

CONSTRAINING THE INITIAL FINAL MASS RELATION: USING WIDE DOUBLE WHITE DWARFS

1. Introduction

White dwarfs (WDs) are formed from the complete evolution of intermediate mass stars from the Main Sequence through hydrogen burning, helium burning, and through mass loss on the Asymptotic Giant Branch (AGB). Since roughly half of all stars are found in some form of binary system, WD binaries are the endpoints of many stellar evolution channels. In particular, binaries composed of stars both massive enough to have completely evolved within a Hubble time, but physically separated enough to have never interacted through mass transfer or tidal effects will evolve into wide double WDs (DWDs).

With quality spectra, wide DWDs can place constraints on the initial-final mass relation (IFMR), providing the mass of a WD a given initial mass star will evolve into. Fitting WD spectra to templates provide measurements of $\log g$ and T_{eff} which can then be mapped into a WD mass and cooling age. Assuming that wide DWDs are coeval and independent implies $\Delta\tau_{\text{cool}} = \Delta\tau_{\text{pre-WD}}$. $\Delta\tau_{\text{cool}}$ is readily determined from the difference in the derived WD cooling ages. For a particular IFMR, the initial mass can be determined from the derived WD masses. Stellar evolution codes provide the pre-WD lifetime as a function of initial mass. Comparing the difference in the pre-WD lifetimes with the cooling age differences then provides a measure of the accuracy of the assumed IFMR used.

Rather than produce constraints on the relation itself, we are interested in constraining physical relevant stellar evolution parameters. In particular, these involve the core mass at the first thermal pulse, and wind mass loss rate, and the efficiency of the so-called third dredge up during the thermally pulsing AGB phase. We construct this problem using a perturbative approach: we use the values of these parameterizations expected from stellar evolution simulations and apply linear perturbations. This methodology allows us to physically motivated models while using relatively few parameters.

2. Statistical Method

We wish to apply a parametric model for the IFMR, where θ are our model parameters and D are the set of observables, in this case the individual WD masses and cooling ages, $M_1, M_2, \tau_{\text{cool},1}$ and $\tau_{\text{cool},2}$. Our method relies on two primary assumptions. First, that the WDs in wide DWDs are coeval and truly independent, and second that the same IFMR can be applied to all the WDs in our sample. We address each of these assumptions in Section **assumptions**. If these assumptions are satisfied, we can conclude that the difference in the WD cooling ages in the pair must be equal to the negative of the difference in the pre-WD lifetimes:

$$\tau_{\text{cool},1} - \tau_{\text{cool},2} = \tau_2 - \tau_1. \quad (1)$$

Here, $\tau_{\text{cool},1}$ and $\tau_{\text{cool},2}$ are directly observed and τ_1 and τ_2 can be obtained by a functional

transformation from $M_{\text{WD},1}$ and $M_{\text{WD},2}$, respectively:

$$\tau_1 = f(M_{\text{ZAMS}}) \quad (2)$$

$$= f[g^{-1}(M_{\text{WD}})] \quad (3)$$

where $f(M_{\text{ZAMS}})$ is the stellar lifetime function, and $g(M_{\text{ZAMS}})$ is the IFMR. We have applied the inverse IFMR, $g^{-1}(M_{\text{WD}})$ to obtain M_{ZAMS} from the observed WD masses. Combining these, we obtain:

$$\tau_{\text{cool},1} - \tau_{\text{cool},2} = f[g^{-1}(M_{\text{WD},2})] - f[g^{-1}(M_{\text{WD},1})]. \quad (4)$$

Since $f(M_{\text{ZAMS}})$ can be determined with accuracy from stellar evolution codes, constraining the IFMR now involves finding the best $g(M_{\text{ZAMS}})$ to satisfy the Equation 4.

We now construct our statistical model to find the optimal parametric model for the IFMR. We begin with Bayes's rule:

$$P(\boldsymbol{\theta} | \mathbf{D}) = \frac{P(\mathbf{D} | \boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\mathbf{D})}, \quad (5)$$

where $P(\boldsymbol{\theta} | \mathbf{D})$ ¹ is the posterior probability we are looking for, $P(D)$ is a constant dependent only on the data, $P(\boldsymbol{\theta})$ are the priors on our model, and $P(\mathbf{D} | \boldsymbol{\theta})$ is the likelihood. The posterior probability over the set of data is a product over individual measurements:

$$P(\mathbf{D} | \boldsymbol{\theta}) = \prod_i P(D | \boldsymbol{\theta}). \quad (6)$$

We now substitute in the individual observables:

$$P(D | \boldsymbol{\theta}) = P(M_1, M_2, \Delta\tau_{\text{cool}} | \boldsymbol{\theta}). \quad (7)$$

To construct our likelihood function, we first marginalize over the true pre-WD lifetimes, τ_1' and τ_2' , and the true WD masses, M_1' and M_2' :

$$P(D | \boldsymbol{\theta}) = \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty P(\tau_1', \tau_2', M_1', M_2', M_1, M_2, \Delta\tau_{\text{cool}} | \boldsymbol{\theta}) d\tau_1' d\tau_2' dM_1' dM_2'. \quad (8)$$

Hereafter, primed quantities refer to true values and unprimed quantities refer to observed parameters. Based on independence, this can be reduced:

$$P(D | \boldsymbol{\theta}) = \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty P(\Delta\tau_{\text{cool}} | \tau_1', \tau_2', \boldsymbol{\theta}) P(\tau_1', M_1' | M_1, \boldsymbol{\theta}) P(\tau_2', M_2' | M_2, \boldsymbol{\theta}) \\ \times P(M_1, M_2) d\tau_1' d\tau_2' dM_1' dM_2'. \quad (9)$$

¹Bold symbols refer to sets of quantities.

Now we will split $P(\tau_1', M_1' | M_1, \boldsymbol{\theta})$ and the corresponding probability for the second WD:

$$P(\tau_1', M_1' | M_1, \boldsymbol{\theta}) = P(\tau_1' | M_1', \boldsymbol{\theta}) P(M_1' | M_1) \quad (10)$$

$$P(\tau_1', M_2' | M_2, \boldsymbol{\theta}) = P(\tau_1' | M_2', \boldsymbol{\theta}) P(M_2' | M_2). \quad (11)$$

These can be combined with equations 9 to produce:

$$\begin{aligned} P(D | \boldsymbol{\theta}) &= P(M_1, M_2) \int_0^\infty \int_0^\infty dM_1' dM_2' P(M_1' | M_1) P(M_2' | M_2) \\ &\times \int_0^\infty \int_0^\infty d\tau_1' d\tau_2' P(\Delta\tau_{\text{cool}} | \tau_1', \tau_2') P(\tau_1' | M_1', \boldsymbol{\theta}) P(\tau_2' | M_2', \boldsymbol{\theta}). \end{aligned} \quad (12)$$

The probabilities $P(\tau_1' | M_1', \boldsymbol{\theta})$ and $P(\tau_2' | M_2', \boldsymbol{\theta})$ are equal to delta functions:

$$P(\tau_1' | M_1', \boldsymbol{\theta}) = \delta \left[M_1' - f \left(g^{-1}(\tau_1') \right) \right] \quad (13)$$

$$P(\tau_2' | M_2', \boldsymbol{\theta}) = \delta \left[M_2' - f \left(g^{-1}(\tau_2') \right) \right]. \quad (14)$$

The inner integrals over τ_1' and τ_2' can then be reduced to:

$$\int_0^\infty \int_0^\infty d\tau_1' d\tau_2' P(\Delta\tau_{\text{cool}} | \tau_1', \tau_2') \delta \left[M_1' - f \left(g^{-1}(\tau_1') \right) \right] \delta \left[M_2' - f \left(g^{-1}(\tau_2') \right) \right]. \quad (15)$$

This expression is a double integral, each of the form:

$$\int dx F(x) \delta[G(x)] = \sum_j F(x_j^*) \left| \frac{\partial G(x_j^*)}{\partial x} \right|^{-1} \quad (16)$$

where the sum is over the roots, x_j^* , of $G(x)$. Since the stellar lifetime as a function of WD mass is a one-to-one correspondence, there is a single root to each of the delta functions in equation 15. The integrals over τ_1' and τ_2' can then be reduced to:

$$P(\Delta\tau_{\text{cool}} | \tau_1^*, \tau_2^*) \left| \frac{\partial \tau_1^*}{\partial M_1'} \right|^{-1} \left| \frac{\partial \tau_2^*}{\partial M_2'} \right|^{-1}. \quad (17)$$

where

$$\tau_1^* = g \left(f^{-1}(M_1) \right) \quad (18)$$

$$\tau_2^* = g \left(f^{-1}(M_2) \right) \quad (19)$$

This can be combined with equation 12 to produce:

$$\begin{aligned} P(D | \boldsymbol{\theta}) &= P(M_1, M_2) \int_0^\infty \int_0^\infty dM_1' dM_2' P(M_1' | M_1) P(M_2' | M_2) \\ &\times P(\Delta\tau_{\text{cool}} | \tau_1^*, \tau_2^*) \left| \frac{\partial \tau_1^*}{\partial M_1'} \right|^{-1} \left| \frac{\partial \tau_2^*}{\partial M_2'} \right|^{-1} \end{aligned} \quad (20)$$

These integrals can be approximated by sums:

$$P(D | \boldsymbol{\theta}) \approx P(M_1, M_2) \frac{1}{N} \frac{1}{M} \sum_{i=1}^N \sum_{j=1}^M P(\Delta\tau_{\text{cool}} | \tau_{1,i}^*, \tau_{2,j}^*) \left| \frac{\partial \tau_1^*}{\partial M_1'} \right|^{-1} \left| \frac{\partial \tau_2^*}{\partial M_2'} \right|^{-1} \quad (21)$$

where the WD masses are randomly drawn from Gaussian distributions around their observed values, M_1 and M_2 , with their associated observational uncertainties, then transformed into $\tau_{1,i}^*$ and $\tau_{2,j}^*$ using equations 18 and 19:

$$\tau_{1,i}^* = g\left(f^{-1}(M_1^*)\right); \quad M_1^* \sim \mathcal{N}(M_1, \sigma_{M_1}) \quad (22)$$

$$\tau_{2,j}^* = g\left(f^{-1}(M_2^*)\right); \quad M_2^* \sim \mathcal{N}(M_2, \sigma_{M_2}). \quad (23)$$

Since the WDs are independent, the sums in equation 21 can be combined:

$$P(D | \boldsymbol{\theta}) \approx P(M_1, M_2) \frac{1}{K} \sum_{k=1}^K P(\Delta\tau_{\text{cool}} | \tau_{1,k}^*, \tau_{2,k}^*) \left| \frac{\partial \tau_1^*}{\partial M_1'} \right|^{-1} \left| \frac{\partial \tau_2^*}{\partial M_2'} \right|^{-1} \quad (24)$$

The first term in the sum is simply a normal distribution centered around the observed cooling age difference, $\Delta\tau_{\text{cool}}$, obtained from the difference of the individual observed cooling ages:

$$P(\Delta\tau_{\text{cool}} | \tau_1', \tau_2') = \frac{\mathcal{N}(\tau_1^* - \tau_2^* | \Delta\tau_{\text{cool}}, \sigma_{\Delta\tau_{\text{cool}}})}{\mathcal{N}(\tau_1^* - \tau_2^* | \Delta\tau_{\text{cool}}, \sigma_{\Delta\tau_{\text{cool}}})}. \quad (25)$$

Here, $\sigma_{\Delta\tau_{\text{cool}}}$ is the quadrature sum of the individual uncertainties on $\tau_{\text{cool},1}$ and $\tau_{\text{cool},2}$:

$$\sigma_{\Delta\tau_{\text{cool}}} = \sqrt{\sigma_{\tau_{\text{cool},1}}^2 + \sigma_{\tau_{\text{cool},2}}^2} \quad (26)$$

Our prior on the WD masses, $P(M_1, M_2)$ should be flat:

$$P(M_1, M_2) \propto 1. \quad (27)$$

Here we must add a factor to ensure that degenerate models are equally likely. This is achieved by dividing the overall equation by the highest probability observational cooling age difference:

$$\frac{1}{\mathcal{N}(\tau_1^* - \tau_2^* | \Delta\tau_{\text{cool}}, \sigma_{\Delta\tau_{\text{cool}}})} \quad (28)$$

Finally, we need to normalize our likelihood. Combining equations 24, 25, and 28, and normalizing, we arrive at our likelihood function:

$$P(D | \boldsymbol{\theta}) \approx \frac{1}{K} \sum_{k=1}^K \frac{\mathcal{N}(\tau_1^* - \tau_2^* | \Delta\tau_{\text{cool}}, \sigma_{\Delta\tau_{\text{cool}}})}{\mathcal{N}(\tau_1^* - \tau_2^* | \Delta\tau_{\text{cool}}, \sigma_{\Delta\tau_{\text{cool}}})} \left| \frac{\partial \tau_1^*}{\partial M_1'} \right|^{-1} \left| \frac{\partial \tau_2^*}{\partial M_2'} \right|^{-1} \quad (29)$$

Finding the model parameters implied by our set of wide DWDs involves maximizing the likelihood function in Equation 29. We are interested in determining how precise of constraints we can place on these model parameters, therefore we choose a Monte Carlo technique rather than employing a maximum likelihood calculation. Specifically, we employ the Markov Chain Monte Carlo algorithm `emcee` (?) which implements an affine invariant, ensemble sampler algorithm to search the parameter space (?).