

A MEASUREMENT OF DIFFUSION IN 47 TUCANAE

JEREMY HEYL¹, HARVEY B. RICHER¹, ELISA ANTOLINI², RYAN GOLDSBURY¹, JASON KALIRAI^{3,4}, JAVIERA PARADA¹, AND PIER-EMMANUEL TREMBLAY³¹ Department of Physics and Astronomy, University of British Columbia, Vancouver, BC V6T 1Z1, Canada; heyjl@phas.ubc.ca, richer@astro.ubc.ca² Dipartimento di Fisica e Geologia, Università degli Studi di Perugia, I-06123 Perugia, Italia³ Space Telescope Science Institute, Baltimore MD 21218, USA⁴ Center for Astrophysical Sciences, Johns Hopkins University, Baltimore MD, 21218, USA; jkalirai@stsci.edu

Received 2015 February 6; accepted 2015 February 20; published 2015 April 30

ABSTRACT

Using images from the *Hubble Space Telescope* Wide-Field Camera 3, we measure the rate of diffusion of stars through the core of the globular cluster 47 Tucanae (47 Tuc) using a sample of young white dwarfs that were identified in these observations. This is the first direct measurement of diffusion due to gravitational relaxation. We find that the diffusion rate $\kappa \approx 10\text{--}13 \text{ arcsec}^2 \text{ Myr}^{-1}$ is consistent with theoretical estimates of the relaxation time in the core of 47 Tuc of about 70 Myr.

Key words: globular clusters: individual (47 Tuc) – Hertzsprung–Russell and C–M diagrams – stars: kinematics and dynamics – stars: Population II

1. INTRODUCTION

Globular clusters have long provided an amazing laboratory for stellar evolution and gravitational dynamics, and the nearby rich cluster, 47 Tuc (47 Tuc), has long been a focus of such investigations. The key point of this investigation is the interplay between these two processes. In particular, in the core of 47 Tuc, the timescale for stellar evolution and the timescale for dynamical relaxation are similar. The relaxation time in the core of 47 Tuc is about 70 Myr (Harris 1996). Meanwhile over a span of about 150 Myr, the most massive stars in 47 Tuc evolve from a red giant star with a luminosity of 2000 times that of the Sun to a white dwarf (WD) with a luminosity less than a tenth that of the Sun. Meanwhile the star loses about 40% of its mass, going from 0.9 to 0.53 solar masses. It is these young WDs that are the focus of this paper.

Although the core of 47 Tuc has been the focus of numerous previous investigations (e.g., McLaughlin et al. 2006; Knigge et al. 2008; Bergbusch & Stetson 2009), this is the first paper that combines the near-ultraviolet filters of the *Hubble Space Telescope* (*HST*) with a mosaic that covers the entire core of the cluster. Probing the core of the cluster in the ultraviolet is advantageous in several ways. First, the young WDs are approximately as bright as the upper main-sequence, giant, and horizontal branch stars at 225 nm, so they are easy to find. In fact the brightest WDs are among the brightest stars in the cluster and are as bright as the blue stragglers. Second, the point-spread function of *HST* is more concentrated in the ultraviolet helping, with confusion in the dense starfield that is the core of 47 Tuc.

In spite of these advantages, for all but the brightest stars, our data set suffers from incompleteness, which presents some unique challenges. We reliably characterize the incompleteness as a function of position and flux in the two bands of interest, F225W and F336W, throughout the color–magnitude diagram (CMD) and especially along the WD cooling sequence through the injection and recovery of about 10^8 artificial stars into the images. How we measure the completeness is described in detail in Section 3.1. The young WDs typically have a mass 40% less than their progenitors, so they are born with less kinetic energy than their neighbors, and two-body interactions

will typically increase the kinetic energy of young WDs over time and change their spatial distribution. We introduce a simple model for the diffusion of the young WDs through the core of the cluster (Section 3.2). To make the most of this unique data set, we have to include stars in our sample whose completeness rate is well below 50%. We have developed and tested statistical techniques to characterize the observational distribution of young WDs in flux and space to understand their motion through the cluster and their cooling (Sections 3.4–3.7) in the face of these potentially strong observational biases. Although these techniques are well-known especially in gamma-ray astronomy, they have never been applied to stellar populations in this way, so Section 3.8 presents a series of Monte Carlo simulations to assess the potential biases of these techniques and verify that these techniques are indeed unbiased in the face of substantial incompleteness within the statistical uncertainties. To establish the time over which the WDs dim we use a stellar evolution model outlined in Section 3.3. Section 4 describes the best-fitting models for the density and flux evolution of the WDs. Section 4.1 looks at the dynamic consequences of these results. Section 5 outlines future directions both theoretical and observational and the broader conclusions of this work.

2. OBSERVATIONS

A set of observations from the Advanced Camera for Surveys (Ford et al. 1998) and the Wide Field Camera 3 (WFC3, MacKenty 2012) on the *HST* of the core of the globular cluster 47 Tuc over 1 yr provides a sensitive probe of the stellar populations in the core of this globular cluster (Cycle 12 GO-12971, PI: Richer), especially the young WDs. Here we will focus on the observations with WFC3 in the UV filters F225W and F336W. The observations were performed over ten epochs from 2012 November to 2013 September. Each of the exposures in F225W was 1080 s, and the exposures in F336W were slightly longer at 1205 s. Each of the overlapping WFC3 images was registered onto the same reference frame and drizzled to form a single image in each band from which stars were detected and characterized, resulting in the CMD depicted in Figure 1.

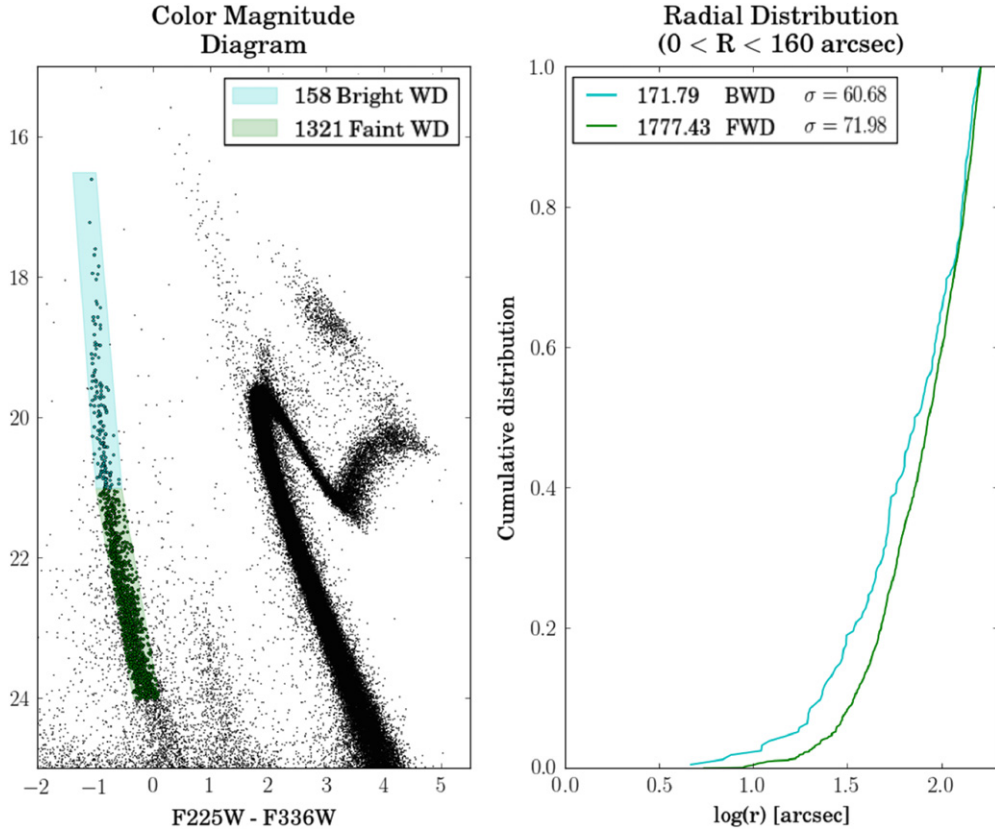


Figure 1. Left: color–magnitude diagram of the core of 47 Tuc in the WFC3 filters F225W and F336W. Right: the radial distribution of young and old white dwarf stars as highlighted in the color–magnitude diagram (completeness corrected numbers appear in the right panel).

What is immediately striking in Figure 1 is that the distribution of young WDs with a median age of 6 Myr is significantly more centrally concentrated than that of the older WDs that have a median age of 127 Myr. The WD distribution appears to become more radially diffuse with increasing age, a signature of relaxation. One concern is immediately apparent. The numbers of observed stars are given in the legend of the left panel and the numbers of stars in the completeness corrected samples are given in the legend of the right panel. The sample of older WDs is only about 75% complete on average. Furthermore, one would expect the completeness of these faint stars to be lower near the center of the cluster, so if the completeness is not accounted for correctly, one could naturally conclude that the WDs are diffusing when in reality they are not. In principle we would like to divide this sample of over 1,300 stars into subsamples, some of which will have even smaller completeness rates. How can we be sure that our analysis techniques are up to the task of measuring this diffusion accurately in the face of completeness rates as low as 20% that vary dramatically with distance from the center of the cluster?

In Section 3 we will characterize the completeness rate through artificial star tests, and develop and test statistical techniques to measure the diffusion of WDs in 47 Tuc without binning the stars at all, thus preserving the maximum information content of these data. We will test these new algorithms on mock data sets that include both the completeness rate and flux error distribution of our sample to verify that they robustly determine the diffusion and flux evolution of the WDs. The subsequent section (Section 4)

explores results of these techniques on the data set depicted in Figure 1.

3. ANALYSIS

3.1. Artificial Star Tests

We inserted $\sim 10^8$ artificial stars into the WFC3 images in F225W and F336W over the full range of observed magnitudes in both bands and a range of distances from the center of the cluster. To determine the completeness rate for the WDs that we have observed, we inserted artificial stars whose F225W and F336W magnitudes lie along the observed WD track in the CMD. The rate of recovering a star along the WD track of a given input magnitude in F336W at a given radius is the completeness rate and is depicted in Figure 2. If an artificial star along the WD track is detected in F336W, it is always detected in F225W as well. The completeness rate is both a strong function of radius and magnitude and is significantly different from unity except for the brightest stars, so accounting for completeness robustly is crucial in the subsequent analysis. The radial bins are 100 pixels in width and the magnitude bins are 0.1358 wide.

The magnitudes of the recovered stars give the error distribution as a function of the input magnitude and position of the star in the field. Furthermore, these distributions are not typically normal and often asymmetric as well. For the analysis in Section 3.5 we use the cumulative distribution of magnitude errors as a function of position and input magnitude, which we obtain by sorting the output magnitudes in a given bin and spline to obtain the cumulative distribution in the form of the

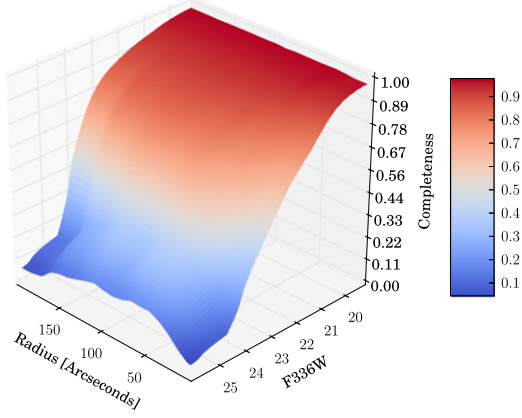


Figure 2. Completeness rate as a function of radius from the center of 47 Tuc and the magnitude of the artificial star.

values of the errors from the first to the ninety-ninth percentile. In the analysis the completeness rate is interpolated over the two dimensions of radius and magnitude with a third-degree spline, and the error distributions are interpolated linearly over the three dimensions (radius, magnitude, and percentile).

3.2. Diffusion and Luminosity Evolution

Sometime during the late evolution of a turn-off star in 47 Tuc, the star loses about 40% of its mass, going from a main-sequence star of 90% of a solar mass to a WD of 53% of a solar mass (Renzini & Fusi Pecci 1988; Renzini et al. 1996; Moehler et al. 2004; Kalirai et al. 2009). These newborn WDs will have the typical velocities of their more massive progenitors, so as they interact gravitationally with other stars, their velocities will increase through two-body relaxation, bringing their kinetic energies into equipartition (e.g., Spitzer 1987). Because the gravitational interaction is long-range and the distance between the stars is small compared to the size of the cluster, the change in velocity will be dominated by distant interactions and small random velocity jumps, i.e., the Coulomb logarithm is large, $\sim \ln N$, where N is the number of stars in the core of the cluster, $\sim 10^{5-6}$. These small jumps in velocity can be modeled as a random walk in velocity so the square of the velocity increases linearly in time and the relaxation time can be defined as $t_r = [d \ln v^2 / dt]^{-1}$. In the center of the cluster, the density of stars is approximately constant, so the gravitational potential has the approximate form

$$\phi = \frac{2}{3}\pi G \rho_c r^2. \quad (1)$$

By the virial theorem the mean kinetic energy of the WDs will equal the mean potential energy,

$$\frac{1}{2}\langle v^2 \rangle = \frac{2}{3}\pi G \rho_c \langle r^2 \rangle \quad (2)$$

so the square of the distance of the WDs from the center of the cluster will also increase linearly with time as a random walk; therefore, let us suppose that newly born WDs diffuse outward through the cluster following the diffusion equation

$$\frac{\partial \rho(r, t)}{\partial t} = \kappa \nabla^2 \rho(r, t), \quad (3)$$

where κ can be related to the relaxation time as $\kappa = r_c^2 / t_r$ because $d \ln r^2 / dt = d \ln v^2 / dt$ from Equation (2). This diffusion equation yields the Green's function

$$u(r, t) = \frac{1}{8(\pi \kappa t)^{3/2}} e^{-r^2 / (4 \kappa t)} \quad (4)$$

if κ is independent of time and position. This gives a cumulative distribution in projected radius

$$C(<R) = 1 - e^{-R^2 / (4 \kappa t)}. \quad (5)$$

The Green's function at $t = 0$ is a delta function centered on the center of the cluster. On the other hand, if the initial distribution is a Gaussian centered on the center of the cluster the density of WD stars near the center of the cluster is a function of age, t , and projected radius, R , of the form

$$\rho = \rho(R, t) = \frac{2R}{4\kappa(t + t_0)} \exp\left[-\frac{R^2}{4\kappa(t + t_0)}\right] \quad (6)$$

where the density distribution is normalized as

$$\int_0^\infty \rho(R, t) dR = 1. \quad (7)$$

The dispersion of the Gaussian at $t = 0$ is simply given by $2\sigma^2 = 4\kappa t_0$. Because the diffusion equation is linear, a sum of several Gaussians with the same value of κ but different normalizations and values of t_0 will also be a solution.

Of course we do not directly observe the ages of the WDs. Rather we observe their fluxes or apparent magnitudes. The cooling curve of the WDs is a relationship between time and the apparent magnitude from the WDs $t(m)$, so the number of WDs that we expect to observe at a given flux and radius is given by

$$f(R, m) = \dot{N} \rho(R, t(m)) \frac{\partial t}{\partial m} C(R, m), \quad (8)$$

where \dot{N} is the birthrate of the WDs (assumed to be constant over the range of ages of the young WDs, i.e., the past 200 Myr) and $C(R, m)$ is the completeness as a function of radius and flux. To this point flux errors have been neglected.

3.3. Cooling Models

To construct the various cooling models here, i.e., $t(m)$ from Equation (8), we used Modules for Experiments in Stellar Astrophysics (MESA; Paxton et al. 2011) to perform simulations of stellar evolution starting with a pre-main-sequence model of 0.9 solar masses and a metallicity of $Z = 3.3 \times 10^{-3}$ appropriate for the cluster 47 Tuc. This is slightly larger than the value for the turnoff mass found by Thompson et al. (2010) for the eclipsing binary V69 in 47 Tuc that is composed of an upper main sequence star of about 0.86 solar masses and a subgiant of 0.88 solar masses. Because we are interested in the stars that have become young WDs just recently, the initial masses of these stars should be slightly larger than the turnoff mass today. We have explicitly assumed that the progenitors of the WDs are a uniform population. Although there is evidence of modest variation in the chemical abundances in 47 Tuc (e.g., Milone et al. 2012), the WD cooling sequence, at least at larger radii, appears uniform (Richer et al. 2013). However, from Figure 1 it is apparent that

the core of 47 Tuc has a substantial population of blue stragglers that will evolve to become more massive WDs. Our sample has about 160 blue stragglers, and if we estimate the duration of the main-sequence for a blue straggler to be 1–2 Gyr (e.g., Sills et al. 2009), we obtain a birth-rate of blue-straggler WDs of about 0.1 Myr^{-1} . The number of giants in our field indicates a birth rate of about eight WDs per million years (see Goldsbury et al. 2012 for further details), so the estimated contamination of the WD cooling sequence is modest at about 1%.

Specifically, we used SVN revision 5456 of MESA and started with the model `1M_pre_ms_to_wd` in the test suite. We changed the parameters `initial_mass` and `initial_z` of the star and adjusted the parameter `log_L_lower_limit` to -6 so the simulation would run well into the WD cooling regime. We also reduced the two values of the wind η to 0.46 (from the default of 0.7) to yield a 0.53 solar mass WD from the 0.9 solar mass progenitor. Interestingly Miglio et al. (2012) argue from *Kepler* asteroseismic measurements of the stars in the metal-rich open cluster NGC 6971 that such values of η are needed to account for the mass loss between the red giant and red clump phases of stars in this metal-rich cluster.

We defined the time of birth of the WD to coincide with the peak luminosity of the model at the tip of the asymptotic giant branch about 10.9 Gyr after the start of the simulation. This is in agreement with the best age of the cluster determined from main-sequence stars of 11.25 ± 0.21 (random) ± 0.85 (systematic) Gyr (Thompson et al. 2010). This age agrees with that derived by Hansen et al. (2013) from WD cooling (9.9 ± 0.7 Gyr at 95% confidence). We choose this definition of the birth so that each observed WD will have a star of similar luminosity in the cooling model. At this point in the evolution we have outputs from the MESA evolution every 100 yr or so; therefore, the cooling curve is well-sampled throughout. At each output time we have the value of the luminosity, radius, effective temperature, and mass of the star. With these values we interpolate the spectral models of Tremblay et al. (2011) in surface gravity and effective temperature, and then scale the result to the radius of the model star. We use a true distance modulus of 13.23 (Thompson et al. 2010) and a reddening of $E(B - V) = 0.04$ (Salaris et al. 2007) to determine the model fluxes in the WFC3 band F336W. We used the standard extinction curve of Fitzpatrick (1999) with $A_V/E(B - V) = 3.1$. We have purposefully used a distance and reddening determined from main-sequence stars to avoid a potential circularity in using the WD models themselves to fix the distance. Woodley et al. (2012) inferred a slightly larger true distance modulus of $13.36 \pm 0.02 \pm 0.06$ from the WD spectral energy distributions.

The brightest WD in our sample has $F336W = 14.92$. According to the models this corresponds to an age of 110,000 yr, an effective temperature of 100,000 K, a luminosity of $1600 L_\odot$, and a radius of $0.13 R_\odot$. Its mass is 0.53 solar masses. The faintest WD in our sample has $F336W = 25.4$, yielding an age of 1.2 Gyr, an effective temperature of 8700 K, a luminosity of $10^{-3} L_\odot$, and a radius of $0.013 R_\odot$, one-tenth of the radius of the brightest WD. Clearly the brightest WD in our sample is not a WD in the usual sense because thermal energy plays an important role in the pressure balance of the star. For this brightest star $\log g = 5.93$, which is less than the minimum of the atmosphere model grid ($\log g = 6$) so we have to

extrapolate slightly off of the grid, but only for this brightest star. For the simulations in Section 3.8 we did not use this particular model, but similar ones of the same WD mass with different neutrino cooling rates or initial metallicities also generated with MESA.

3.4. Likelihood Function

The model outlined in Section 3.2 predicts the number of WDs as a function of magnitude and position. Let us divide the space of position and magnitude into bins of width ΔR and Δm where the bins are numbered with indices j and k , respectively. The probability of finding n stars in a particular bin is given by

$$P(n; f(R_j, m_k)) = \frac{[f(R_j, m_k) \Delta R \Delta m]^n e^{-f(R_j, m_k) \Delta R \Delta m}}{n!}. \quad (9)$$

Now we imagine dividing the sample into so many bins that there is either a single star in a bin or no stars at all; we have

$$P(n; f(R_j, m_k)) = e^{-f(R_j, m_k) \Delta R \Delta m} \times \begin{cases} f(R_j, m_k) \Delta R \Delta m & \text{one star} \\ 1 & \text{no star} \end{cases}. \quad (10)$$

We can define the likelihood as the logarithm of the product of the probabilities of observing the number of stars in each bin. Since the bins are so small we can replace R_j and m_k for the bins with stars in them with the measured values for that particular star R_i and m_i . This gives the so-called “unbinned likelihood” of observing the sample as follows (Cash 1979; Mattox et al. 1996; Davis 2014):

$$\log L = \sum_i \log f(R_i, m_i) - \sum_{j,k} f(R_j, m_k) \Delta R \Delta m. \quad (11)$$

We have dropped the constant widths of the bins from the first term, which is a sum over the observed stars; consequently, the absolute value of the likelihood is not important, just differences matter. The second term is a sum over the really narrow (and arbitrary) bins that we have defined, so we have

$$\sum_{j,k} f(R_j, m_k) \Delta R \Delta m = \int f(R, m) dR dm = N_{\text{pred}}, \quad (12)$$

where N_{pred} is the number of stars that the model predicts that we will observe, so finally we have

$$\log L = \sum_i \log f(R_i, m_i) - N_{\text{pred}}, \quad (13)$$

where the summation is over the observed stars. The integral for N_{pred} when combined with Equation (8) yields

$$N_{\text{pred}} = \dot{N} \int_0^{t_1} \int_0^{r_{\text{max}}} \rho(R, t) C(R, m) dR dt \quad (14)$$

or

$$N_{\text{pred}} = \dot{N} \int_{m_0}^{m_1} \int_0^{r_{\text{max}}} \rho(R, t) C(R, m) \frac{dt}{dm} dR dm. \quad (15)$$

If we take the luminosity function as fixed and try to maximize the likelihood with respect to the diffusion model

$$\begin{aligned} \log L &= \sum_i \log \left[\dot{N} \rho(R, t(m_i)) \frac{\partial t}{\partial m} \Big|_i C(R_i, m_i) \right] - N_{\text{pred}} \quad (16) \\ &= \sum_i \log [\dot{N} \rho(R, t(m_i))] \\ &\quad + \sum_i \log \left[\frac{\partial t}{\partial m} \Big|_i C(R_i, m_i) \right] - N_{\text{pred}}, \quad (17) \end{aligned}$$

where the second summation does not depend on the diffusion model, so it is constant with respect to changes in the diffusion model and can be dropped from the logarithm of the likelihood. However, it must be included if one wants to compare different cooling curves, $t(m)$.

3.5. Magnitude Errors

An important complication to the analysis is that the measured magnitudes are not the same as the actual magnitudes of the stars; in particular the error distribution is not normal or even symmetric. This transforms the model distribution function via a convolution,

$$\begin{aligned} f'(R, m) &= \int_{-\infty}^{\infty} f(R, m') g(R, m', m - m') dm' \quad (18) \\ &= \int_{-\infty}^{\infty} f(R, m - \Delta m) g(R, m - \Delta m, \Delta m) d(\Delta m) \\ &= \int_0^1 f(R, m - \Delta m) dG, \quad (19) \end{aligned}$$

where

$$G(R, m, \Delta m) = \int_{-\infty}^{\Delta m} g(R, m - \Delta m, \Delta m') d(\Delta m') \quad (20)$$

is the cumulative distribution of magnitude errors with the observed radius and magnitude fixed. If we calculate the percentiles of the magnitude error distribution as Δm_j we can approximate the integral as the sum

$$f'(R, m) = \frac{1}{100} \sum_j f(R, m - \Delta m_j), \quad (21)$$

so for a given star i we have

$$\begin{aligned} f'(R_i, m_i) &= \dot{N} \sum_j \rho(R, t(m_i - \Delta m_j)) \\ &\quad \times t'(m_i - \Delta m_j) C(R_i, m_i - \Delta m_j), \quad (22) \end{aligned}$$

where $t' = \partial t / \partial m$. This new function $f'(R_i, m_i)$ can be substituted into Equation (13) to yield a likelihood including the magnitude errors. We will assume that the magnitude errors do not affect our estimate of N_{pred} ; this simplifies the analysis. We will verify our technique with Monte Carlo simulations in Section 3.8.

3.6. Constraining the Luminosity Function

We can construct a maximum likelihood estimator of the luminosity function of the WDs or alternatively the cooling

curve as follows

$$f(R, m) = \frac{dN}{dm dR} = \dot{N} \rho(R, t) \frac{\partial t}{\partial m} \quad (23)$$

$$= \dot{N} \rho(R, t) \sum_i A_i \delta(m - m_i), \quad (24)$$

where i is an index that runs over the observed stars. With this model we can define a likelihood function for the stars that we observe

$$\log L = \sum_i \log [\dot{N} A_i \rho(R_i, t_i)] - N_{\text{pred}}, \quad (25)$$

where a multiplicative constant (infinite in this case) and the completeness for each star have been dropped from the logarithm.

Substituting the trial luminosity function Equation (23) yields

$$N_{\text{pred}} = \sum_i A_i \dot{N} \int_0^{r_{\text{max}}} \rho(R, t_i) C(R, m_i) dR. \quad (26)$$

If we maximize the likelihood with respect to the values of A_k we obtain

$$\frac{\partial \log [\rho(R_i, t_i) C(R_i, m_i)]}{\partial A_k} = \frac{\partial \log [\rho(R_i, t_i)]}{\partial t_i} \frac{\partial t_i}{\partial A_k}, \quad (27)$$

where

$$\frac{\partial t_i}{\partial A_k} = \begin{cases} 1 & \text{if } k \leq i \\ 0 & \text{if } k > i \end{cases} \quad (28)$$

so

$$\begin{aligned} &\sum_i \frac{\partial \log [A_i \rho(R_i, t_i) C(R_i, m_i)]}{\partial A_k} \\ &= \frac{1}{A_k} + \sum_{i \geq k} \frac{\partial \log [\rho(R_i, t_i)]}{\partial t_i} \end{aligned} \quad (29)$$

and

$$\frac{\partial \log [\rho(R, t)]}{\partial t} = \frac{R^2}{4\kappa(t + t_0)^2} - \frac{1}{t + t_0}. \quad (30)$$

Taking the derivative of N_{pred} yields the second part of the variance in $\log L$,

$$\begin{aligned} \frac{\partial N_{\text{pred}}}{\partial A_k} &= \dot{N} \int_0^{r_{\text{max}}} \rho(R, t_k) C(R, m_k) dR \\ &\quad + \sum_i A_i \dot{N} \int_0^{r_{\text{max}}} \frac{\partial \rho(R, t_i)}{\partial t_i} \frac{\partial t_i}{\partial A_k} C(R, m_i) dR \quad (31) \\ &= \dot{N} \int_0^{r_{\text{max}}} \rho(R, t_k) C(R, m_k) dR \\ &\quad + \sum_{i \geq k} A_i \dot{N} \int_0^{r_{\text{max}}} \frac{\partial \rho(R, t_i)}{\partial t_i} C(R, m_i) dR. \quad (32) \end{aligned}$$

Combining these results with $\partial \log L / \partial A_k = 0$ yields an equation of the form

$$\frac{1}{A_k} = \dot{N} \int_0^{r_{\max}} \rho(R, t_k) C(R, m_k) dR - \sum_{i \geq k} \left[A_i \dot{N} \int_0^{r_{\max}} \frac{\partial \rho(R, t_i)}{\partial t_i} C(R, m_i) dR - \frac{\partial \log[\rho(R_i, t_i)]}{\partial t_i} \right] \quad (33)$$

or a matrix equation of the form

$$\sum_i M_{ik} A_i = b_k + \frac{1}{A_k}, \quad (34)$$

where

$$M_{ik} = \begin{cases} M_i & \text{if } i \geq k \\ 0 & \text{if } i < k \end{cases} \quad (35)$$

and

$$M_i = \dot{N} \int_0^{r_{\max}} \frac{\partial \rho(R, t_i)}{\partial t_i} C(R, m_i) dR. \quad (36)$$

The vector b_k is given by

$$b_k = \sum_{i \geq k} \frac{\partial \log[\rho(R_i, t_i)]}{\partial t_i} - \dot{N} \times \int_0^{r_{\max}} \rho(R, t_k) C(R, m_k) dR. \quad (37)$$

Although this matrix equation has as many rows as there are stars in the sample, it is straightforward to solve at least formally in two ways. The first is

$$A_k = \frac{1}{\sum_i M_{ik} A_i - b_k} \quad (38)$$

and the second is

$$A_i = \frac{b_i + (A_i)^{-1}}{M_i} - \frac{b_{i+1} + (A_{i+1})^{-1}}{M_{i+1}}, \quad (39)$$

$$A_n = \frac{b_n + (A_n)^{-1}}{M_n}.$$

The values of b_i and M_i depend on the values of A_i through the parameter t_k , so the solution must proceed iteratively while minimizing with respect to the other parameters of the model κ and t_0 .

For each value of the diffusion parameters, we chose to iterate Equation (38) three times to determine the values of A_k within a loop of two iterations where M_i (Equation (36)) and b_k (Equation (37)) vary. Given this new trial luminosity function, the diffusion parameters are varied to find the maximum likelihood, and the iterative solution of the luminosity function is repeated. These two steps are repeated until the values of the diffusion parameters from one iteration to the next have changed by less than one part per hundred.

An interesting limit is when the density distribution is independent of time. This understandably yields a simpler

solution for A_i . In particular, $M_i = 0$ so

$$A_i = -\frac{1}{b_i} = \left[\dot{N} \int_0^{r_{\max}} \rho(R, m_i) C(R, m_i) dR \right]^{-1}, \quad (40)$$

where the underlying density distribution is normalized. The weight is not the reciprocal of the completeness for star i but rather the reciprocal of the mean of the completeness of a star with the flux of star i over the density distribution. The latter could be evaluated by taking the mean of the completeness measured for all the stars in the sample in a magnitude range about star i sufficiently wide to sample the density distribution. It is important to note that the weight is the reciprocal of the mean of the completeness not the mean of the reciprocal. If the completeness does not depend strongly on radius, these two will approximately coincide. Finally if the density distribution is not known a priori and is not modeled, the weight for a particular star is simply given by $A_i = [C(R_i, m_i)]^{-1}$. We call this ‘‘Inv Comp’’ in Figures 6 and 10.

The likelihood is invariant under changes in the birth rate of the WDs (\dot{N}) if one also changes the values of κ , A_i and t_0 as follows:

$$\dot{N} \rightarrow \alpha \dot{N}, A_i \rightarrow \alpha^{-1} A_i, t_0 \rightarrow \alpha^{-1} t_0 \text{ and } \kappa \rightarrow \alpha \kappa. \quad (41)$$

That is, the timescale cannot be fixed without some additional input such as a theoretical cooling curve or an independent estimate of the WD birthrate. The quantities $A_i \dot{N}$, κ / \dot{N} , and $\dot{N} t_0$ are invariant with respect to this transformation. In our data set when we use this modeling technique, we fix the value of \dot{N} to the value inferred by the number of giants in our field as in Goldsbury et al. (2012).

3.7. Constraining the Luminosity Function with Errors

We start the analysis including magnitude errors with Equations (18) and (24), which when combined yield

$$f'(R, m) = \int_{-\infty}^{\infty} \dot{N} \rho(R, t) C(R, m') \sum_i A_i \delta(m' - m_i) \times g(R, m', m - m') dm' \quad (42)$$

$$= \dot{N} \sum_i A_i \rho(R, t_i) g(R, m_i, m - m_i). \quad (43)$$

With this model we can define a likelihood function for the stars that we observe

$$\log L = \sum_j \log \left[\dot{N} \sum_i A_i \rho(R_j, t_i) g(R_j, m_i, m_j - m_i) \times C(R_j, m_i) \right] - N_{\text{pred}}. \quad (44)$$

Note how the magnitude error essentially translates into a spread in the age of the observed stars

$$\frac{\partial \log L}{\partial A_k} = \sum_{j, i \geq k} \left\{ \frac{\rho(R_j, t_i) g(R_j, m_i, m_j - m_i) C(R_j, m_i)}{\sum_l A_l \rho(R_j, t_l) g(R_j, m_l, m_j - m_l) C(R_j, m_l)} \times \left[\delta_{ik} + A_i \frac{\partial \log[\rho(R_i, t_i)]}{\partial t_i} \right] \right\} - \frac{\partial N_{\text{pred}}}{\partial A_k}. \quad (45)$$

Although we have included this additional complication in the derivations for completeness, we have found that the inclusion of error convolution in modeling simulated data does not affect the fitting results, so we did not include this in the modeling of the dynamics while simultaneously determining the luminosity functions.

3.8. Monte Carlo Simulations

To test these techniques in the face of the challenges of incompleteness and magnitude errors present in our data, we typically simulated on the order of 10,000 catalogs of the same size as our data set with a known luminosity function and a known diffusion model and attempted to recover the input parameters. In both cases, the age of the star is first selected to be between zero and 1.5 Gyr. Given this age the model cooling curve determines the F336W magnitude. Second, a radius is selected from the cumulative distribution in projected radius (Equation (5)). Given the radius and magnitude of the candidate for the catalog, the completeness for this star is calculated and the star is included in the sample with this probability. Finally, the magnitude errors are applied by drawing from the magnitude error distribution. We created a sample of 3167 stars—the same as in the WFC3 WD sample. The fitting procedure followed two different strategies.

The first was to assume a fixed cooling curve and try to find the density evolution to determine whether the process is biased in determining the diffusion parameters and the typical errors. Finally, we performed simulations where we did not convolve the models with the error distribution to calculate the likelihood (in all cases errors were applied to the simulated data) to see whether the omission of this step introduced biases. The second strategy did not assume a cooling curve and determined the cooling curve as a part of the process of determining the diffusion. We did not convolve the cooling curve with the error distribution while fitting the model; however, the fake catalogs were created in the same way as in the first strategy. In this technique the resulting cooling curve can be multiplied by a constant factor (Equation (41)), so we determine the values of κ and t_0 by fixing the value of \dot{N} to the one used to build the catalog. This also fixes the age estimates of all of the WDs in the sample.

The results of these simulations are depicted in Figures 3 and 4 and in Table 1. The key results of the simulations are that the likelihood fitting of the diffusion model results in an unbiased estimate of the diffusion parameters regardless of whether the fitting technique includes the magnitude errors (Section 3.5). Furthermore, even when one fits for the luminosity function as well one can obtain reliable estimates of the diffusion model without prior knowledge of the cooling curve; of course, in this latter case the timescales of the diffusion rely on an independent estimate of the birth rate of the WDs \dot{N} . Observationally, this is determined from a sample of giant stars numbering in the thousands (see Figure 1) so the statistical error in this determination is small. Typically the birth rate is recovered with an uncertainty of less than 1% and the diffusion rate with an uncertainty of 10% and t_0 with an uncertainty of 15%. The errors in t_0 and κ are correlated so the error in $t_0\kappa$ is typically less than 10%.

In the second type of simulation, we found the density evolution along with an estimate of the cooling curve, so this cooling curve can be compared with the input cooling curve for the simulations. Furthermore, the determination of the cooling

curve is iterative, so we have to give an initial guess of the curve. The input, the initial guess and the results are given in Figure 5. We can also fit for just the cooling curve and assume that the density distribution does not evolve or not assume a density model at all and use the per star completeness as outlined in Section 3.6. Figure 6 highlights the difference between the model age and the inferred age with the various likelihood techniques. For young WDs the uncertainties are large (because there are few young WDs in the sample), but for old WDs there is a small bias of order of 10% in the inferred age, the sign of which depends on the technique. Again this is on the order of the relative errors in the diffusion parameters.

4. RESULTS

The results of the diffusion model fitting are given in Table 2. The results do not depend strongly on the modeling technique, especially the assumed cooling curve for the WDs. The inferred relaxation times are also in good agreement with the value tabulated by Harris (1996). Figure 7 shows the posterior probability distribution for the various parameters and how the uncertainties are correlated with each other. An important conclusion is that the no-diffusion model (i.e., $\kappa = 0$) is excluded at high confidence.

To find whether we could better fit the radial distribution with a sum of Gaussians, we performed the fitting with two and three Gaussians. We did not include the error convolution in the fitting models. The fit with two Gaussians has a value of $\log L$ that is lower by 75 from a fit with a single Gaussian. From Figure 4 we can see that this is a significantly better fit. However, the decrease in $\log L$ by adding a third Gaussian is only 2; furthermore, the third Gaussian has a very low value of \dot{N} so it does not affect the resulting distributions strongly. Table 2 shows that the diffusion parameters from the two-Gaussian fit only differ slightly from the one-Gaussian fits. In any case, these differences lie within the statistical errors. We can compare the best-fitting model density distributions as a function of time with the observed (completeness corrected) density distributions for several age ranges of WD. The diffusion model for the median age of the WDs in each bin is depicted with a solid line for the one-Gaussian model and a dot-dashed line for the two-Gaussian model. The two-Gaussian model does a better job at following the distribution of the WDs, especially at smaller radii.

4.1. Two-body Relaxation

Figure 8 depicts the radial distribution of WDs of various ages. Each bin is 50 Myr wide, and the bins are centered on 25, 125, and 225 Myr. The evolution at up to a few core radii (about 60 arcsec) is dramatic from 25 to 125 Myr and modest thereafter. Outside 60 arcsec the cumulative distributions are nearly parallel indicating little evolution in this region at early times. The simple diffusion models used here assume that the diffusion coefficient is constant in space and in time, so the models continue to evolve at late times and for all radii. At the smaller radii the WDs reach the distribution corresponding to their masses after about 100 Myr and stop diffusing.

Figure 9 focuses on the outer half of the WFC3 field. Here we see more evolution between the second and third epochs, with little early evolution. This indicates the increase in the relaxation time as the stellar density decreases. The WDs diffuse modestly over the first 100 Myr and more dramatically

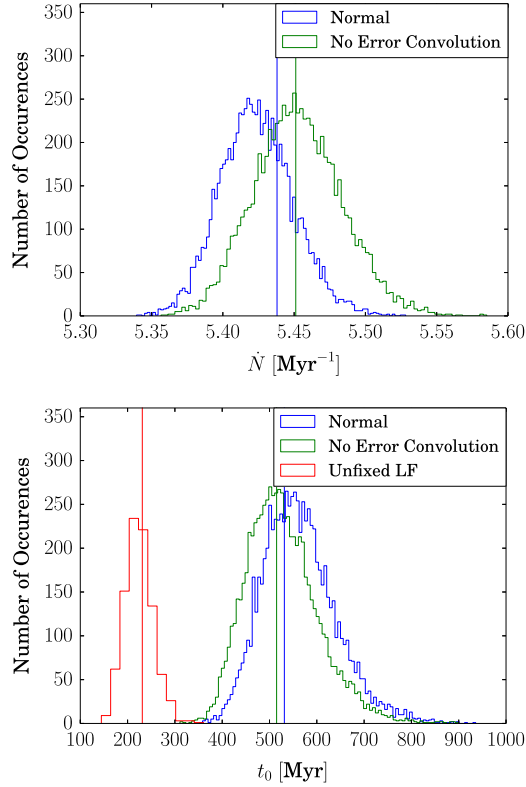


Figure 3. Upper panel: the values of the fitted parameters are typically unbiased with respect to the input values in the simulations. Here \dot{N} is depicted. The input values for the two types of simulations are given by the vertical lines. Lower panel: the values of t_0 .

during the second 200 Myr. The WDs as expected from theoretical considerations suffer diffusion that is a function of radius and time and beyond the scope of the simple model used to quantify the diffusion in this paper. However, this model does capture the diffusion within a few core radii for a few core relaxation times.

5. CONCLUSIONS

5.1. Further Analysis

In this paper we used the Green’s function (Equation (4)) to model the diffusion of the stars through the cluster. We simply took the initial conditions to be a Gaussian or a sum of Gaussians centered on the center of the cluster. This allowed for a simple closed-form expression for the density function in spherical coordinates and in projection as well. Without relaxing the spherical symmetry one could imagine much more general initial conditions. In fact we have an estimate of the initial conditions in the form of the projected radial distribution of the stars on the upper main-sequence. This distribution could be possibly deprojected as a lowered-isothermal distribution in phase space (Michie 1963; King 1966) and convolved with the Gaussian Green’s function, Equation (4), to give the expected density distribution as a function of time. This technique shares the advantage of the technique used in this paper that the density distribution can be guaranteed to be positive because the convolution of the positive kernel with a positive distribution is necessarily positive; however, the density distribution even in spherical coordinates is not available in closed form.

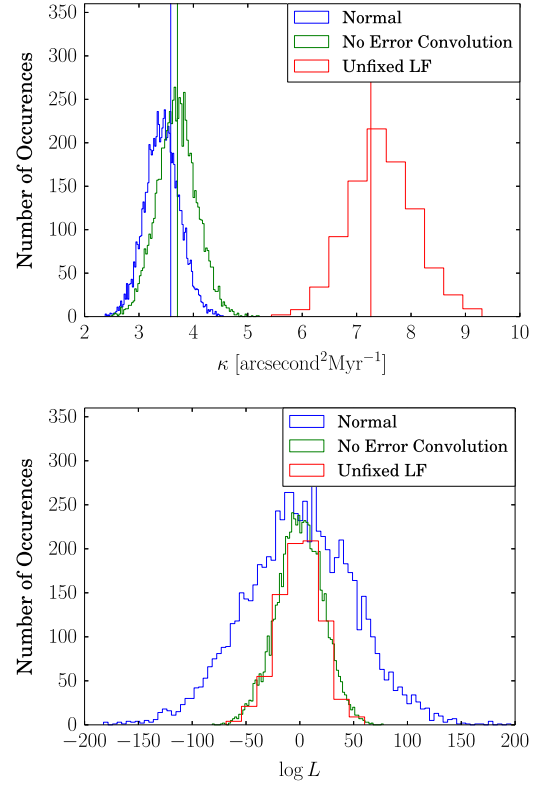


Figure 4. Upper panel: the values of the fitted parameters are typically unbiased with respect to the input values in the simulations. Here κ is depicted. The input values for the two types of simulations are given by the vertical lines. Lower panel: the distribution of $\log L$ is significantly wider when error convolution is included in the fitting process.

Table 1

Parameters from the Monte Carlo Simulations: Means, Standard Derivations, Input Values

Technique	κ	Input	t_0	Input
Full Modeling	3.4 ± 0.3	3.58	560 ± 80	531
No-error Convolution	3.7 ± 0.4	3.71	530 ± 80	515
Unfixed LF	7.5 ± 0.6	7.26	220 ± 30	231
	\dot{N}	Input	$\log L$	
Full Modeling	5.42 ± 0.03	5.44	± 50	
No-error Convolution	5.45 ± 0.03	5.45	± 22	
Unfixed LF	...	± 20	...	

A second strategy would be to expand the initial density distribution in terms of spherical Bessel functions and spherical harmonics. If we restrict ourselves to an initially spherical distribution we have

$$\rho(r, t) = \int_0^\infty dk a(k) e^{-\kappa k^2 t} \frac{\sin(kr)}{kr}, \quad (46)$$

where the coefficients $a(k)$ are determined from the initial density distribution

$$a(k) = \frac{2}{\pi} \int_0^\infty dr (kr)^2 \rho(r, 0) \frac{\sin(kr)}{kr}. \quad (47)$$

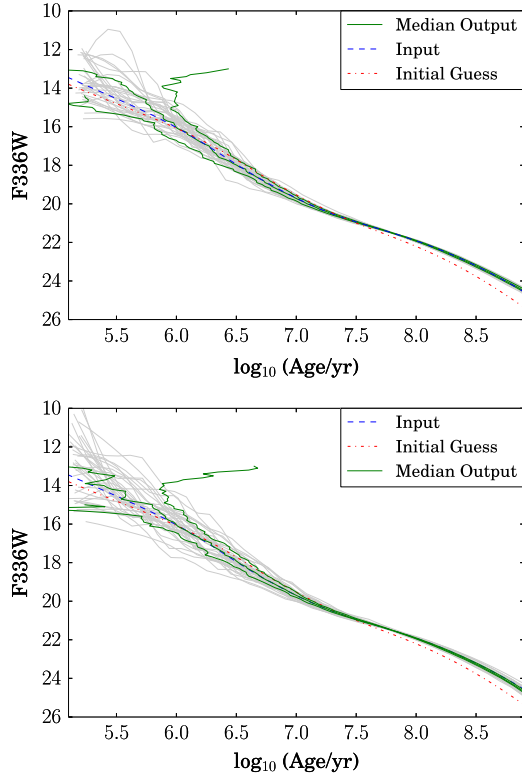


Figure 5. Upper panel: Monte Carlo simulations of an evolving density distribution that are fit by an evolving distribution yield an unbiased estimate of the cooling curve. Lower panel: if the evolving density distribution is not included in the fitting the resulting cooling curve is typically steeper than the input cooling curve; however, this difference is subtle. The initial guess is the starting point for the iterative solution of the cooling curve.

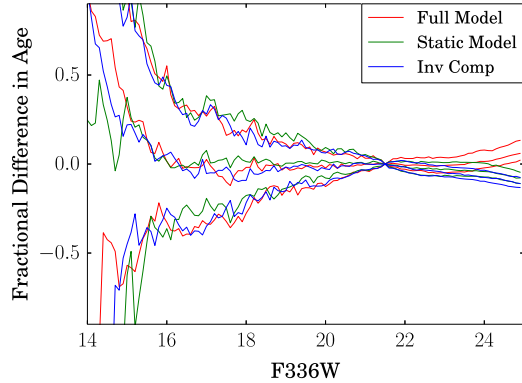


Figure 6. Relative differences between the first, second, and third quartiles and the input model for the different luminosity function fitting techniques.

Table 2
Diffusion Parameters from the Likelihood Fitting

Model	κ ($''^2 \text{ Myr}^{-1}$)	t_0 (Myr)	σ_0 ($''$)	\dot{N} (Myr^{-1})	$t_r(r_c)$ (Myr)
Full	13.1	166	66	7.07	37
No Errors	13.1	166	66	7.07	37
2 Gaussians	12.8	14.9	19.5	1.90	38
(No Errors)	...	260	82	5.32	...
Free LF	9.80	241	69	8.10	49

Note. The values of σ_0 and t_r are given by $\sqrt{2\kappa t_0}$ and r_c^2/κ , where $r_c = 22''$. The posterior probabilities for the full model are given in Figure 7.

If the initial density distribution can be well represented with a few values of k , then the density evolution is straightforward to evolve forward and backward in time; however, it is no longer guaranteed to be positive even at the initial time if only a range of values of k are considered in $a(k)$.

From the point of view of the likelihood analysis, a natural next step would be to use the additional information available with the current observations, i.e., the flux in the F225W band. This would provide an additional constraint on the ages of the WD stars or alternatively constrain the cooling curve in both bands. In the first case one would perhaps get better constraints on the dynamical evolution and could also fit for the distance and reddening to the cluster, and possibly the mass of the WDs or specifics of the cooling mechanism. In the second case one would get a cooling curve in a second band. It is straightforward to see that the weights for the cooling curve in F225W would be the same as in F336W, so simply plotting the inferred ages of the WDs from Figure 10 against the F225W magnitude would yield the cooling curve in F225W. The agreement with the F336W model is poorer at early times, but improves with age and lasts until nearly 1 Gyr. In the context of this paper, we obtain similar diffusion parameters whether we fit a luminosity function or assume a theoretical model.

5.2. Theoretical Directions

As argued in Section 3.2 the interactions with other stars cause the WDs to diffuse in velocity, not in radius. However, using the virial theorem we argued that this diffusion in velocity would be manifest as a diffusion in radius as well. Furthermore, our simple model assumes that the diffusion coefficient is constant in space and time when in fact with time the WD distribution approaches that of stars of similar mass so the diffusion must cease and also at larger radii the diffusion must happen more slowly. We see both of these effects in Figures 8 and 9. How this diffusion actually manifests itself could be simulated in two possible ways.

The first is direct numerical simulation of on order of one million stars that form the central regions of the globular cluster 47 Tuc. This appears to be a Herculean labor because the state-of-the-art direct calculation of the two body interactions in a globular cluster involve merely $\sim 10^5$ stars and the simulation in question would normally take 100 times longer. However, we are only interested in the dynamical evolution of the young WDs over about one hundredth of the age of the cluster (100–150 Myr out of 10 Gyr). Second, because we are not interested in the long-term evolution of the cluster, neither stellar evolution nor the dynamics of binaries should play an important role in this process. These two simplifications result in a factor of a thousand speed up to obtain results and these calculations are already underway.

The second direction would be to model the diffusion in phase space using a Fokker–Planck or Monte Carlo scheme (Giersz & Heggie 2011; Pattabiraman et al. 2013; Hong et al. 2013). Such simulations would be more rapid than a direct N -body simulations and possibly yield more physical insights.

5.3. Further Observations

Following the arguments of Section 5.2 a natural direction would be to measure the proper motions of the WDs in the core of 47 Tuc with a second epoch of observations. Because we already have the colors of the WDs, only observations in a

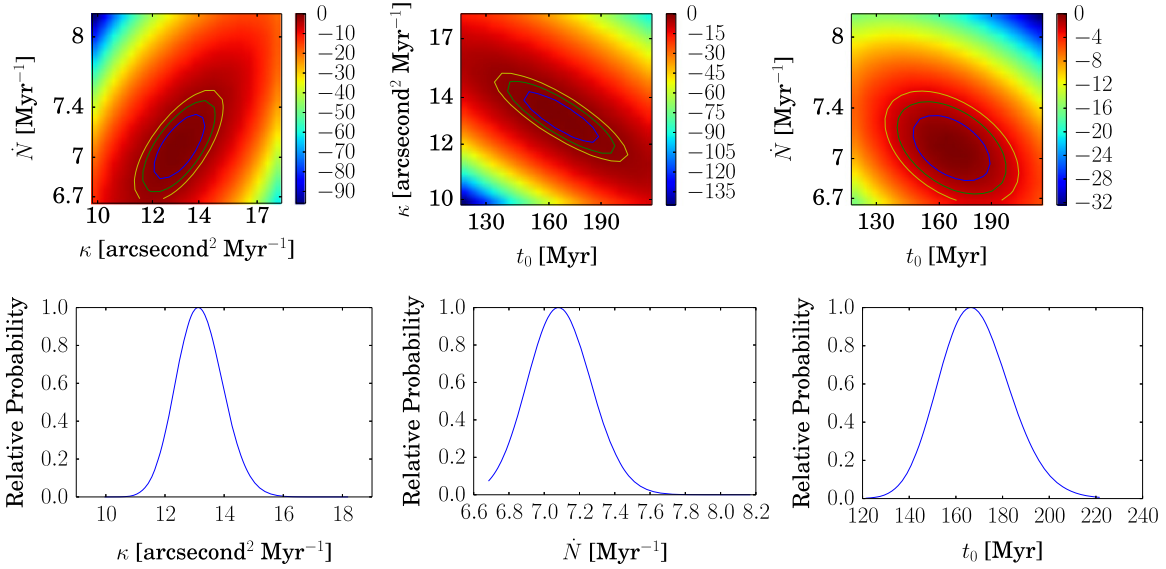


Figure 7. Posterior probabilities of the best-fitting parameters. The upper panels depict the covariance among the various parameters. The contours trace probabilities a factor of e , e^2 , and e^3 smaller than the maximum likelihood. The color scale also gives the natural logarithm of the relative likelihood. The lower panels give the likelihood as a function of a single parameter with the other parameters integrated out.

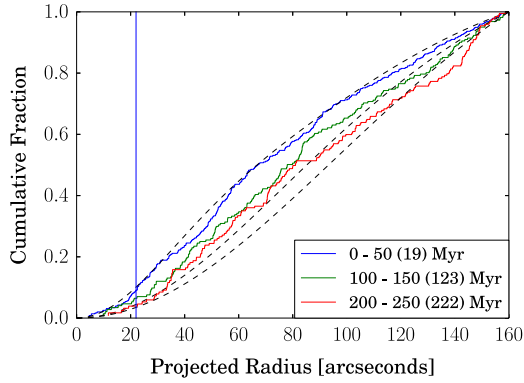


Figure 8. Radial distribution of white-dwarf samples of various age ranges and median ages (given in parenthesis) and the best-fitting two-Gaussian diffusion models superimposed. The core radius from Harris (1996) is depicted by the vertical line.

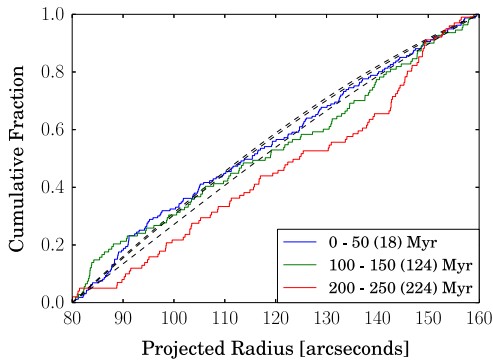


Figure 9. Radial distribution of white-dwarf samples of various age ranges and median ages (given in parenthesis) and the best-fitting two-Gaussian diffusion models superimposed for the outer half of the region.

single band would be required and possibly not as deep as the present set of observations because the stars have already been detected. To obtain the most precise positions and to minimize the crowding, the bluest band would be best, i.e., F225W, and possibly over only a portion of the field of the current data,

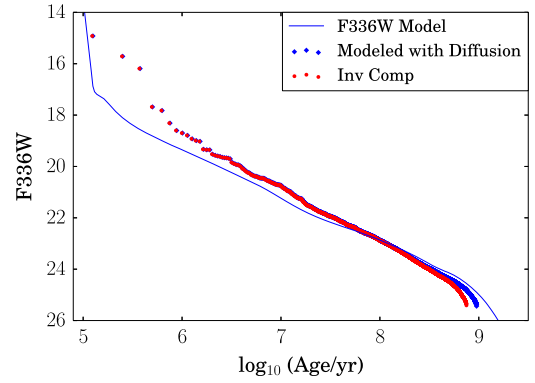


Figure 10. Flux in the WFC3 band F336W as a function of time since the peak luminosity of the star that we define to be the birth of the “white dwarf.” The model curve assumes a true distance modulus of 13.23 (Thompson et al. 2010) and a reddening of $E(B - V) = 0.04$ (Salaris et al. 2007). The “Inv Comp” technique ignores the effects of diffusion in modeling the stars and uses the completeness rate corresponding to the observed magnitude and radius of each star.

because here the goal would be to verify the current result by finding the corresponding signal in velocity space, so a full sample of 3000 plus WDs may not be required.

5.4. Final Remarks

We have measured directly for the first time the dynamical relaxation of stars in a globular cluster. To do this we have introduced new statistical techniques for the characterization of stellar populations. These techniques can robustly and straightforwardly account for high incompleteness and non-Gaussian magnitude errors. They can be applied to a wide variety of questions from globular cluster dynamics to galaxy luminosity functions. There are many avenues for further investigation, such as a more thorough analysis of the existing data using the information from the second band, the simulation of the relaxation of young WDs in numerical models, and measuring the proper motions of the young WDs to search for signatures of relaxation in their velocities as well.

This research is based on NASA/ESA *Hubble Space Telescope* observations obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy Inc. under NASA contract NAS5-26555. These observations are associated with proposal GO-12971 (PI: Richer). This work was supported by NASA/*HST* grant GO-12971, the Natural Sciences and Engineering Research Council of Canada, the Canadian Foundation for Innovation, and the British Columbia Knowledge Development Fund. This project was supported by the National Science Foundation (NSF) through grant AST-1211719. It has made use of the NASA ADS and arXiv.org.

REFERENCES

- Bergbusch, P. A., & Stetson, P. B. 2009, *AJ*, 138, 1455
- Cash, W. 1979, *ApJ*, 228, 939
- Davis, D. 2014, Likelihood Tutorial, http://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/likelihood_tutorial.html, accessed: 2014-08-11
- Fitzpatrick, E. L. 1999, *PASP*, 111, 63
- Ford, H. C., Bartko, F., Bely, P. Y., et al. 1998, in SPIE Conf. Ser. 3356, Space Telescopes and Instruments V, ed. P. Y. Bely, & J. B. Breckinridge (Bellingham, WA: SPIE), 234
- Giersz, M., & Heggie, D. C. 2011, *MNRAS*, 410, 2698
- Goldsbury, R., Heyl, J. S., Richer, H. B., et al. 2012, *ApJ*, 760, 78
- Hansen, B. M. S., Kalirai, J. S., Anderson, J., et al. 2013, *Natur*, 500, 51
- Harris, W. E. 1996, *AJ*, 112, 1487
- Hong, J., Kim, E., Lee, H. M., & Spurzem, R. 2013, *MNRAS*, 430, 2960
- Kalirai, J. S., Saul Davis, D., Richer, H. B., et al. 2009, *ApJ*, 705, 408
- King, I. R. 1966, *AJ*, 71, 64
- Knigge, C., Dieball, A., Maíz Apellániz, J., et al. 2008, *ApJ*, 683, 1006
- MacKenty, J. W. 2012, in SPIE Conf. Ser. 8442, Space Telescopes and Instrumentation: Optical, Infrared and Millimeter Wave, ed. M. C. Clampin, G. G. Fazio, H. A. MacEwen, & J. M. Oschmann (Bellingham, WA: SPIE), 1
- Mattox, J. R., Bertsch, D. L., Chiang, J., et al. 1996, *ApJ*, 461, 396
- McLaughlin, D. E., Anderson, J., Meylan, G., et al. 2006, *ApJS*, 166, 249
- Michie, R. W. 1963, *MNRAS*, 125, 127
- Miglio, A., Brogaard, K., Stello, D., et al. 2012, *MNRAS*, 419, 2077
- Milone, A. P., Piotto, G., Bedin, L. R., et al. 2012, *ApJ*, 744, 58
- Moehler, S., Koester, D., Zoccali, M., et al. 2004, *A&A*, 420, 515
- Pattabiraman, B., Umbreit, S., Liao, W.-k., et al. 2013, *ApJS*, 204, 15
- Paxton, B., Bildsten, L., Dotter, A., et al. 2011, *ApJS*, 192, 3
- Renzini, A., Bragaglia, A., Ferraro, F. R., et al. 1996, *ApJL*, 465, L23
- Renzini, A., & Fusi Pecci, F. 1988, *ARA&A*, 26, 199
- Richer, H. B., Goldsbury, R., Heyl, J., et al. 2013, *ApJ*, 778, 104
- Salaris, M., Held, E. V., Ortolani, S., Gullieuszik, M., & Momany, Y. 2007, *A&A*, 476, 243
- Sills, A., Karakas, A., & Lattanzio, J. 2009, *ApJ*, 692, 1411
- Spitzer, L. 1987, Dynamical Evolution of Globular Clusters, Princeton Series in Astrophysics (Princeton, NJ: Princeton Univ. Press)
- Thompson, I. B., Kaluzny, J., Rucinski, S. M., et al. 2010, *AJ*, 139, 329
- Tremblay, P.-E., Bergeron, P., & Gianninas, A. 2011, *ApJ*, 730, 128
- Woodley, K. A., Goldsbury, R., Kalirai, J. S., et al. 2012, *AJ*, 143, 50