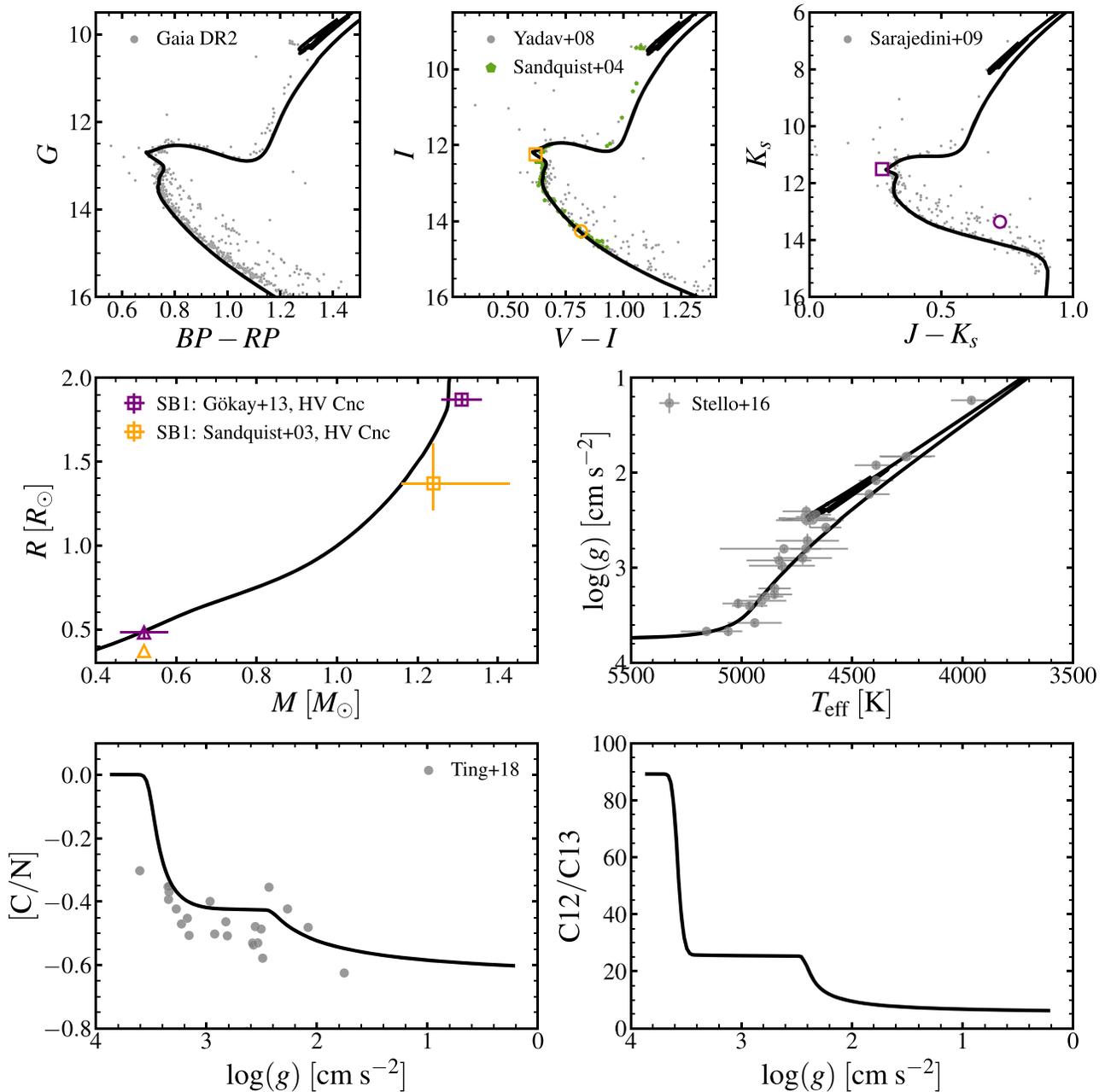


Figure 9. Same as Figure 8, except now for M67, a solar-metallicity and solar-age (4 Gyr) open cluster. The top left panel shows photometry of likely cluster members selected using the *Gaia* DR2 data, while the middle and right panels show stars with $>50\%$ membership probabilities as determined by Yadav et al. (2008). The Sandquist (2004) sample contains likely single-star members selected based on their proper motions. The MIST CMDs, shown in black, adopt $m - M = 9.73$, $[\text{Fe}/\text{H}] = -0.01$, $\log(\text{Age})[\text{yr}] = 9.58$, and $A_V = 0.18$ (see text for more details). The left panel in the middle row shows two measurements of the stellar parameters of HV Cnc (Sandquist & Shetrone 2003; Gökay et al. 2013) with the MIST mass–radius relation. The open symbols indicate that the system is a single-lined (SB1) spectroscopic binary system. When available, the individual EB components are also shown in the CMDs, where the squares, triangles, and circles correspond to the primary, secondary, and tertiary components, respectively. The right panel shows the asteroseismic $\log g$ inferred from the scaling relations (Stello et al. 2016), with the effective temperatures from the Casagrande & Vandenberg (2014) color–temperature relations. The surface abundances are obtained by reanalyzing the publicly available APOGEE DR14 spectra (Holtzman et al. 2015; SDSS Collaboration et al. 2017) with *the Payne* (Ting et al. 2018). The measurement of $^{12}\text{C}/^{13}\text{C}$ in the right panel is currently being explored.

contraction phase along the Henyey hook. Gökay et al. (2013) provided an updated set of parameters by adding in the JHK_s photometry, confirming the spectroscopic detection of a third component. They combined the radial velocity solution of the primary with the mass ratio inferred from the light curves in order to obtain the full solution of the binary system, shown in purple points. Again, we use open symbols to indicate that HV Cnc is a single-lined binary system. When available, the

individual EB components are also shown in the CMDs, where the squares, triangles, and circles correspond to the primary, secondary, and tertiary components, respectively.

The right panel shows the asteroseismic $\log g$ from the analysis of the *K2* photometry (Stello et al. 2016). The authors computed T_{eff} using the optical and Two Micron All Sky Survey (2MASS) photometry with the color–temperature relations from Casagrande et al. (2010), assuming $[\text{Fe}/\text{H}] = 0$

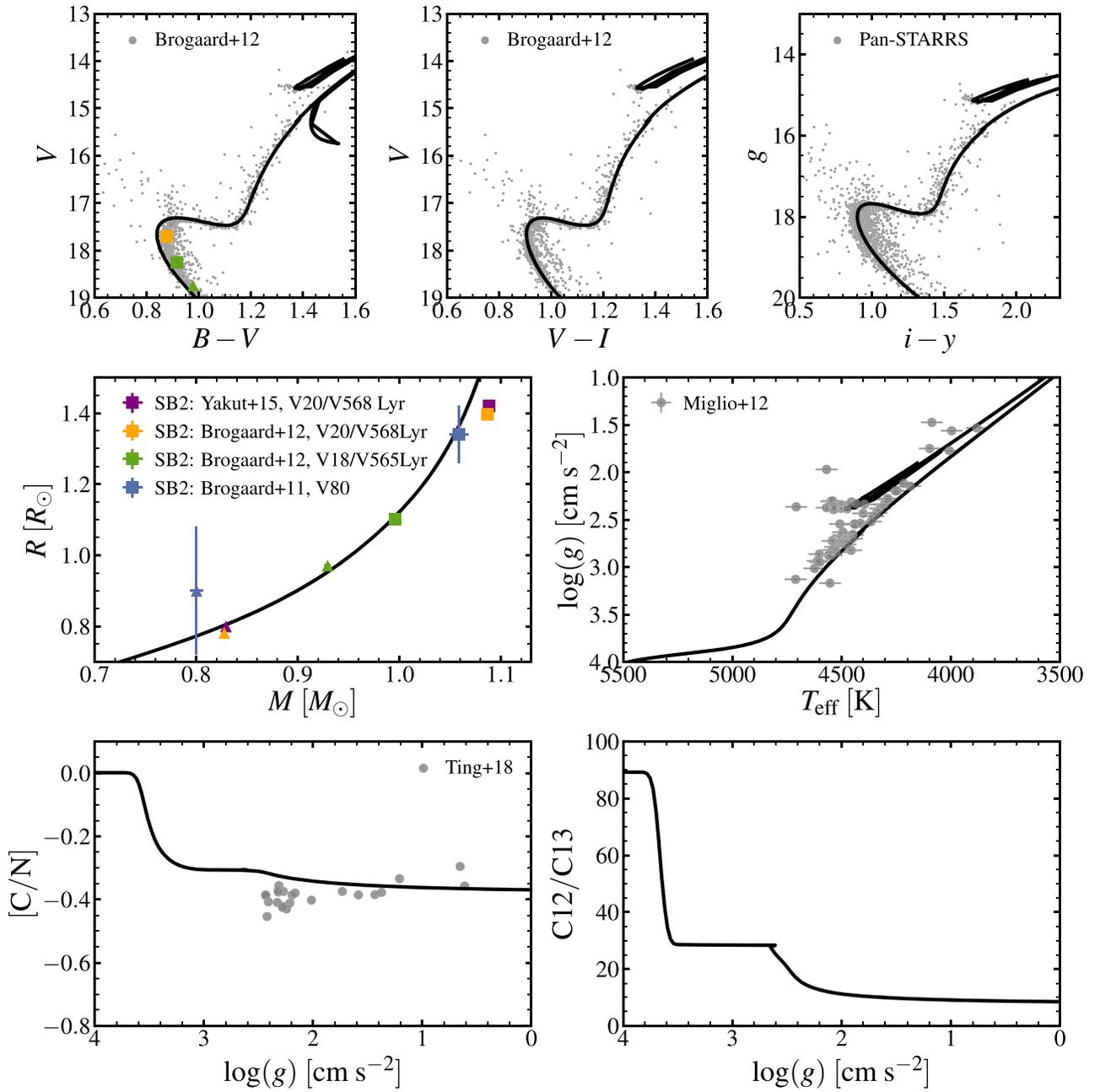


Figure 10. Same as Figure 8, except now for NGC 6791, a metal-rich ($[\text{Fe}/\text{H}] \approx +0.3$) and old (8 Gyr) open cluster. The top left panel shows photometry of likely cluster members selected using the *Gaia* DR2 data. The MIST CMDs assume $m - M = 13.05$, $[\text{Fe}/\text{H}] = +0.28$, $\log(\text{Age})[\text{yr}] = 9.95$, and $A_V = 0.434$ and show interesting tension with the data (see the text for more details). The left panel in the middle row compares the mass and radius measurements of three DEB systems (Brogaard et al. 2011, 2012; Yakut et al. 2015) with the MIST mass–radius relation. The filled symbols indicate that the system is a double-lined (SB2) spectroscopic binary system. When available, the individual EB components are also shown in the CMDs, where the squares and triangles correspond to the primary and secondary, respectively. The right panel shows the asteroseismic $\log g$ inferred from the scaling relations (Miglio et al. 2012), with the effective temperatures from the Ramírez & Meléndez (2005) color–temperature relations. The surface abundances are obtained by reanalyzing the publicly available APOGEE DR14 spectra (Holtzman et al. 2015; SDSS Collaboration et al. 2017) with *the Payne* (Ting et al. 2018). The measurement of $^{12}\text{C}/^{13}\text{C}$ in the right panel is currently being explored.

and $E(B - V) = 0.03$. Their T_{eff} uncertainties were estimated by the scatter in the T_{eff} inferred from different combinations of the photometric systems, plus an additional 20 K to account for the T_{eff} zero-point uncertainty. Overall, the T_{eff} and $\log g$ are in excellent agreement with the model predictions including the RC magnitude, albeit the T_{eff} uncertainties are quite large.

Finally, we plot the $[\text{C}/\text{N}]$ and $^{12}\text{C}/^{13}\text{C}$ surface abundance evolution on the RGB. *The Payne* $[\text{C}/\text{N}]$ abundances are shown as gray circles for comparison. The spectroscopic $[\text{C}/\text{N}]$

abundances fall ≈ 0.1 dex below the MIST prediction, suggesting a weak preference for a younger age and/or higher metallicity (see Figure 1). Measurement of $^{12}\text{C}/^{13}\text{C}$ from the APOGEE spectra is currently being explored.

4.4. NGC 6791

NGC 6791 is an exceptionally old (8 Gyr) and metal-rich ($[\text{Fe}/\text{H}] \approx 0.3\text{--}0.5$) open cluster (e.g., Stetson et al. 2003; Gratton et al. 2004; Carney et al. 2005; King et al. 2005;

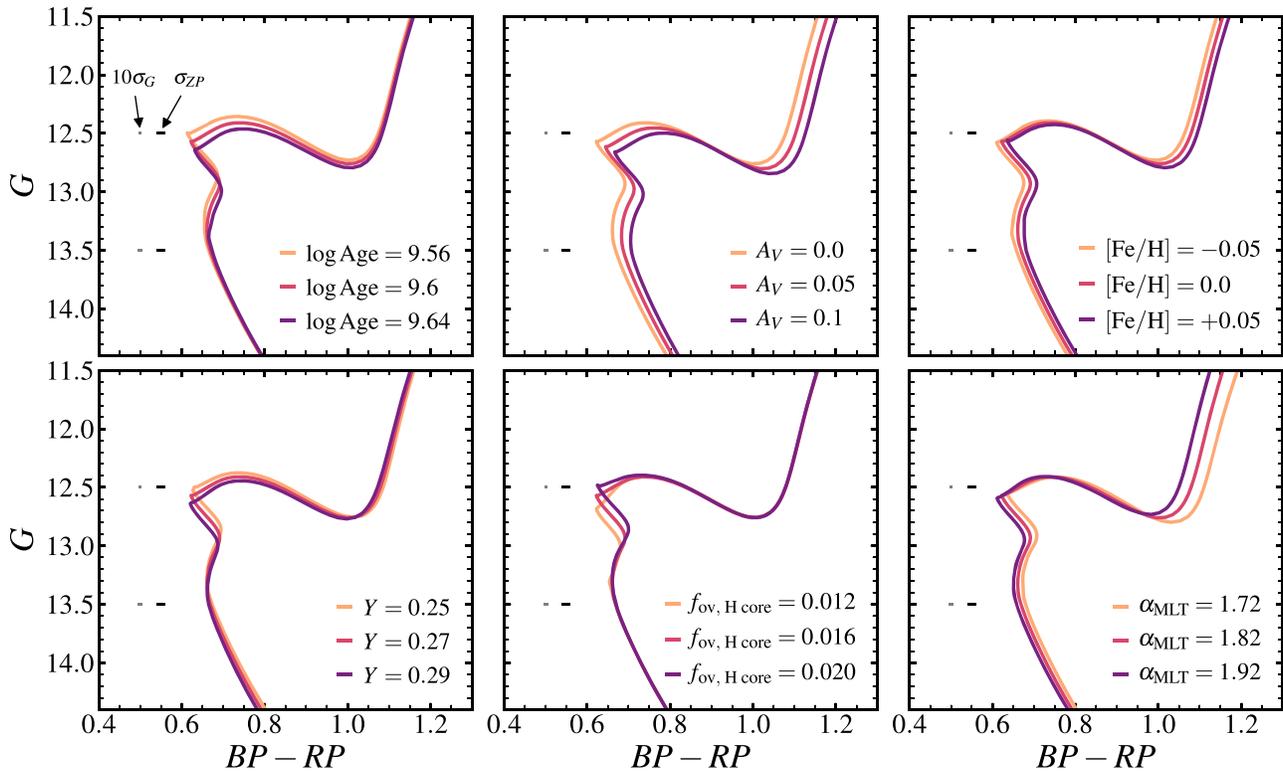


Figure 11. MIST isochrones that illustrate the effects of uncertain parameters on various parts of the *Gaia* CMD. A distance modulus of $m - M = 9.7$ is applied to all models, and extinction is not included unless noted otherwise. The $\log \text{Age}$, $[\text{Fe}/\text{H}]$, initial helium abundance, convective overshoot efficiency in the hydrogen-burning core, and mixing length α are held fixed to 9.6, 0.0, 0.2703, 0.016, and 1.82 unless noted otherwise. While these parameters indeed have only a subtle influence on the CMD morphology, they change the CMD in unique ways and thus should be separable with high-quality models, data, and fitting tools. What these panels do not explicitly show is the effect of these parameters on the lifetimes. The representative *Gaia* end-of-mission median-stray-light photometric standard errors assuming 70 visits per field are shown in gray. For display purposes, we multiply the uncertainties in each band by a factor of 10. We also show the absolute photometric accuracy due to zero-point uncertainty (≈ 0.014 mag) in black. Top left: varying the stellar age. Top middle: varying the amount of reddening assuming the $R_V = 3.1$ reddening law from Cardelli et al. (1989). Top right: varying the metallicity. Bottom left: varying the initial helium abundance. Bottom middle: varying the convective overshoot mixing in the hydrogen core. Bottom right: varying the mixing length α_{MLT} .

Origlia et al. 2006; Linden et al. 2017). It is also well known for its puzzling double white dwarf (WD) cooling sequence, both of which imply cluster ages that are nominally inconsistent with the MSTO age (Bedin et al. 2005, 2008a; but see also García-Berro et al. 2010). Several explanations have been put forth, including the presence of a secondary population of massive helium WDs (Hansen 2005; Kalirai et al. 2007) and WD binaries (Bedin et al. 2008b).

Figure 10 shows the multipanel plot summary of NGC 6791. The left and middle panels show *BVI* photometry from Brogaard et al. (2012), which includes differential reddening corrections to the Stetson et al. (2003) photometry, while the right panel shows Pan-STARRS photometry (Chambers et al. 2016; Flewelling et al. 2016). All three panels show likely members selected by the *Gaia* DR2 proper-motion and parallax data. We apply additional quality cuts in the Brogaard et al. (2012) data, restricting the sample to those with $\sigma_V < 10$ mmag. We do not show the *Gaia* DR2 photometry because it is too shallow and shows a much larger scatter. The reported photometric uncertainties are $\lesssim 1$ –10 mmag and 6 mmag in V and g_{PS} near the MSTO, respectively.¹² As before, the metallicity is fixed to *the Payne* measurement. The distance modulus estimated from the *Gaia* DR2 parallaxes is uncertain,

which is unsurprising given the large distance (~ 4 kpc). If we compute the distance modulus by adopting the 10% precision cut (21 stars total) and calculating the median parallax as described in Section 4.1, we obtain $m - M = 13.29$. However, if we relax the precision cut and/or calculate the mean or the weighted median rather than the median, the resulting distance modulus can range from ~ 13.03 to ~ 13.45 . This is expected to improve with more data over the next several years, but for now, we start with $m - M = 13.29$ and allow flexibility in the final adopted value. In summary, the MIST models, shown in black, assume $m - M = 13.05$ (requiring a zero-point offset of 0.025 mas), $[\text{Fe}/\text{H}] = +0.28$, $\log(\text{Age})[\text{yr}] = 9.95$, and $A_V = 0.434$. Interestingly, these models cannot simultaneously fit the MSTO in all three CMDs, and they also show moderate discrepancy in the RGB color. A simple increase in the reddening value improves the agreement in $B - V$, but not without introducing tension in $V - I$ and $i - y$. An increase in metallicity to $+0.4$ dex results in an overall improved fit at the MSTO in the three panels, but at the cost of increasing tension on the RGB, not to mention introducing inconsistency with *the Payne* metallicity.

In the middle row, we compare the mass–radius measurements of three systems, V18, V20, and V80, with the MIST model predictions. Brogaard et al. (2012) updated the analysis for V18 and V20 from Brogaard et al. (2011) using a new photometric reduction procedure and an improved analysis of the V20 secondary: the contribution to the light curve from the

¹² The photometric errors in V near $V \approx 18$ as reported by Brogaard et al. (2012) show a bimodal distribution with peaks at $\sigma_V \approx 0.3$ and 5 mmag, respectively, after we limit the sample to $\sigma_V < 10$ mmag. The MS is not much different when we restrict the sample to $\sigma_V < 1$ mmag.

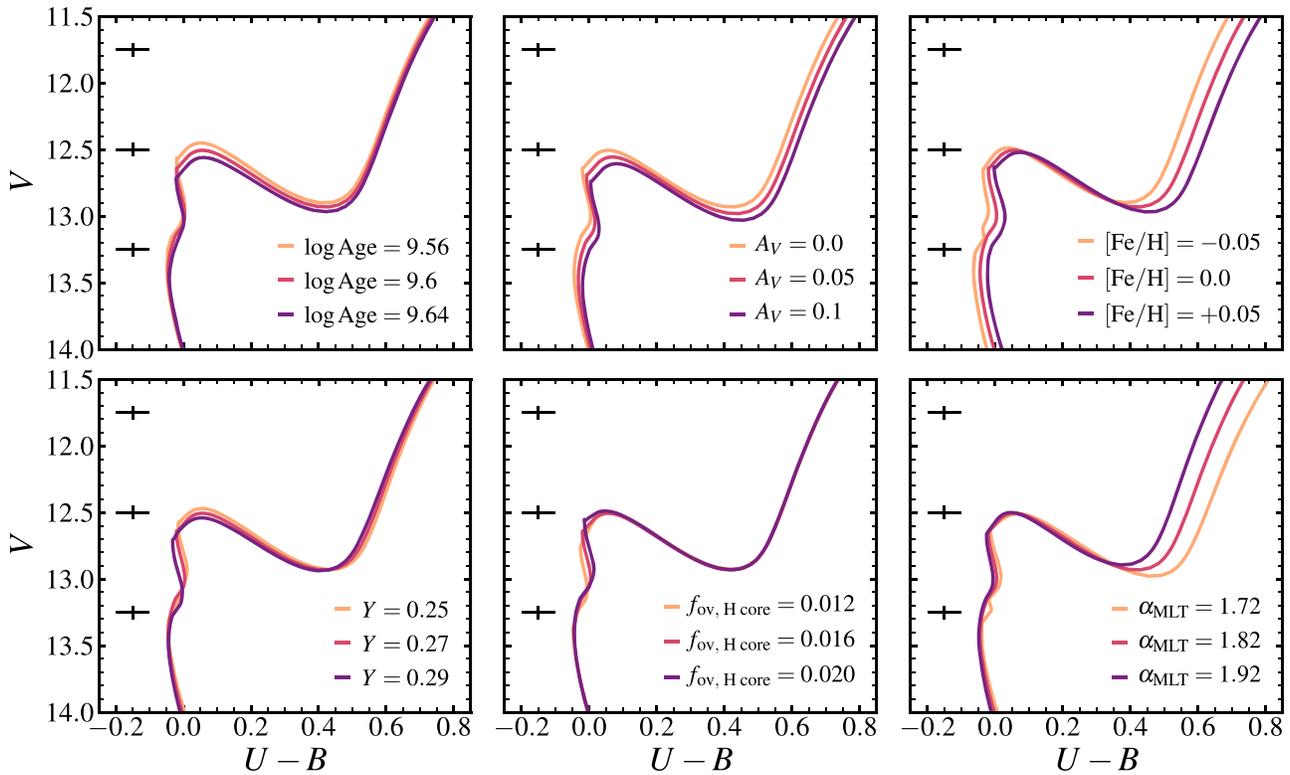


Figure 12. Same as Figure 11, except now showing V vs. $U - B$. We also show error bars, 0.05 mag in both color and magnitude, to help guide the eye.

third component in V20 was accounted for using four “twin stars” that were identified in a much higher resolution *HST*/ACS image. According to Brogaard et al. (2011), photometric and radius measurements of V80 are very uncertain owing to magnetic activity possibly induced by its close-in orbit. Nevertheless, all three systems are double-lined binaries: there are mass measurements for all three systems and radius measurements for V18 and V20 to within 1%. Additionally, there is an updated measurement for V20 from Yakut et al. (2015), where the authors utilized very precise *Kepler* light curves to obtain more accurate estimates of the stellar parameters. The most massive EB shows tension with the isochrone, requiring either a younger age at fixed $[\text{Fe}/\text{H}]$ or a higher $[\text{Fe}/\text{H}]$ at fixed age to resolve the discrepancy. When available, the individual EB components are also shown in the CMDs, where the squares and triangles correspond to the primary and secondary, respectively.

In the next panel, we plot the asteroseismic $\log g$ measurements from Miglio et al. (2012). The authors adopted effective temperatures calculated from the Ramírez & Meléndez (2005) $V - K$ color–temperature relation assuming $[\text{Fe}/\text{H}] = +0.3$ and $E(B - V) = 0.16 \pm 0.02$. Following Hekker et al. (2011), they assume an uncertainty of 50 K, though they caution that systematic uncertainties due to color–temperature calibrations and reddening could result in a number closer to ~ 110 K. We adopt 50 K for computing uncertainties in $\log g$. There is a mild offset of $\lesssim 50$ K between the predicted and “observed” temperatures, though note that the latter was estimated using a reddening value that is larger than what is adopted for the CMDs.

In the last row, we plot the $[\text{C}/\text{N}]$ and $^{12}\text{C}/^{13}\text{C}$ surface abundance evolution on the RGB. For comparison, the *Payne* $[\text{C}/\text{N}]$ abundances are shown as gray circles. Finally, we plot the $[\text{C}/\text{N}]$ and $^{12}\text{C}/^{13}\text{C}$ surface abundance evolution on the RGB.

The *Payne* $[\text{C}/\text{N}]$ abundances are shown as gray circles for comparison, which are generally in good agreement with the MIST prediction. Measurement of $^{12}\text{C}/^{13}\text{C}$ from the APOGEE spectra is currently being explored.

5. What We Can Expect from *Gaia* in the Future

The *Gaia* mission was designed to obtain μas astrometry and proper motions for a billion Milky Way stars along with high-precision photometry consisting of both broadband G and blue/red (B/R) spectrophotometry (Jordi et al. 2010). While this work already benefited greatly from *Gaia* DR2, especially in cluster membership, there should be notable improvement with future data releases in, e.g., zero-points and uncertainties in the astrometric solutions. In this section, we illustrate what we might expect from end-of-mission *Gaia* data using M67 as a fiducial case.

In Figure 11, we show example CMDs representative of M67, where each of the six panels shows a series of MIST models illustrating the effects of uncertain parameters. The gray error bars represent the end-of-mission (assuming 70 visits to each field) photometric standard errors estimated according to a performance model made available by the *Gaia* mission.¹³ For display purposes, we inflate the errors in each band by a factor of 10. We emphasize that these errors are representative of *relative* photometric precision only, because the *absolute* photometric accuracy is still dominated by the photometric zero-point uncertainty. The photometric zero-point measurement is tied to the 1% calibration of Vega’s spectra (see Carrasco et al. 2016; see also the discussion in Section 2.1), which ultimately yields ≈ 0.014 mag absolute photometric uncertainty in color, shown as black error bars (0.01 mag added

¹³ <https://www.cosmos.esa.int/web/gaia/science-performance>

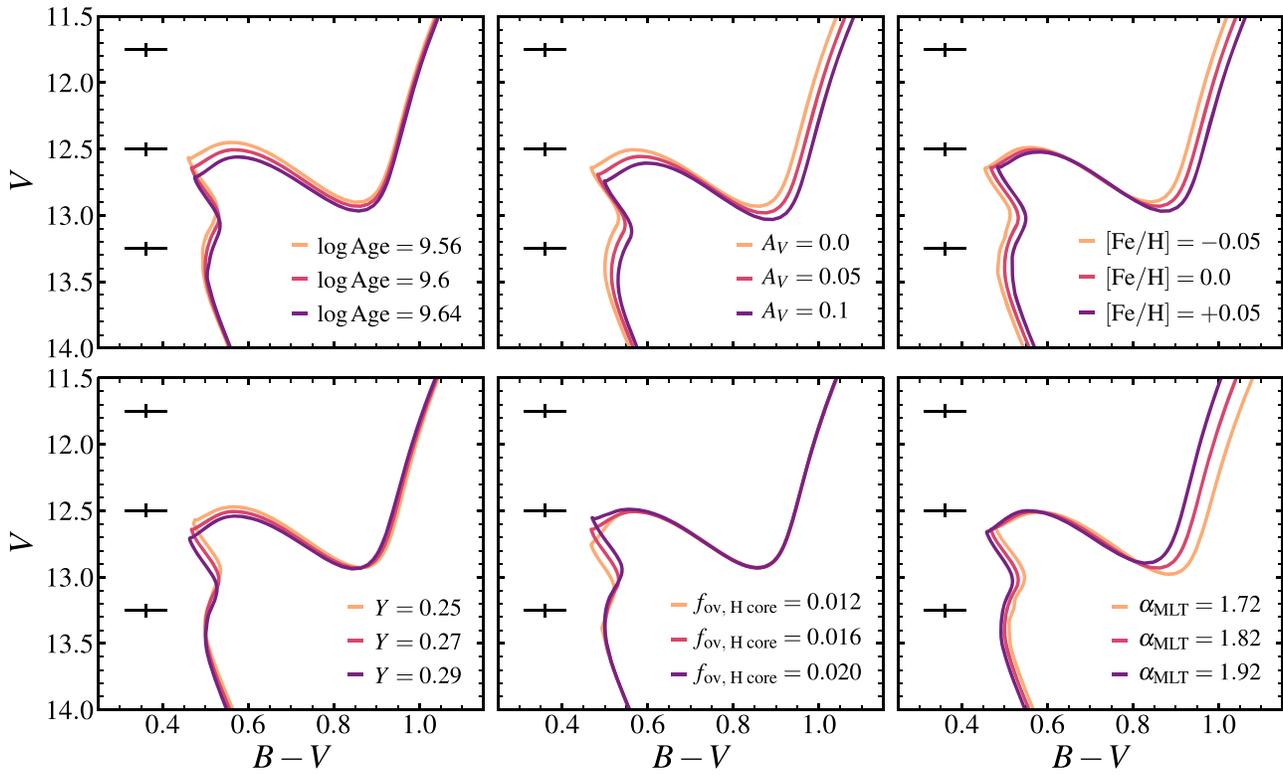


Figure 13. Same as Figure 12, except now showing V vs. $B - V$.

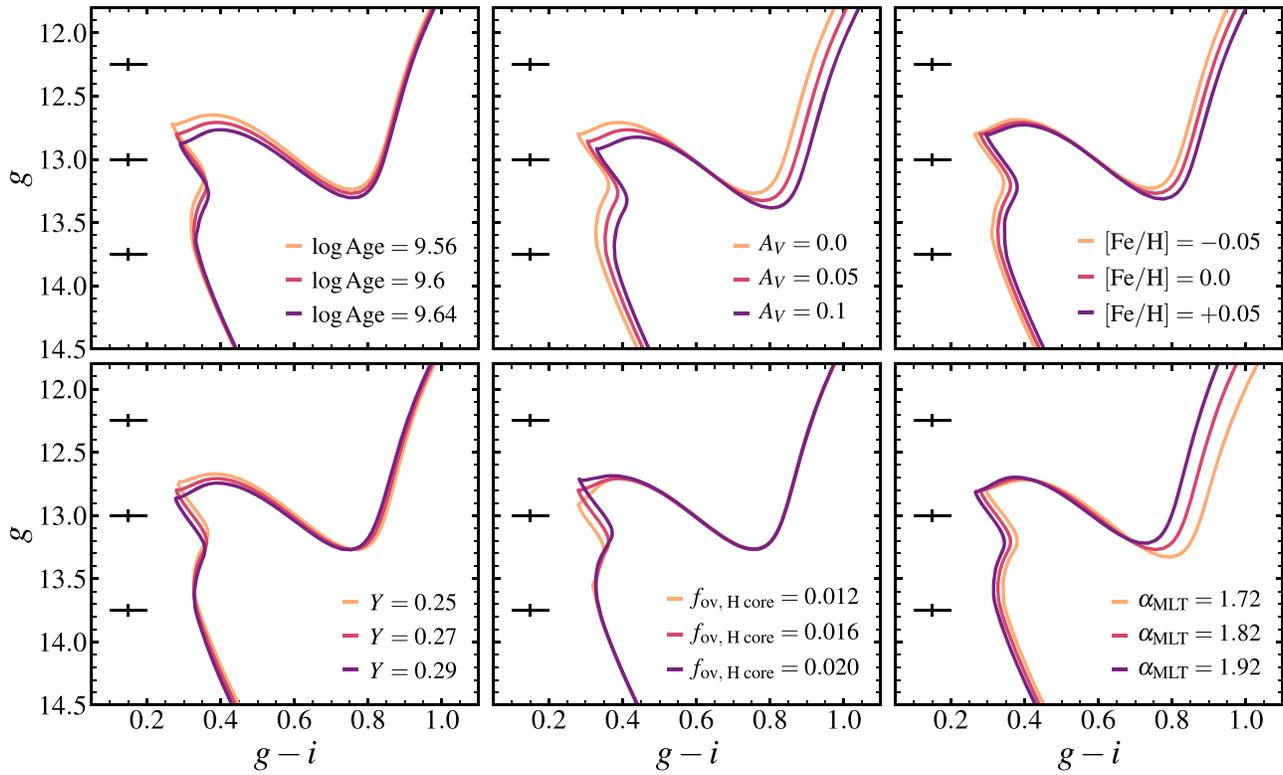


Figure 14. Same as Figure 12, except now showing Pan-STARRS g vs. $g - i$.

in quadrature). However, these uncertain model parameters change the CMD morphology in qualitatively distinct ways and thus should be separable with high-quality models, data, and

fitting tools even in the presence of zero-point uncertainties. What these panels do not explicitly show is the effect of these parameters on the lifetimes—this information is encoded in the

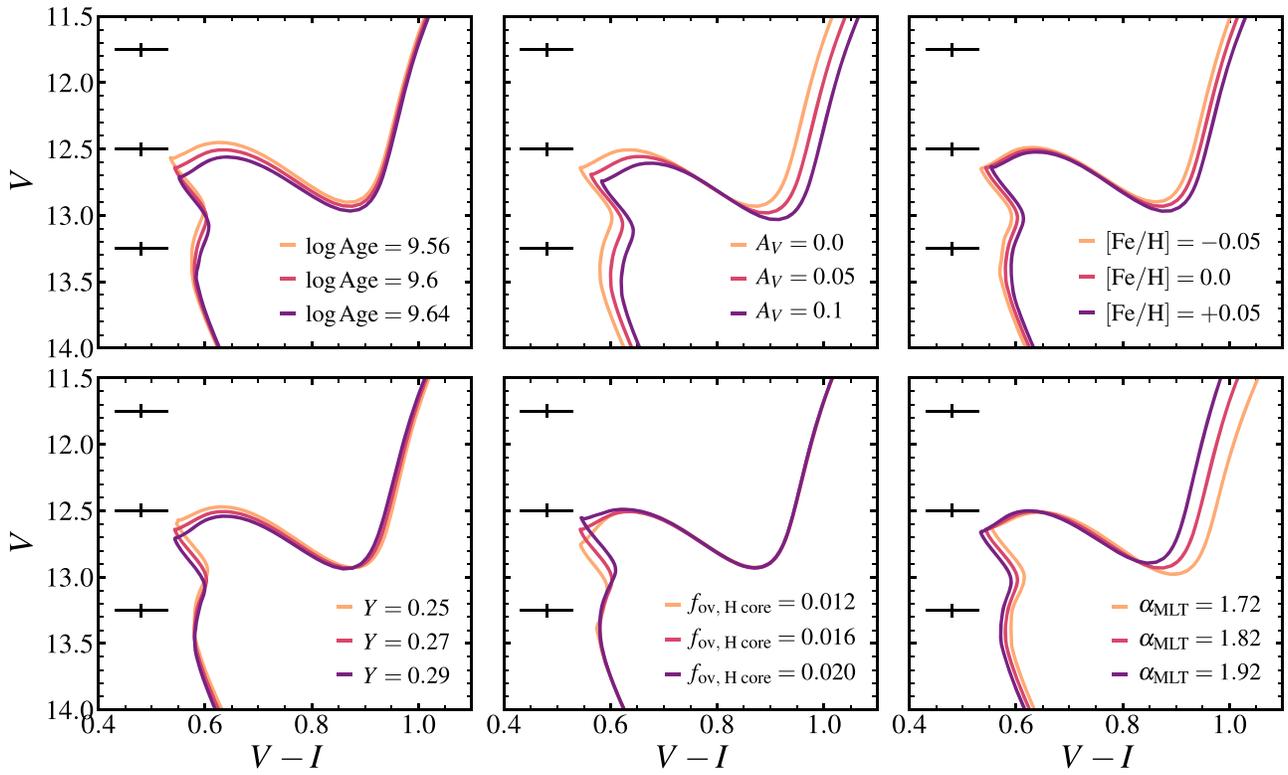


Figure 15. Same as Figure 12, except now showing V vs. $V - I$.

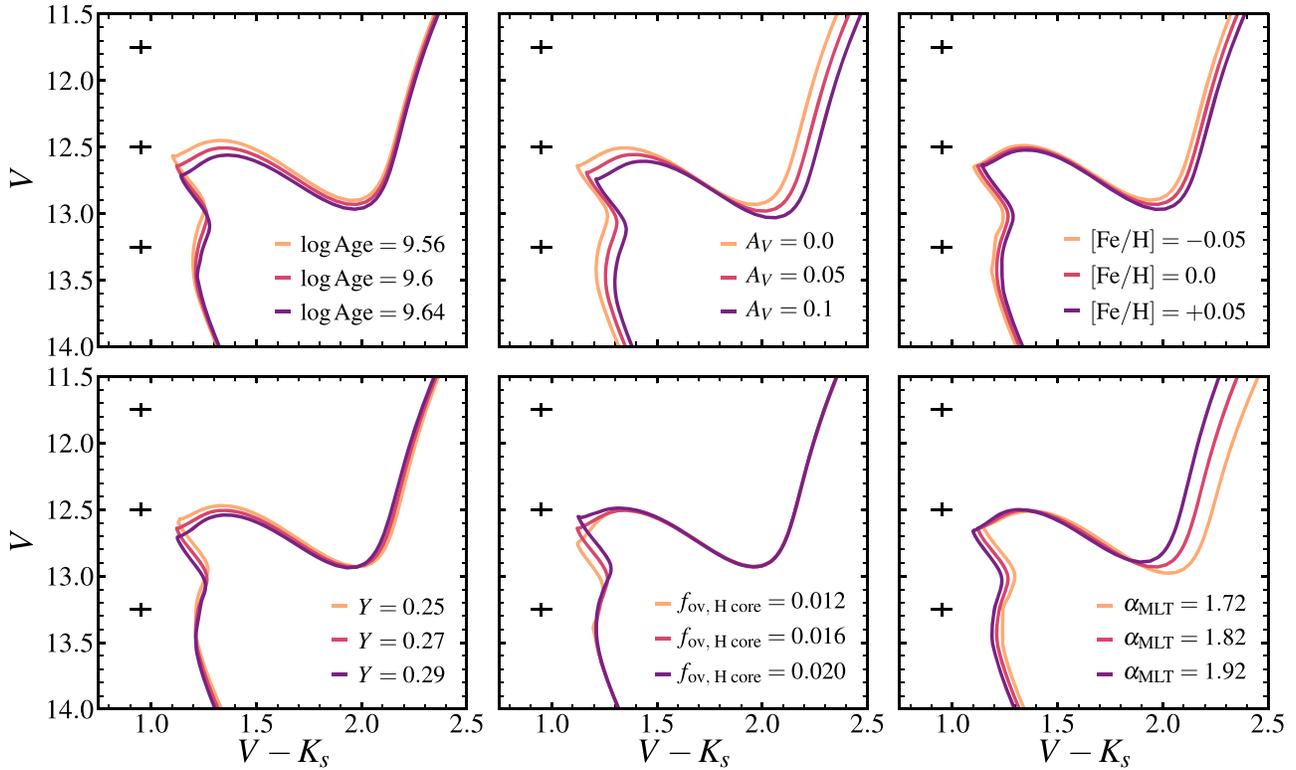


Figure 16. Same as Figure 12, except now showing V vs. $V - K_s$.

density distribution or number ratios of stars in various parts of the CMD. The subtle differences illustrated in Figure 11 are difficult to distinguish using traditional techniques, e.g., fitting empirical MS ridelines, but we will soon be able to leverage

the exquisite *Gaia* photometry, proper motions, and parallax, in combination with a diverse data set including spectroscopy and asteroseismology. These data sets will place very stringent constraints on the models in Figures 8, 9, and 10,

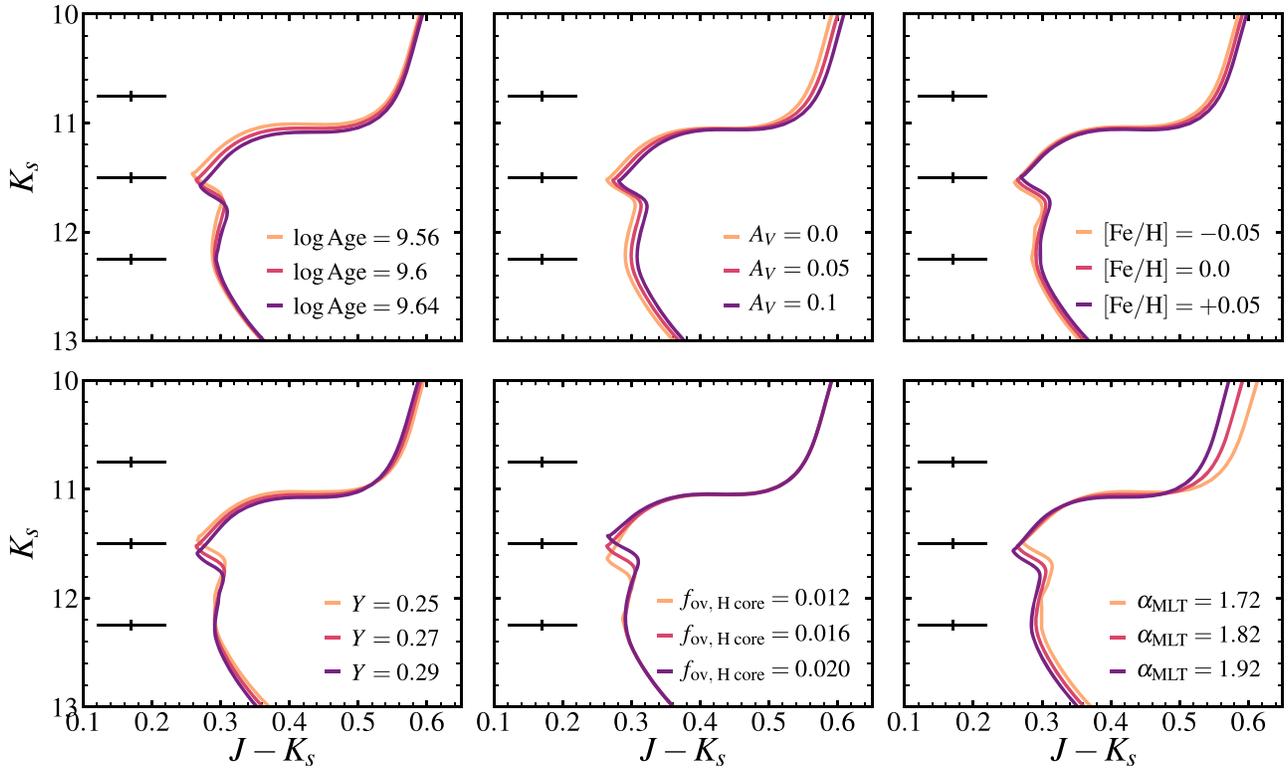


Figure 17. Same as Figure 12, except now showing K_s vs. $J - K_s$.

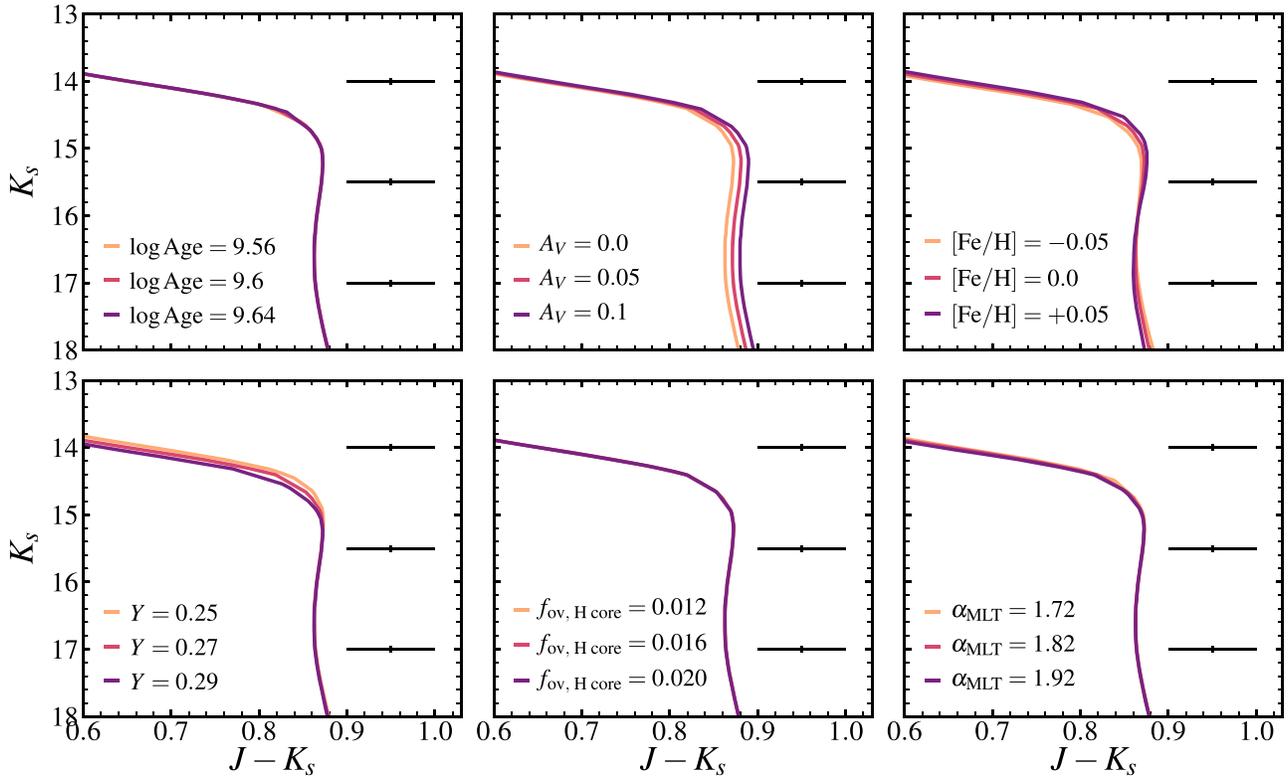


Figure 18. Same as Figure 17, except now zoomed in around the kink on the lower MS.

which show isochrones that were “fit” by eye for illustrative purposes. A future direction in this area includes a quantitative and objective determination of the best likelihood

parameters (e.g., MATCH, Dolphin 2002; BASE9, von Hippel et al. 2006; MINESweeper, P. Cargile et al. 2018, in preparation).

6. Summary

In this work, we provided an overview of the currently available and future data sets that can be leveraged simultaneously to both improve our constraints on the uncertain stellar model parameters and infer the properties of open clusters. We first explored the effects of key parameters—age, metallicity, helium content, mixing-length parameter, convective boundary mixing efficiency in hydrogen and helium cores, thermohaline-mixing efficiency, and mass-loss efficiency—on the various observational diagnostics. Next, we identified pairs of observables that are sensitive to each parameter of interest and stellar age, taking into account the observational feasibility. The key plot that summarizes the results is shown in Figure 7.

There are several important caveats. At this level of scrutiny, photometric/parallax zero-points and differential reddening (see the discussion in Sections 3.9 and 4.1) may well dominate the observational uncertainties. However, the zero-points induce an overall shift in the CMD, while the key parameters considered in this work shape the CMD morphologies in qualitatively distinct ways, and thus the two types should be separable. On the theoretical modeling side, a proper treatment of the detailed abundance patterns (see, e.g., Dotter et al. 2015 where the authors analyzed NGC 6752 using self-consistent stellar interior and atmosphere models computed according to the detailed spectroscopic abundances), the effects of atomic diffusion on the surface abundances (Dotter et al. 2017), and the surface boundary conditions (e.g., Salaris & Cassisi 1996; Chabrier & Baraffe 1997; VandenBerg et al. 2008; Choi et al. 2018) will likely be important.

We also evaluated the current status of the various observational data sets using three well-studied open clusters—NGC 6819, M67, and NGC 6791—as case studies. Although we find no obvious discrepancies between the existing data and the MIST models for NGC 6819 (Figure 8), M67 shows a mild tension in the RGB colors (Figure 9). NGC 6791 appears to prefer a slightly higher metallicity than what is inferred from the *Payne* analysis of the APOGEE spectra, though this would lead to a tension in the RGB colors (Figure 10). More precise observations (e.g., parallax-based distance for NGC 6791) and robust fitting will help to conclusively identify and quantify the discrepancies. *Gaia* parallax measurements, with careful modeling of the zero-point offsets, should remove distance as a source of uncertainty, and the accompanying *Gaia* photometry (*B*, *R*, and *G*; see Figure 11) and proper-motion memberships will immensely improve the quality of the CMDs, as already demonstrated with the DR2 data. CMDs contain a tremendous amount of information, and thus the combination of exquisite photometry, flexible and robust stellar models, and objective fitting tools will allow us to measure stellar ages and disentangle the effects of key stellar model parameters in the near future.

We thank the anonymous referee for their feedback, which greatly improved the quality of this manuscript. We also thank Lars Bildsten, Harshil Kamdar, Guillermo Torres, and Dan Weisz for helpful discussions and comments. Finally, we thank Bill Paxton and the rest of the MESA community, who have made this project possible.

Appendix

Here we present a series of figures illustrating the effects of uncertain parameters—logAge, extinction, [Fe/H], initial helium abundance, convective overshoot efficiency in the hydrogen-burning core, and mixing length α —on various parts of the CMD. Figures 13–18 are analogous to Figure 11 and show qualitatively the same behavior, but they demonstrate that the sizes of the effects can vary significantly depending on the combination of filters. We also include error bars, 0.05 mag in both color and magnitude, to guide the eye in each panel and assist with the direct comparison of CMDs plotted on different axis scales. We caution the reader against overinterpreting the minor blemishes in these figures, e.g., the $\alpha_{\text{MLT}} = 1.72$ curve in the bottom right panel of Figure 12.

ORCID iDs

Jieun Choi  <https://orcid.org/0000-0002-8822-1355>
 Charlie Conroy  <https://orcid.org/0000-0002-1590-8551>
 Yuan-Sen Ting  <https://orcid.org/0000-0001-5082-9536>
 Aaron Dotter  <https://orcid.org/0000-0002-4442-5700>

References

- Ak, T., Bostancı, Z. F., Yontan, T., et al. 2016, *Ap&SS*, 361, 126
 Alonso, A., Arribas, S., & Martínez-Roger, C. 1996, *A&A*, 313, 873
 Andersen, J. 1991, *A&ARv*, 3, 91
 Angelou, G. C., Bellinger, E. P., Hekker, S., & Basu, S. 2017, *ApJ*, 839, 116
 Angelou, G. C., D’Orazi, V., Constantino, T. N., et al. 2015, *MNRAS*, 450, 2423
 Anthony-Twarog, B. J., Deliyannis, C. P., & Twarog, B. A. 2014, *AJ*, 148, 51
 Arenou, F., Luri, X., Babusiaux, C., et al. 2018, arXiv:1804.09375
 Arentoft, T., Brogaard, K., Jessen-Hansen, J., et al. 2017, *ApJ*, 838, 115
 Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, 47, 481
 Baglin, A., Auvergne, M., Barge, P., et al. 2006, in ESA Special Publication 1306, The CoRoT Mission Pre-Launch Status—Stellar Seismology and Planet Finding, ed. M. Fridlund et al. (Paris: ESA), 33
 Bastian, N., & Lardo, C. 2017, arXiv:1712.01286
 Basu, S., & Antia, H. M. 2004, *ApJL*, 606, L85
 Bedding, T. R., Huber, D., Stello, D., et al. 2010, *ApJL*, 713, L176
 Bedin, L. R., King, I. R., Anderson, J., et al. 2008a, *ApJ*, 678, 1279
 Bedin, L. R., Piotto, G., Anderson, J., et al. 2004, *ApJL*, 605, L125
 Bedin, L. R., Salaris, M., Piotto, G., et al. 2005, *ApJL*, 624, L45
 Bedin, L. R., Salaris, M., Piotto, G., et al. 2008b, *ApJL*, 679, L29
 Boesgaard, A. M., Lum, M. G., & Deliyannis, C. P. 2015, *ApJ*, 799, 202
 Bohlin, R. C. 2012, Flux Calibration of the ACS CCD Cameras IV. Absolute Fluxes, STScI, Tech. Rep., ACS 2012-01
 Bohlin, R. C. 2014, *AJ*, 147, 127
 Böhm, K.-H. 1963, *ApJ*, 138, 297
 Böhm-Vitense, E. 1958, *ZAp*, 46, 108
 Bonaca, A., Tanner, J. D., Basu, S., et al. 2012, *ApJL*, 755, L12
 Bond, H. E., Nelan, E. P., VandenBerg, D. A., Schaefer, G. H., & Harmer, D. 2013, *ApJL*, 765, L12
 Bovy, J., Rix, H.-W., Liu, C., et al. 2012, *ApJ*, 753, 148
 Boyajian, T. S., McAlister, H. A., van Belle, G., et al. 2012a, *ApJ*, 746, 101
 Boyajian, T. S., von Braun, K., van Belle, G., et al. 2012b, *ApJ*, 757, 112
 Boyajian, T. S., von Braun, K., van Belle, G., et al. 2013, *ApJ*, 771, 40
 Bragaglia, A., Sneden, C., Carretta, E., et al. 2014, *ApJ*, 796, 68
 Bressan, A., Marigo, P., Girardi, L., et al. 2012, *MNRAS*, 427, 127
 Brewer, L. N., Sandquist, E. L., Mathieu, R. D., et al. 2016, *AJ*, 151, 66
 Brogaard, K., Bruntt, H., Grundahl, F., et al. 2011, *A&A*, 525, A2
 Brogaard, K., VandenBerg, D. A., Bruntt, H., et al. 2012, *A&A*, 543, A106
 Campello, R. J. G. B., Moulavi, D., & Sander, J. 2013, in *Advances in Knowledge Discovery and Data Mining*, ed. J. Pei et al. (Berlin: Springer), 160
 Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, 345, 245
 Carney, B. W., Lee, J.-W., & Dodson, B. 2005, *AJ*, 129, 656
 Carraro, G., Villanova, S., Demarque, P., et al. 2006, *ApJ*, 643, 1151

- Carrasco, J. M., Evans, D. W., Montegriffo, P., et al. 2016, *A&A*, 595, A7
- Carretta, E., Bragaglia, A., & Gratton, R. G. 2007, *A&A*, 473, 129
- Casagrande, L., Ramírez, I., Meléndez, J., Bessell, M., & Asplund, M. 2010, *A&A*, 512, A54
- Casagrande, L., & VandenBerg, D. A. 2014, *MNRAS*, 444, 392
- Castellani, V., Chieffi, A., Tornambe, A., & Pulone, L. 1985, *ApJ*, 296, 204
- Catelan, M. 2009, *Ap&SS*, 320, 261
- Chabrier, G., & Baraffe, I. 1997, *A&A*, 327, 1039
- Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560
- Charbonnel, C. 1994, *A&A*, 282, 811
- Charbonnel, C., & Lagarde, N. 2010, *A&A*, 522, A10
- Charbonnel, C., & Zahn, J.-P. 2007, *A&A*, 467, L15
- 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
- Claret, A., & Torres, G. 2016, *A&A*, 592, A15
- Correnti, M., Gennaro, M., Kalirai, J. S., Brown, T. M., & Calamida, A. 2016, *ApJ*, 823, 18
- Corsaro, E., De Ridder, J., & García, R. A. 2015, *A&A*, 579, A83
- Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, 2MASS All Sky Catalog of point sources, The IRSA 2MASS All-Sky Point Source Catalog, NASA/IPAC Infrared Science Archive, <http://irsa.ipac.caltech.edu/applications/Gator/>
- Dalton, G., Trager, S. C., Abrams, D. C., et al. 2012, *Proc. SPIE*, 8446, 84460P
- de Jong, R. S., Barden, S. C., Bellido-Tirado, O., et al. 2016, *Proc. SPIE*, 9908, 99081O
- De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015, *MNRAS*, 449, 2604
- Denissenkov, P. A. 2010, *ApJ*, 723, 563
- Dolphin, A. E. 2002, *MNRAS*, 332, 91
- Dotter, A. 2008, *ApJL*, 687, L21
- Dotter, A. 2016, *ApJS*, 222, 8
- Dotter, A., Chaboyer, B., Ferguson, J. W., et al. 2007, *ApJ*, 666, 403
- Dotter, A., Conroy, C., Cargile, P., & Asplund, M. 2017, *ApJ*, 840, 99
- Dotter, A., Ferguson, J. W., Conroy, C., et al. 2015, *MNRAS*, 446, 1641
- Draine, B. T. 2003, *ARA&A*, 41, 241
- Dupree, A. K., & Avrett, E. H. 2013, *ApJL*, 773, L28
- Eggleton, P. P., Dearborn, D. S. P., & Lattanzio, J. C. 2006, *Sci*, 314, 1580
- Eggleton, P. P., Dearborn, D. S. P., & Lattanzio, J. C. 2008, *ApJ*, 677, 581
- Ekström, S., Georgy, C., Eggenberger, P., et al. 2012, *A&A*, 537, A146
- Evans, D. W., Riello, M., De Angeli, F., et al. 2018, arXiv:1804.09368
- Feiden, G. A., & Chaboyer, B. 2013, *ApJ*, 779, 183
- Fitzpatrick, E. L. 1999, *PASP*, 111, 63
- Flewelling, H. A., Magnier, E. A., Chambers, K. C., et al. 2016, arXiv:1612.05243
- Freeman, K., & Bland-Hawthorn, J. 2002, *ARA&A*, 40, 487
- Freytag, B., Ludwig, H.-G., & Steffen, M. 1996, *A&A*, 313, 497
- Gaia Collaboration, Babusiaux, C., van Leeuwen, F., et al. 2018a, arXiv:1804.09378
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018b, arXiv:1804.09365
- García-Berro, E., Torres, S., Althaus, L. G., et al. 2010, *Natur*, 465, 194
- Gaulme, P., McKeever, J., Jackiewicz, J., et al. 2016, *ApJ*, 832, 121
- Geisler, D., Villanova, S., Carraro, G., et al. 2012, *ApJL*, 756, L40
- Georgy, C., Ekström, S., Meynet, G., et al. 2012, *A&A*, 542, A29
- Gilliland, R. L., Brown, T. M., Christensen-Dalsgaard, J., et al. 2010, *PASP*, 122, 131
- Gilmore, G., Randich, S., Asplund, M., et al. 2012, *Msngr*, 147, 25
- Girardi, L., Bressan, A., Bertelli, G., & Chiosi, C. 2000, *A&AS*, 141, 371
- Gökay, G., Gürol, B., & Derman, E. 2013, *AJ*, 146, 123
- González Hernández, J. I., & Bonifacio, P. 2009, *A&A*, 497, 497
- Gratton, R., Bragaglia, A., Carretta, E., & Tosi, M. 2006, *ApJ*, 642, 462
- Gratton, R., Sneden, C., & Carretta, E. 2004, *ARA&A*, 42, 385
- Gratton, R. G., Carretta, E., & Bragaglia, A. 2012, *A&ARv*, 20, 50
- Gratton, R. G., Carretta, E., Bragaglia, A., Lucatello, S., & D'Orazi, V. 2010, *A&A*, 517, A81
- Gratton, R. G., Sneden, C., Carretta, E., & Bragaglia, A. 2000, *A&A*, 354, 169
- Handberg, R., Brogaard, K., Miglio, A., et al. 2017, *MNRAS*, 472, 979
- Hansen, B. M. S. 2005, *ApJ*, 635, 522
- Harmanec, P. 1988, *BAICz*, 39, 329
- Hekker, S., Basu, S., Stello, D., et al. 2011, *A&A*, 530, A100
- Hekker, S., Elsworth, Y., Mosser, B., et al. 2013, *A&A*, 556, A59
- Hendricks, B., Stetson, P. B., VandenBerg, D. A., & Dall'Ora, M. 2012, *AJ*, 144, 25
- Herwig, F. 2000, *A&A*, 360, 952
- Herwig, F., Pignatari, M., Woodward, P. R., et al. 2011, *ApJ*, 727, 89
- Hole, K. T., Geller, A. M., Mathieu, R. D., et al. 2009, *AJ*, 138, 159
- Holtzman, J. A., Shetrone, M., Johnson, J. A., et al. 2015, *AJ*, 150, 148
- Huber, D., Bedding, T. R., Stello, D., et al. 2011, *ApJ*, 743, 143
- Huber, D., Silva Aguirre, V., Matthews, J. M., et al. 2014, *ApJS*, 211, 2
- Jordi, C., Gebran, M., Carrasco, J. M., et al. 2010, *A&A*, 523, A48
- Juric, M. 2012, LSD: Large Survey Database Framework, Astrophysics Source Code Library, ascl:1209.003
- Kalirai, J. S., Bergeron, P., Hansen, B. M. S., et al. 2007, *ApJ*, 671, 748
- Kallinger, T., Mosser, B., Hekker, S., et al. 2010a, *A&A*, 522, A1
- Kallinger, T., Weiss, W. W., Barban, C., et al. 2010b, *A&A*, 509, A77
- Kaluzny, J., Thompson, I. B., Rozycka, M., et al. 2013, *AJ*, 145, 43
- Kharchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., & Scholz, R.-D. 2013, *A&A*, 558, A53
- King, I. R., Bedin, L. R., Piotto, G., Cassisi, S., & Anderson, J. 2005, *AJ*, 130, 626
- Kippenhahn, R., Ruschenplatt, G., & Thomas, H.-C. 1980, *A&A*, 91, 175
- Kirk, B., Conroy, K., Prša, A., et al. 2016, *AJ*, 151, 68
- Kochanek, C. S., Beacom, J. F., Kistler, M. D., et al. 2008, *ApJ*, 684, 1336
- Kraus, A. L., Tucker, R. A., Thompson, M. I., Craine, E. R., & Hillenbrand, L. A. 2011, *ApJ*, 728, 48
- Kroupa, P. 2001, *MNRAS*, 322, 231
- Kurucz, R. L. 1970, SAOSR, 309
- Kurucz, R. L. 1993, SYNTHES Spectrum Synthesis Programs and Line Data, Kurucz CD-ROM (Cambridge, MA: Smithsonian Astrophysical Observatory)
- Kurucz, R. L., & Avrett, E. H. 1981, SAOSR, 391
- Lagarde, N., Robin, A. C., Reylé, C., & Nasello, G. 2017, *A&A*, 601, A27
- Lee, Y.-W., Demarque, P., & Zinn, R. 1990, *ApJ*, 350, 155
- Lee-Brown, D. B., Anthony-Twarog, B. J., Deliyannis, C. P., Rich, E., & Twarog, B. A. 2015, *AJ*, 149, 121
- Lindgren, L., Hernandez, J., Bombrun, A., et al. 2018, arXiv:1804.09366
- Linden, S. T., Pryal, M., Hayes, C. R., et al. 2017, *ApJ*, 842, 49
- Luri, X., Brown, A. G. A., Sarro, L. M., et al. 2018, arXiv:1804.09376
- Magic, Z., Serenelli, A., Weiss, A., & Chaboyer, B. 2010, *ApJ*, 718, 1378
- Magic, Z., Weiss, A., & Asplund, M. 2015, *A&A*, 573, A89
- Martig, M., Fouesneau, M., Rix, H.-W., et al. 2016, *MNRAS*, 456, 3655
- Martig, M., Rix, H.-W., Silva Aguirre, V., et al. 2015, *MNRAS*, 451, 2230
- Masseron, T., & Gilmore, G. 2015, *MNRAS*, 453, 1855
- Mathieu, R. D., Latham, D. W., & Griffin, R. F. 1990, *AJ*, 100, 1859
- Mathur, S., Huber, D., Batalha, N. M., et al. 2017, *ApJS*, 229, 30
- Melo, C. H. F., Pasquini, L., & De Medeiros, J. R. 2001, *A&A*, 375, 851
- Miglio, A., Brogaard, K., Stello, D., et al. 2012, *MNRAS*, 419, 2077
- Miglio, A., Chaplin, W. J., Brogaard, K., et al. 2016, *MNRAS*, 461, 760
- Milliman, K. E., Mathieu, R. D., Geller, A. M., et al. 2014, *AJ*, 148, 38
- Milone, A. P., Marino, A. F., Bedin, L. R., et al. 2014, *MNRAS*, 439, 1588
- Montalbán, J., Miglio, A., Noels, A., et al. 2013, *ApJ*, 766, 118
- Muraveva, T., Delgado, H. E., Clementini, G., Sarro, L. M., & Garofalo, A. 2018, arXiv:1805.08742
- Ness, M., Hogg, D. W., Rix, H.-W., et al. 2016, *ApJ*, 823, 114
- Netopil, M., Paunzen, E., Heiter, U., & Soubiran, C. 2016, *A&A*, 585, A150
- O'Donnell, J. E. 1994, *ApJ*, 422, 158
- Önehag, A., Gustafsson, B., & Korn, A. 2014, *A&A*, 562, A102
- Origlia, L., Valenti, E., Rich, R. M., & Ferraro, F. R. 2006, *ApJ*, 646, 499
- Palacios, A., Charbonnel, C., Talon, S., & Siess, L. 2006, *A&A*, 453, 261
- 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
- Pietrinferni, A., Cassisi, S., Salaris, M., & Castellani, F. 2004, *ApJ*, 612, 168
- Piotto, G., Bedin, L. R., Anderson, J., et al. 2007, *ApJL*, 661, L53
- Popper, D. M. 1967, *ARA&A*, 5, 85
- Pulone, L., De Marchi, G., Paresce, F., & Allard, F. 1998, *ApJL*, 492, L41
- Ramírez, I., & Meléndez, J. 2005, *ApJ*, 626, 465
- Recio-Blanco, A., de Laverny, P., Allende Prieto, C., et al. 2016, *A&A*, 585, A93
- Reese, D. R., Chaplin, W. J., Davies, G. R., et al. 2016, *A&A*, 592, A14
- Reimers, D. 1975, in *Problems in Stellar Atmospheres and Envelopes*, ed. B. Baschek, W. H. Kegel, & G. Traving (New York: Springer), 229
- Renzini, A., & Fusi Pecci, F. 1988, *ARA&A*, 26, 199
- Rood, R. T. 1973, *ApJ*, 184, 815
- Ross, T. L., Holtzman, J. A., Anthony-Twarog, B. J., et al. 2014, *AJ*, 147, 4
- Russell, H. N. 1912, *ApJ*, 36, 133
- Salaris, M., & Cassisi, S. 1996, *A&A*, 305, 858
- Salaris, M., & Cassisi, S. 2015, *A&A*, 577, A60
- Salaris, M., Cassisi, S., & Pietrinferni, A. 2016, *A&A*, 590, A64
- Salaris, M., Cassisi, S., Schiavon, R. P., & Pietrinferni, A. 2018, *A&A*, 612, A68
- Sandquist, E. L. 2004, *MNRAS*, 347, 101
- Sandquist, E. L., Mathieu, R. D., Brogaard, K., et al. 2013, *ApJ*, 762, 58

- Sandquist, E. L., & Shetrone, M. D. 2003, *AJ*, 126, 2954
- Sarajedini, A., Dotter, A., & Kirkpatrick, A. 2009, *ApJ*, 698, 1872
- Schaller, G., Schaerer, D., Meynet, G., & Maeder, A. 1992, *A&AS*, 96, 269
- SDSS Collaboration, Albareti, F. D., Allende Prieto, C., et al. 2017, *ApJS*, 233, 25
- Sharma, S., Stello, D., Bland-Hawthorn, J., Huber, D., & Bedding, T. R. 2016, *ApJ*, 822, 15
- Shaviv, G., & Salpeter, E. E. 1973, *ApJ*, 184, 191
- Smartt, S. J. 2009, *ARA&A*, 47, 63
- Smiljanic, R., Korn, A. J., Bergemann, M., et al. 2014, *A&A*, 570, A122
- Stassun, K. G., & Torres, G. 2018, arXiv:1805.03526
- Steinmetz, M., Zwitter, T., Siebert, A., et al. 2006, *AJ*, 132, 1645
- Stello, D., Bruntt, H., Preston, H., & Buzasi, D. 2008, *ApJL*, 674, L53
- Stello, D., Huber, D., Bedding, T. R., et al. 2013, *ApJL*, 765, L41
- Stello, D., Vanderburg, A., Casagrande, L., et al. 2016, *ApJ*, 832, 133
- Stetson, P. B., Bruntt, H., & Grundahl, F. 2003, *PASP*, 115, 413
- Sweigart, A. V., & Mengel, J. G. 1979, *ApJ*, 229, 624
- Tayar, J., Somers, G., Pinsonneault, M. H., et al. 2017, *ApJ*, 840, 17
- Taylor, B. J. 2007, *AJ*, 133, 370
- Ting, Y.-S., Conroy, C., Rix, H.-W., & Cargile, P. 2018, arXiv:1804.01530
- Torres, G. 2013, *AN*, 334, 4
- Torres, G., Andersen, J., & Giménez, A. 2010, *A&ARv*, 18, 67
- Torres, G., Fischer, D. A., Sozzetti, A., et al. 2012, *ApJ*, 757, 161
- Trampedach, R., Stein, R. F., Christensen-Dalsgaard, J., Nordlund, Å., & Asplund, M. 2014, *MNRAS*, 445, 4366
- Traxler, A., Garaud, P., & Stellmach, S. 2011, *ApJL*, 728, L29
- Ulrich, R. K. 1972, *ApJ*, 172, 165
- Unno, W. 1957, *ApJ*, 126, 259
- VandenBerg, D. A., & Bell, R. A. 2001, *NewAR*, 45, 577
- VandenBerg, D. A., Bergbusch, P. A., Dotter, A., et al. 2012, *ApJ*, 755, 15
- VandenBerg, D. A., Bergbusch, P. A., & Dowler, P. D. 2006, *ApJS*, 162, 375
- VandenBerg, D. A., Edvardsson, B., Eriksson, K., & Gustafsson, B. 2008, *ApJ*, 675, 746
- Viani, L. S., Basu, S., Chaplin, W. J., Davies, G. R., & Elsworth, Y. 2017, *ApJ*, 843, 11
- Vink, J. S., & Cassisi, S. 2002, *A&A*, 392, 553
- von Hippel, T., Jefferys, W. H., Scott, J., et al. 2006, *ApJ*, 645, 1436
- Wachlin, F. C., Vauclair, S., & Althaus, L. G. 2014, *A&A*, 570, A58
- White, T. R., Bedding, T. R., Stello, D., et al. 2011, *ApJ*, 743, 161
- Wu, X. S., Alexeeva, S., Mashonkina, L., et al. 2015, *A&A*, 577, A134
- Yadav, R. K. S., Bedin, L. R., Piotto, G., et al. 2008, *A&A*, 484, 609
- Yakut, K., Eggleton, P. P., Kalomeni, B., Tout, C. A., & Eldridge, J. J. 2015, *MNRAS*, 453, 2937
- Yang, S.-C., Sarajedini, A., Deliyannis, C. P., et al. 2013, *ApJ*, 762, 3
- Zasowski, G., Cohen, R. E., Chojnowski, S. D., et al. 2017, *AJ*, 154, 198