

# Combining Computer Models to Account for Mass Loss in Stellar Evolution

Nathan M. Stein

Statistics Department, Harvard University, Cambridge, MA 02138-2901, USA  
*nmstein@fas.harvard.edu*

David A. van Dyk

Statistics Section, Department of Mathematics, Imperial College London, SW7 2AZ, UK  
*dvandyk@imperial.ac.uk*

Ted von Hippel

Physical Sciences Department, Embry-Riddle Aeronautical University  
*ted.vonhippel@erau.edu*

Steven DeGennaro

Department of Astronomy, University of Texas at Austin, Austin, TX, USA  
*deg@astro.as.utexas.edu*

Elizabeth J. Jeffery

James Madison University, Harrisonburg, VA, USA  
*jeffe2ej@jmu.edu*

William H. Jefferys

Department of Astronomy, University of Texas at Austin, Austin, TX, USA  
Department of Mathematics and Statistics, The University of Vermont, Burlington, VT, USA  
*bill@astro.as.utexas.edu*

## Abstract

Intricate computer models can be used to describe complex physical processes in astronomy such as the evolution of stars. Like a sampling distribution, these models typically predict observed quantities as a function of a number of unknown parameters. Including them as components of a statistical model, however, leads to significant modeling, inferential, and computational challenges. In this article, we tackle these challenges in the study of the mass loss that stars experience as they age. We have developed a new Bayesian technique for inferring the so-called initial-final mass relation (IFMR), the relationship between the initial mass of a Sun-like star and its final mass as a white dwarf. Our model incorporates several separate computer models for various phases of stellar evolution. We bridge these computer models with a parameterized IFMR in order to embed them into a statistical model. This strategy allows us to apply the full force of powerful statistical tools to build, fit, check, and improve the statistical models and their computer model components. In contrast to traditional techniques for inferring the IFMR, which tend to be quite ad hoc, we can estimate the uncertainty in our fit and ensure that our model components are internally coherent. We analyze data from three star clusters: NGC 2477, the *Hyades*, and M35 (NGC 2168). The results from NGC 2477 and M35 suggest different conclusions about the IFMR in the mid- to high-mass range, raising questions for further astronomical work. We also compare the results from two different models for the primary

hydrogen-burning stage of stellar evolution. We show through simulations that misspecification at this stage of modeling can sometimes have a severe effect on inferred white dwarf masses. Nonetheless, when working with observed data, our inferences are not particularly sensitive to the choice of model for this stage of evolution.

## 1 Introduction

Complex models that cannot be expressed using analytical forms are becoming more and more common in the physical, engineering, and social sciences. In many cases the models are the implicit solutions to non-linear simultaneous equations that can only be solved using sophisticated computational algorithms. Such *computer models* are used for example to model weather and climate, the Earth's interior and seismology, genetic drift, and many other specific scientific phenomena. Computer models are particularly useful to describe complex astronomical processes, for example, to model stellar evolution, describe the properties of planetary and stellar atmospheres, simulate chemical reactions in interstellar clouds, calculate the emergence of clusters and superclusters of galaxies in the early Universe, and determine the yield of different elements during the Big Bang. In this article we explore how such computer models can be embedded into a principled statistical analysis and how we can use the resulting statistical methods for model building, fitting, checking, and improvement. Our ultimate goal is to enhance both the computer models and our understanding of the underlying astronomical phenomena.

We focus on developing statistical methods that take advantage of a set of computer models for stellar evolution with the aim of improving our understanding of the composition, age, mass, and location of clusters of stars. Stars are formed when the dense parts of a molecular cloud collapse into a ball of hot plasma. If the mass of this ball is sufficient, its core will ignite in a thermonuclear reaction that is powered by the fusion of hydrogen into helium, forming a so-called *main sequence* star. This reaction can continue for millions or billions of years, depending on the original mass and composition of the star. More massive stars are denser, hotter, and burn their fuel more quickly. When the hydrogen at the core has been depleted the core collapses, heating up the inner star and igniting the same reaction in regions surrounding the core. At the same time, the diameter of the star increases dramatically and the surface temperature cools, resulting in a *red giant* star. This period of the star's life is much shorter, lasting only about one tenth as long as the initial period. The temperature of the core continues to rise and for more massive stars will become hot enough to fuse helium into carbon, oxygen, neon, and perhaps heavier elements. During this time, stars with mass less than about 8 times that of the Sun undergo mass loss leading to the formation of a very short lived *planetary nebula*. The dense core of such a star eventually stabilizes as the outer layers of the star blow away, leaving a small hot dense ember, known as a *white dwarf* star. Nuclear reactions cease in a white dwarf star, which slowly cools via radiation from its surface.

From Earth, main sequence, red giant, white dwarf, and other types of stars can be distinguished by the relative intensity of the light emitted by the star at different wavelengths. We use *photometric* data which records the *luminosity* or *magnitude* of each star under observation in a particular range of wavelengths. The luminosity is a direct measurement of the amount of energy an astronomical object radiates per unit time. A magnitude is a negative logarithmic transformation of the luminosity, thus smaller magnitudes correspond to brighter objects. The computer models that we employ predict the photometric magnitudes as a function of a number of parameters that describe the composition, size, distance, and age of the star. Because the physical processes that power white dwarf stars on the one hand and main sequence and red giant stars on the other are completely different (i.e., latent heat and thermonuclear fusion, respectively), separate computer models are used for the white dwarf stars and for main sequence and red giant stars. Since white dwarfs are evolved main sequence stars, predicting the colors of white dwarfs involves first running a computer model for the main sequence and red giant

stars (MS/RG computer model) and then running a white dwarf model (WD computer model), using a transformation of the outputs from the MS/RG model as inputs.

Our goal is to build principled statistical models and methods that embed these computer models into a likelihood function as components of a statistical model that can be used to learn about the underlying parameters of scientific interest. The basic components of our model are described in von Hippel *et al.* (2006), DeGennaro *et al.* (2009), and van Dyk *et al.* (2009) and applied in Jeffery *et al.* (2011). These articles employ a Gaussian measurement error model for multiple photometric magnitudes per star, allow for contaminated data, and account for binary star systems composed of two separately evolving stars that appear as a single star from Earth. This article starts with this baseline and develops the linkage between the component computer models for stellar evolution and describes how we evaluate them and their fit to data. In particular, we propose a parametric model to link the MS/RG model with the WD computer model and include the parameters as part of our likelihood function. This enables us to evaluate and improve existing models that predict the mass of a white dwarf star from the initial mass of its progenitor star, the so-called *initial-final mass relationship* (IFMR). This is done within the context of a fully Bayesian model that allows us to quantify the uncertainty in the fitted IFMR.

The primary scientific goal of this article is to study the total mass loss between the initial mass of a newly formed main sequence star and the final mass of a white dwarf. Mass loss occurs as a main sequence star fuses atoms to produce energy and more dramatically as the outer layers of a red giant blow away. This later process is not well understood (Catelan, 2009) but drives the relationship between stellar ages, masses, and composition and observed photometric magnitudes for the brightest stars of a cluster. This relationship is quantified by a MS/RG computer model and is particularly important in the study of extra-galactic clusters, especially those outside our local group of galaxies. For such distant clusters, photometric data can only be obtained for the brightest stars. Thus, an understanding of red giant mass loss may be our only clue to the ages of such distant stellar systems. Yet this clue is fraught with uncertainty because we lack a properly determined mass loss formula for stellar models. Our study is aimed directly at improving our understanding of stellar mass loss.

The hallmark of our approach is principled statistical modeling of as many components of the data generation and stellar evolution processes as is practicable. This is in contrast to the standard approaches to fitting computer models for stellar evolution in astronomy. A standard strategy is to overlay the data with the predictions under the model with a specific parameter value and to manually adjust the parameters in order to visually improve the fit (e.g., Caputo *et al.*, 1990; Montgomery *et al.*, 1993; Dinescu *et al.*, 1995; Chaboyer *et al.*, 1996; Rosvick and VandenBerg, 1998; Sarajedini *et al.*, 1999; VandenBerg and Stetson, 2004). It is difficult to be sure the best fit has been found in this way or to assess the uncertainty in the fitted values. While such methods tend to leverage the knowledge and intuition of astronomers, they are suboptimal and are not reproducible. A more systematic approach involves searching for values of the parameter that are most consistent with the data (e.g., Raftery *et al.*, 1995), perhaps with the help of an *emulator* that speeds computation by approximating the computer model with an easy to evaluate function. This process of fitting the parameter values of the computer model is known as *calibration* and results in a fitted or calibrated computer model that is often used to run additional experiments under different conditions (e.g., different values of covariates). When collecting actual data under different experimental conditions is expensive or impossible researchers may use the calibrated computer model in place of actual experiments (e.g., Higdon *et al.*, 2008; Kennedy and O’Hagan, 2001; O’Hagan, 2006; Rougier, 2008).

While these methods provide principled statistical analyses that account for multiple sources of uncertainty, they are difficult to extend to situations where the computer model is but one component of a multilevel model or where multiple computer models are used in conjunction as components of a statistical model. In such situations there may be a large number of parameters

that cannot be easily manipulated manually. Our strategy is to treat the computer models as any other component of a statistical model. Like a likelihood function, computer models can be used to predict observations as a function of unknown parameters. As such, they can be embedded in a statistical model as the mean function of a conditional distribution in the same way as any conditional distribution would be incorporated. This strategy allows us to apply the full force of powerful statistical tools to build, fit, check, and improve the statistical models and their computer model components. The primary challenge becomes computational. Insofar as the computer models are time consuming to evaluate standard computational techniques will be slow to run. Moreover, the complex nature of the computer model can lead to an overall likelihood function or posterior distribution with complex nonlinear structure and/or multiple modes. Exploring these functions may involve serious computational challenges. Nonetheless, we view such computational issues as fundamentally more tractable than establishing the statistical properties of *ad hoc* numerical procedures. Our strategy involves Markov chain Monte Carlo (MCMC) fitting in a Bayesian context. While we build on the computational methods described in van Dyk *et al.* (2009) and DeGennaro *et al.* (2009), fitting the IFMR involves significant new challenges.

Our method strives for an internal coherence lacking in many previous attempts to fit the IFMR. Salaris *et al.* (2009) provide an excellent criticism of the inconsistencies in physical model assumptions that can plague IFMR studies. Most studies start with the goal of obtaining accurate initial and final mass estimates for as many white dwarfs as possible, and then derive an IFMR by treating the estimated masses as observed data. See for instance Weidemann (2000); Kalirai *et al.* (2008); Kalirai *et al.* (2005); Williams *et al.* (2009); and Williams *et al.* (2004). Spectroscopic observations paired with WD computer models can yield estimates of the current (final) mass and cooling age of each white dwarf. The cooling age is the length of time the star has been a white dwarf. Separately, the overall age of the cluster can be derived by comparing photometric observations of the main sequence stars to MS/RG models. Once a cluster age is determined, each white dwarf progenitor’s lifetime as a main sequence star is set equal to the cluster age minus the cooling age. This progenitor age can then be used to derive the progenitor’s initial mass, again in concert with a set of MS/RG computer models. Calculating reliable measures of the uncertainty for the estimate of the initial mass computed using this string of correlated calculations would not be a simple task. The typical strategy is to propagate observational errors through the calculation using delta-method-like techniques, but ignoring the uncertainty on the cluster age (Williams *et al.*, 2009). There are numerous statistical challenges beyond accounting for the star-by-star uncertainties on the initial and final masses. For example, the masses derived for white dwarfs in the same cluster rely on common estimated cluster characteristics and common MS/RG models, and are therefore correlated. A typical analysis, however, combines white dwarfs from different clusters as if they were independent observations. Moreover, internal coherence requires that the two steps that use MS/RG models—deriving the cluster age from the main sequence stars and deriving the initial masses given the cluster and cooling ages—should use the same MS/RG models, but this is not a universal practice (see for example Ferrario *et al.* (2005)). We address these concerns by embedding the entire procedure for each cluster into a statistical model, rather than combining separate star-by-star analyses, which may rely on different evolution models, in an *ad hoc* fashion. In this way we aim to arrive at a coherent statistical analysis of the IFMR.

The remainder of the article is divided into four sections. We begin with a detailed description of our statistical model and likelihood function in Section 2. In Section 3 we describe the computational challenges and solutions for fitting the model. Numerical results including simulations, data analyses, and model checking appear in Section 4. Finally, conclusions and discussion are in Section 5.

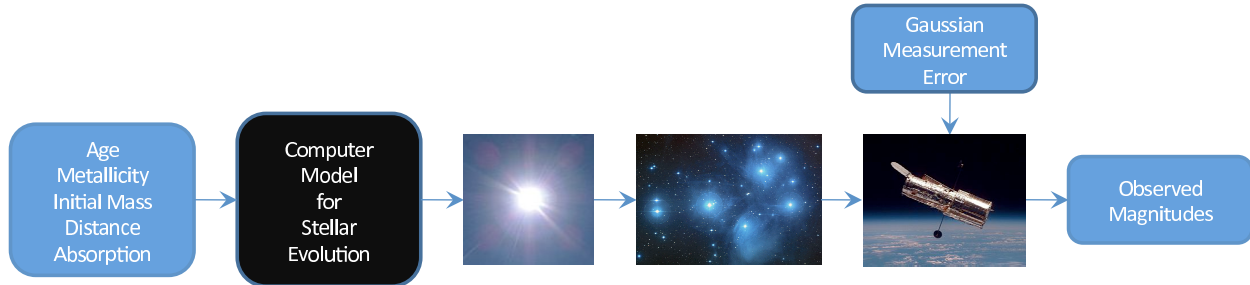


Figure 1: Schematic of the basic computer-model embedded statistical model for stellar evolution. The five input parameters for the “black box” computer model are listed in the first node. The second node is the “black box” computer model that predicts the photometric magnitudes of a single star (Node 3). We apply the computer model to each star in a physical cluster of stars; the Pleiades is pictured in Node 4. The magnitudes are recorded using filters that allow only light of specified wavelengths through. Both ground-based and space-based telescopes are used; the Hubble Space Telescope appears in Node 5. The actual observed photometric magnitudes are observed with Gaussian measurement errors.

## 2 Statistical Model

### 2.1 The Basic Models

**MS/RG Computer Models** We use a set of computer models that predict the observed photometric magnitudes of a star given a set of parameters that describe the physical characteristics of the star, its distance from us, and the density of matter along the line of sight between the star and the Earth. Figure 1 gives an overview of the basic computer model for stellar evolution. Age, metallicity, and initial mass describe the star itself and affect the photometric magnitudes of the star as it evolves. The spectra of all stars change as they evolve, very slowly as main sequence stars and with much larger and more abrupt changes as the star becomes and evolves as a giant and eventually a white dwarf. More massive stars are denser, hotter, bluer, and burn their fuel much more quickly. Thus, massive stars evolve into giants and dwarfs much more quickly. (The most massive stars end their lives as neutron stars or even black holes rather than as white dwarfs. These stars are quite rare and would not be in our datasets, since they evolve so quickly. See van Dyk *et al.* (2009) for discussion.) The composition of a star also affects the spectrum. Metals (in astronomy any element heavier than helium) absorb more blue light, and excess helium at the core reduces the efficiency of the nuclear reaction. The composition of the star is summarized by the metallicity parameter, which is the logarithm of the ratio of a star’s iron abundance compared to that of the Sun. (Iron is used as a proxy for all elements heavier than helium because it is relatively easy to identify from the optical spectrum of a star.) The final two parameters, distance and absorption, describe not the star itself but rather our ability to observe the star. More distant stars appear fainter, and some of the emitted light is absorbed by interstellar material.

The computer models (Node 2 of Figure 1) predict the observed magnitudes as a function of the five parameters. For the MS/RG model this involves iteratively solving a set of coupled differential equations. The solution is a static physical model of a star that provides a snapshot of how a star of a particular mass and radial abundance profile would appear in terms of its luminosity and color. The radial abundance profile is essentially the stellar metallicity and helium abundance along its radius, with the assumption that the stellar composition is angularly homogeneous. Stars are evolved by updating the mass and radial abundance profile to account

for elements newly produced by the thermonuclear reaction. The model is run forward in this way from the time of the stellar birth to the current time, at which point absorption and distance are used to convert absolute magnitudes into apparent magnitudes. There are a number of competing MS/RG models that differ in their implementation of the underlying physics and give somewhat different predictions. In this article we implement the Yi *et al.* (2001) and Dotter *et al.* (2008) MS/RG models. A more detailed discussion of the computer models can be found in van Dyk *et al.* (2009). Another set of models must be used for white dwarfs. A careful embedding of the WD models into our statistical model is one of the primary contributions of this article and is the topic of Section 2.2.

**Stellar Clusters** Statistical analysis is complicated by the fact that we typically observe only three or four magnitudes per star but have five unknown parameters per star. To address this we use data from stellar clusters which are groups of stars that reside physically close together and were formed together out of the same molecular cloud; see Node 4 of Figure 1. As such these stars have the same or very similar ages, metallicities, distances, and absorption parameters. Only their masses vary. The magnitudes are obtained with a mixture of ground-based and space-based telescopes (depending on how bright the stars are) using filters that remove the light outside of a particular band of wavelengths. Each magnitude is observed with Gaussian errors, the variances of which are known from observational and telescopic conditions.

**Binary Star Systems** Between one third and one half of stars are actually binary or multiple-star systems (Jao *et al.*, 2009; Sollima *et al.*, 2010), most of which cannot be resolved into their component stars. The luminosities of the component stars sum, resulting in observed magnitudes that differ from what is expected under the computer models with parameters set at the cluster values. Thus, we explicitly model all main sequence and red giant stars as binary systems. In this way the observed luminosities are modeled as the sum of those expected from two separate runs of the MS/RG model. Single-star systems can still be reasonably modeled in this way by setting their secondary masses to be near zero, with a negligible effect on the modeled luminosity. Because white dwarfs are so dim, a multiple-star system involving a white dwarf and a main sequence (or red giant) star would appear as a main sequence (or red giant) and be modeled as a binary system, presumably with a very small secondary mass. Thus, what appears as a white dwarf star could only be a system composed solely of white dwarfs. We assume all such systems are single white dwarf systems because white dwarf + white dwarf binaries are relatively rare (perhaps 10% of the cluster white dwarfs) and because such systems often have a complex history beyond the scope of our models.

**Field Stars** Stars that appear from the Earth’s vantage point to be in or near a stellar cluster but are not actually part of the physical cluster are called *field stars*. Even in the best studied star clusters, there is uncertainty about whether some stars are cluster members or field stars. These stars may be in the foreground or background of the cluster and typically have different characteristics than the cluster stars. For example, there is no reason to assume they share common values of metallicity, age, distance, or absorption with the cluster stars. Our statistical model is not designed to physically model the characteristics of field stars, but includes the capability to determine whether a star can adequately be modeled as a cluster member. In essence we use a mixture model to give our statistical model an escape hatch if a star cannot be satisfactorily modeled by the computer models in accordance with the parameter preferences of the cluster stars. For simplicity, field stars are assumed uniformly distributed in magnitude space. While this is certainly not a realistic assumption from an astrophysical perspective, in practice it accomplishes our goal of reducing the influence on parameter estimation of poorly modeled stars. Such poorly modeled stars include white dwarf + white dwarf binaries, objects known as blue stragglers that are not well understood but thought to be the result of merged

stars, as well as a subset of binary star systems that are so physically close to each other that mass is currently flowing or once did flow from one star to the other.

**Notation** We refer to the observed magnitude in color band  $j$  of star  $i$  as  $x_{ij}$  for  $j = 1, \dots, n$  and  $i = 1, \dots, N$ , with  $N$  the number of stars in the dataset and  $n$  the number of magnitudes observed for star  $i$ . The vector of observed magnitudes for star  $i$ ,  $\mathbf{X}_i = (x_{i1}, \dots, x_{in})^\top$ , has known variance-covariance matrix  $\Sigma_i$ . The predicted magnitudes for cluster stars are determined by the computer models. We express the MS/RG model as

$$G_{\text{MS/RG}}(M_i, \Theta),$$

where  $M_i$  is the stellar mass and  $\Theta = (\theta_{\text{age}}, \theta_{[\text{Fe}/\text{H}]}, \theta_{m-M_V}, \theta_{A_V})$  consists of the four parameters common to all cluster stars, i.e., *cluster parameters*, see Table 1. The distance is quantified by comparing how bright a star would appear at a particular fixed distance to how bright it actually appears. In particular, the *absolute magnitude* is the magnitude that the star would have if it were 10 parsecs (32.6 light years) away as opposed to the *apparent magnitude* it has when viewed from Earth. The distance modulus,  $\theta_{m-M_V}$ , is the difference between the apparent and absolute magnitudes.<sup>1</sup> The output from the computer model is a vector of predicted apparent magnitudes for a single main sequence or red giant star.

Since main sequence and red giant stars are presumed binary, each is modeled as a primary (more massive) star plus a secondary (less massive) star. We denote the mass of the primary star in observed star system  $i$  as  $M_i$  and the ratio of the secondary mass to the primary mass as  $R_i$ . The luminosities of the component stars sum to yield the observed luminosity. Since luminosity is related to magnitude by

$$\text{magnitude} = -2.5 \log_{10}(\text{luminosity}),$$

for main sequence and red giant stars, in the absence of noise,  $x_{ij}$  is predicted by

$$\mu_{ij} = -2.5 \log_{10} \left( 10^{-G_{\text{MS/RG},j}(M_i, \Theta)/2.5} + 10^{-G_{\text{MS/RG},j}(M_i R_i, \Theta)/2.5} \right), \quad (1)$$

where  $G_{\text{MS/RG},j}(M_i R_i, \Theta)$  is the predicted magnitude in the  $j$ th filter for the secondary star in the binary system, since  $M_i R_i$  is the secondary mass. We let  $\boldsymbol{\mu}_i = (\mu_{i1}, \dots, \mu_{in})^\top$  be the vector of predicted values of  $\mathbf{X}_i$  if star  $i$  is a main sequence or red giant star. Table 1 reviews the stellar and cluster parameters of the basic model.

## 2.2 Embedding the WD Computer Models and Fitting the IFMR

In addition to the stellar and cluster parameters our full model also includes *evolution parameters*, by which we mean parameters describing the general physical process involved in stellar evolution as opposed to parameters describing the specific stars or cluster of stars under investigation. In principle, evolution parameters should be common to all stars and all clusters because they describe underlying physical processes. In practice, however, we expect enough variation among datasets and disagreement between our models and reality that we fit these parameters separately for each star cluster. Combining the resulting fits and measures of uncertainty will be approached as a second stage in the analysis in order to formulate a global description of the underlying physical processes across a range of stars and clusters.

The evolution parameters that we aim to fit are involved with relating the MS/RG model to the WD model in the context of embedding both computer models into a statistical model. While the MS/RG model simulates the effects of the thermonuclear reaction by stepping forward

---

<sup>1</sup>We can compute the distance from the distance modulus. In particular, in the absence of absorption,  $\theta_{m-M_V} = 5 \log_{10}(d) - 5$ , where  $d$  is the distance measured in parsecs.

Table 1: Cluster and Stellar Parameters

Cluster parameters	
$\theta_{\text{age}}$	$\log_{10}$ of the cluster age
$\theta_{[\text{Fe}/\text{H}]}$	metallicity (chemical composition of the cluster)
$\theta_{m-M_V}$	distance modulus (difference between apparent and absolute magnitudes)
$\theta_{A_V}$	absorption (loss of light due to interstellar dust)
Stellar parameters	
$M_i$	primary initial mass of a binary star system in units of solar mass, $M_{\text{Sun}}$
$R_i$	the ratio of secondary mass to primary mass
$Z_i$	indicator for cluster membership

through time, the WD models simulate the energy flow from the interior and its radiation into space, by which the cinders of the nuclear reaction cool. The actual physical processes that occur in the transformation of a giant star into a white dwarf are complex and turbulent. In the last stage, a red giant fuses progressively heavier elements at different shells of its interior, begins to pulsate, contracting and expanding, finally losing its outer layer in a planetary nebulae and forming a white dwarf. All of the computer models break down at this stage and we must use a parametric model to link the inputs and outputs of the MS/RG model to the inputs of the WD model to form a coherent model of the evolution of the white dwarf stars. To see how this works, we must go into some detail of how the computer models for the white dwarfs work.

The MS/RG model simulates the life of a main sequence star by stepping forward through time either until the present age of the star or until the star evolves into a giant and then a white dwarf, and the nuclear reaction ceases. In the latter case, the MS/RG model returns how long it lived as a main sequence and red giant star. For a white dwarf, we refer to this as the *progenitor age*,

$$\phi_{\text{prog age}} = F_{\text{MS/RG}}(\theta_{[\text{Fe}/\text{H}]}, M), \quad (2)$$

where  $F_{\text{MS/RG}}$  is the MS/RG computer model, but evaluating the progenitor age rather than the expected magnitudes. Here we suppress the subscript  $i$  because the models are all for individual stars.

Since the white dwarf is essentially the cooling ember of its progenitor, its age as a white dwarf will affect its temperature and colors. Thus, the progenitor age is needed to compute the cooling age of the white dwarf, which is simply the total age of the star,  $\theta_{\text{age}}$ , less the progenitor age,  $\phi_{\text{prog age}}$ . The white dwarf cooling model is part of the WD computer model and returns the effective temperature,  $\phi_{T_{\text{eff}}}$ , and radius of the white dwarf as a function of its age and white dwarf mass:

$$(\phi_{T_{\text{eff}}}, \phi_{\text{radius}}) = \mathbf{F}_{\text{cooling}}(\theta_{\text{age}} - \phi_{\text{prog age}}, M_{\text{WD}}). \quad (3)$$

The cooling model uses the (final) mass and cooling age of a hydrogen-dominated white dwarf<sup>2</sup>

<sup>2</sup>We only consider white dwarfs with hydrogen-dominated atmospheres in our analyses. These stars make up about 70%-75% of white dwarfs and they seem to be even more prevalent in clusters (Kleinman *et al.*, 2004). Most other white dwarfs have helium-dominated atmospheres, though there are also hybrid types and rare peculiar types, and some stellar scientists argue that a white dwarf can change its atmospheric type at certain points in its cooling history. The atmospheric type must be accounted for because it affects the relationship between the stellar characteristics and the observed magnitudes. Because there are no deterministic models that predict the atmospheric type, we focus on the dominant type and obtain detailed spectra and/or a combination of optical and infrared photometry to distinguish white dwarfs of this type from those of other types.



at any point in its history to derive its latent heat, which it then numerically passes to the surface via some combination of conduction, convection, and radiation, depending on the local temperature throughout the white dwarf. This is used to compute the effective (surface) temperature and radius of the white dwarf.

With  $\phi_{\text{radius}}$  in hand, we can compute the base-10 logarithm of the gravity experienced at the surface of the white dwarf, denoted  $\phi_{\log g}$ . Using Newton’s law, the surface gravity is

$$\phi_{\log g} = \log_{10}(GM_{\text{WD}}/\phi_{\text{radius}}^2), \quad (4)$$

where  $G$  is Newton’s gravitational constant.

Finally, the expected magnitudes rely on a second component of the WD computer model that simulates the star’s atmosphere,

$$\boldsymbol{\mu} = \mathbf{F}_{\text{atmosphere}}(\phi_{T_{\text{eff}}}, \phi_{\log g}). \quad (5)$$

The white dwarf atmosphere model uses the surface gravity, which is dictated by the white dwarf mass, and the effective temperature, given by the cooling model, to derive the emergent spectrum of the hydrogen-dominated atmosphere as a function of wavelength. The atmosphere model further integrates the emergent spectrum over the wavelength range of the appropriate filter(s) in order to calculate the magnitude(s) of the model white dwarf. The missing piece in these calculations is the white dwarf mass  $M_{\text{WD}}$ , which is needed in (3) and (4). The initial-final mass relation (IFMR) aims to fill this gap by computing  $M_{\text{WD}}$  as a function of the initial mass of the progenitor star,  $M$ . When viewed in this way, the IFMR links the MS/RG models (which depend on  $M$ ) to the WD cooling and atmosphere models (which depend on  $M_{\text{WD}}$ ). DeGennaro *et al.* (2009) and van Dyk *et al.* (2009) used a piecewise-linear functional form (with no unknown parameters) for the IFMR given by Weidemann (2000). The main contribution of this article is to parameterize the link between the two sets of computer models, to determine how much information is in the data to constrain this relationship, and in this way to evaluate the model assumptions inherent in previous analyses. Figure 2 extends Figure 1 to illustrate how we open up the black box computer model to fit the IFMR that connects the MS/RG model with the WD model. In our analyses we use the white dwarf cooling model of Wood (1992) and the white dwarf atmosphere model of Bergeron *et al.* (1995).

We consider three parameterized forms for the IFMR, which we denote

$$M_{\text{WD},i} = f(M_i, \boldsymbol{\alpha}).$$

First, following Williams *et al.* (2009) and Salaris *et al.* (2009), we consider a linear relationship,

$$M_{\text{WD},i} = f(M_i, \boldsymbol{\alpha}) = \alpha_0 + \alpha_1(M_i - M^*), \quad (6)$$

where  $\boldsymbol{\alpha} = (\alpha_0, \alpha_1)$  are parameters that will be fitted to the data and  $M^*$  is a pre-selected fixed value for approximately centering the initial masses,  $M_i$ , of the white dwarfs. (Approximately centering  $M_i$  tends to reduce the posterior correlation between  $\alpha_0$  and  $\alpha_1$ .) Although we do not expect that the linear form given in (6) will hold over a wide range of  $M_i$ , we believe that it is a reasonable default in the context of applying our model to a cluster of stars. This is because cluster stars are all the same age and the more massive stars are the first to evolve into white dwarfs. This means that the cluster stars that are now white dwarfs have a lower bound on their initial mass. The white dwarfs in our dataset will also have an upper bound on their initial mass because stars more massive than about 8 solar masses will explode as supernovae and become neutron stars or black holes, instead of white dwarfs. Additionally, the most massive white dwarfs are rare because the number of stars of all types drops off inversely with about the square of the mass of the star. Thus, the white dwarfs in any one cluster will typically contain only white dwarfs that span a fairly narrow range of initial mass, where the linear assumption

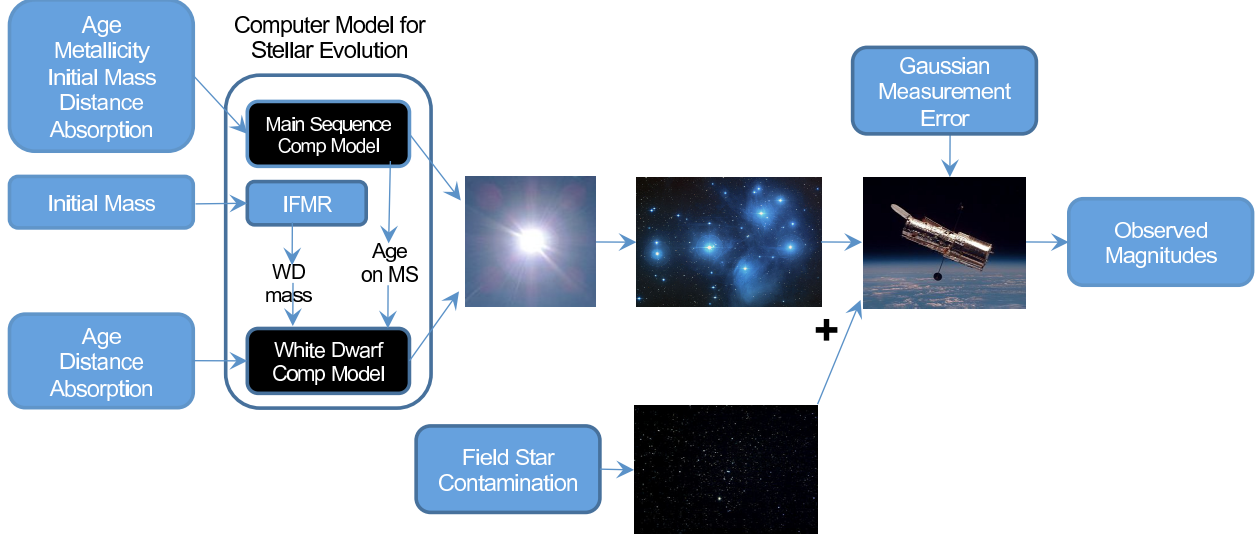


Figure 2: Embedding the IFMR model into a statistical model. This schematic extends the one in Figure 1. We use a finite mixture model to separate field stars that contaminate the data from the stars that are members of a physical star cluster. Furthermore, we open up the “black box” computer models and parameterize the relationship between the input mass of the MS/RG computer model and the input mass of the WD computer model,  $M$  and  $M_{\text{WD}}$ , respectively. In this way we include parameters that describe the evolutionary process in addition to parameters that describe the particular stars and cluster under study. In the interest of space we have included only one node for the computer model, whereas we actually apply the model twice and sum the expected luminosities to account for binary systems involving main sequence and red giant stars.

in (6) may be reasonable. Even if there are nonlinearities in the *global* initial-final mass relation, *local* linearity is a much weaker assumption and more likely to be an acceptable approximation.

Although a linear model is our default, we also consider a quadratic IFMR and a broken linear IFMR, a continuous model with a point where the slope changes. In all three cases we use  $\alpha$  to parameterize the IFMR, though  $\alpha$  may have two or three components (we fix the breakpoint of the broken linear model in advance). We compare the three IFMR models for a particular star cluster in Section 4.2, in the discussion of NGC 2477.

With the IFMR in place, the parameters we use to describe white dwarf evolution are  $\phi = (\phi_{\text{proage}}, \phi_{T_{\text{eff}}}, \phi_{\text{radius}}, \phi_{\log g})$  and  $M_{\text{WD},i}, i \in \mathcal{I}_{\text{WD}}$ , where  $\mathcal{I}_{\text{WD}}$  is the pre-specified set of indices corresponding to the white dwarf stars. The components of  $\phi$  are deterministic functions of the cluster parameters,  $\theta_{\text{age}}$  and  $\theta_{[\text{Fe}/\text{H}]}$ ; the stellar parameters  $M_i, i \in \mathcal{I}_{\text{WD}}$ ; and the evolution parameters,  $\alpha$ . As such, we do not consider  $\phi$  to be a model parameter, and do not fit  $\phi$  in our formal analyses. It is only used to compute  $\mu$  for the white dwarfs. Thus, we can express the WD computer model as  $\mathbf{G}_{\text{WD}}(M_i, \Theta, \alpha)$  which is a compilation of  $F_{\text{MS/RG}}$ , the IFMR, and the white dwarf cooling and atmosphere models. The output from  $\mathbf{G}_{\text{WD}}$  is a vector of predicted magnitudes for a single white dwarf star.

### 2.3 The Likelihood Function

Because the data are a mixture of cluster and field stars, we formulate the statistical model as a finite mixture of two populations. In particular, we define  $\mathbf{Z} = (Z_1, \dots, Z_N)$ , where  $Z_i$  equals one if star  $i$  is a cluster star and  $Z_i$  equals zero if star  $i$  is a field star. We then specify the

conditional distribution of the photometric magnitudes for cluster stars and field stars separately.

For cluster stars, we assume that observed photometric magnitudes,  $\mathbf{X}_i$ , are independent  $n$ -dimensional multivariate Gaussian distribution with mean  $\boldsymbol{\mu}_i$  and diagonal variance-covariance matrix  $\boldsymbol{\Sigma}_i$ . (The observation errors are small enough in magnitude to justify an assumption of symmetric errors, even though strictly speaking they are not.) While this appears to be a simple model, it is actually quite complex owing to the dependence of  $\boldsymbol{\mu}_i$  on the cluster, stellar, and evolution parameters. Because this dependence is quantified via the MS/RG and WD computer models, the relationship between  $\boldsymbol{\mu}$  and the model parameters is highly non-linear. In particular, the components of  $\boldsymbol{\mu}_i$  are given by (1) for  $i \in \mathcal{I}_{\text{MS/RG}}$ , where  $\mathcal{I}_{\text{MS/RG}}$  is the pre-specified set of indices corresponding to the main sequence and red giant stars, and

$$\boldsymbol{\mu}_i = \mathbf{G}_{\text{WD}}(M_i, \boldsymbol{\Theta}, \boldsymbol{\alpha}) \text{ for } i \in \mathcal{I}_{\text{WD}}. \quad (7)$$

The MS/RG computer model does not depend on the IFMR, so the IFMR parameters  $\boldsymbol{\alpha}$  do not appear as arguments to  $\mathbf{G}_{\text{MS/RG}}$ , whereas the WD computer model explicitly depends on  $\boldsymbol{\alpha}$ . Recall that our statistical model for white dwarf cluster members differs in two ways from that for main sequence cluster members: (i) the deterministic models involve more components, including the IFMR, and (ii) we do not model binaries involving white dwarf stars so that the functional form of (7) is simpler than that in (1).

For field stars, we simply assume that each magnitude is uniformly distributed over a finite range that corresponds to the range of the data,

$$p_{\text{field}}(\mathbf{X}_i) = c \text{ for } \min_j \leq x_{ij} \leq \max_j, j = 1, \dots, n,$$

and is zero elsewhere, where  $(\min_j, \max_j)$  is the range of values for magnitude  $j$ , and  $c = \left[ \prod_{j=1}^n (\max_j - \min_j) \right]^{-1}$ . This is equivalent to assuming that a field star is equally probable at any point in the luminosity space spanned by the cluster stars. In the future, we may incorporate a more realistic model and even tune the model for specific Galactic fields, where field star properties are often available. To date, we have found this additional effort unnecessary, as most field stars are easily identified because their luminosities and colors are inconsistent with models of cluster stars.

Finally, the likelihood function can be written,

$$\begin{aligned} & L(\mathbf{M}, \mathbf{R}, \boldsymbol{\Theta}, \mathbf{Z}, \boldsymbol{\alpha} | \mathbf{X}, \boldsymbol{\Sigma}) \\ &= \prod_{i \in \mathcal{I}_{\text{MS/RG}}} \left[ \frac{Z_i}{\sqrt{(2\pi)^n |\boldsymbol{\Sigma}_i|}} \exp \left( -\frac{1}{2} (\mathbf{X}_i - \boldsymbol{\mu}_i)^\top \boldsymbol{\Sigma}_i^{-1} (\mathbf{X}_i - \boldsymbol{\mu}_i) \right) + (1 - Z_i) p_{\text{field}}(\mathbf{X}_i) \right] \\ &\times \prod_{i \in \mathcal{I}_{\text{WD}}} \left[ \frac{Z_i}{\sqrt{(2\pi)^n |\boldsymbol{\Sigma}_i|}} \exp \left( -\frac{1}{2} (\mathbf{X}_i - \mathbf{G}_{\text{WD}}(M_i, \boldsymbol{\Theta}, \boldsymbol{\alpha}))^\top \boldsymbol{\Sigma}_i^{-1} (\mathbf{X}_i - \mathbf{G}_{\text{WD}}(M_i, \boldsymbol{\Theta}, \boldsymbol{\alpha})) \right) \right. \\ &\quad \left. + (1 - Z_i) p_{\text{field}}(\mathbf{X}_i) \right] \end{aligned} \quad (8)$$

where  $\mathbf{M} = (M_1, \dots, M_n)$ ,  $\mathbf{R} = \{R_i, i \in \mathcal{I}_{\text{MS/RG}}\}$ ,  $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_n)$ ,  $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_n)$ , and  $\boldsymbol{\mu}_{ij}$  is given by (1) for  $i \in \mathcal{I}_{\text{MS/RG}}$ .

## 2.4 Prior Distributions

We generally take a Bayesian perspective for model fitting both because it allows us to use sophisticated computational tools to explore the complex structures in the parameter space resulting from the highly non-linear character of the computer models in (8) and because it gives us a mechanism for inference in this complex non-Gaussian space, where point estimates

and confidence intervals simply do not suffice (see Figure 6). A Bayesian analysis also allows us to handle hundreds of nuisance parameters in a principled manner and to include prior information based on substantive scientific knowledge. For example, we put an informative prior distribution on the stellar masses based on what astronomers know to be the population distribution of stellar masses in the Milky Way. Likewise, since we know the age of the Galaxy and we know that the cluster is part of the Galaxy, we know something about the cluster age that can be formulated into a prior distribution. In particular, we use a uniform prior distribution on the  $\log_{10}$  of cluster age in years. This prior distribution quantifies the fact that younger clusters are more common than older clusters. Similarly, for all of the cluster parameters, we use substantive external information to set prior distributions when it is available, and relatively non-informative prior distributions when we have no such information. In particular, for  $\theta_{[\text{Fe}/\text{H}]}$ ,  $\theta_{m-M_V}$ , and  $\theta_{A_V}$ , we use Gaussian prior distributions (truncated to the positive real line in the case of  $\theta_{A_V}$ ), with means set according to astrophysical prior knowledge and variances chosen to be reasonably non-informative.

For well studied clusters, we may have multi-decadal deep imaging that can be used to determine which stars are likely to be cluster stars and which are likely to be field stars. Because field stars are typically some distance from the cluster and in different orbits in the Galaxy, their apparent position on the sky relative to the cluster will slowly change as they orbit. Cluster stars, on the other hand, are gravitationally bound to one another and so orbit the Galaxy with a common velocity. This enables us to distinguish field stars from cluster stars via their proper motions or radial velocities, which are their three-dimensional space velocities projected onto the plane of the sky and along the line of sight, respectively. When such information is available, we use it to set the prior probability of cluster membership for each star. When it is unavailable, we simply use the likely fraction of cluster stars in the dataset for all prior probabilities of cluster membership.

For each of the three models that we consider for the IFMR, we use flat priors on  $\alpha$  over the region such that the parameterized IFMR is monotonically increasing. Thus, we assume that more massive main sequence stars evolve into more massive white dwarfs, which is well supported by our understanding of stellar evolution (Salaris *et al.*, 2009).

Our joint prior distribution on  $(\Theta, \alpha, \mathbf{M}, \mathbf{R}, \mathbf{Z})$  is specified so that

$$p(\Theta, \alpha, \mathbf{M}, \mathbf{R}, \mathbf{Z}) = p(\Theta)p(\alpha)p(\mathbf{M}, \mathbf{R}, \mathbf{Z}). \quad (9)$$

We denote the prior probability of cluster membership for star  $i$ ,  $p(Z_i = 1)$ , by  $\pi_i$  and note that given that  $Z_i = 0$ , the conditional posterior distribution  $p(M_i, R_i | X_i, Z_i = 0)$ , is the same as the conditional prior distribution  $p(M_i, R_i | Z_i = 0)$  because the computer models are not involved in the likelihood function for field stars. We specify a prior distribution for  $M_i$  and  $R_i$  that depends on  $Z_i$  partially because  $X_i$  is uninformative for the stellar masses for field stars so that a proper prior distribution must be used for the masses given  $Z_i = 0$  (but not necessarily given  $Z_i = 1$ ) and partially for computational reasons; see van Dyk *et al.* (2009). As discussed in the next section, these choices will simplify our computational techniques.

### 3 Statistical Computation

There are three stellar parameters per star, four cluster parameters, and two or three evolution parameters (depending on the choice of IFMR) in model (8), for a total of  $3N+4+2$  or  $3N+4+3$  unknown parameters. A typical cluster that we analyze has at least  $N = 150$  stars, yielding a posterior distribution on hundreds of dimensions.

Because of the high number of parameters and because the computer evolution models are not available in closed form, this model defies many standard computational techniques. DeGenaro *et al.* (2009) and van Dyk *et al.* (2009) rely on MCMC with dynamic reparameterization.

That is, a Metropolis-within-Gibbs sampler is implemented entirely sequentially: first the cluster indicators are updated, then the primary masses, mass ratios, and finally each cluster parameter, one parameter at a time conditioning on the current values of all the other parameters. The dynamic reparameterization is based on a sequence of initial runs that explore the posterior correlations among the parameters and automatically derive transformations to reduce the correlations and improve convergence. This strategy has been successful with simulated data and with several observed clusters, including the *Hyades* (DeGennaro *et al.*, 2009) and NCG 2477 (Jeffery *et al.*, 2011). However, some datasets frustrate this approach due to a combination of high posterior correlations and multiple modes. The mixture model for cluster membership can be particularly challenging computationally because local posterior modes can correspond to particular configurations of stars selected as cluster members. Once the MCMC sampler arrives in one of these modes, it may be all but impossible to escape and select a different cluster membership configuration, even if the latter would yield a higher joint posterior density.

To address these MCMC convergence issues we collapse the parameter space (Liu *et al.*, 1995) by marginalizing over the parameter vectors  $\mathbf{M}$ ,  $\mathbf{R}$ , and  $\mathbf{Z}$ , which for most stellar evolution analyses are nuisance parameters. The marginal posterior distribution of  $\Theta$  and  $\alpha$  is

$$p(\Theta, \alpha | \mathbf{X}) = \int \cdots \int \left( \sum_{Z_1} \cdots \sum_{Z_N} p(\Theta, \alpha, \mathbf{M}, \mathbf{R}, \mathbf{Z} | \mathbf{X}) \right) d\mathbf{M}d\mathbf{R} \quad (10)$$

$$\propto p(\Theta, \alpha) \prod_{i=1}^N \left[ \pi_i \int \int p_{\text{clust}}(\mathbf{X}_i | M_i, R_i, \Theta, \alpha) p_{\text{clust}}(M_i, R_i) dM_i dR_i + (1 - \pi_i) p_{\text{field}}(\mathbf{X}_i) \right], \quad (11)$$

where  $p_{\text{clust}}(\mathbf{X}_i | M_i, R_i, \Theta, \alpha) = p(\mathbf{X}_i | M_i, R_i, \Theta, \alpha, Z_i = 1)$  is the likelihood function for a cluster star and  $p_{\text{clust}}(M_i, R_i) = p(M_i, R_i | Z_i = 1)$  is the prior distribution of  $(M_i, R_i)$  for a cluster star. The functional form of (11) takes advantage of the simplifications in the prior distribution described in Section 2.4. If the integral in (11) can be evaluated, the dimension of the posterior distribution is reduced from hundreds to just six or seven. This integral cannot be evaluated analytically because  $p_{\text{clust}}(\mathbf{X}_i | M_i, R_i, \Theta, \alpha)$  depends on  $M_i$  and  $R_i$  through the computer stellar evolution models. However, numerical integration is possible.

Conditional independence allows us to reduce the  $2N$ -dimensional integration and  $N$  summations in (10) to the product of  $N$  2-dimensional integrations in (11). We take advantage of this factorization of the marginal posterior distribution by parallelizing the evaluation of the  $N$  conditionally independent factors. Each processor can receive a subset of the  $N$  stars and return the appropriate integral evaluations. For numerical integration, we evaluate the posterior density over a grid of values of the primary and secondary masses and use a reliable but computationally expensive Riemann sum approximation. The grid spacing is a refinement of the mass spacing in the tabulated MS/RG models,<sup>3</sup> so that for instance the grid spacing is extremely fine over the red giant initial mass range, since for these stars small changes in mass lead to large changes in predicted photometry.

The parameter vector  $(\Theta, \alpha)$  is six- or seven-dimensional. We implement a Metropolis algorithm to update all of these parameters together in a single jump within each iteration. We divide the burn-in period in half, saving draws during the second half to estimate the posterior covariances. After the burn-in, we use a multivariate  $t$  proposal distribution with 6 degrees of freedom, centered at the current value of  $(\Theta, \alpha)$  and with scale equal to the estimated covariance matrix multiplied by  $2.4^2/6$  (Gelman *et al.*, 1996), which tends to work well in practice.

Given a set of draws  $\{(\Theta^{(l)}, \alpha^{(l)})\}_{l=1}^L$  from the marginal posterior distribution, we can compute cluster membership probabilities and draw initial mass parameters for each white dwarf.

---

<sup>3</sup>Instead of solving the coupled differential equations for MS/RG evolution at each parameter value, which would be computationally prohibitive, we interpolate among tabulated values computed in advance.

We perform both of these steps by evaluating each conditional posterior distribution on a grid. In particular, we compute the conditional probability that white dwarf  $i$  is a cluster member as

$$\Pr(Z_i^{(l)} = 1 \mid \Theta^{(l)}, \alpha^{(l)}, \mathbf{X}) = \frac{\pi_i \int p_{\text{clust}}(\mathbf{X}_i \mid M_i, \Theta^{(l)}, \alpha^{(l)}) p_{\text{clust}}(M_i) dM_i}{(1 - \pi_i) p_{\text{field}}(\mathbf{X}_i) + \pi_i \int p_{\text{clust}}(\mathbf{X}_i \mid M_i, \Theta^{(l)}, \alpha^{(l)}) p_{\text{clust}}(M_i) dM_i}$$

where we have dropped the dependence on  $R_i$  because we only implement this step for white dwarfs, which are all modeled as single star systems ( $R_i = 0$ ). We use the Rao-Blackwellized estimate of the marginal posterior probability of cluster membership,

$$\widehat{\Pr}(Z_i = 1 \mid \mathbf{X}) = \frac{1}{L} \sum_{l=1}^L \Pr(Z_i^{(l)} = 1 \mid \Theta^{(l)}, \alpha^{(l)}, \mathbf{X}).$$

Finally, we sample initial masses  $M_i^{(l)}$  from their conditional posterior distribution  $p(M_i^{(l)} \mid Z^{(l)} = 1, \Theta^{(l)}, \alpha^{(l)}, \mathbf{X})$  and compute final masses  $M_{\text{WD},i}^{(l)}$  as

$$M_{\text{WD},i}^{(l)} = f(M_i^{(l)}, \alpha^{(l)}),$$

where  $f$  is the IFMR model.

## 4 Numerical Results

### 4.1 Simulations

As an initial test of our method, we simulated and fitted eight stellar clusters under our model. Each cluster was simulated with solar metallicity ( $\theta_{[\text{Fe}/\text{H}]} = 0$ ), no absorption ( $\theta_{A_V} = 0$ ), and distance equal to 1 kiloparsec, ( $\theta_{m-M_V} = 10$  magnitudes). We use a distance of 1 kpc because many clusters of interest are within this distance. The stellar masses were generated from the prior distribution for mass, which is a representation of the population distribution of stellar masses in our Galaxy. The clusters were generated with no field star contamination and all stars were generated as single-star systems. The ages of the clusters varied between approximately 250 million years and 4 billion years (see Figure 3). This range was chosen because it covers most of the age range of clusters we hope to analyze. The age parameter is the most important to vary because it affects the initial and final masses of the stars that are currently white dwarfs. The older the cluster, the less massive were the typical progenitor stars of the current white dwarfs. Thus, varying age allows us to explore a wider range of the IFMR. Because stellar masses were generated from the prior distribution, a random number of simulated stars would not have sufficient mass to be detectable when compared with the realistic threshold signal-to-noise ratio and were therefore dropped from the simulated cluster. Simulating a total of 600 stars resulted in each of the eight clusters having about 150 detectable stars, a typical size for an actual observed cluster. The first set of simulations was run using the MS/RG computer model of Yi *et al.* (2001), the white dwarf cooling model of Wood (1992), the white dwarf atmosphere model of Bergeron *et al.* (1995), and the IFMR of Williams *et al.* (2009):

$$M_{\text{WD}} = 0.339 + 0.129 M_{\text{init}},$$

so the true simulated IFMR was linear. For each of the eight simulations, we generated three magnitudes, corresponding to the filters B (blue), V (visual), and I (infrared) filters.

We fit each of the eight simulated clusters using our Bayesian MCMC code, using all of the same model components as in the simulation models, except we fit the parameters of the linear IFMR instead of fixing them at the Williams *et al.* (2009) values. Each star was given

a 99% prior probability of cluster membership, and the model allowed MS/RG binary systems. The prior distributions for  $\theta_{[\text{Fe}/\text{H}]}$ ,  $\theta_{m-M_V}$ , and  $\theta_{A_V}$  were centered at the true values used in simulating data, with variances 0.09, 0.04, and 0.01, respectively, which are typical values of the prior variances for clusters we analyze.

The MCMC fitting code behaved well for all eight clusters. We ran one chain for 25000 iterations for each cluster, saving every 25th iteration. Inspection of the time series and autocorrelation functions showed that the chains converged quickly to their apparent stationary distributions. Results from the simulations appear in the first row of Figure 3, which plots the true initial mass  $M$  (left column) and the true final mass  $M_{\text{WD}}$  (right column), both against their 95% central posterior intervals for each of the white dwarfs generated in the eight simulations. As is evident in Figure 3, our method performs extremely well—at least when computer models for stellar evolution are correctly specified. The posterior distributions of the initial and final masses are all very close to their true values.

One of the primary goals of this article is to investigate the effects of model misspecification and to propose model diagnostics for the IFMR. As an initial test, we investigate the effect of using one MS/RG computer model for cluster simulation and another for model fitting. In particular we exactly replicated the simulation study illustrated in the first row of Figure 3 except that the clusters were simulated using the MS/RG computer model of Dotter *et al.* (2008); the model of Yi *et al.* (2001) was still used in model fitting. The results appear in the second row of Figure 3. Conversely, the third row of Figure 3 shows the results when the model of Yi *et al.* (2001) was used for simulations and that of Dotter *et al.* (2008) was used for model fitting. Figure 3 clearly shows the systematic error in the fitted masses resulting from model misspecification. This is troubling since both of these computer models are viewed as viable by the scientific community. Of course ultimately we are interested in the bias introduced by the use of our models in place of the actual physical processes that govern the birth, evolution, and transformations of stars. Insofar as we believe that the range of available models has a broad enough span to include these physical processes, we may be willing to accept that the magnitude of the actual systematic error is on the order of the errors we see in Figure 3. This, however, may be hard to justify given the overlap among the physical assumptions and numerical techniques used by the computer models. There are likely correlations among the models and the actual systematic error could quite possibly be larger than what is illustrated in rows two and three of Figure 3.

Nonetheless, by quantifying these systematic errors we illustrate the challenges involved with these models for the stellar model creators and users. Insofar as we can precisely quantify the effect of the choice of MS/RG computer model on the IFMR, we not only provide the best current estimate of the IFMR, but also point the way toward possible improvement in the underlying computer models. For example, if external data can be used to rule out a fitted statistical model, it will correspondingly call into question the computer model upon which that statistical model was based.

## 4.2 Analyses of Observed Star Clusters

In this section we apply our method to three stellar clusters in order to fit the IFMR. Our aim is to provide a more coherent analysis than has been conducted to date and to evaluate existing models for the IFMR. When comparing our fits with those in the literature it is important to bear in mind that we use only two or three photometric magnitudes per star, whereas the existing models are based on much more costly high resolution spectra that can be viewed as hundreds of photometric magnitudes per star. Despite the fact that our data is coarser by a factor of hundreds, our results agree very well with existing models for the IFMR and we are able to provide principled measures of the uncertainty in our fitted model. Because we are interested in the shape of the IFMR over a wide range of initial mass ( $M$ ) and believe that a

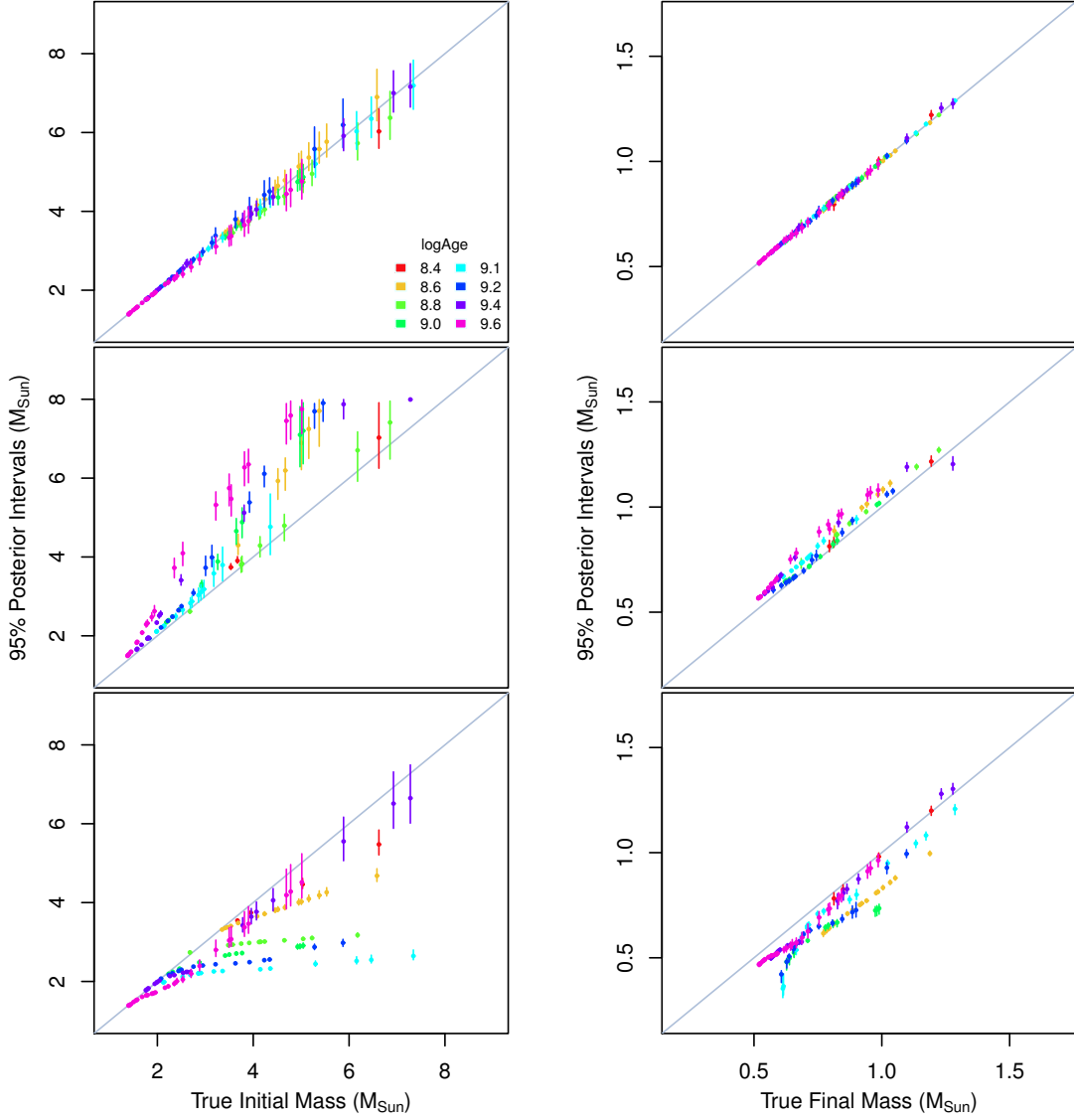


Figure 3: Results of a simulation study designed to test the recovery of parameters from clusters simulated under the model. In the top row, both simulations and model fitting were performed using the MS/RG computer model of Yi *et al.* (2001) and the linear IFMR of Williams *et al.* (2009). We fit the intercept and slope of the linear IFMR. The panels compare the true initial mass (left column) and the true final mass (right column) with 95% posterior intervals of these quantities for each white dwarf generated in the simulation study. In the middle row, the eight clusters were simulated using the Dotter *et al.* (2008) model and fitted with the Yi *et al.* (2001) model. The roles of the two MS/RG computer models were reversed for the bottom row.



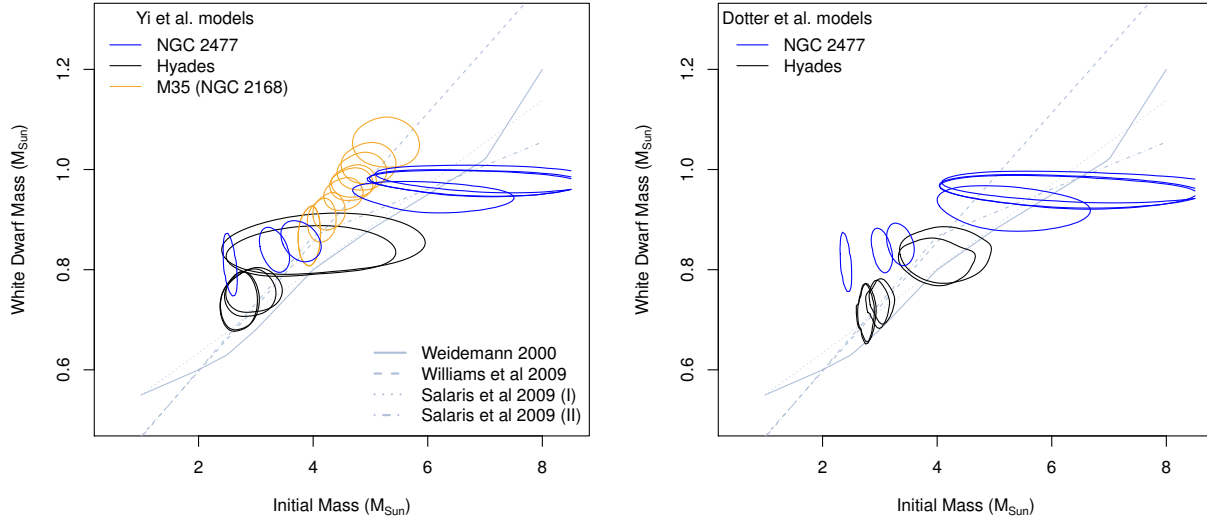


Figure 4: Results for NGC 2477, the *Hyades*, and M35 white dwarfs, using the Yi *et al.* (2001) and Dotter *et al.* (2008) models (left and right panels, respectively). M35 does not appear in the right panel because the Dotter *et al.* (2008) models do not have predictions for clusters as young as M35. Contours are 95% highest posterior density regions of each white dwarf’s joint initial and final mass posterior distribution. The IFMR model enforces linearity within each cluster, but not between clusters.

simple functional form may only be appropriate locally, we chose clusters of varying age with the hope of getting a hint of a global fit for the IFMR. (Recall that the younger the cluster, the more massive the progenitor stars that have evolved into the current bright white dwarfs.) We detail the analysis of each of the three clusters in turn, starting with NGC 2477, followed by the nearby and well studied *Hyades*, and conclude with the young cluster known as M35.

**NGC 2477** At approximately one billion years old, NGC 2477 is a moderately old cluster (Jeffery *et al.*, 2011). It is composed of hundreds of stars, and in the night sky it is about the size of the full moon. We analyze the dataset compiled in Jeffery *et al.* (2011), which consists of observations in the V (visual) and I (infrared) photometric filters. Because the MS/RG computer models tend to do a poor job at fitting faint main sequence stars (see the discussion in DeGennaro *et al.* (2009)), we removed from our dataset MS/RG stars fainter than 15.5 magnitudes in the V band. The dataset then contained 7 white dwarfs and 223 MS/RG stars. We use the default linear IFMR and prior distributions described in Section 2.4 for the cluster age, IFMR intercept, and IFMR slope and fix the other three cluster parameters at the mean values of the prior distributions used in Jeffery *et al.* (2011), namely  $\theta_{[\text{Fe}/\text{H}]} = -0.1$ ,  $\theta_{m-M_V} = 11.46$  magnitudes, and  $\theta_{A_V} = 0.75$ .<sup>4</sup> By fixing the values for these three parameters, we avoid a full study of the posterior distribution of this cluster’s parameters, which is affected by variations in the amount of absorption along the line-of-sight to different stars, which astronomers call differential reddening. We expect this component of the analysis to be presented in a future article in an astronomy journal.

<sup>4</sup>These values are also close to the posterior inferences for these parameters in Jeffery *et al.* (2011).

The MCMC code was run for 30000 iterations, discarding the first 5000 iterations as burn-in, for two separate chains with different starting values. (We followed this same procedure for the *Hyades* and M35 as well.) Inspection of the time series and autocorrelation plots indicated that the chains converged quickly to their apparent stationary distributions. With 8 cores, the MCMC code took approximately 49 hours to run using the Dotter *et al.* (2008) models; with 16 cores, it took 19.5 hours using the Yi *et al.* (2001) models. Results using the Yi *et al.* (2001) and Dotter *et al.* (2008) computer models are shown in the first and second panel of Figure 4, respectively. The blue solid contours correspond to 95% highest posterior density regions of the joint posterior distribution of the initial and final masses of each white dwarf in the dataset. For comparison, the IFMRs from Weidemann (2000), Williams *et al.* (2009), and Salaris *et al.* (2009) are displayed as grey lines of varying line type. The impact of the choice of MS/RG model can be seen by comparing the solid blue contours in the left and right panels of Figure 4, corresponding to the models of Yi *et al.* (2001) and of Dotter *et al.* (2008). The contours plotted with different colors correspond to other clusters and are discussed below.

The initial masses of the NGC 2477 white dwarfs appear to span a fairly large range, which makes it more difficult to support the default linear IFMR for this cluster. However, much of the apparent age range is due to large uncertainties in the initial mass of the high mass white dwarfs, a common phenomenon in IFMR studies (see Salaris *et al.* (2009) for more examples). To test the sensitivity of our inferences for the NGC 2477 white dwarfs to the linear IFMR, we also fit the piecewise linear IFMR and a quadratic IFMR; see Figure 5. For the piecewise linear IFMR, we fixed one breakpoint at 4 solar masses. As is evident in Figure 5, the initial and final masses posterior distributions are somewhat sensitive to the choice of IFMR model. However, in the context of the posterior distributions for the other clusters (Figure 4), the results from the three IFMR models are qualitatively similar. Under any of the three IFMRs, the high mass NGC 2477 white dwarfs appear to agree more with the Salaris *et al.* (2009) and Weidemann (2000) IFMRs than with the Williams *et al.* (2009) IFMR, but there is enough uncertainty to prevent any definitive statement.

**The *Hyades*** The *Hyades* is the nearest star cluster to our Solar System and is visible to the unaided eye as the nose of Taurus the Bull. As one of the best studied clusters in the sky, it is ideal for checking the results of our method against established knowledge. The *Hyades* is the only cluster fit in DeGennaro *et al.* (2009) and van Dyk *et al.* (2009). Here we extend their results by fitting the IFMR within a single coherent analysis.

The dataset includes U (ultraviolet), B (blue), and V (visual) photometry from the General Catalogue of Photometric Data (Mermilliod *et al.*, 1997). As in our analysis of NGC 2477, we removed the faintest main sequence stars, using a cut-off of 4.5 magnitudes in the V band, leaving 93 MS/RG stars and 6 white dwarf stars. We fit our model using the MS/RG computer models of Yi *et al.* (2001) and Dotter *et al.* (2008) with a linear IFMR to the *Hyades* data. The prior distributions, except where noted below, were

$$\begin{aligned}\theta_{[\text{Fe}/\text{H}]} &\sim N(0.07, 0.05^2), \\ \theta_{m-M_V} &\sim N(0, 0.03^2), \text{ and} \\ \theta_{A_V} &\sim TN(0.009, 0.006^2; 0),\end{aligned}$$

where the three parameters are *a priori* independent, with  $TN(\mu, \sigma^2; 0)$  representing a Gaussian distribution with mean  $\mu$  and variance  $\sigma^2$ , truncated to be greater than 0. The prior distributions were chosen to reflect astronomers' knowledge of the *Hyades*, with somewhat conservative uncertainties. See the Appendix for a discussion of why the prior on  $\theta_{m-M_V}$  is centered at 0. Figure 4 summarizes the resulting posterior distributions for the initial and final masses from each separate analysis of the *Hyades*. Typical runtimes were 6.6 hours with the Dotter *et al.* (2008) models and 3.7 hours with the Yi *et al.* (2001) models, each using 16 cores.

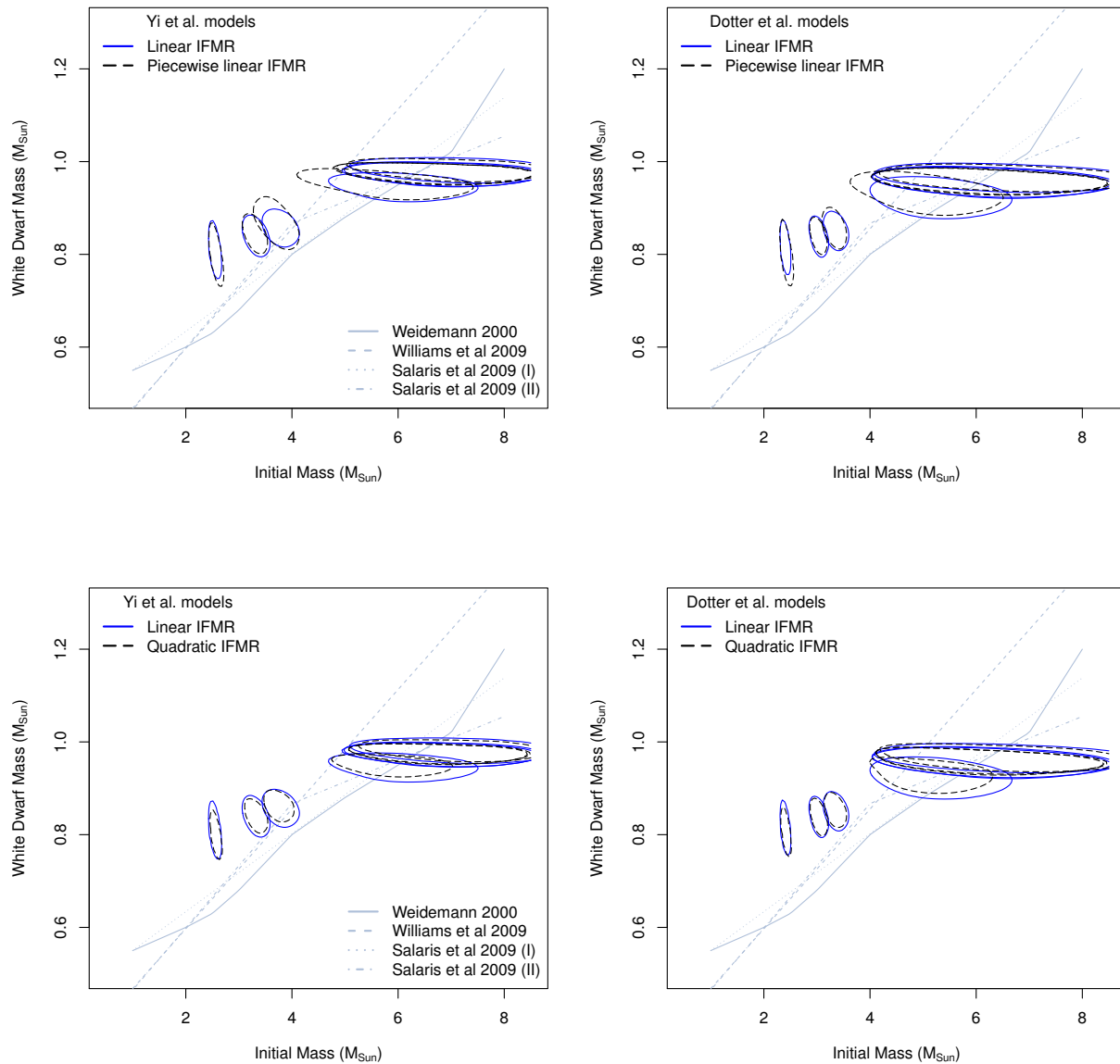


Figure 5: Initial and final mass inferences for NGC 2477 white dwarfs under three different models for the IFMR. The blue contours corresponding to the linear IFMR are the same in the top and bottom rows.

Ultra-precise distance estimates obtained with the Hipparcos Satellite are available for the *Hyades* (Perryman *et al.*, 1997). Although such distance measurements are not available for a typical cluster, they do allow us to check the fitted value of  $\theta_{m-M_V}$  we obtain for the *Hyades*. While the estimate of  $\theta_{m-M_V}$  fit using the Dotter *et al.* (2008) model matched the Hipparcos value well, that obtained with Yi *et al.* (2001) is lower than expected. To address this we reran both models with  $\theta_{m-M_V}$  fixed at 0.0, corresponding to an adjusted distance of 10 parsecs.

(See the Appendix and DeGennaro (2009) for more on the distance-correction procedure.) In Figure 6 we compare the resulting posterior distribution of  $\theta_{\text{age}}$  under the two MS/RG models. The solid red curve corresponds to the MS/RG model of Dotter *et al.* (2008) and the solid black to that of Yi *et al.* (2001), without the constraint on  $\theta_{m-M_V}$ . The color-coded dashed lines give the posterior distributions for  $\theta_{\text{age}}$  with  $\theta_{m-M_V}$  fixed at 0. The posterior distribution for  $\theta_{\text{age}}$  is multimodal in all four cases and in three cases is younger than the generally accepted value of the Hyades age,  $625 \pm 50$  million years (Perryman *et al.*, 1998), shaded in gray in Figure 6. As expected, conditioning on the distance had little effect under the model of Dotter *et al.* (2008) but the effect is rather pronounced under the model of Yi *et al.* (2001).

The two models yield somewhat discrepant estimates of  $\theta_{\text{age}}$ . In this case, the model of Dotter *et al.* (2008) appears more trustworthy because it was able to match the external and reliable estimate of  $\theta_{m-M_V}$  obtained with Hipparcos. Although its estimate of  $\theta_{\text{age}}$  is somewhat less than that obtained by Perryman *et al.* (1998) it is based on more, newer, and more precise data and models. The uncertainty in the Perryman *et al.* (1998) estimate was obtained by running a progenitor computer model with different input ages and with other parameter values fixed at their fitted values. The resulting expected magnitudes were then plotted and compared with the observed magnitudes. Visual inspection of this plot was used to propose an initial measure of uncertainty that was then inflated based on intuition about the magnitude of other sources of error in the models and in data processing. This stands in stark contrast to our approach that delivers principled estimates of uncertainty under a coherent statistical model.

The joint posterior distributions of the white dwarf initial and final masses conditioning on  $\theta_{m-M_V} = 0$  appear in Figure 7. Comparing the *Hyades* fits in Figures 4 and 7 shows that fixing  $\theta_{m-M_V}$  results in much better agreement in terms of the posterior distribution of the initial and final masses between the two MS/RG computer models (despite their discrepancy in terms of  $\theta_{\text{age}}$ ). Given this improvement, we fixed  $\theta_{m-M_V}$  in all our final analyses of the *Hyades*.

Contrary to the assumption in (8), the errors in different bands are not independent for this data set. This issue is unique to the *Hyades*, which occupies such a large area of the sky that the data were collected one star at a time. That is, each star was observed through three filters before moving on to the next star, likely introducing correlations in photometric errors across filters. To address the possibility of correlated errors in the *Hyades*, we also fit this data set using a multivariate Student’s  $t$  distribution with 3 degrees of freedom, centered at  $\boldsymbol{\mu}_i$  and with scale  $\boldsymbol{\Sigma}_i$ , for the cluster star component of the likelihood. This can be viewed as first drawing a covariance matrix  $\mathbf{V}_i$  from an inverted Wishart distribution  $W^{-1}(\boldsymbol{\Sigma}_i, 5)$  with expectation  $E(\mathbf{V}_i) = 3\boldsymbol{\Sigma}_i$ , and then conditional on  $\mathbf{V}_i$ , drawing observations  $\mathbf{X}_i \mid \mathbf{V}_i \sim N(\boldsymbol{\mu}_i, \mathbf{V}_i)$ . Thus, in addition to treating observational errors conservatively, this model allows some flexibility to fit correlations among filters. As seen in Figure 8, while the inferences are sensitive to the choice of likelihood, our substantive conclusions regarding the agreement between the *Hyades* white dwarfs and previously published IFMRs change very little.

**M35 (NGC 2168)** M35 is a young cluster, and hence the progenitors of all of its white dwarfs must have had relatively high initial masses. (While the M35 results appear in the middle of the initial mass range plotted in Figure 4, this range is considered “high mass” by astronomers.) To analyze this cluster, we combined the white dwarf data of Williams *et al.* (2009) with the upper main sequence data from Sung and Bessell (1999). For both datasets, we used the U (ultraviolet), B (blue), and V (visual) photometric filters. Williams *et al.* (2009) provide a thorough study of the white dwarfs, including spectroscopic observations of 14 of them. We analyze all of these white dwarfs except LAWDS 11 because its cluster membership is in doubt (see Williams *et al.*, 2009). For the main sequence data, we included those stars that Sung and Bessell (1999) were able to match with the proper-motion study of McNamara and Sekiguchi (1986), which yields probabilities of cluster membership for each star system that we use as prior probabilities in our statistical model. We set the prior probabilities of cluster

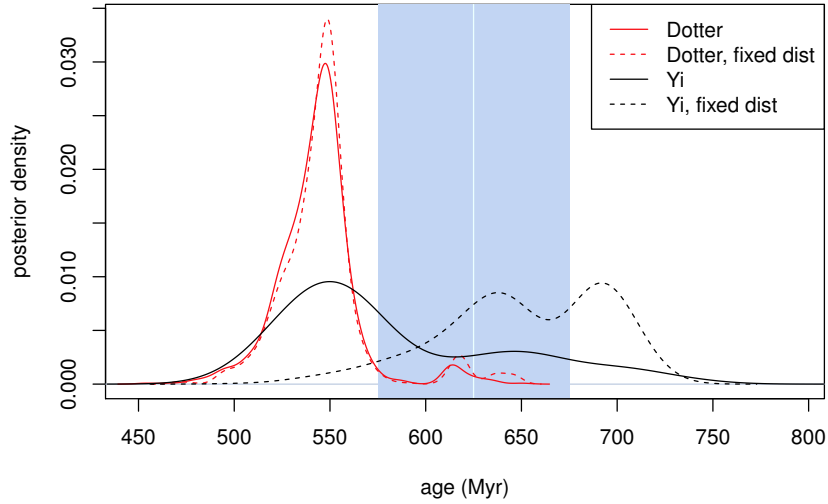


Figure 6: Sensitivity of the *Hyades* cluster age inferences to fixing the distance modulus. “Dotter, fixed dist” and “Yi, fixed dist” are results from fitting the remaining parameters while fixing the adjusted cluster distance to 10.0 parsecs. The generally accepted age of the *Hyades* along with its somewhat questionable uncertainties (Perryman *et al.*, 1998) are shown in gray.

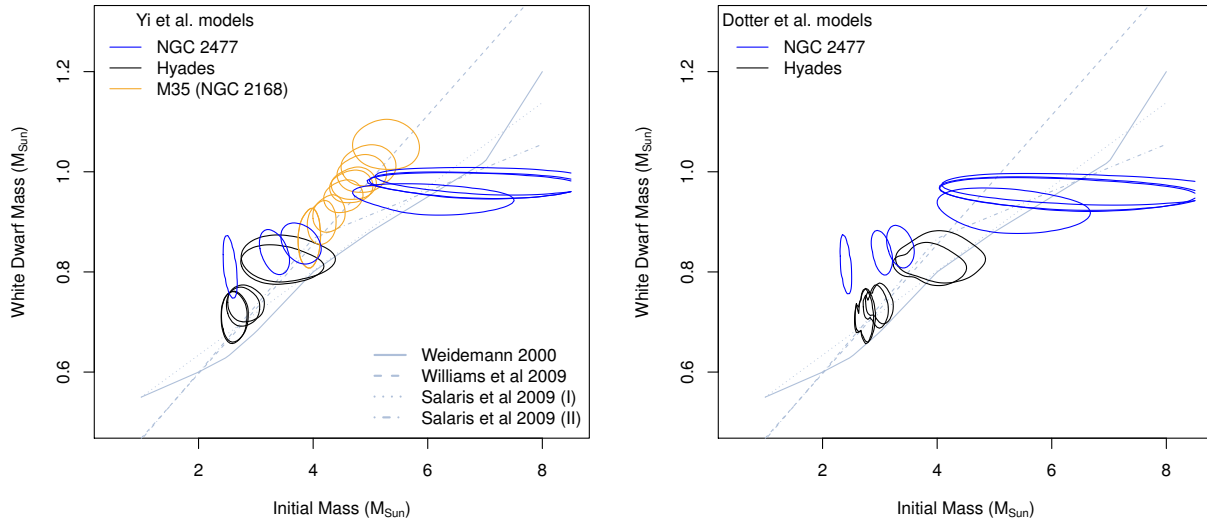


Figure 7: An alternative version of Figure 4, with initial-final mass posterior distributions for the *Hyades* computed conditional on its adjusted distance,  $\theta_{m-M_V} = 0$ , i.e., 10 parsecs.

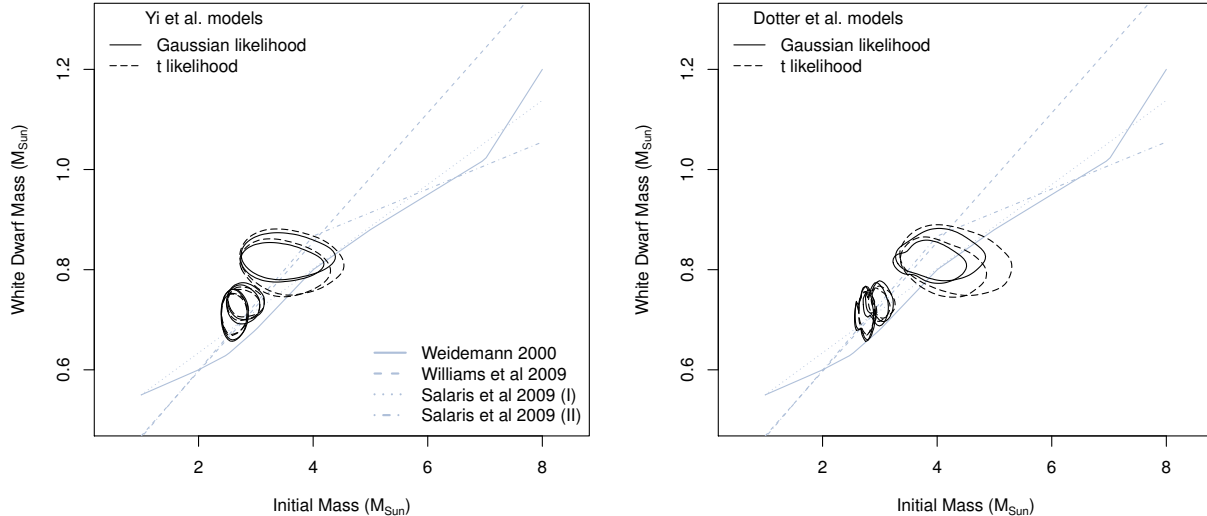


Figure 8: Sensitivity of *Hyades* white dwarf initial and final masses to the choice of likelihood. The solid contours show the results from the likelihood in (8). The dashed contours are the results from replacing the multivariate Gaussian cluster star component of the likelihood with a multivariate  $t$  distribution with 3 degrees of freedom.

membership equal to one for all of the white dwarfs included in our dataset. Allowing the white dwarfs to become field stars resulted in numerical difficulties. For cluster parameter prior distributions, we used

$$\begin{aligned}\theta_{[\text{Fe}/\text{H}]} &\sim N(-0.2, 0.3^2), \\ \theta_{m-M_V} &\sim N(10.3, 0.1^2), \text{ and} \\ \theta_{A_V} &\sim TN(0.682, 0.1^2; 0),\end{aligned}$$

which are the values adopted in Williams *et al.* (2009), with a more conservative uncertainty on  $\theta_{[\text{Fe}/\text{H}]}$ .

For M35, we cannot compare the results of the different MS/RG models because, at approximately 200 million years, M35 is too young for the Dotter *et al.* (2008) models. Results using the Yi *et al.* (2001) models are displayed in Figures 4 and 7. The MCMC code took approximately 10.2 hours to run, using 16 cores. Our results for M35 agree remarkably well with the IFMR of Williams *et al.* (2009), while diverging from the results for NGC 2477 and from the IFMRs of Weidemann (2000) and Salaris *et al.* (2009). This discrepancy may be due to concerns with the NGC 2477 white dwarf data or possible cluster differences not accounted for by the IFMR. See Section 5 for discussion.

While it is perhaps unsurprising that our M35 results agree with the IFMR of Williams *et al.* (2009) because we are using observations of the same white dwarfs as Williams *et al.* (2009) to constrain this high-mass portion of the IFMR, we have only used their photometric magnitudes and cluster membership information, whereas their work determined white dwarf final masses spectroscopically. That is, our method of coherently linking the physical computer models for different stages of stellar evolution agrees with results that rely on both additional data and additional physical models. This is a powerful demonstration of the value that principled

modeling of computer models can add: we can obtain sensible results, even with less data than traditional methods may demand. We interpret this as empirical evidence that our requirement of coherence among the model assumptions at different evolutionary stages allows us to extract more information out of the physical models than would be possible by performing separate analyses of the different evolutionary stages and using *ad hoc* methods to combine them.

## 5 Conclusion

Our ultimate goal for the IFMR is to model the conditional distribution of the final mass given the initial mass in a robust non-parametric manner, perhaps including other covariates that affect the conditional distribution such as measures of the star’s composition. In this first analysis we focus on the simple deterministic linear model given in (6). Our hope is that this linear form will suffice over the relatively narrow ranges of the initial mass that are typically observed in a single cluster and that by analyzing a number of clusters of differing ages we can begin to piece together the conditional distribution, or more simply, the overall deterministic relationship. Unfortunately the situation is a bit more complicated. The disagreement that we see between our analysis of the *Hyades* and M35 on the one hand and of NGC 2477 on the other mirrors the disagreement between the published IFMR of Williams *et al.* (2009) and those of Weidemann (2000), and Salaris *et al.* (2009). NGC 2477 poses particular difficulties because we know less about its white dwarfs than we do about those of the *Hyades* or M35. In particular the cluster membership and atmosphere type (hydrogen versus helium dominated; see Footnote 2) are much better established for the *Hyades* and M35 white dwarfs. Thus one explanation for the discrepancy is that one of the supposed NGC 2477 white dwarfs is actually a field star or has a helium dominated atmosphere. Another possibility is that there may be peculiarities of the particular clusters (e.g., metallicity or other measures of composition) that affect the IFMR and that the IFMR parameters should be allowed to vary as random effects between clusters.

Ideally, the analyses from multiple clusters could be combined into a single analysis of a hierarchical model. This would allow for shrinkage of the IFMR parameters and the sharing of information for the IFMR across clusters. Furthermore, as pointed out by a reviewer, there may also be an advantage to embedding both the Dotter *et al.* (2008) and Yi *et al.* (2001) components in a larger statistical model that accounts for uncertainty in the underlying deterministic computer models. In either case, the appropriate formulation and parameterization requires careful intermediate analyses of the within-cluster IFMR and how the IFMR varies among clusters. We hope that the methods proposed in this article will help in the formulation of such an omnibus model.

## Acknowledgments

We are grateful to two anonymous reviewers and the review editor for many helpful comments that greatly improved this paper. Professors van Dyk and von Hippel’s research was supported in part by NASA under grant 10-ADAP10-0076. In addition Professor van Dyk was supported in part by NSF grant DMS-09-07522, by a British Royal Society Wolfson Research Merit Award, and by the STFC (UK). N. Stein was supported in part by NSF grant DMS-09-07185. We are grateful to Kurtis Williams for helpful discussions and for providing us with M35 white dwarf data. The computations in this article were run on the Odyssey cluster supported by the FAS Science Division Research Computing Group at Harvard University.

## Appendix

Because the stars in a cluster are spatially near each other, one of the assumptions we make is that their distances from us are all the same. Thus  $\theta_{m-M_V}$  need not be subscripted by  $i$ . Of course, the distances to individual stars within a cluster are not identical, yet we typically assume that any such differences are negligible relative to the overall distance from our Solar System. Because the *Hyades* is so close to us, however, the variation in the distances from the Solar System to the individual cluster members is an appreciable fraction (approximately 20-40%) of the distance from the Solar System to the center of the cluster (DeGennaro, 2009). Thus, we cannot assume that all of the distances are equal. Stars that are closer to us appear too bright and stars that are farther away appear too dim when compared to the predictions of the stellar evolution computer models. To address this issue, DeGennaro (2009) adjusted the magnitudes of each star for its distance using the ultra-precise distance estimates obtained with the Hipparcos Satellite (Perryman *et al.*, 1997). In particular, the differences in the perceived brightness that the various individual distances would imply were computed and the observed magnitudes were adjusted for these differences.

Rather than offsetting each individual *Hyades* star to the distance of the cluster center ( $46.34 \pm 0.27$  pc or  $\theta_{m-M_V} = 3.33 \pm 0.01$ , Perryman *et al.* (1998)), each *Hyades* star was offset to a nominal distance modulus,  $\theta_{m-M_V} = 0.0$ , equivalent to 10 pc. This was partially done for convenience and partially because Perryman *et al.* (1998) measured the cluster center distance from Hipparcos trigonometric parallaxes whereas DeGennaro (2009) determined individual stellar distances from Hipparcos proper motions via the moving cluster method. This allowed DeGennaro (2009) to avoid adding in the small uncertainty in the trigonometric parallax distances as well as including any possible unknown systematic from these data.

## References

- Bergeron, P., Wesemael, F., and Beauchamp, A. (1995). Photometric calibration of hydrogen- and helium-rich white dwarf models. *Publications of the Astronomical Society of the Pacific* **107**, 1047–1054.
- Caputo, F., Chieffi, A., Castellani, V., Collados, M., Roger, C. M., and Paez, E. (1990). CCD photometry of stars in the old open cluster NGC 188. *The Astronomical Journal* **99**, 261–272.
- Catelan, M. (2009). Horizontal branch stars: The interplay between observations and theory, and insights into the formation of the Galaxy. *Astrophysics and Space Science* **230**, 261.
- Chaboyer, B., Demarque, P., and Sarajedini, A. (1996). Globular cluster ages and the formation of the galactic halo. *The Astrophysical Journal* **459**, 558–569.
- DeGennaro, S. (2009). *White Dwarfs and the Ages of Stellar Populations*. Ph.D. thesis, The University of Texas at Austin, Austin, Texas.
- DeGennaro, S., von Hippel, T., Jefferys, W. H., Stein, N., van Dyk, D. A., and Jeffery, E. (2009). Inverting color-magnitude diagrams to access precise cluster parameters: A new white dwarf age for the Hyades. *The Astrophysical Journal* **696**, 1, 12–23.
- Dinescu, D. I., Demarque, P., Guenther, D. B., and Pinsonneault, M. H. (1995). The ages of the disk clusters NGC 188, M67, and NGC 752, using improved opacities and cluster membership data. *The Astronomical Journal* **109**, 2090–2095.



- Dotter, A., Chaboyer, B., Jevremovic, D., Kostov, V., Baron, E., and Ferguson, J. W. (2008). The Dartmouth stellar evolution database. *The Astrophysical Journal Supplement Series* **178**, 89–101.
- Ferrario, L., Wickramasinghe, D., Liebert, J., and Williams, K. A. (2005). The open-cluster initial-final mass relationship and the high-mass tail of the white dwarf distribution. *Monthly Notices of the Astronomical Society* **361**, 1131–1135.
- Gelman, A., Roberts, G., and Gilks, W. (1996). Efficient Metropolis jumping rules. In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, eds., *Bayesian Statistics*, vol. 5, 599. Oxford University Press.
- Higdon, D., Gattiker, J., Williams, B., , and Rightley, M. (2008). Computer model calibration using high-dimensional output. *Journal of the American Statistical Association* **103**, 482, 570–583.
- Jao, W.-C., Mason, B. D., Hartkopf, W. I., Henry, T. J., and Ramos, S. N. (2009). Cool subdwarf investigations (csi) II: Multiplicity. *The Astronomical Journal* **137**, 3800.
- Jeffery, E. J., von Hippel, T., DeGennaro, S., van Dyk, D. A., Stein, N., and Jefferys, W. H. (2011). The white dwarf age of NGC 2477. *The Astrophysical Journal* **730**, 1, 35.
- Kalirai, J. S., Hansen, B. M. S., Kelson, D. D., Reitzel, D. B., Rich, R. M., and Richer, H. B. (2008). The initial-final mass relation: Direct constraints at the low-mass end. *The Astrophysical Journal* **676**, 594.
- Kalirai, J. S., Richer, H. B., Reitzel, D., Hansen, B. M. S., Rich, R. M., Fahlman, G. G., Gibson, B. K., and von Hippel, T. (2005). The initial-final mass relationship: Spectroscopy of white dwarfs in NGC 2099 (M37). *The Astrophysical Journal* **618**, L123.
- Kennedy, M. and O’Hagan, A. (2001). Bayesian calibration of computer models (with discussion). *Journal of the Royal Statistical Society, Series B* **68**, 425–464.
- Kleinman, S. J., Harris, H. C., Eisenstein, D. J., Liebert, J., Nitta, A., Krzesiński, J., Munn, J. A., Dahn, C. C., Hawley, S. L., Pier, J. R., Schmidt, G., Silvestri, N. M., Smith, J. A., Szkody, P., Strauss, M. A., Knapp, G. R., Collinge, M. J., Mukadam, A. S., Koester, D., Uomoto, A., Schlegel, D. J., Anderson, S. F., Brinkmann, J., Lamb, D. Q., Schneider, D. P., and York, D. G. (2004). A Catalog of Spectroscopically Identified White Dwarf Stars in the First Data Release of the Sloan Digital Sky Survey. *The Astrophysical Journal* **607**, 426–444.
- Liu, J. S., Wong, W. H., and Kong, A. (1995). Correlation structure and convergence rate of the Gibbs sampler with various scans. *Journal of the Royal Statistical Society, Series B, Methodological* **57**, 157–169.
- McNamara, B. and Sekiguchi, K. (1986). A proper-motion analysis of the cluster M35. *The Astronomical Journal* **91**, 557–562.
- Mermilliod, J.-C., Mermilliod, M., and Hauck, B. (1997). The general catalogue of photometric data (gcpd). II. *Astronomy and Astrophysics Supplement Series* **124**, 349.
- Montgomery, K. A., Marschall, L. A., and Janes, K. A. (1993). CCD photometry of the old open cluster M67. *The Astronomical Journal* **106**, 181–219.
- O’Hagan, A. (2006). Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering and System Safety* **91**, 1290–1300.

- Perryman, M. A. C., Brown, A. G. A., Lebreton, Y., Gomez, A., Turon, C., de Strobel, G. C., Mermilliod, J. C., Robichon, N., Kovalevsky, J., and Crifo, F. (1998). The Hyades: distance, structure, dynamics, and age. *Astronomy and Astrophysics* **331**, 81–120.
- Perryman, M. A. C., Lindegren, L., Kovalevsky, J., Hog, E., Bastian, U., Bernacca, P. L., Creze, M., Donati, F., Grenon, M., Grewing, M., van Leeuwen, F., van der Marel, H., Mignard, F., Murray, C. A., Le Poole, R. S., Schrijver, H., Turon, C., Arenou, F., Froeschle, M., and Petersen, C. S. (1997). The Hipparcos catalogue. *Astronomy and Astrophysics* **323**, L49.
- Raftery, A. E., Givens, G. H., and Zeh, J. E. (1995). Inference from a deterministic population dynamics model for bowhead whales (with discussion). *Journal of the American Statistical Association* **90**, 402–430.
- Rosvick, J. M. and Vandenberg, D. A. (1998). BV photometry for the  $\sim 2.5$  Gyr open cluster NGC 6819: More evidence for convective core overshooting on the main sequence. *The Astronomical Journal* **115**, 1516–1523.
- Rougier, J. C. (2008). Efficient emulators for multivariate deterministic functions. *Journal of Computational and Graphical Statistics* **17**, 4, 827–843.
- Salaris, M., Serenelli, A., Weiss, A., and Bertolami, M. M. (2009). Semi-empirical white dwarf initial-final mass relationships: A thorough analysis of systematic uncertainties due to stellar evolution models. *The Astrophysical Journal* **692**, 1013–1032.
- Sarajedini, A., von Hippel, T., Kozhurina-Platais, V., and Demarque, P. (1999). WIYN open cluster study. II. UBVR CCD photometry of the open cluster NGC 188. *The Astronomical Journal* **118**, 2894–2907.
- Sollima, A., Carballo-Bello, J. A., Beccari, G., Ferraro, F. R., Pecci, F. F., and Lanzoni, B. (2010). The fraction of binary systems in the core of five Galactic open clusters. *Monthly Notices of the Royal Astronomical Society* **401**, 577.
- Sung, H. and Bessell, M. S. (1999). UBVRI CCD photometry of M35 (NGC 2168). *Monthly Notices of the Royal Astronomical Society* **306**, 361–370.
- van Dyk, D. A., DeGennaro, S., Stein, N., Jefferys, W. H., and von Hippel, T. (2009). Statistical analysis of stellar evolution. *The Annals of Applied Statistics* **3**, 1, 117–143.
- Vandenberg, D. A. and Stetson, P. B. (2004). On the old open clusters M67 and NGC 188: Convective core overshooting, color-temperature relations, distances, and ages. *Publications of the Astronomical Society of the Pacific* **116**, 997–1011.
- von Hippel, T., Jefferys, W. H., Scott, J., Stein, N., Winget, D. E., DeGennaro, S., Dam, A., and Jeffery, E. (2006). Inverting color-magnitude diagrams to access precise star cluster parameters: A Bayesian approach. *The Astrophysical Journal* **645**, 2, 1436.
- Weidemann, V. (2000). Revision of the initial-to-final mass relation. *Astronomy and Astrophysics* **363**, 647–656.
- Williams, K. A., Bolte, M., and Koester, D. (2004). An empirical initial-final mass relation from hot, massive white dwarfs in NGC 2168 (M35). *The Astrophysical Journal* **615**, L49.
- Williams, K. A., Bolte, M., and Koester, D. (2009). Probing the lower mass limit for supernova progenitors and the high-mass end of the initial-final mass relation from white dwarfs in the open cluster M35 (NGC 2168). *The Astrophysical Journal* **693**, 355–369.

- Wood, M. A. (1992). Constraints on the age and evolution of the Galaxy from the white dwarf luminosity function. *The Astrophysical Journal* **386**, 539–561.
- Yi, S., Demarque, P., Kim, Y.-C., Lee, Y.-W., Ree, C. H., Lejeune, T., and Barnes, S. (2001). Toward better age estimates for stellar populations: The  $Y^2$  isochrones for solar mixture. *The Astrophysical Journal Supplement Series* **136**, 417–437.