A Novel Method for the Determination of the Binary Period Distribution in Globular Clusters

by

Maigan Devries

A Thesis Submitted to Saint Mary's University, Halifax, Nova Scotia in Partial Fulfillment
of the Requirements for the Degree of Master of Science in Astronomy
(Department of Astronomy and Physics)

August 2023 Halifax, Nova Scotia

Approved: __________________________
Dr. Vincent Hénault-Brunet
Supervisor

Approved: __________________________
Dr. Ivana Damjanov
Examiner

Approved: __________________________
Dr. Gregory Christian
Examiner

Date: August 2023

Acknowledgements

I would like to express my gratitude to all those who have contributed to the completion of this thesis. A special thanks to my supervisor Dr. Vincent Hénault-Brunet for his support, guidance, and encouragement. It has been invaluable during this project and this work would not have been possible without it. I am truly grateful for his mentorship. A big thanks to the clusters research group as well for all of the helpful discussions and insights.

I would also like to thank my thesis committee, Dr. Ivana Damjanov and Dr. Gregory Christian, for their time and expertise when reviewing this thesis.

Additionally, thank you to the BGO for being a great place to work and to Tiffany Fields for being a great boss and mentor. I appreciated all the support and advice you provided.

Finally, thank you to my family, especially Tia, for your continual love and support throughout my schooling journey.

Table of Contents

List of Figures

List of Tables

Abstract

A Novel Method for the Determination of the Binary Period Distribution in Globular Clusters

by

Maigan Devries

Globular cluster binaries are dynamically affected by interactions within their host cluster, leaving a signature on their orbital period distribution. As most disrupting interactions occur during the early evolution of the cluster (when the cluster is densest), the present-day intrinsic binary period distribution, in particular the hard-soft limit, can provide insight into the early cluster environment. However, constraining the period distribution is difficult due to observational biases. To overcome this, we developed a hierarchical method for fitting the period distribution. This method takes multi-epoch radial velocity data for stars in the cluster, fits the radial velocity curve for stars detected as radial velocity variable, and then uses these posterior samples to recover the intrinsic orbital period distribution. We validate the method using mock data and present the first results on the globular cluster NGC 3201. This will provide information about binaries, globular clusters, and topics such as gravitational waves.

Date: August 2023

Chapter 1

Introduction

1.1 Globular Clusters

Globular clusters are ancient self-gravitating collections of stars and stellar remnants that are not only fascinating in their own right, but can also be used to gain valuable insight into other astronomical phenomena. Within our own galaxy, there are approximately 150 globular clusters (Harris, 2010), with the majority residing within the Milky Way halo. One such example is NGC 3201, which can be seen in Figure 1.1. Globular clusters have ages of around 10 to 12 Gyrs, signifying their formation in the early universe and during the early stages of galaxy formation (Ashman & Zepf, 1998).

A defining characteristic of globular clusters is their dense concentration of stars as they typically host around 10^5 stars within a compact region. The cluster's size is commonly measured using the following radii: the half-light radius (radius enclosing half the light of the cluster in projection), the half-mass radius (radius enclosing half the mass of the cluster), and the tidal radius (the radius where the stars are no longer bound to the cluster). These radii are typically <10 pc, around 3 pc, and around 35 pc respectively (Heggie & Hut, 2003).

Figure 1.1: Globular Cluster NGC 3201. Image taken from European Southern Observatory public images. Image was obtained using the Wide Field Imager (WFI) on the ESO/MPG 2.2m telescope with B,V, and I filters.

As these clusters are dense, their evolution is driven by the gravitational interactions between stellar objects, where stellar objects includes stars, stellar systems such as binaries, and stellar remnants (white dwarfs, neutron stars, and black holes). The interactions between these stellar objects cause perturbations to their orbits within the cluster as well as to their kinetic energies. The cumulative effect of these perturbations is referred to as two-body relaxation. The time it takes for this cumulative effect to cause a change to the velocity on the order of the velocity itself for a typical star in the cluster is the relaxation time, T_{relax}. This is important as it signifies the time it takes for the typical star to forget its initial conditions. This can be found by considering the number of crossings that are required for this to occur, n_{cross}, and the time it takes to cross the cluster, t_{cross}, where $t_{\mathrm{cross}} = \frac{r_{\mathrm{hm}}}{v_{disp}}$. As expected, n_{cross} depends on the number of interactions it has as it crosses the cluster and thus on the number of stars in the cluster, N. It is found to be approximately $\frac{0.1N}{\ln N}$. Putting this together, the half-mass relaxation time can be approximated as:

$$T_{\mathrm{relax}} = n_{\mathrm{cross}} t_{\mathrm{cross}} \approx \frac{0.1N}{\ln N} \frac{r_{\mathrm{hm}}}{v_{\mathrm{disp}}} \tag{1.1}$$

(Binney & Tremaine, 2008). For globular clusters, the median half-mass radius and velocity dispersion are 3.08 pc and 5.50 km/s respectively, making the crossing time, T_{cross}, approximately 0.6 Myr. Assuming $N \approx 10^5$, the relaxation time is on the order of 1 Gyr. As globular clusters are highly collisional environments and typically have ages that exceed their two-body relaxation time, two-body relaxation is important in their evolution (Binney & Tremaine, 2008).

During these gravitational interactions, the cluster evolves towards kinetic energy equipartition, thus objects with larger kinetic energies will transfer energy to systems with less kinetic energy. This leads to heavier objects slowing down and moving towards the cluster core and lighter objects gaining speed and moving to the cluster halo. This process is called mass segregation (Heggie & Hut, 2003).

As globular clusters are self-gravitating systems, they have a net negative energy and a negative heat capacity. This negative energy stems from the virial theorem,

$$2K + U = 0, \tag{1.2}$$

where K is kinetic energy and U is the gravitational potential. From this we can see that the kinetic energy is given by $K = -\frac{1}{2}U$. Putting this into the net energy equation we have, $E_{\text{net}} = K + U = -\frac{1}{2}U + U = \frac{1}{2}U$. As gravitational potential energy is negative, the net energy of the system must also be negative and, consequently, the heat capacity. Therefore, they do not evolve towards thermal equilibrium as expected and instead, as a star loses energy, it will gain speed. Through two-body relaxation, the cluster core loses energy to the stars in the halo, causing the core temperature to increase and for it to contract. Consequently, stars sink towards the core even more, leading to a runaway process of core contraction. This process is referred to as gravothermal catastrophe, better known as core collapse in globular clusters, and results in a very dense cluster core (Heggie & Hut, 2003). However, around the Milky Way, there are more clusters that have not yet undergone core-collapse than expected

(Goodman & Hut, 1989). This suggests the presence of mechanisms that delay core-collapse. One such mechanism involves stellar binaries and binary black holes within the cluster. These binary systems are able to harden (shrinkage of the binary orbit) as a result of interactions and thus lose gravitational potential energy (see Section 1.3) and provide kinetic energy to nearby cluster stars, delaying their fall into the cluster core (Goodman & Hut, 1989). This process has the additional effect of expanding the globular cluster as this energy from the binaries radiates outwards through two-body relaxation when stars gain kinetic energy through interactions with binaries and thus move to the outer regions of the cluster (Gieles et al., 2010).

Globular clusters are important astronomical objects not just due to their interesting nature as stated above, but also for their broader importance in various fields of astronomy. For example, due to their high stellar density, globular clusters are key systems for studying stellar dynamics. Stellar evolution can also be studied as each globular cluster is host to a population of stars with similar ages and compositions, but with differing initial masses. Moreover, globular clusters play an important role in understanding their host galaxies. For example, evidence suggests a connection between globular clusters and the assembly process of their host galaxy, and they can also be used to infer episodes of star formation in the galaxy (Brodie & Strader, 2006). More recently, work has been done to use globular clusters in the Milky Way to infer information about the Milky Way's assembly history and formation by comparing cosmological zoom-in models and globular cluster age-metallicity distributions (e.g. Kruijssen et al., 2019).

Within globular clusters, one interesting component is their binary population. As stated earlier, these binaries have dynamical importance within the cluster due to their role in delaying core-collapse alongside black-hole binaries. However, there are many other interesting reasons for studying binaries, some of which can be found in the preceding section and will be explored further in this work.

1.2 Binaries

Binaries are ubiquitous in both the galactic field and in globular clusters. The properties of binaries can be described using three main parameters: orbital eccentricity (e), orbital period (P), and the mass ratio between the two stars (q), or the probability distributions of these three orbital parameters when considering populations of binaries. We discuss these distributions in Section 1.4; here we note that our knowledge of binary properties is limited by a detection bias due to the limited time sampling of the observations (e.g. Giesers et al., 2019). An example of a binary orbit can be seen in Figure 1.2.

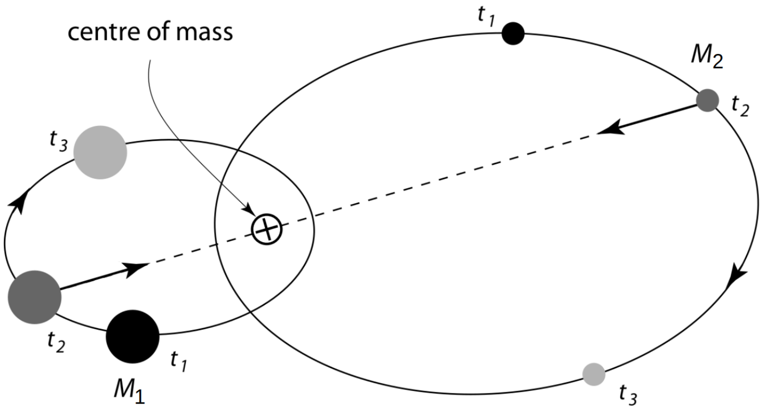


Figure 1.2: The orbit of two stars around a common centre of mass, the barycenter, and their corresponding orbital positions at three different times. Figure adapted from Perryman (2011).

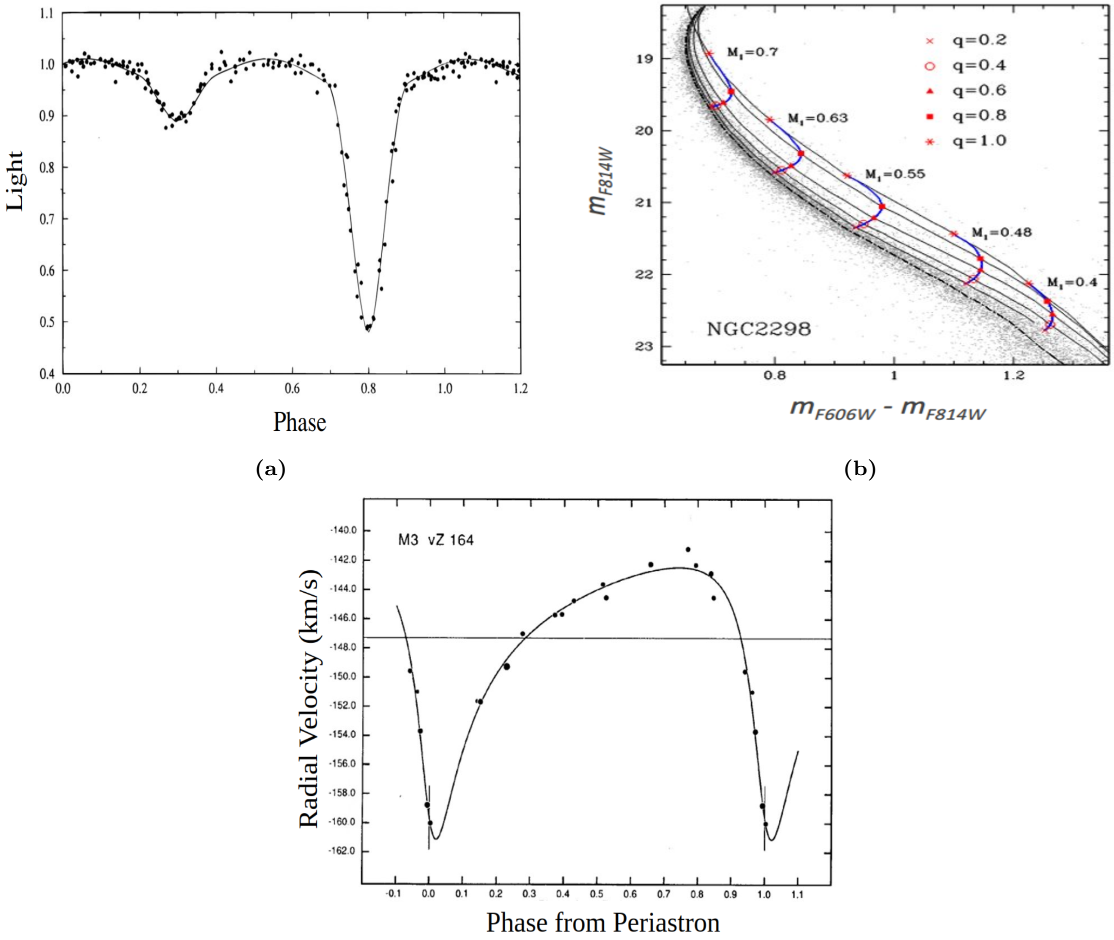
To detect binaries in globular clusters, there are three main methods. The first is by looking for eclipsing binaries. An eclipsing binary occurs when one star moves behind the other one, causing a dip in the light curve as shown in Figure 1.3a (McVean et al., 1997). However, in the case of a globular cluster, high spatial resolution and a large number of observations are required to adequately sample the light curve. Additionally, this method is limited to binaries with inclination angles close to 90 degrees (binary orbital plane seen edge-on), making this an ineffective method for detecting binaries.

Another detection method uses the colour-magnitude diagrams of globular clusters and the binary sequence. This binary sequence occurs as the light from both stars in the binary contribute to the observed light - causing the system to appear brighter. The end result is stars appearing above and to the right of the main sequence as can be seen in Figure 1.3b. Using this method on 54 globular clusters, Milone et al. (2012) found that the cluster core has a higher concentration of binaries than the surrounding region. They found that, typically, the binary fraction doubles from two core radii to the cluster core. However, the major shortcomings of this method are that it requires high precision photometry, such as that done with the Hubble Space Telescope, and it does not provide information regarding the binary orbital parameters, with the exception of the binary mass ratios. It is also limited to binaries with a mass ratio above 0.5 as binaries below this threshold cannot be distinguished from single main sequence stars.

The final method of binary detection is through repeated radial velocity measurements. This is done by looking at the spectra for each star and extracting the radial velocities from the (Doppler-shifted) observed central wavelength of the spectral lines. In terms of their spectra, binaries appear as either single-lined spectroscopic binaries, where the spectral lines are seen for one of the stars, or double-lined spectroscopic binaries, where the spectral lines of both stars are seen in the spectra.

In the case of a single star, the radial velocity is expected to remain unchanged, within measurement uncertainties, when viewed at multiple epochs. However, in the case of a star within a binary, its radial velocity will vary depending on where the star is in its orbit and its motion along the line of sight to the observer. An example of a radial velocity curve for a binary within a globular cluster is shown in Figure 1.3c. From this method, the orbital parameters can be determined from the radial velocity curve as described below.

As a star in a binary follows an elliptical orbit (such as that seen in Figure 1.4),



(a)

(b)

(c)

Figure 1.3: Detection methods for binaries in globular clusters. The first panel (a) shows the light curve of an eclipsing binary. Figure from McVean et al. (1997). Panel (b) shows the colour-magnitude diagram for the globular cluster NGC2298. The red symbols give the various binary mass ratios and the blue lines give the positions of the binaries. Figure adapted from Milone et al. (2012). Panel (c) is an example of a radial velocity curve for a binary within a globular cluster fitted to the data points. The horizontal line shows the velocity of the barycenter. For this radial velocity curve, the orbital parameters are as follows: period of 2657 days, eccentricity of 0.58, velocity semi-amplitude of 9.3 km s^{-1}, argument of periastron of 147°, and barycenter velocity of -147.3 km s^{-1}. Figure from Hut et al. (1992).

its position relative to the barycenter is given by the equation for an ellipse:

$$r = \frac{a(1-e^2)}{1+e\cos(\nu)} \tag{1.3}$$

where a is the semi-major axis, e is the eccentricity, and ν is the true anomaly, and both r and ν vary with time. The true anomaly describes the angle between the pericentre and the current position of the star and can be related to the eccentricity through the equation $\nu = \cos^{-1}(\frac{\cos(E)-e}{1-e\cos(E)})$ where E is the eccentric anomaly, an angle that describes the position of the orbiting body (see Figure 1.4). The time dependence of the eccentric anomaly, E, can be seen in the mean anomaly, M, (which describes the mean motion of the star in its orbit) which is given by $M = \frac{2\pi}{P}(t - t_P) = E - e\sin(E)$, where t_p is the time from pericentre passage and P is the period of the binary. In Figure 1.4, F1 represents the barycenter of the binary system in this figure and is a focus of the ellipse.

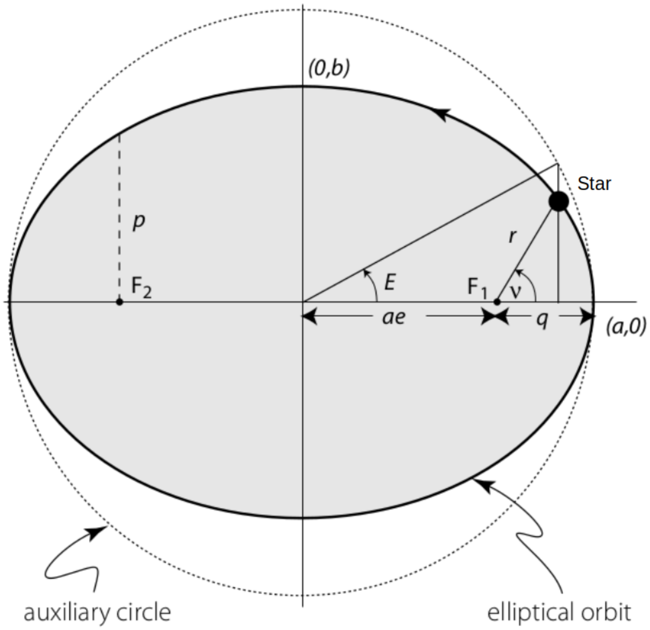


Figure 1.4: The elliptical orbit for one of the stars in the binary. The elliptical orbit is the true stellar path and the auxiliary circle is a circle with radius equal to the semi-major axis, a. The distance to the pericentre is q, the angle between the pericentre and the position of the star as seen from the barycentre (F1) is the true anomaly ν, and the angle between the pericentre, the orbit's centre, the projection of the star position onto the auxiliary circle is the elliptical anomaly, E and the distance from the star to the focus of the ellipse is r. This figure is adapted from Perryman (2011).

To calculate the radial (line-of-sight) velocity as a function of time, the inclination of the orbit also needs to be quantified. This can be seen in Figure 1.5 where i represents the inclination angle, ω is the argument of the pericentre (the angle between the ascending node, the point where the orbit crosses the reference plane, and the pericentre), and $\nu(t)$ is the true anomaly. From the geometry of this problem, it can be seen that z, the line-of-sight position of the star, is given by:

$$z = r\sin(i)\sin(\omega + \nu), \tag{1.4}$$

where r is the distance from the star to the ellipse focus.

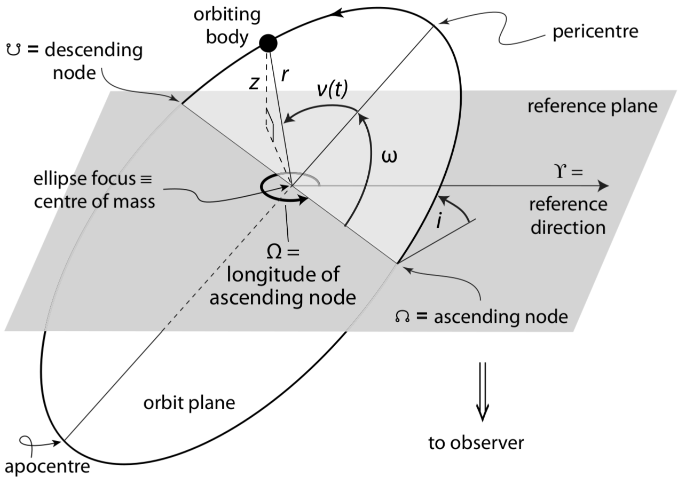


Figure 1.5: The orbital parameters describing the geometry of a binary orbit with inclination i. Figure from Perryman (2011).

Since we want to find the radial velocity, we can take the derivative of this equation to get:

$$v = \dot{z} = \sin(i)[\dot{r}\sin(\omega + \nu) + r\dot{\nu}\cos(\omega + \nu)]. \tag{1.5}$$

Therefore, in order to find the radial velocity of the star relative to the barycenter of the binary, $\dot{r}$ and $\dot{\nu}$ are needed. In Equation 1.3, both r and v depend on time. Therefore, taking the derivative using the reciprocal rule gives:

$$\dot{r} = \frac{a(1-e^2)}{1+e\cos(\nu)} \times \frac{e\sin(\nu)\dot{\nu}}{1+e\cos(\nu)} = \frac{r\dot{\nu}e\sin(\nu)}{1+e\cos(\nu)}. \tag{1.6}$$

To find $\dot{\nu}(t)$, we need to use Kepler's second law. Kepler's second law states that the area, A, that is swept out (by the line connecting the focus of the ellipse and the position of the object) during the orbital motion is equal in equal time intervals. In order for this to hold, the speed of the orbiting body must change depending on how close it is to the focus of the ellipse. Kepler's second law gives $\frac{dA}{dt} = \frac{1}{2}r^2\frac{d\nu}{dt}$. Taking the time interval to be one full period, the area is the entire ellipse and is given by $A = \pi ab$ or $\pi a^2\sqrt{1-e^2}$ in terms of just the eccentricity and the semi-major axis. Therefore, from Kepler's second law we have:

$$\frac{\pi a^2\sqrt{1-e^2}}{Pr} = \frac{1}{2}r\dot{\nu}. \tag{1.7}$$

Putting Equation 1.3 into 1.7 and solving for $r\dot{\nu}$ gives:

$$r\dot{\nu} = \frac{2\pi a(1+e\cos(\nu))}{P\sqrt{1-e^2}}. \tag{1.8}$$

Using Equation 1.6 with 1.8, we obtain the following expression for $\dot{r}$:

$$\dot{r} = \frac{2\pi ae\sin(\nu)}{P\sqrt{1-e^2}}. \tag{1.9}$$

Substituting the expressions for $\dot{r}$ and $r\dot{\nu}$ into Equation 1.5 and using sine and cosine addition trigonometric identities gives the equation for the line-of-sight velocity of a star in a binary relative to its barycenter:

$$v = \frac{2\pi a\sin(i)}{P\sqrt{1-e^2}}[\cos(\nu+\omega)+e\cos(\omega)] = K[\cos(\nu+\omega)+e\cos(\omega)], \tag{1.10}$$

where K is defined as $\frac{2\pi a\sin(i)}{P\sqrt{1-e^2}}$ and is the velocity semi-amplitude.

To describe the entire radial velocity equation as a function of time for a star in a binary, we can add in the velocity contribution from the barycenter's motion within the cluster (v_0) and put the true anomaly in terms of the eccentric anomaly to get

$$v(t; P, e, \omega, \phi_0, K, v_0) = v_0 + K[\cos\left(\omega + \cos^{-1}(\frac{\cos(E) - e}{1 - \cos(E)})\right) + e\cos(\omega)]. \tag{1.11}$$

In this equation, there is no obvious time dependence of the velocity, however, this time dependence is implicit in the argument of the pericentre (ω) and in the eccentric anomaly. There is no analytic form of this equation in terms of time, however, it can be solved numerically.

By utilizing radial velocities to detect binaries and constrain their orbital parameters, we can better understand the binary populations within globular clusters. As binaries are influenced by gravitational interactions with other objects within the cluster, the surviving binary population can provide insight into the dynamical environment of the cluster. Therefore, it is important to understand the binary dynamics within a globular cluster, as we discuss next.

1.3 Binary Dynamics

Binaries in globular clusters are interesting to study because their dynamics differ from those in the galactic field. This is due to the mutual dynamical effects between the binaries and their host globular cluster (e.g. Ivanova et al., 2005). For example, as stated earlier, as binaries interact with other stars, they can provide energy to these stars. This energy comes from the binaries' gravitational potential energy and is converted to kinetic energy, pushing stars towards the outer cluster regions, thus delaying core-collapse (Heggie & Hut, 2003).

The majority of binaries in globular clusters are thought to have formed during the cluster formation and these binaries are referred to as primordial binaries. It is also possible for binaries to form dynamically as the cluster evolves, however the main mechanisms of binary evolution are ionization and evaporation, not formation (Sollima, 2008). There are two mechanisms through which binary formation is known to occur. The first is through three-body interactions. This occurs when three single stars interact and exchange energy resulting in two of them becoming bound. While this is likely to form soft binaries (binaries with a larger semi-major axis) that will quickly be destroyed, it is also possible for them to harden over time to produce hard binaries (binaries with a shorter semi-major axis) (Spitzer, 1987).

The second formation mechanism is through tidal capture. This process involves two stars that are gravitationally perturbed when they interact, causing the stars to gain orbital energy. If the total gain of orbital energy is larger than the kinetic energy of the system, given by $E_K = \frac{1}{2} m_r V^2$ where m_r is the reduced mass and V is the velocity of the system's center of mass, then the two stars will become bound. Tidal capture is more likely to lead to short period binaries so is a more effective method of increasing the binary fraction of the cluster compared to three-body encounters (Spitzer, 1987).

While binaries can form during the clusters evolution, it is also possible for them to escape the cluster (evaporation) or for the stars in the binary to separate (ionization). Evaporation is not as common since heavier masses, such as binary systems, preferentially segregate to the cluster core. On the other hand, ionization does significantly affect binaries. It is also possible for interactions with single stars to lead to an exchange or the formation of a three-body system (Heggie et al., 1996).

To understand the dynamics of binary interactions, it is important to first understand the distinction between hard and soft binaries. A hard binary refers to a binary whose binding energy is larger than the average kinetic energy of a star in the cluster. The opposite is true for a soft binary (Heggie, 1975). The binding energy of a binary is given by

$$E_b = \frac{G m_1 m_2}{2a}, \tag{1.12}$$

where m_1 and m_2 are the masses of the two stars, G is the gravitational constant, and a is the semi-major axis. The average kinetic energy of a star in the cluster is given by $\frac{1}{2}\langle m\rangle V^2$ where $\langle m\rangle$ is the average mass of a star and V is the typical velocity of a star in the cluster or velocity dispersion.

When a binary interacts with a single star, the possible interactions are a flyby, an exchange, or a capture. For a flyby, the single star approaches the binary and leaves intact, however an exchange of energy occurs during the encounter. If the binary is hard, it is more likely to transfer energy to the third body and its orbit will shrink while the single star's speed will increase. If the binary is soft, it is more likely to receive energy from the star causing the star to slow down and for the binary's orbit to widen. In the case of the soft binary, it is possible that this interaction can cause the orbit of the binary to widen to the point where the stars in the binary become unbound (e.g. Spitzer, 1987, Heggie & Hut, 2003). Since the binding energy is inversely proportional to the semi-major axis, the net result of these flybys is that hard binaries get harder and soft binaries get softer with time. This is referred to as Heggie's Law (Heggie, 1975).

The velocity required to cause ionization of the binary can be quantified by taking the point where the kinetic energy of the interacting system is equal to the binding energy of the binary and solving this for the critical velocity. This gives

$$V_{crit}^2 = \frac{2E_b}{m_r}, \tag{1.13}$$

where m_r is the reduced mass and is given by $m_r = \frac{m_3(m_1+m_2)}{m_1+m_2+m_3}$ (m_1 and m_2 are the masses of the binary components and m_3 is the mass of the single star). If the velocity of the single star is larger than the critical velocity, then the binary will be ionized. As this velocity is proportional to the binding energy, soft binaries are more likely to be ionized, as expected. In Figure 1.6, the semi-major axis for a soft

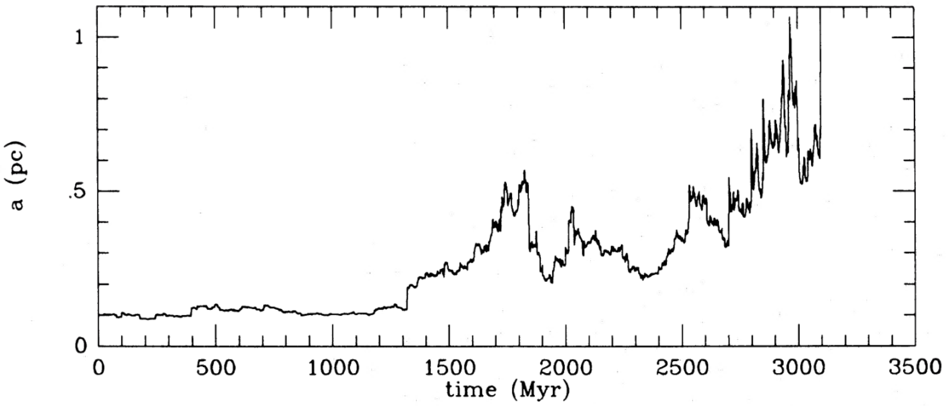


Figure 1.6: The evolution of the semi-major axis of a soft binary in the galactic field. Data is from numerical gravitational three-body scattering simulations. Binary components are both 1 $M_\odot$ and the initial semi-major axis is 0.1 pc. The start time represents the formation of the binary. Figure from Bahcall et al. (1985).

binary in the galactic field can be seen as it interacts with passing stars based on the numerical simulations of Bahcall et al. (1985). While this dataset is from a non-cluster environment, a similar trend for the semi-major axis would be expected in a globular cluster, however it would be expected to occur at a faster pace in a globular cluster given the more frequent interactions. From this, it is clear that the semi-major axis slowly increases as the soft binary undergoes encounters. However, it is important to note that not all interactions result in the widening of the binary orbit as it is also possible for individual interactions to shrink the orbit. Nonetheless, we expect the long-term trend to be an increase in the semi-major axis of a soft binary.

In the case where an interaction between a single star and a binary causes an exchange, this means that the single star replaced one of the stars in the binary. This would affect the orbital parameters of the binary system; however, a binary is present both before and after the encounter. The final binary-single star encounter possibility is the capture. This occurs when the single star approaches the binary and then ends up bound to the stars in the binary, forming a three-body system. This capture process is only possible if the velocity of the single star is less than the critical velocity. The resulting three-body system is typically unstable and one of the stars quickly becomes unbound. Therefore, the end result of a capture appears as either a flyby or an exchange (Heggie & Hut, 2003).

It is also possible for two binaries to interact. If two hard binaries have an encounter, it is possible that one of the binaries will dissociate, making it a mechanism for hard binary destruction. It is also possible that the encounter can result in a three-body system. While this is a method for hard binary destruction, it is rare compared to encounters like flybys (Spitzer, 1987).

These dynamical interactions suggest that over time, the binaries with large semi-major axes or long periods are preferentially going to be lost as the cluster evolves, resulting in period distributions that are progressively skewed towards short-period binaries (Gratton et al., 2019, Spitzer, 1987). However, what classifies as a soft or hard binary depends on the cluster properties, i.e. the average kinetic energy of a star within the cluster. Therefore, more binaries will be ionized in denser cluster environments as the stars in these regions have higher velocities. Since globular clusters are the densest when they initially form (before they expand due to two-body relaxation), the majority of binary disruptions occur early on in the evolution, allowing us to relate the binary disruptions and period distribution to the initial cluster density.

An example of how the initial density of a globular cluster affects the surviving binary distribution is shown in the study by Ivanova et al. (2005). In this study, they took a flat $\log_{10} P$ distribution for the initial binary period distribution and then evolved the cluster using Monte Carlo dynamical models. They tracked the evolution of the stars in the cluster and looked at the remaining binaries after the dynamical interactions and stellar evolution (stellar mass loss and the formation of remnants). This was done for four different models, each of which started with the same number of stars (2.5×10^5) and with the same binary fraction ($f_{\rm bin}$=1[1]) but

---

[1] The binary fraction is defined as $\mathrm{f_{bin}} \equiv \frac{n_{\rm b}}{n_{\rm s}+n_{\rm b}}$ where $n_{\rm b}$ is the number of binaries and $n_{\rm s}$ is the number of single stars.

differed in their initial central cluster densities. From this process, they found that the binary period distribution evolves over time and this evolution depends on the initial central density. Their results for the present-day (after 14 Gyr) distributions for their models is shown in Figure 1.7. This figure shows the fraction of binaries in each period bin relative to the total number of binaries for four different initial cluster densities. Looking at the top panel (the lowest density case), the orbital period distribution has a gradual decline towards long-period binaries, but a substantial portion of long-period binaries have survived. Looking at the lower panels, as the initial density increases, the period distribution becomes skewed towards short-period binaries as more long-period binaries are ionized. While the classification of a long/short period binary depends on the mass ratio and the binary's environment, from this figure we can see from the tail of the distributions that long-period binaries are on the order of 10^2 days and beyond, as these are the binaries most affected by the dynamics of the cluster.

Using this idea, the orbital period distributions for the binaries inferred from multi-epoch radial velocity surveys can provide information about the initial density required to give rise to the present-day binary period distributions. For example, if a globular cluster has retained more long-period binaries, it would imply a lower initial density compared to a globular cluster lacking long-period binaries.

1.4 Orbital Period Distribution

Currently, binary orbital parameter distributions are well constrained for field binaries composed of solar-type stars. For example, Raghavan et al. (2010) and Duquennoy & Mayor (1991) used radial velocity measurements over up to 14 and 13 years, respectively to determine field binary period distributions. In each of these cases, they used the observed binaries to find the period distributions and both found distributions

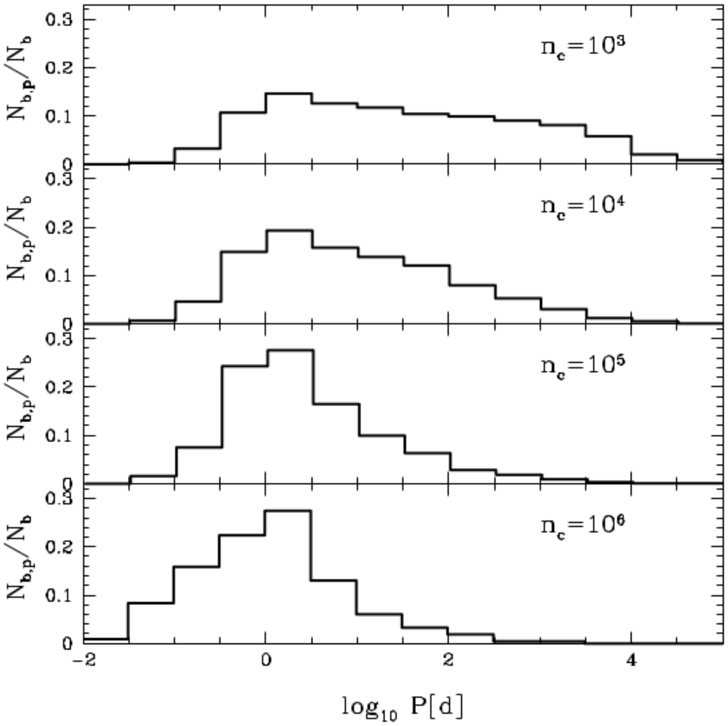


Figure 1.7: Number of binaries in each period bin relative to the total number of binaries for various initial cluster core number densities after 14 Gyr. Densities are given in terms of stars per cubic parsec. Figure from Ivanova et al. (2005).

that follow a log-normal distribution. The Raghavan distribution has a mean $\log_{10} P$ of 5.03 with P in days and a standard deviation $\sigma_{\log P}$ of 2.28 resulting in a median period of 252 years. The Duquennoy & Mayor distribution is similar with a mean of $\log_{10} P = 4.8$ and a standard deviation of $\sigma_{\log P} = 2.28$ giving a median period of 180 years.

For globular clusters, there is less consensus on the best assumptions for the binary period distribution because globular cluster binaries are influenced by the cluster dynamics and their period distribution evolves with time, as discussed in Section 1.3. One example of a distribution sometimes used is a uniform distribution in log P. This distribution is believed to be a good assumption for cases where there are few dynamical effects to influence the distribution, such as in a low density cluster or in the early stages of cluster evolution. However, this assumption breaks down for higher density globular clusters (Ivanova et al., 2005).

Another commonly used initial binary period distribution within a globular cluster is the Kroupa initial binary period distribution. This is given by:

$$f_p(\log_{10} P) = 2.5 \frac{\log_{10} P - 1}{45 + (\log_{10} P - 1)^2} \tag{1.14}$$

for binaries with a primary mass less than 5 $M_\odot$ and

$$f_p(\log_{10} P) = 0.23 \times (\log_{10} P)^{-0.55} \tag{1.15}$$

for binaries with a primary mass greater than or equal to 5 $M_\odot$ where the period, P is in days (Marks et al., 2011). These distributions are similar to those found for O-type stars by Sana et al. (2012) as this primordial distribution is before stellar evolution has occurred.

While these initial distributions do not directly describe the current distributions, they can be used as initial cluster conditions and then evolved with the cluster to the present day (e.g. Ivanova et al., 2005). However, there is no universal form for the binary period distribution for all globular clusters as different initial cluster conditions and dynamical evolution histories affect the distribution as shown in Figure 1.7. Even so, better constraining the binary period distribution for a given globular cluster is important for understanding the abundance of retained soft binaries and therefore the initial cluster density.

One issue with trying to better constrain the binary period distribution is the effect of the detection bias. This affects binaries with longer periods as it is difficult to sample enough of the orbit to detect radial velocity variations in these systems. With many epochs of data over multiple years, this effect decreases, however it is always present to some extent. Therefore, this needs to be accounted for when fitting binary population parameters.

1.5 This Thesis: Globular Cluster Binaries

The objective of this work is to develop a method that takes multi-epoch radial velocity measurements for a globular cluster and use these to constrain their true binary period distributions while taking into account the effect of the detection bias.

Currently, there are a few globular clusters with multiple epochs of radial velocity measurements available for a large number of stars, in particular thanks to observations with the MUSE integral-field unit spectrograph on the Very Large Telescope (VLT). However, there has not yet been any work done to take the binary periods that are observed and use this information to statistically infer the true period distribution. This will need to be done while accounting for limited time sampling and the detection bias. Therefore, this thesis will involve using hierarchical Bayesian statistics on the multi-epoch radial velocity measurements for individual stars in globular clusters to constrain the distribution of orbital parameters of the cluster's binaries, in particular, the period distribution.

Constraining the intrinsic period distribution for binaries is important for a number of reasons. One is that globular cluster models (such as N-body models) can include binaries and by doing so, assumptions need to be made about the binary period distribution. By better constraining this distribution and mitigating the effects of the detection bias, we can improve the assumptions used in globular cluster models.

Another reason for our interest in the period distribution is that understanding the depletion of long-period binaries can give us insight into the initial density of the host globular cluster. As was discussed in Section 1.3, the dynamical interactions within the cluster result in a correlation between the initial density of the cluster and the remaining soft binaries. Therefore, by understanding the relative abundance of long-period binaries, this can provide insight into the initial cluster environment.

Understanding the initial cluster density can also improve our understanding of the black hole population in globular clusters. For example, N-body simulations were done for the cluster Pal 5 that looked at the black hole population in the cluster (Gieles et al., 2021). The initial density of the cluster was varied and one model was computed with black holes and the other without. The goal was to determine what initial conditions give rise to what is seen today in Pal 5, such as its large size (half light radius of ≈ 20 pc) and its tidal tails. This is done by comparing the kinematics and spatial distribution of stars from the N-body models with the observed kinematics and spatial distribution. From this, it was found that two possible models exist: one with a high initial density with black holes and one with no black holes and a low initial density. The initial density and black hole populations are related as a higher density will result in the ejection of more black holes. This means that the presence of long period binaries (and therefore a low initial density) would be an argument against the presence of a black hole population while the absence of long period binaries could be used as evidence to support the model with black holes. Therefore, combined with multi-epoch radial velocities for this cluster, the method developed in this work can eventually help us to gain insight into the initial density of Pal 5 to determine which one of these two models is more probable. As well, our knowledge of gravitational waves can be enhanced by improving our understanding of black holes in Pal 5. These are connected as it is believed that the dynamical formation of black hole binaries in the cores of globular clusters can produce binary black hole mergers, sources of gravitational waves (e.g. Antonini & Gieles, 2020).

Overall, binaries and globular clusters are dynamically tied and by improving our understanding of binaries, we can further our understanding of their host globular clusters. This work will use this idea to better constrain the intrinsic binary orbital period distribution using multi-epoch radial velocity data and hierarchical Bayesian statistics and then use this to infer the initial density of NGC 3201.

This thesis is organized as follows. Chapter 2 describes how mock data is generated, the procedure used to determine which stars, from a sample of stars observed in a given globular cluster with multi-epoch radial velocities, are classified as radial velocity variable and the fitting procedure used to fit the radial velocity curves. Chapter 3 describes the method developed to take the posterior samples from the radial velocity curve fits and use these to constrain the intrinsic binary period distribution for the binaries in the cluster. In this chapter we also discuss how to fit on a flexible form of the distribution to minimize assumptions about the distribution shape. Chapter 4 then takes the method developed in the prior two chapters and applies it to data for the globular cluster NGC 3201. NGC 3201 was chosen as it has a rich multi-epoch MUSE dataset available, making it a good test case for this work. In this chapter we also discuss what this intrinsic period distribution means for the initial density of NGC 3201. In the final chapter, Chapter 5, we present our conclusions and ideas for future work.

Chapter 2

Fitting Radial Velocity Curves

While the ultimate goal of this work is to develop and apply a method to find the intrinsic binary orbital period distribution of stars in a globular cluster, this population distribution depends on the inferred parameters for each cluster binary - making this a hierarchical process. Therefore, in order to determine the period distribution, we first need to determine which of the stars in the cluster are detected as radial velocity variable (binaries) and then, for each of these, find the probability distribution for its period. These probability distributions can be found by fitting the radial velocity (RV) curve for each star as this provides posterior probability distributions for each of the RV curve parameters. These parameters are the period (P), eccentricity (e), pericentre phase (ϕ_0), pericentre argument (ω), velocity semi-amplitude (K), and the barycenter velocity (v_0) as was described in Section 1.2 and more specifically in Equation 1.11. In this chapter, we discuss how we generate mock data for model testing in Section 2.1, the method in general in Sections 2.2 to 2.4, the sampling methods that were tested to probe the posterior probability distributions of the model parameters in Sections 2.5 to 2.7, and then the method is illustrated using mock datasets in Section 2.8.

2.1 Generating Mock Multi-Epoch Radial Velocity Data

Before we can begin to formulate a method and test it, we need a way to generate mock datasets with known binary orbital parameters. To do so, we use the python package `velbin`[1] (Cottaar & Hénault-Brunet, 2014). This package allows us to generate radial velocities for a sample of stars and binaries given input period, eccentricity, and mass ratio distributions for a population of binaries, as well as the desired fraction of binaries, cluster systemic velocity, cluster velocity dispersion, the mass of the primary stars, and the epochs of observations. This returns the multi-epoch radial velocities and the adopted radial velocity uncertainties for the stars (and binaries) by solving the radial velocity equation (Equation 1.11) for each of the generated binaries. From this, the true values for the radial velocity curve parameters can be extracted.

We generated a large array of mock data sets by tweaking the various input parameters such as the period distribution, number of stars, binary fraction, velocity uncertainties, velocity dispersion, and time sampling of the observations. This allows us to see the effect of changing different quantities (and thus our ability to constrain the period distribution for different clusters) and to ensure that the method works in a variety of cases. For example, by changing the number of epochs of data in the dataset, we can see how the three sampling methods, to fit RV curves, (see Sections 2.5 to 2.7) compare in different cases for fitting radial velocity curves.

For these generated stars, the method that will be discussed in this chapter can be applied to multi-epoch radial velocity data to fit radial velocity curves for each star identified as radial velocity variable and ensure that the true solution is recovered. Once we have this mock multi-epoch data and the posterior samples from the radial velocity curve fits, we can then use this data to fit the binary period distribution as

---

[1] https://github.com/MichielCottaar/velbin

will be discussed in Chapter 3.

Thus, `velbin` allows us to generate mock datasets that fulfill our needs and allow us to compare the results from both the radial velocity curve fits and the period distribution fits with known true values - allowing us to confirm that the method developed in this work is accurate.

2.2 Binary Star Classification

Before the radial velocity curves can be fit, we first need to determine which stars in a sample are radial velocity variable and thus classified as binaries using multi-epoch radial velocity data. To determine if a star is considered variable or non-variable, a χ^2 test is applied (Equation 2.1) to the radial velocity data of each star. In this case, the null hypothesis is that the radial velocity is constant (non-variable) and the alternative hypothesis is that the radial velocity is variable. This allows us to determine if the variations in the velocity can be explained just by measurement errors and noise or if orbital motions are required. In the equation below, indices i correspond to the epochs of measurement, v_i is the velocity at each epoch i, $\bar{v}$ is the weighted mean velocity, and σ_i is the measurement uncertainty on the velocity at that epoch:

$$\chi^2 = \sum_i \frac{(v_i - \bar{v})^2}{\sigma_i^2}. \tag{2.1}$$

Once we have the χ^2 value for a star, we can find the probability, $P\left(\chi^2, df\right)$, of observing a χ^2 value equal to or larger than this value under the null hypothesis given the number of degrees of freedom, df. In this case, df is one less than the number of radial velocity epochs for the star (i.e. under the null hypothesis, the model only has one parameter, the value of the constant mean velocity). This probability, $P\left(\chi^2, df\right)$, represents the cumulative probability in the right tail of the χ^2 distribution. We can see the effect of the number of epochs of data available on the χ^2 distribution in the left panel in Figure 2.1.

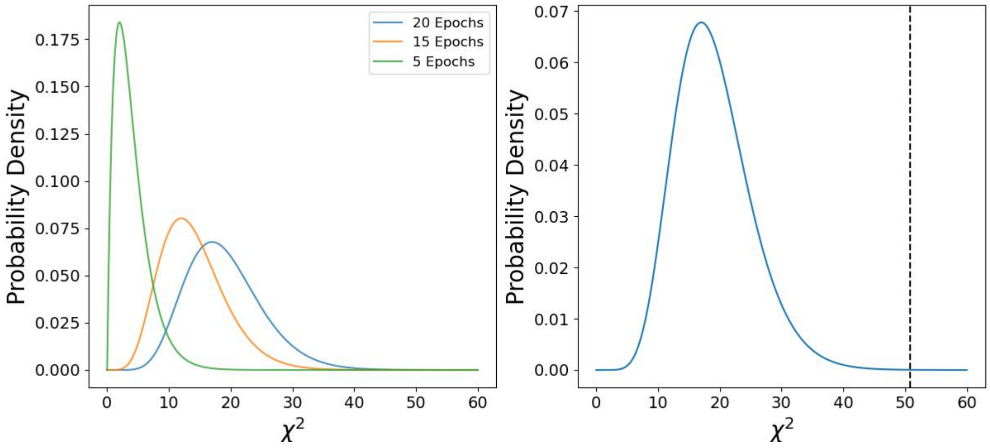


Figure 2.1: The panel on the left shows the χ^2 distributions for different degrees of freedom (number of epochs of data) and the panel on the right shows our threshold for detection, p, with the dashed black line. This example is for the 20 epoch case. Systems with a χ^2 to the left of this line are classified as non-variable and systems to the right as variable.

In order to determine if we can confidently reject the null hypothesis, and therefore consider the star to be radial velocity variable, we need to compare the probability that this χ^2 could be due to only measurement uncertainties to a probability threshold, p (see the right panel of Figure 2.1). Any stars with a probability P greater than this threshold p have sufficient evidence that they are consistent with the null hypothesis and the stars are classified as non-variable. Any stars with a probability P smaller than the threshold p do not have sufficient evidence that they are well described by the model so the null hypothesis is rejected and the stars are classified as variable.

In line with the method used by Cottaar & Hénault-Brunet (2014), this threshold p is taken to be 10^{-4} in this work. This threshold was chosen as it results in a negligible rate of false positives (labelling a non-variable star as variable due to apparent variability from measurement noise) for samples of a few hundred to a few thousand stars, like the typical datasets for globular clusters. Based on this criteria, we can see the stars that are detected as variable (binaries) and non-variable (apparently single stars) in the right panel of Figure 2.1 for an example with 20 epochs of data.

One thing to note is that while the variable stars are expected to all be binaries (due to the low expected rate of false positives and the small chance of seeing variability due to pulsations or atmospheric variability), it is expected that some (and possibly a significant fraction) of the stars classified as non-variable are undetected binaries. This is especially true for stars with sparse/limited radial velocity measurements. This is because binaries with long periods ($P \approx 10^2$ for a typical cluster) are harder to detect due to their variations occurring over much longer time scales and their smaller RV semi-amplitude, meaning that they often appear to be non-variable. As a result, in a real data set we cannot distinguish between these long-period binaries and single stars and we cannot fit radial velocity curves to these undetected binaries. For a discussion on how we get around this issue when inferring the intrinsic period distribution of binaries in a cluster, see Section 3.1.

Now that stars can be classified as either variable or non-variable depending on their radial velocity measurements, radial velocity curves can be fit to the velocity measurements for each star classified as variable.

2.3 Likelihood

To fit the model radial velocity curve to the velocity measurements of a star, we need to fit on the set of free parameters that describe the radial velocity curve. To find the best-fitting parameters and their credibility intervals, Bayesian inference is used. In general, Bayesian inference is a way of finding the posterior probability distribution for a set of parameters (linked to our degree of belief in the true value of the parameter) by using the likelihood of the data given those parameters, and prior knowledge of those parameters. This is formulated by Bayes' rule which is given by the following equation where $\mathcal{M}$ is the model and $\mathcal{D}$ is the data:

$$P\left(\mathcal{M}|\mathcal{D}\right) = \frac{P\left(\mathcal{D}|\mathcal{M}\right)P\left(\mathcal{M}\right)}{P\left(\mathcal{D}|\mathcal{M}\right)} = \frac{\mathcal{L}\left(\mathcal{M}\right)\pi\left(\mathcal{M}\right)}{\mathcal{Z}}. \tag{2.2}$$

In this equation, $P\left(\mathcal{M}|\mathcal{D}\right)$ is the posterior and describes the probability of the model given the data. $P\left(\mathcal{D}|\mathcal{M}\right) = \mathcal{L}\left(\mathcal{M}\right)$ is the likelihood and is the probability of the data given the model, $P\left(\mathcal{M}\right) = \pi\left(\mathcal{M}\right)$ is the prior and is the probability of the model, and $P\left(\mathcal{D}\right) = \mathcal{Z}$ is the evidence and is the probability of the data.

We typically describe the model using a series of parameters (such as the set of parameters that describe a radial velocity curve in the case of fitting RV curves) so Bayes' Rule can be rewritten as:

$$P\left(\vec{\theta}|\mathcal{D}\right) = \frac{\mathcal{L}\left(\vec{\theta}\right)\pi\left(\vec{\theta}\right)}{\mathcal{Z}} \text{ with } \vec{\theta} = \{\theta_1, \theta_2, ..., \theta_n\}, \tag{2.3}$$

where $\vec{\theta}$ is the set of all model parameters θ_i that describe the model. When doing Bayesian model comparisons, the evidence ($\mathcal{Z}$) is important, however, for parameter estimation the posterior probability can be calculated up to a normalization constant (the evidence), allowing us to ignore the evidence.

Thus, when using Bayesian parameter estimation for a given problem, a functional form needs to be chosen for the likelihood as well as for the prior (one for each model parameter when assuming independent priors).

In the case of fitting radial velocity data, the errors on the velocities are assumed to be Gaussian so we can use a Gaussian likelihood (Equation 2.4) for the velocity of each star at each epoch. This Gaussian is centered on the predicted radial velocity at a given epoch, $v_{i,m}$, and the width parameter of the Gaussian is the error on the velocity, σ_i. It is then evaluated at the observed velocity, $v_{i,o}$ at that epoch. This Gaussian likelihood is evaluated for every star that is classified as a binary and the total likelihood for the set of radial velocity measurements of a star is given by:

$$\mathcal{L}_n(\vec{\theta}) = \prod_{i}^{\#\text{of epochs}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(v_{i,o} - v_{i,m}(\vec{\theta}))^2}{\sigma_i^2}\right) \tag{2.4}$$

where $\mathcal{L}_n$ is the likelihood for star n, i are the epochs of measurements for the star, $v_{i,o}$ are the observed velocities, $v_{i,m}$ is the model radial velocity calculated using the radial velocity equation parameters (period, eccentricity, pericentre phase, pericentre argument, velocity semi-amplitude, and barycenter velocity) at each epoch (i) in the observed dataset, and σ_i are the measurement errors associated with each velocity.

As the values of these total likelihoods result from the product of several potentially small numbers, they can be quite small, leading to underflow issues with the numerical precision of computers. To overcome this, likelihoods are often treated as log-likelihoods. Thus, the total log-likelihood for the RV data of star n given the model parameters is:

$$\log\left(\mathcal{L}_n\right) = -\frac{1}{2}\sum_i \left(\log\left(2\pi\sigma_i^2\right) + \frac{\left(v_{i,o} - v_{i,m}\right)^2}{\sigma_i^2}\right). \tag{2.5}$$

In order to calculate the model radial velocities, v_m, based on the radial velocity curve parameters, the python package `twobody` [2] is used, which essentially follows the physics described in Section 1.2 (see Equation 1.11). The assumption that goes into this calculation is that the radial velocity variations are caused by the gravitational orbit of two point masses and that the radial velocity curve is not impacted by any other bodies or any stellar oscillations. With this, we now have all of the components needed to compute the likelihood term in Equation 2.3.

---

[2] https://github.com/adrn/TwoBody

2.4 Priors

Now that we have a likelihood function, we need to define priors for each model parameter. Once we have a prior for each parameter, these can be combined to form the overall prior using the following equation, with the assumption that the prior probabilities for each parameter are independent:

$$\pi\left(\vec{\theta}\right) = \prod_i \pi\left(\theta_i\right). \tag{2.6}$$

When working with priors, the priors can either be informative and based on our previous knowledge of a parameter, or they can be uninformative and reflect our level of ignorance of the parameter value. An example of a common uninformative prior is the uniform prior. The uniform prior is a distribution with equal probability density at every point within two bounds and zero probability density outside of those bounds. As the prior is a probability distribution, it must be normalized so the uniform prior is given by:

$$\pi\left(\theta\right) = \begin{cases} \frac{1}{U-L} & L \leq \theta \leq U \\ 0 & \text{otherwise} \end{cases} \tag{2.7}$$

where L is the lower bound and U is the upper bound. These bounds are chosen such that all valid parameter values are contained within. This prior is used for the $\ln\left(P\right)$, ϕ_0, and ω parameters in the case of our fits to the radial velocity curves. For $\ln\left(P\right)$, the lower bound is chosen such that any binaries with a period less than this would result in a binary with one star within the radius of the other. The upper bound is chosen such that it is sufficiently large enough that we do not expect any binaries with a period longer than this to remain bound to the present day, which is dependent on the present-day hard-soft limit of a given cluster. From Giesers et al. (2019), the period bounds that satisfy these conditions for the cluster of interest in this work (NGC 3201) are 0.2 and 4096 days. For the two angles, ϕ_0 and ω, the bounds range from 0 to 2π so that the full range of allowed angles is encompassed.

In other cases, such as for the eccentricity, velocity semi-amplitude, and the barycentric velocity, we have a bit more information about the parameters so we can instead use informative priors. Following the work done by Price-Whelan et al. (2020) on stellar binaries, the prior on the eccentricity is given by the beta distribution:

$$\pi_\beta\left(e;a,b\right)=\frac{\Gamma\left(a+b\right)}{\Gamma\left(a\right)\Gamma\left(b\right)}e^{a-1}\left(1-e\right)^{b-1}, \tag{2.8}$$

with parameters $a = 0.867$ and $b = 3.03$ as was found by Kipping (2013) (through exoplanet observations) and where the gamma function, $\Gamma\left(x\right)$ is given by:

$$\Gamma\left(x\right)=\int_0^\infty t^{x-1}e^{-t}dt. \tag{2.9}$$

While this prior was found for exoplanets, it is also expected to hold for stellar binaries as the physics governing the orbits of exoplanets and stellar binaries is the same.

For the velocity parameters, again using the prescription in Price-Whelan et al. (2020) for stellar binaries, their priors can be described using normal distributions given by

$$\pi_{\mathcal{N}}\left(x;\sigma,\mu\right)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \tag{2.10}$$

where x is the parameter of interest (either K or v_0 in the case of the RV fit), σ is the width of the distribution, and μ is where the distribution is centered (the mean of the distribution). In the case of the velocity semi-amplitude, $\mu = 0$ km s^{-1} and σ is given by:

$$\sigma^2 = \sigma_{K,0}^2 \left(\frac{P}{P_0}\right)^{-\frac{2}{3}} \left(1 - e^2\right)^{-1}, \tag{2.11}$$

where $\sigma_{K,0}$ was taken to be 30 km/s and P_0 is 365 days. This was determined by Price-Whelan et al. (2020) as this prior results in a companion mass independent of the period or eccentricity for a fixed primary mass. For the binaries' barycentric velocity, v_0, the normal distribution is centered on the systemic velocity of the cluster and therefore needs to be modified depending on which cluster is being investigated. As we will later apply this method to NGC 3201 (see Chapter 4) and the literature value for the systemic velocity is 494 km s^{-1} (Giesers et al., 2019), the normal distribution for the systemic velocity is centered on μ = 494 km s^{-1} in this case. The width for the prior is taken to be σ = 100 km s^{-1} in alignment with Price-Whelan et al. (2017).

A summary of the priors for each of the parameters that describe the radial velocity curves is given in Table 2.1.

Parameter	Prior	Unit
$\ln P$	$\mathcal{U}\,(0.3, 4096)$	days
e	$\mathcal{B}\,(0.867, 3.03)$	
ϕ_0	$\mathcal{U}\,(0, 2\pi)$	rad
ω	$\mathcal{U}\,(0, 2\pi)$	rad
K	$\mathcal{N}\,(0, \sigma_K^2)$	km s^{-1}
v_0	$\mathcal{N}\,(v_{sys}, 100^2)$	km s^{-1}

Table 2.1: The parameters used to fit the radial velocity curves (see Equation 1.11) and their corresponding priors. $\mathcal{U}$ refers to a uniform distribution with a specified minimum and maximum. $\mathcal{B}$ refers to the Beta distribution with the necessary shape parameters and $\mathcal{N}$ is the normal distribution with a specified mean (the cluster systemic velocity) and width. For the velocity semi-amplitude, K, the width of the normal distribution, σ_K is given by Equation 2.11 where $\sigma_{K,0}$ is 30 km s^{-1} and P$_0$ is 365 days.

With these priors and the likelihood (see Section 2.3) for the radial velocity curves, the only remaining component left in Bayes' rule to find the posterior is the evidence. However, it turns out that the evidence is computationally intensive to calculate as it involves a high dimensional integral and, as mentioned earlier, actually not needed for parameter estimation. Thus, instead of calculating the evidence, we can calculate the posterior up to a normalization constant using

$$P\left(\vec{\theta}|\mathcal{D}\right) \propto \mathcal{L}\left(\mathcal{D}|\vec{\theta}\right)\pi\left(\vec{\theta}\right). \tag{2.12}$$

It is still computationally intensive to evaluate this posterior probability distribution over a grid when working with a multitude of model parameters. Instead, we need to utilize a sampling method in order to approximate and recover the posterior probability distribution for the sets of parameters and thus fit on the radial velocity curve.

By fitting the radial velocity curve, we obtain posterior samples for each of the radial velocity curve parameters, including the desired parameters - the orbital period and systemic velocity (see Section 3.3 for a discussion on why the systemic velocity posteriors are required). These posterior samples are important as these are what will later be used to fit the binary period distribution for a given globular cluster based on multi-epoch radial velocity measurements of a large number of stars. While there are a variety of sampling methods that can be used, we need a method that works efficiently and also generates the desired number of posterior samples. Three methods that were considered in this work and their advantages and disadvantages are outlined in the following sections.

2.5 Markov Chain Monte Carlo (MCMC)

One of the methods considered was Markov Chain Monte Carl (MCMC), and in particular, `emcee`[3] (Foreman-Mackey et al., 2013) which is the python implementation of the affine-invariant ensemble sampler proposed by Goodman & Weare (2010).

---

[3] https://emcee.readthedocs.io/en/stable/

MCMC methods are a class of algorithms for sampling from a probability distribution, which involves constructing a Markov chain that evolves towards an equilibrium distribution, and for which the density of samples in a region of parameter space is proportional to the value of the target (posterior) probability distribution at that location. It works by exploring the parameter space by taking a series of steps starting from an initial set of points. From the current point in the chain, the next step is always taken based on some sort of stochastic process. The specific process used is determined by the type of MCMC algorithm being used. This process then continues where each following step is chosen based on taking a step from the prior position. By repeating this random process, each step only depends on the one before it, making this series of steps (referred to as the chain) a Markov process. Regardless of the MCMC implementation used, the stochastic process of accepting or rejecting the next step is always based on favouring an area of higher posterior probability (although moves to regions of lower posterior probability are also possible, albeit with a lower acceptance probability). Each step in the chain represents a possible set of model parameters and collectively, the density of samples from the chain (once it has reached the target equilibrium distribution) at a given position in parameter space can give us the posterior probability distribution as it is proportional to the posterior probability density at that point in parameter space (e.g. Ivezić et al., 2020).

However, we do not want to consider the first part of the chain when looking at the posterior distribution as this part still has a memory of the initial starting condition and might not be an accurate representation of the distribution. In other words, we are interested in the target equilibrium distribution of the chain and not where it starts. Thus, when considering the posterior, we first remove a 'burn-in' portion (the part before the chain has lost the memory of the initial starting point) and then look at the remaining chain elements.

Additionally, we need to ensure that the chain has reached the equilibrium state.

To do this, we test for chain convergence. There are a few ways of testing for chain convergence and while none of them guarantee that the chain has fully converged, they provide us with an estimate for how many iterations the chain should have. One common way this is done is by using an auto-correlation test within `emcee`. This test looks at how many iterations are required before the chain begins to be correlated with itself, referred to as the auto-correlation time. Within the python package `emcee`, it is recommended to run the chain for at least 50 times as many iterations as the auto-correlation time to ensure convergence.

As stated above, for this work we are specifically considering `emcee`. `emcee` involves using multiple walkers and steps are taken by using stretches defined by the current positions of the walkers. The result of this in an affine-invariant ensemble method, meaning that it can efficiently explore distributions with parameters that are linearly correlated. Due to the use of multiple walkers, this process is easily parallelizable to reduce its run time. In addition, it can be more efficient than other MCMC methods as if one walker finds an area of higher posterior probability, it has a chance to 'drag' the other walkers to this point and better sample this area, making it a potentially efficient MCMC method for cases like fitting a radial velocity curve.

While MCMC is a good method for sampling parameter space in a lot of cases, one of its drawbacks is that it can miss peaks in multi-modal distributions as if the walkers find a local peak in the posterior, they can get 'stuck' there and miss other peaks. This can lead to misleading results. Additionally, as mentioned above, it can be difficult to determine if the MCMC chain has fully converged. There are other sampling methods where convergence is better defined, such as nested sampling (see Section 2.7). Due to this, MCMC is often run for a longer period of time than other sampling methods to ensure convergence. For example, in the case of using MCMC for fitting radial velocity curves, we found that it took substantially longer than when using the other sampling methods described below.

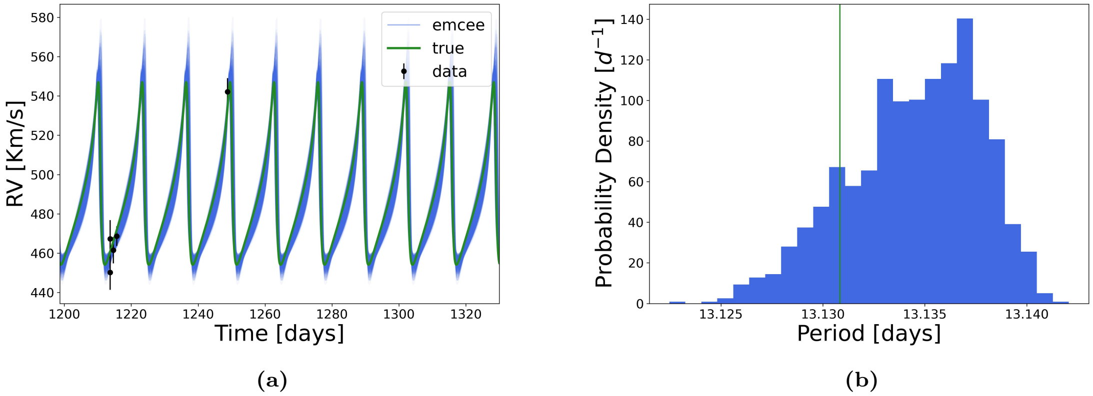


Figure 2.2: An example of a radial velocity curve fit using `emcee`. Blue curves in the left panel and the blue histogram in the right panel show the posterior samples. The true value for the period and curve are shown in green. Panel (a) shows the radial velocity curve with the mock radial velocity measurements shown in black and panel (b) shows the histogram of posterior period samples.

We can see an example of a fit to a radial velocity curve using `emcee` in Figure 2.2. This is an example of a fit to the mock data of a short-period binary with a period of 13.13 days with 20 epochs of radial velocity measurements over 1605 days. In the left panel, we can see the model fit to the radial velocity measurements in blue and the true solution in green. Note that this does not show all of the velocity measurements as this is a short-period binary so for easier visualization, only ten full periods are shown. However, we can see from the data points included in this range that the best-fitting model does agree with the measurements and recovers the true solution. On the right, we can see the histogram of the period posterior samples with the true value shown in green, again showing that the true solution is recovered.

Diagnostic plots for this fit are shown in Figures 2.3 and 2.4. In Figure 2.3, we can see the post burn-in chain (where the burn-in was taken to be 500) and from this it looks like the chain is fully converged. In Figure 2.4, we can see the corner plot. Corner plots show the 2D joint posterior probability distribution for each pair of parameters as well as the marginalized posterior probability distribution for each parameter. The grey and black regions within each box shows the 2D contours for the combinations of parameters. The confidence regions $(1\sigma, 2\sigma, 3\sigma)$ are shown and

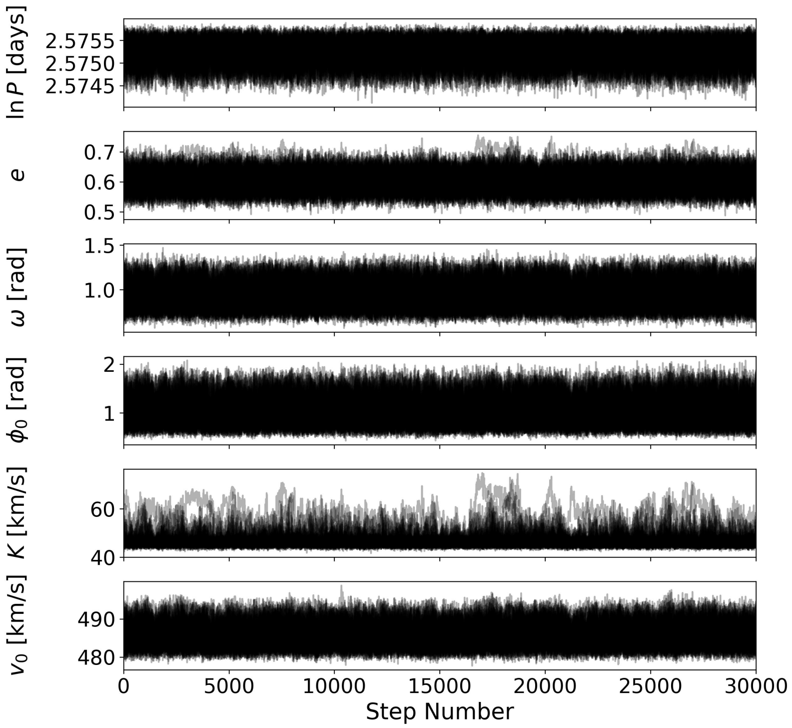


Figure 2.3: The continuation of Figure 2.2 showing the MCMC chain for the fit.

beyond that it shows the individual samples. This is a useful diagnostic tool as it provides information about the inferred value and credibility intervals for each parameter but also information about any co-variances between parameters. For example, if two parameters are independent then the contours will be circular, so any non-circular contours can provide information about the type/strength of any correlations. We can see an example of this for the two angle parameters, ϕ_0 and ω, as their 2D contour shows a clear correlation.

While this example is a case of a good fit, there are also cases where `emcee` does not recover the full posterior in multi-modal cases, as will be seen in Section 2.8. Moreover, we found that `emcee` had long run times compared to other sampling methods. Overall, while MCMC can be a good method for sampling parameter space, due to some of its limitations, it is not an ideal method for attempting to fit radial velocity curves in the context of this work as radial velocity curves can have multi-modal solutions and it is important to ensure convergence in a timely manner given the number

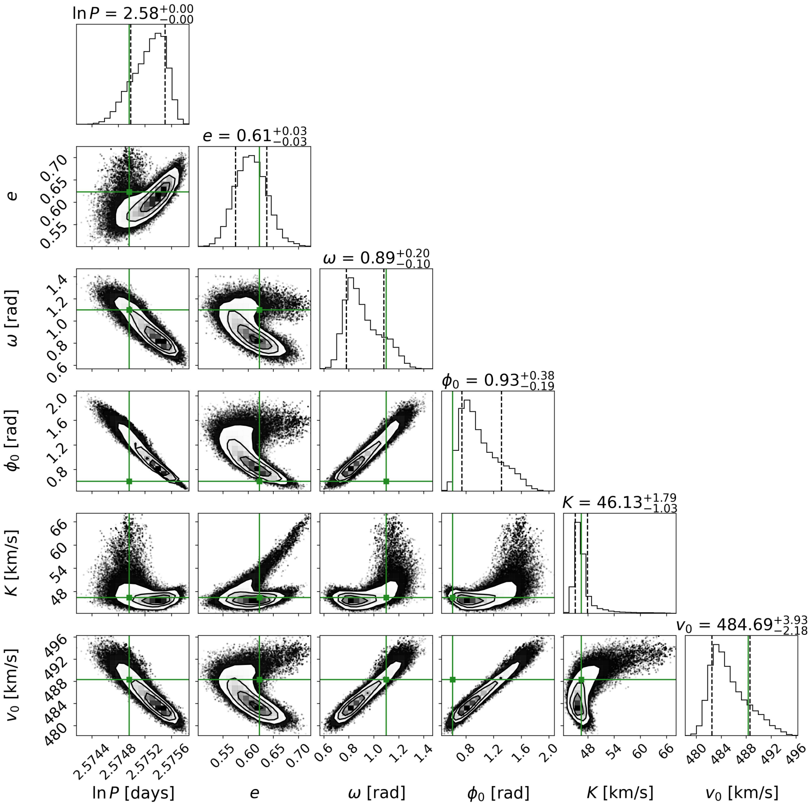


Figure 2.4: The continuation of Figure 2.2 and 2.3 showing the corner plot for the fit. The true values are shown with the solid green lines and the 1σ interval is shown with the dashed black lines in the marginalized posterior probability distributions.

of stars (few hundreds to few thousands) in our typical samples.

2.6 Rejection Sampling

Another method considered in this work for fitting radial velocity curves is rejection sampling and in particular, `TheJoker` [4] by Price-Whelan et al. (2017). Instead of starting with an initial set of parameters and then generating posterior samples from there (MCMC approach), rejection sampling begins with all of the potential samples and then rejects samples until the final sample is achieved. This initial set of prior samples is obtained by drawing a large number of samples from the prior distributions for each parameter (see Section 2.1) and then the likelihood for each set of parame-

---

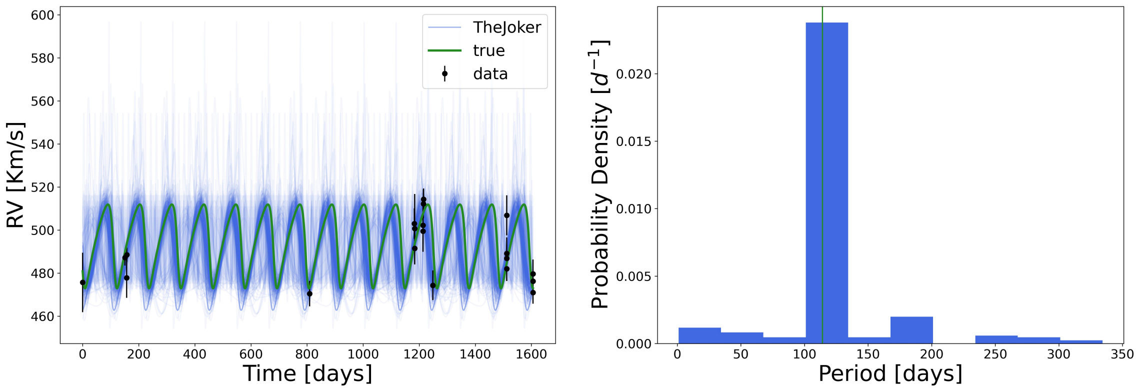
[4] https://github.com/adrn/thejoker

ters is determined. Based on the likelihoods for each set of parameters, the maximum likelihood across all the samples can be found. At this point, each sample is compared to the maximum likelihood and a random value is chosen between zero and the maximum likelihood, and if the likelihood of the sample is less than this random value, it is rejected. Through this process, higher likelihood solutions are more likely to remain, although, rejection sampling still remains a stochastic process.

Based on this process, the sample with the maximum likelihood is guaranteed to survive. However, it is possible that all other solutions will be rejected, or that not enough samples will be retained to be considered a fair representation of the posterior. In these cases where not enough samples are retained, the typical course of action is to either take the remaining samples and then use another sampling method to sample around these points, or to generate a larger number of prior samples and repeat the rejection sampling procedure.

For this work, `TheJoker` was the rejection sampler considered as it was developed specifically for the goal of fitting the radial velocity curves of exoplanets and/or binary stars, particularly when only sparse radial velocity measurements are available (e.g. Giesers et al., 2019, Price-Whelan et al., 2020).

In general, the benefit of this method is that it is unlikely to miss multi-modal peaks in the posterior distribution, unlike MCMC, since it starts with prior samples and works its way down from there so it cannot get stuck in a local peak of the posterior. In the case of sparsely sampled radial velocity curves, there is a significant potential that they will have multi-modal posterior probability distributions so the ability to recover this is key. However, in cases of well-sampled radial velocity curves (such is generally the case for the NGC 3201 data used in this thesis; see Chapter 4), `TheJoker` often fails to produce an adequate number of samples. This is because, in the cases with a well-sampled curve, the likelihood is more sharply peaked and the posterior is dominated by the likelihood. This makes it less likely that a significant number of the prior samples generated for rejection sampling will fall within a region of relatively high likelihood. Thus, either a larger number of prior samples is required, or else the use of another sampling method may be preferred if significantly increasing the number of prior samples for which the likelihood is to be evaluated is prohibitively computationally intensive. This is not an issue in the case of sparse radial velocity data as the likelihood is typically broader.



(a) **(b)**

Figure 2.5: An example of a radial velocity curve fit using `TheJoker`. Blue curves in the left panel and the blue histogram in the right panel show the posterior samples. The true value for the period and curve are shown in green. Panel (a) shows the radial velocity curve with the mock radial velocity measurements shown in black and panel (b) shows the histogram of posterior period samples.

An example of a radial velocity curve that is well fit by `TheJoker` and has the desired number of posterior samples (`TheJoker` recommended number is 256) is shown in Figure 2.5. This is a fit to the mock data of a binary with a period of 114.10 days and with 20 epochs of measurements over 1605 days. As with Figure 2.2, this figure shows the fit to the RV curve and the histogram of the posterior samples for the period. From this we can see that the true RV curve and period are recovered. The corner plot for this fit is shown in Figure 2.6. Due to the limited number of samples, this corner plot still uses darker colors for areas of higher density, but the contour lines are not drawn. From this we can still see that the areas of higher density do agree with the true value for each model parameter.

However, looking at another binary with the same number of measurements and

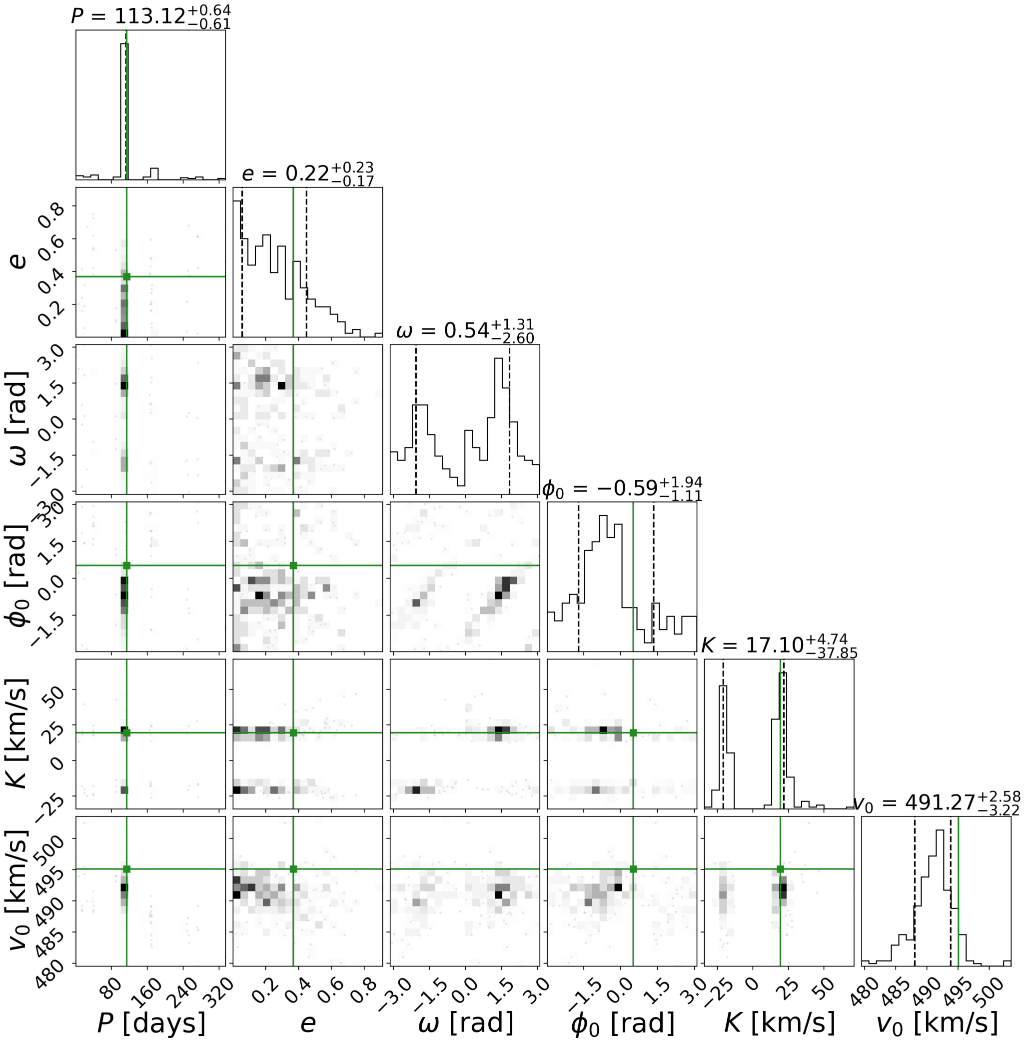
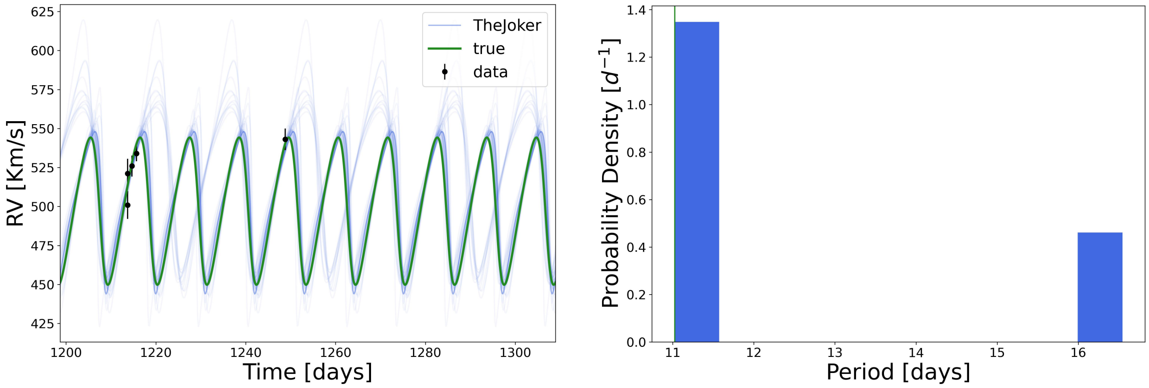


Figure 2.6: Continuation of Figure 2.5 showing the corner plot for this binary.

time sampling but with a period of 11.03 days in Figure 2.7, we can see an example of a fit where not enough samples were retained. In this case, only 51 samples were retained out of the desired 256. While this binary also appears to have a good fit to its radial velocity curve and to each of the model parameters (see Figure 2.8), we can see that not enough samples have been retained to have confidence in the posterior distribution. Therefore, in the cases where the true values of the parameters are unknown, we cannot be certain of these results.

One thing to note in Figure 2.8 is that we do see some parameters with multiple peaks in this corner plot for the marginalized posterior distributions. This is not unexpected when fitting RV curves as there can be multiple solutions that provide good fits to the data. As well, in the case of the angle parameters ω and ϕ, these both determine where the RV curve starts so it is not uncommon for them to have multiple peaks in their posterior distributions and this is not a concern.



(a) (b)

Figure 2.7: Same as Figure 2.5 for another binary.

Repeating the rejection sampling with a larger number of prior samples can quickly become quite inefficient as in some cases, it can take billions of prior samples in order to well sample the posterior. Using another sampling technique (like MCMC) in these problematic cases can be effective but this raises two questions. What are the limitations of the alternative sampler being used and, if an alternative sampler is found that works in these cases, why can this sampler not be used in all cases rather than using a mix of methods to analyze all the stars in a dataset? With this in mind, in this work we decided to try and find a sampler that would work at least as well as `TheJoker` for multi-modal posteriors but with the ability to provide more posterior samples when not enough samples are produced by `TheJoker`.

In summary, while rejection sampling and `TheJoker` is a good method for fitting sparsely sampled radial velocity curves with multi-modal posterior probability distributions, in cases with well-sampled radial velocity curves it is likely to fail to produce enough samples without relying on a very significant increase in computational times

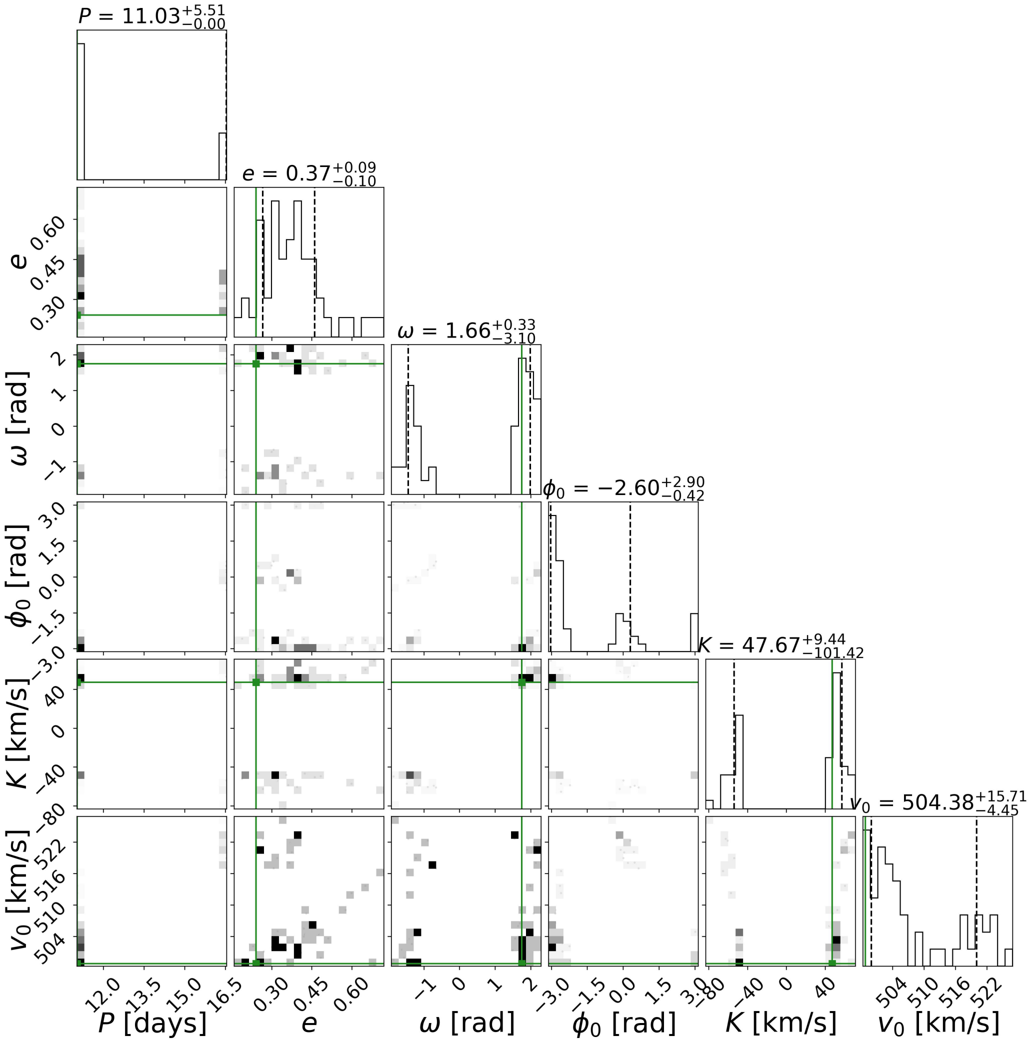


Figure 2.8: The corner plot for the binary fit in Figure 2.7.

or alternative sampling methods.

2.7 Nested Sampling

The third posterior sampling technique for Bayesian parameter estimation considered in this work for fitting the radial velocity curves is nested sampling (Skilling, 2004) and in particular, the python package `Dynesty`[5] (Speagle, 2020). Nested sampling is a method that was developed to estimate the evidence, $\mathcal{Z}$, in Equation 2.3 and

---

[5] https://dynesty.readthedocs.io/en/stable/index.html

works by starting with a prior volume and then integrating over the shells of constant likelihood within this prior volume. While this method was originally developed to find the evidence, it can also provide the posterior distribution. In implementations such as `Dynesty`, the importance of the posterior vs. evidence estimation can be chosen by the user, making it useful not only for evidence estimation, but for posterior sampling and parameter estimation as well.

In general, nested sampling works by breaking the prior volume into nested slices, sampling within each slice and then recombining the slices. Sampling within each slice is done by randomly generating points within each slice and comparing the likelihoods of these points. The worst point is removed and a new volume is constructed around the set of points and a replacement point is sampled from this new volume. As this process continues, the volume will shrink until the remaining estimated evidence left to calculate has reached the desired threshold. This process and how it compares to MCMC can be seen in Figure 2.9. The benefit of this method is that instead of trying to sample from the entire posterior distribution at once, it breaks it into smaller pieces that are easier to sample from.

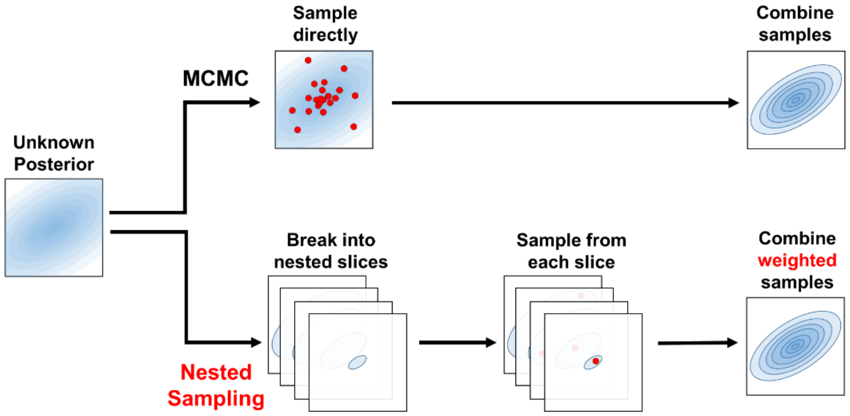
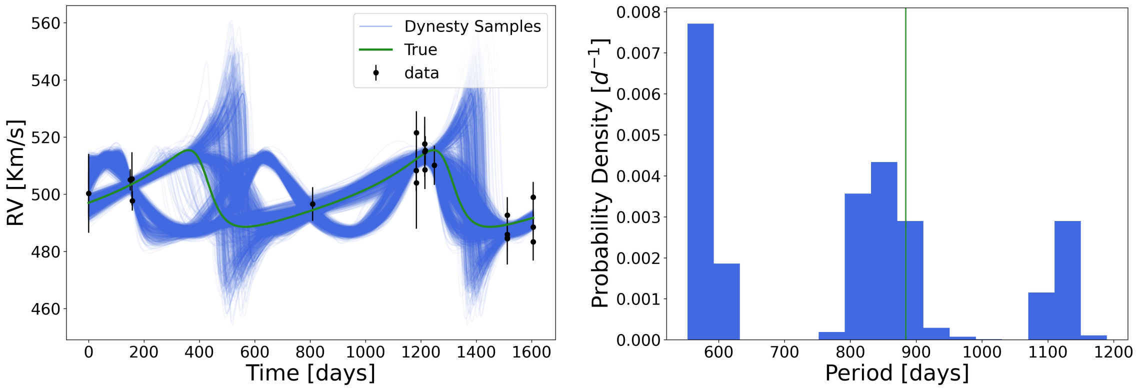


Figure 2.9: A comparison between the sampling approaches taken in MCMC and nested sampling. Figure from Speagle (2020).

Since nested sampling starts with the entire prior volume and works its way to the areas of high posterior probability, it is able to recover multi-modal distributions and is immune to the choice of starting conditions. Moreover, since it continues to draw samples at each step, nested sampling can well sample areas of higher posterior probability, unlike rejection sampling which is limited to the initial draw of parameters. Another advantage of nested sampling is that is has a more concrete stopping point than methods like MCMC. This stopping point looks at how much evidence is estimated to be remaining and stops when only a desired fraction of evidence is left and is given by the following equation:

$$\Delta \ln \mathcal{Z}_i \equiv \ln\left(\mathcal{Z}_i + \Delta\mathcal{Z}\right) - \ln\left(\mathcal{Z}_i\right), \tag{2.13}$$

where $\mathcal{Z}_i$ is the estimated evidence (found by numerically approximating the evidence integral) at iteration i and $\Delta\mathcal{Z}$ is the estimated remaining evidence (found using $\Delta\mathcal{Z} \approx \mathcal{L}_{max} X_i$, where X_i is the prior volume). In this case, $\Delta\mathcal{Z}$ can be defined to set the stopping condition for the nested sampling and the default value for this within `Dynesty` is $\Delta \ln \mathcal{Z} = 0.01$.



(a) **(b)**

Figure 2.10: An example of a radial velocity curve fit using `Dynesty`. Blue curves in the left panel and the blue histogram in the right panel show the posterior samples. The true value for the period and curve are shown in green. Panel (a) shows the radial velocity curve with the mock radial velocity measurements shown in black and panel (b) shows the histogram of posterior period samples.

The advantage of using `Dynesty` over other nested sampling algorithms is that it has the option for dynamic nested sampling. Dynamic nested sampling proceeds sim-

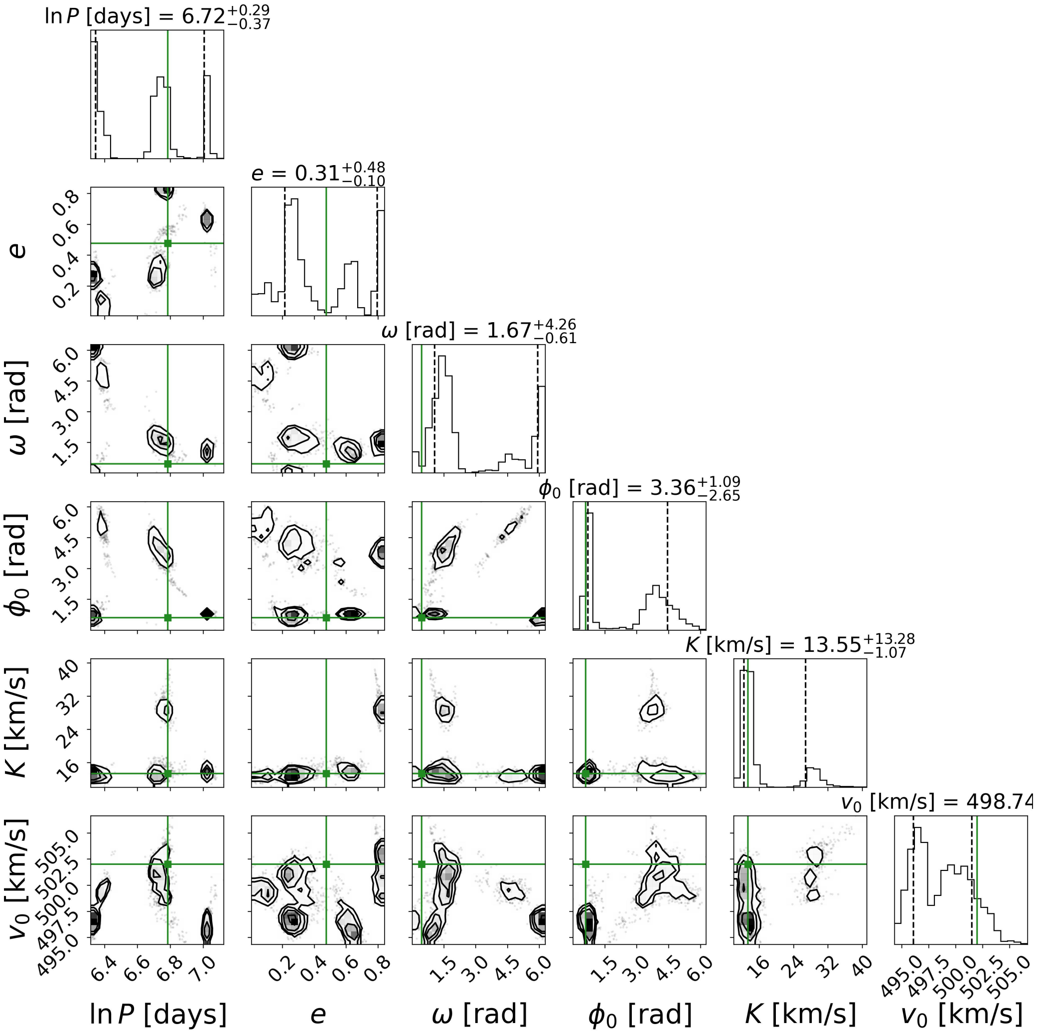


Figure 2.11: Corner plot for the binary fit in Figure 2.10.

ilarly to typical nested sampling, but with the ability to readjust during the sampling process. Instead of spending an equal amount of time and points across all sections of the prior volume, dynamic nested sampling initially samples the prior volume using less points. As areas of higher likelihood are found, more points are added in to increase the resolution in these areas. This does decrease the accuracy of the evidence since not all of the prior volume is sufficiently sampled, but, it does efficiently sample the posterior.

An example of a radial velocity curve fit using `Dynesty` can be seen in Figure 2.10 for a binary with a period of 884.15 days with 20 epochs of data over 1605 days. As with the fits shown for `TheJoker` and `emcee`, the right panel shows the histogram of the period samples and the left panel shows the radial velocity curve. As we can see in both, the radial velocity curve and the true period is recovered. Another thing to note is that there are multiple peaks in the period posterior. This is not a concern as it is possible for radial velocity curves to have multiple well-fit solution, as long as the true solution is included. However, we can also see in this example (both in the corner plot and the RV curve fit) that the eccentricity is not well recovered. This can be seen in the RV curve fit as the eccentricity is the parameter that determines the RV curve shape. As with `TheJoker` and `emcee`, `Dynesty` results can also be shown in a corner plot as we can see in Figure 2.11.

Another diagnostic plot for `Dynesty` is the trace plot as shown in Figure 2.12. This is a diagnostic plot unique to nested sampling and it shows the progression of the sampling as a function of the prior volume ($\ln x$). The colors represent the importance weight of the samples where blue represents the samples with the lowest weight and yellow are the samples with the highest weight. From Figures 2.10, 2.11, and 2.12 we can see that `Dynesty` is able to recover the true radial velocity curve as well as other possible solutions that need to be accounted for.

Based on this, `Dynesty` should be able to efficiently produce posterior probability distributions for the parameters of radial velocity curves in cases where the posterior is multi-modal and in cases where `TheJoker` is unable to retain enough samples. This potentially makes `Dynesty` the ideal method for fitting radial velocity curves in the context of this thesis, especially in cases with well-sampled radial velocity curves as will be seen in Section 2.8.

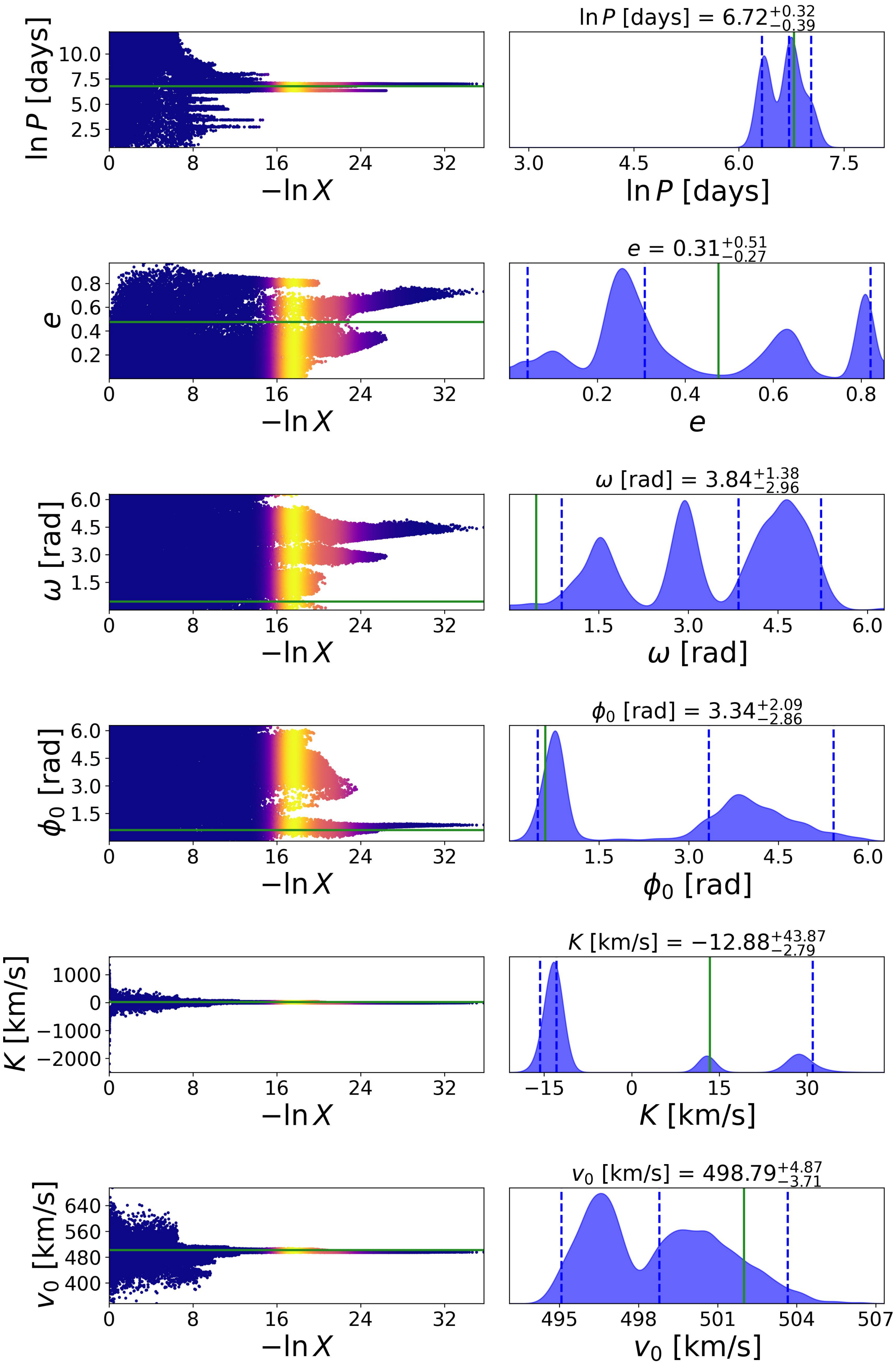


Figure 2.12: Trace plot for the binary fit in Figure 2.10 where the green lines in each plot represent the true value and the dashed blue lines on the right show the 16, 50, and 86% percentiles. In the left columns, the colors represent the importance weights with yellow having the highest weight. The x-axis shows the prior volume.

2.8 Testing Radial Velocity Curve Fitting Methods on Mock Data

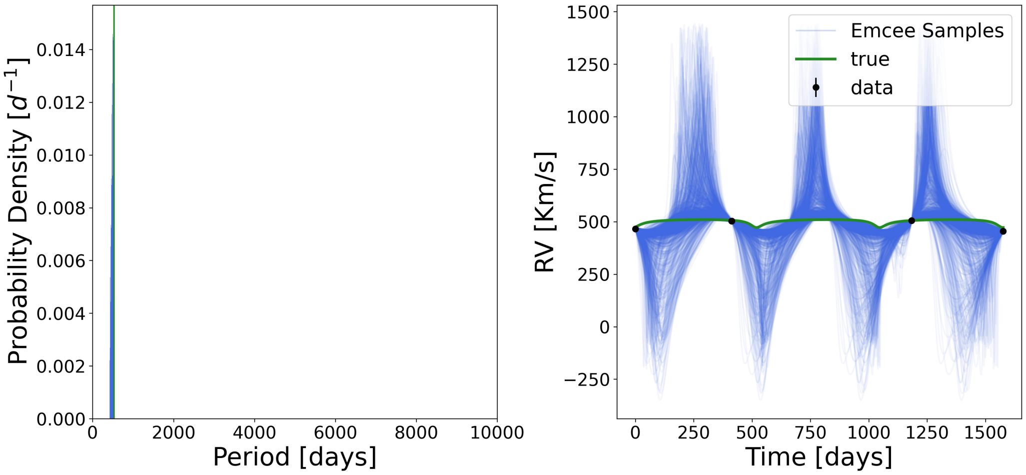
To further assess which sampling method works best in the case of fitting RV curves for the typical stars found in the NGC 3201 dataset (see Chapter 4), the three methods were tested on various mock datasets and compared.

One such test involved looking at the performance of each method in the case of a sparsely sampled RV curve as we can see in Figure 2.13. This is an example of a long-period binary (period of 527.74 days) with five epochs of data over 1575 days (time sampling taken from a typical star in the NGC 3201 dataset). The spacing between each consecutive epoch is approximately 412 days, 1 day, 771 days, and 392 days. The periods for the binaries in this mock dataset were drawn from a log-normal distribution centered at 2.5 and with a standard deviation of 0.8, the angle parameters were drawn from uniform distributions from 0 to 2π, the eccentricities were drawn from a thermal distribution, the mass ratio was drawn from a flat distribution (this determines the semi-velocity amplitude) and and barycenter velocity was drawn using scatter around the cluster systemic velocity (490 km/s in this case). In this figure we can see the posterior samples for the period on the left side and the RV curve fits for each of the sampling methods on the right. The top panel shows the fit using `emcee`, the middle panel shows the fit using `TheJoker`, and the bottom panel shows the fit using `Dynesty`. By comparing these fits, we can see that the true solution for the period is recovered for all three methods but with lower probability in the case of `TheJoker` and `emcee`. The RV curve for every binary in this dataset (151) was then fit using `TheJoker` and `Dynesty`. `emcee` was not run on the full dataset as we quickly found that it had a larger fail rate and it took substantially longer to run. We then looked at the success of `TheJoker` and `Dynesty` where success was defined as recovering the true binary period within 1σ. Doing so, we found that `TheJoker` successfully recovered the binary period for 72.0% binaries and `Dynesty` for 69.2% binaries for this particular example, showing that these two samplers are comparable for sparsely sampled RV curves.

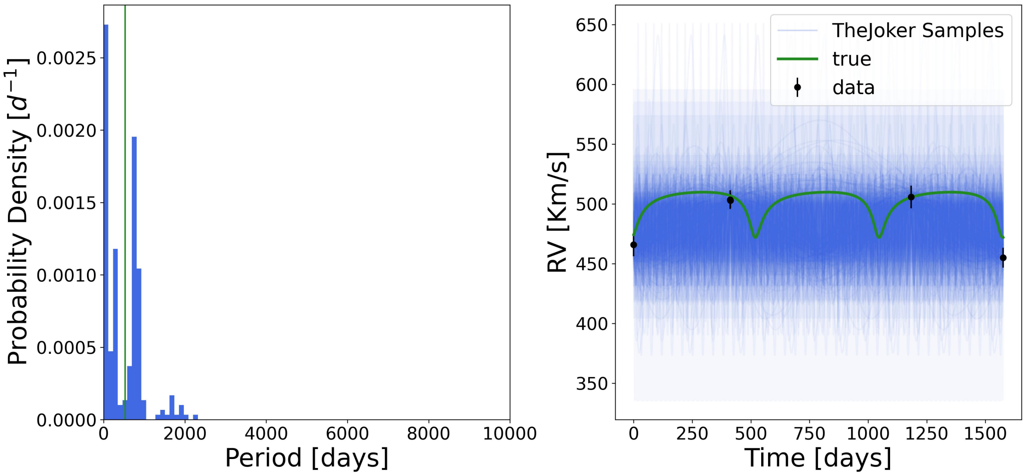
Another such test was done for the case of a binary with 20 epochs of data over 1605 days, again with time sampling chosen to represent a typical star in NGC 3201 and orbital parameters drawn from the same distributions as the five epoch case. An example of a binary from this dataset with a period of 11.03 days (the same binary as in Figure 2.7 but now showing the fits with `emcee` and `Dynesty` as well) is shown in Figure 2.14. In this figure, only part of the RV curve is shown so that the fit can clearly be seen. Comparing the three methods in this case, we can see that while all three methods are able to recover the true radial velocity curve, `TheJoker` does not retain enough samples whereas this is not an issue for `emcee` and `Dynesty`. As with the five epoch case, we also fit the RV curves for every detected binary in this dataset (280) with `TheJoker` and `Dynesty`. In this case, we found that `TheJoker` successfully recovered the period for 64.7% of binaries and `Dynesty` for 81.3% of binaries. Additionally, for 16 of the binaries in the dataset, `TheJoker` did not retain enough posterior samples. This highlights how `TheJoker` works well for sparsely sampled RV curves, but not in cases of well-sampled RV curves.

Another example of a binary with the same measurement times and spacing as the one in Figure 2.14 is shown in Figure 2.15. This example shows a binary with a period of 50.95 days and again the range shown is restricted to best show the fit. In this example, `Dynesty` and `TheJoker` are able to recover the true solution for the period, however it is completely missed by `emcee`.

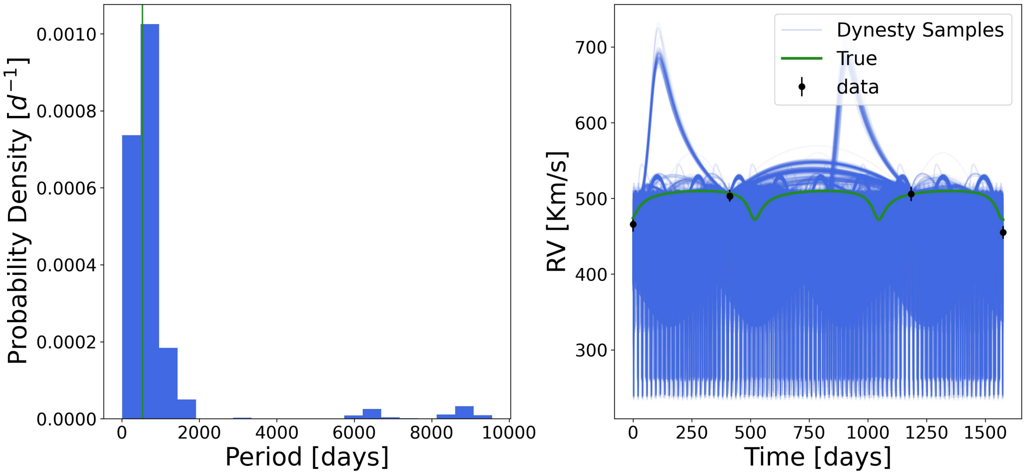
Overall it was found that `emcee` resulted in the longest run times and did not always recover multi-modal solutions or the true solutions found by the other sampling methods. `TheJoker` was found to be more accurate than `emcee` and worked well in cases of poorly sampled radial velocity curves, however, in cases with well-sampled radial velocity curves, there were cases where not enough samples were retained. This remained true even when large numbers of prior samples were used and increasing the number of prior samples further would be prohibitively slow. For `Dynesty`, it was found that it worked just as well as `TheJoker` in cases with sparsely sampled curves, but it was also found that it was just as efficient at sampling in the cases of well-sampled curves. It was also found that `Dynesty` took less time to run in our setup, making it the ideal method overall. Based on the tests and discussion above, `Dynesty` was chosen as the optimal method for fitting the RV curves for this work.



(a)



(b)



(c)

Figure 2.13: An example of a radial velocity curve fit using all three methods. The left shows the histogram of the period samples and the right shows the radial velocity curve for each fit. (a) shows the fit with `emcee`, (b) shows the fit with `TheJoker` and (c) shows the fit with `Dynesty`.

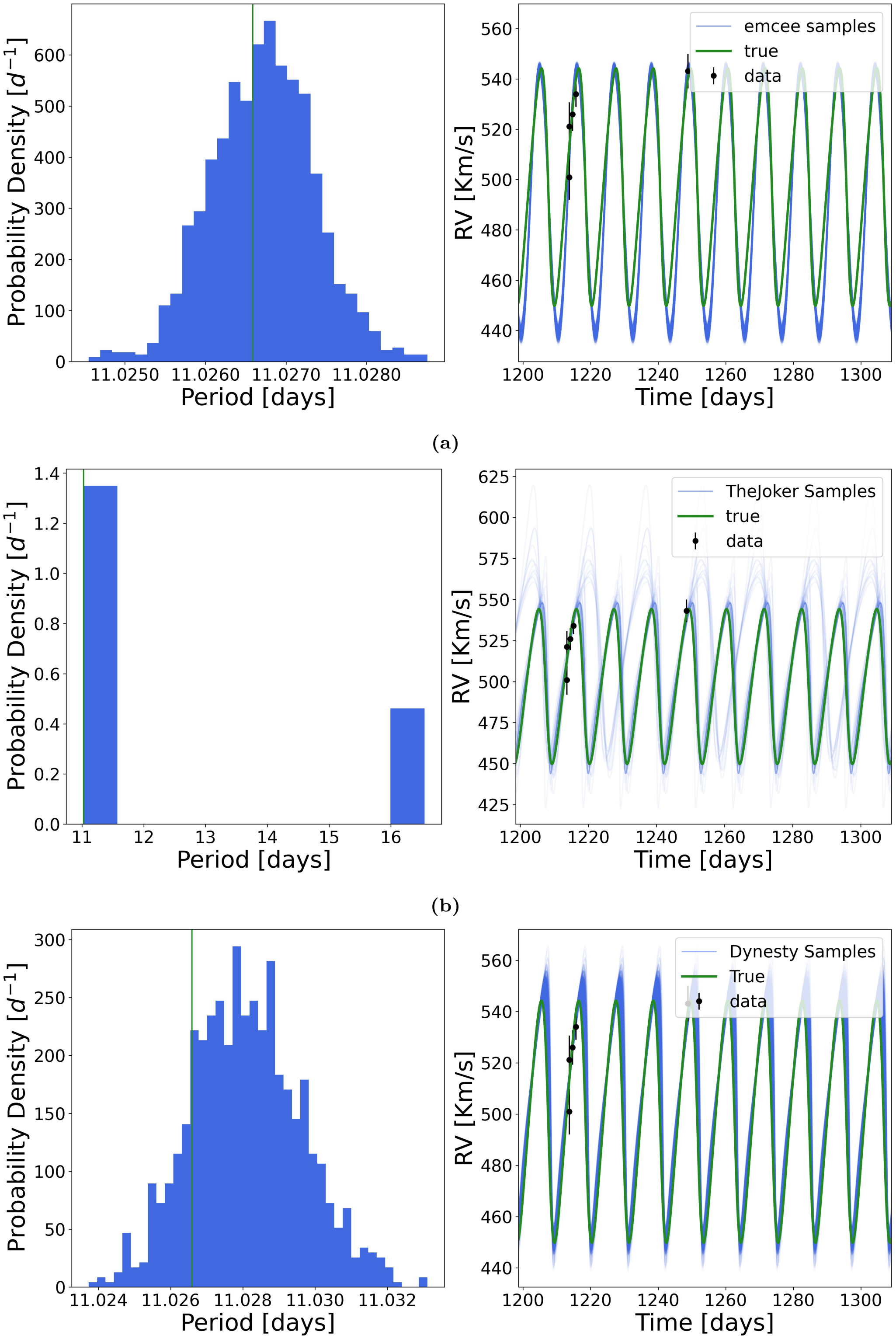
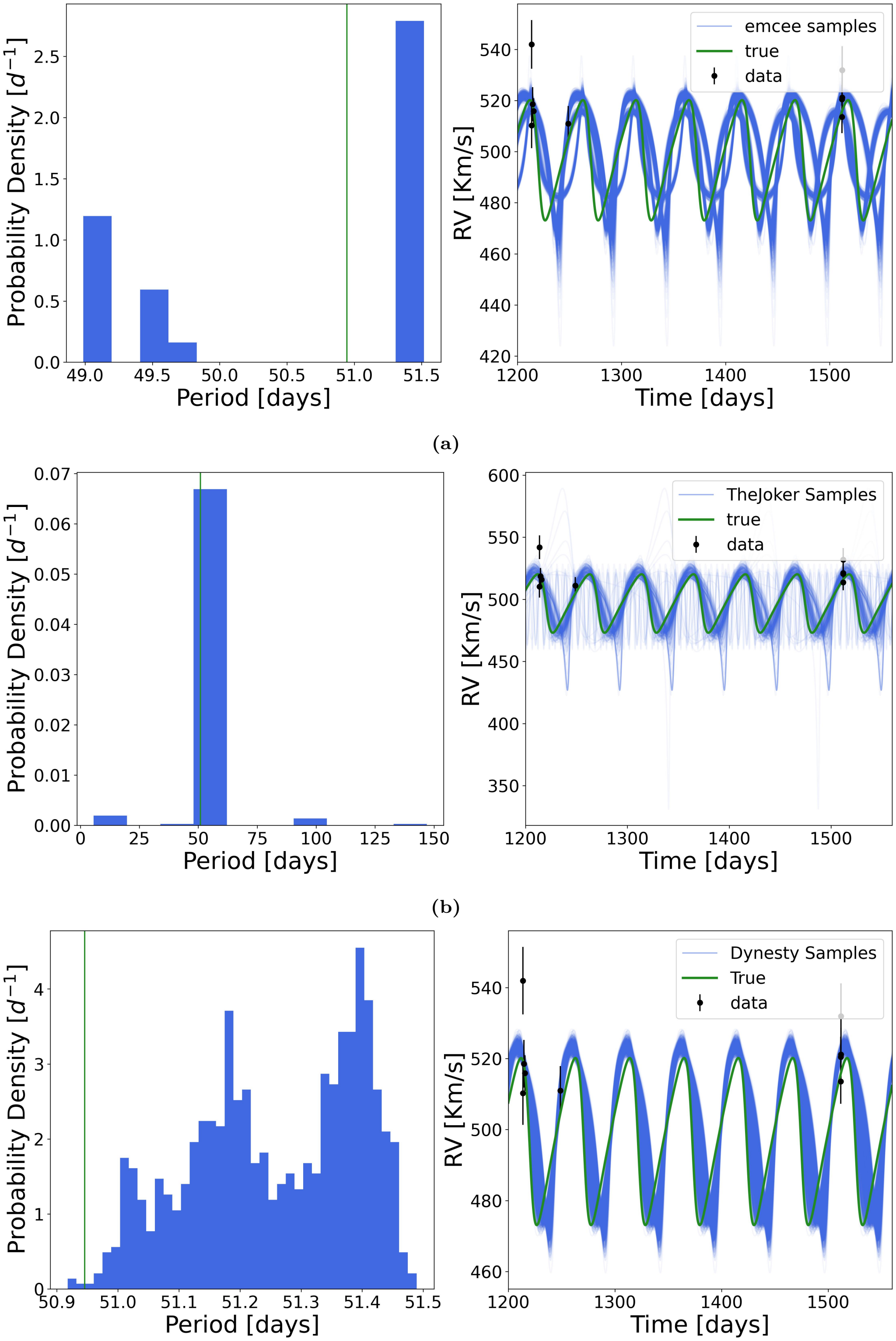


Figure 2.14: Same as Figure 2.13 for another binary.



(a)

(b)

(c)

Figure 2.15: Same as Figure 2.13 and 2.14 for another binary.

Chapter 3

Recovering the Intrinsic Binary Period Distribution

Now that we have a method for fitting the radial velocity curves and obtaining posterior samples for each detected radial velocity variable star, we use these samples to obtain the intrinsic orbital period distribution of binaries in a globular cluster.

As we will see in this chapter, this process is not as straightforward as taking the histogram of the period posterior samples or using the posterior samples to directly fit on the intrinsic period distribution as we need to be able to overcome the detection bias. Therefore, we need to construct a method that takes in the posterior samples obtained using the method in Chapter 2 and determines what intrinsic distribution would be needed to give rise to this population of detected RV variables. This also needs to be done while minimizing the assumptions made about the shape of the intrinsic period distribution as this is not well constrained.

In this chapter, we will look at how the posteriors are used to construct the period distribution (Section 3.1), what additional parameters are required to better constrain the period distribution (Sections 3.2 and 3.3), and examples of this method applied to mock data sets (Sections 3.4 and 3.5).

3.1 Constraining the Period Distribution Using the Results of the RV Curve Fits

As finding the period distribution for a population of binaries is a hierarchical process, we need to construct a likelihood that takes the data from the fits of individual RV curves for each RV variable star and uses these to determine the likelihood of the data given the parameters that describe the period distribution model.

The likelihoods for individual stars can be combined to find the total likelihood across all stars by taking the product of each individual likelihood:

$$\mathcal{L} \equiv \prod_{n=1}^{N} \mathcal{L}_n, \tag{3.1}$$

where N is the total number of RV variable stars, $\mathcal{L}_n$ is the likelihood for star n and $\mathcal{L}$ is the total likelihood.

As we are trying to find the total likelihood of the data for the population of stars given our model parameters ($\vec{\alpha}$), this means that we find the likelihood of the orbital period for each individual star given the parameters describing the period distribution of the population of binaries, and take the product. However, the posterior samples for each star were obtained by fitting all of the radial velocity curve parameters, most of which we no longer require. Therefore, we need to marginalize over the individual binary parameters. This is given by :

$$\mathcal{L}_{\vec{\alpha},n} = \int p(\mathcal{D}_n|\vec{\omega_n})p(\vec{\omega_n}|\vec{\alpha})\mathrm{d}\tilde{\omega}_\mathrm{n}, \tag{3.2}$$

where $\mathcal{L}_{\vec{\alpha},n}$ is the likelihood of the data given the period distribution model parameters $\vec{\alpha}$ for star n, $\vec{\omega}$ are the parameters that describe the radial velocity curve, $\mathcal{D}_n$ is the radial velocity data for star n, $p(\mathcal{D}_n|\vec{\omega_n})$ is the probability of the radial velocity data given the radial velocity curve parameters, and $p(\vec{\omega_n}|\vec{\alpha})$ is the probability of the radial velocity curve parameters given the period distribution parameters. This last term can be re-expressed in terms of the function of the period probability distribution given $\vec{\alpha}$, $f_{\vec{\alpha}}(P_n)$, the priors on all the radial velocity parameters, $p_0(\vec{\omega}_n)$, and the prior on the parameter of interest (the period in our case), $p_0(P_n)$, and is given by

$$p(\vec{\omega_n}|\vec{\alpha}) \equiv \frac{f_{\vec{\alpha}}(P_n)p_0(\vec{\omega}_n)}{p_0(P_n)}. \tag{3.3}$$

The probability in this form is based on the assumption that the prior on the other radial velocity curve parameters and the prior on the period are separable.

To avoid computing the seven-dimensional integral in Equation 3.2, this can be approximated as the sum over all posterior samples and thus the likelihood for each RV variable star can be approximated as:

$$\mathcal{L}_{\vec{\alpha},n} \approx \frac{1}{K}\sum_{k=1}^{K}\frac{f_{\vec{\alpha}}(P_{nk})}{p_0(P_{nk})}, \tag{3.4}$$

where K is the number of samples and P_{nk} is the period sample for star n and sample k. See Hogg et al. (2010) for details.

Putting this together with Equation 3.1 gives the total likelihood of the data given the parameters describing the orbital period distribution:

$$\mathcal{L}_{\vec{\alpha}} = \prod_{n=1}^{N}\frac{1}{K}\sum_{k=1}^{K}\frac{f_{\vec{\alpha}}(p_{nk})}{p_0(p_{nk})}. \tag{3.5}$$

This likelihood was derived and used by Hogg et al. (2010). In their work they looked at fitting the intrinsic eccentricity distribution for populations of binary stars and exoplanets given individual likelihoods, similar to the process in this work.

As the $\vec{\alpha}$ parameters will depend on the model chosen for the binary orbital period distribution, the priors will also vary. However, they are chosen to be uninformative priors such that the distribution includes the minimum and maximum allowed periods.

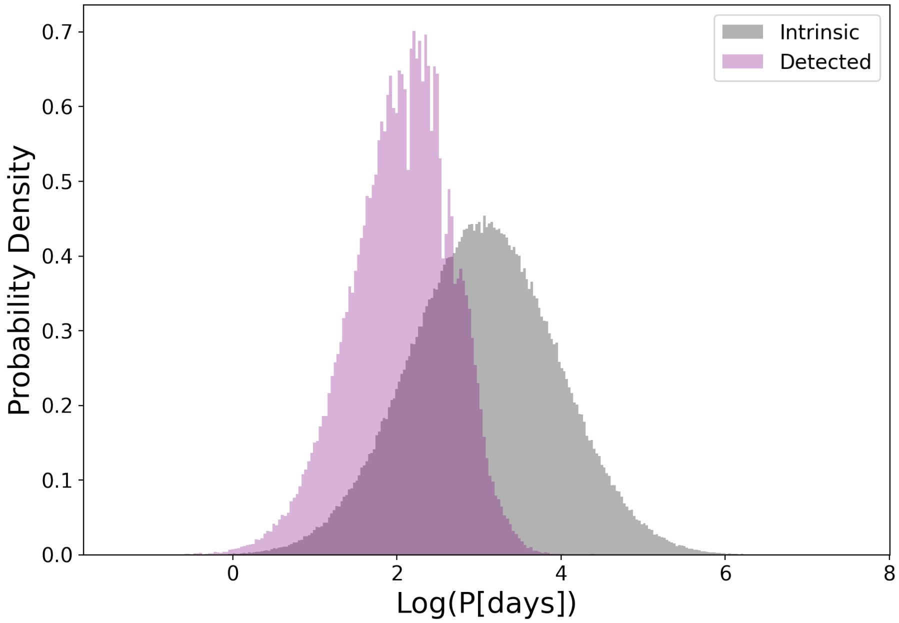
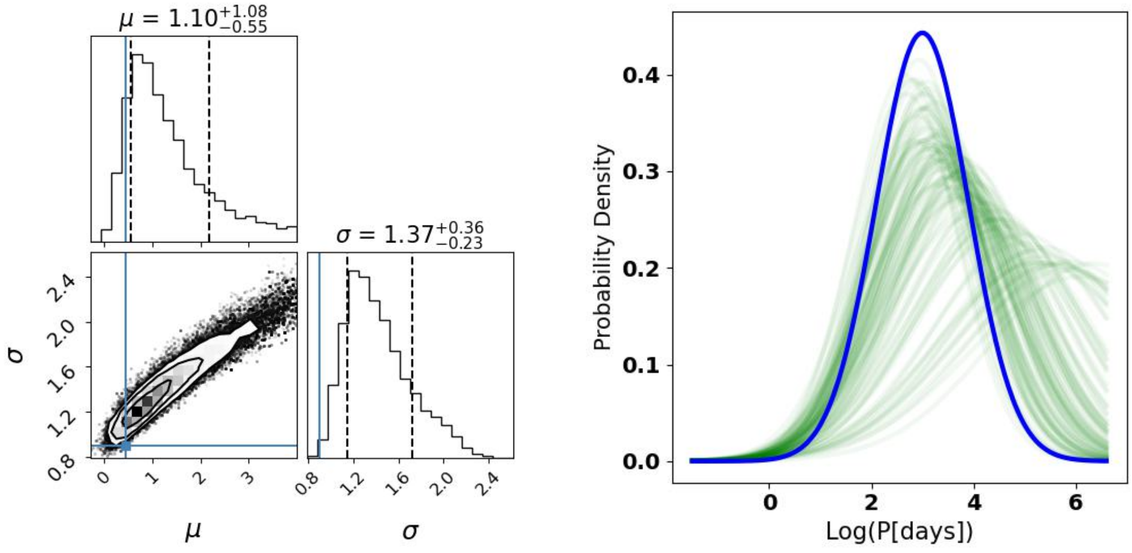


Figure 3.1: The comparison between the true intrinsic binary orbital period distribution (grey) and distribution of periods for the binaries that would be detected as radial velocity variable (purple) for a mock dataset.

However, due to the detection bias, we do not want to directly evaluate the intrinsic period probability distribution function at the value of each of the period posterior since it does not represent the likelihood of the period for the detected binaries. In Figure 3.1, we can see an example of the intrinsic and detected (as RV variable) binary period distributions for a mock dataset generated from a log-normal distribution with a mean of $\log(P[\text{days}]) \approx 2.7$ and a standard deviation of 0.9. This shows a 'best-case' situation as the mock dataset has 23 epochs spanning over 4 years with errors on each velocity measurement ranging from 4 km s^{-1} to 9 km s^{-1} which is better sampling than we see for most stars in the real dataset of interest in this work (see Chapter 4 and in particular Figure 4.2). As we can see in Figure 3.1, there can be (and is expected to be) a substantial difference between the true period distribution for all of the binaries (grey histogram in Figure 3.1) in the population and the period distribution for the binaries that we observe as binaries (through detecting them as RV variable; purple histogram in Figure 3.1). This is most evident in the missing long-period binaries in the 'detected' case. Thus, even with ideal radial velocity curve coverage (given state-of-the-art observations), our sample of detected RV variables is likely to be missing long-period binaries and this would further be amplified in cases with fewer RV measurements.



(a) **(b)**

Figure 3.2: An example of a fit to the period distribution when only fitting on the parameters describing the period distribution (the mean, μ, and the width, σ). This example is for a case with a mock data set generated from a log-normal orbital period distribution. The left panel shows the corner plot with the blue lines representing the true values and the vertical dashed lines representing the 1σ intervals. The right panel shows the period distribution fit with 200 random samples from the posterior probability distribution for the model parameters shown in green and the true solution shown in blue.

To overcome this, we use the following procedure to map the intrinsic period distribution into a probability distribution function for the orbital period of the binaries detected as RV variable. First, we construct the intrinsic distribution using the parameters $\vec{\alpha}$ and drawing a large number of binaries from this distribution using `velbin`, as outlined in Section 2.1. For each of these mock binaries, the true period is known. A χ^2 test (Equation 2.1) is then applied to these binaries in the same manner as was done in Section 2.2 to determine which of the binaries would have been observed as radial velocity variable given the time sampling for the stars in the sample. From this, we construct an 'observed' period distribution using the known periods for each of the drawn binaries that were detected as RV variable. This new distribution now becomes f_α in Equation 3.5 and is what we use to calculate the likelihood of the data given the period distribution model parameters using the posterior samples. Through this process, we are using what we can observe to learn about the intrinsic period distribution and thus take into account the detection bias to learn more about the long-period end of the distribution.

An example of the fit to the period distribution using the likelihood from Equation 3.5 is shown in Figure 3.2. This is an example of a fit done to a mock dataset with 2000 stars. In this case, the true intrinsic distribution is a log-normal centered on $\mu = \log(P[\text{days}]) = 3$ and with a standard deviation of $\sigma = 0.9$ and a binary fraction of 0.3. At this point, we are fitting just on the parameters describing the period distribution (μ and σ). Looking at this figure we can see that while we are able to recover the distribution in the short-period regime, it is unconstrained in the long-period regime. Consequently, we need to find an additional way to better constrain the period distribution.

3.2 Fitting on the Binary Fraction and Velocity Distribution

Although the method presented in the previous section does allow us to gain some information about the intrinsic binary orbital period distribution, this process on its own does not use enough information to fully constrain the long-period end of the distribution. This is because a distribution with more probability at longer periods and with a larger binary fraction can give rise to the same ‘observed’ distribution as a case with less long-period binaries and a smaller binary fraction, as we can see in Figure 3.3.

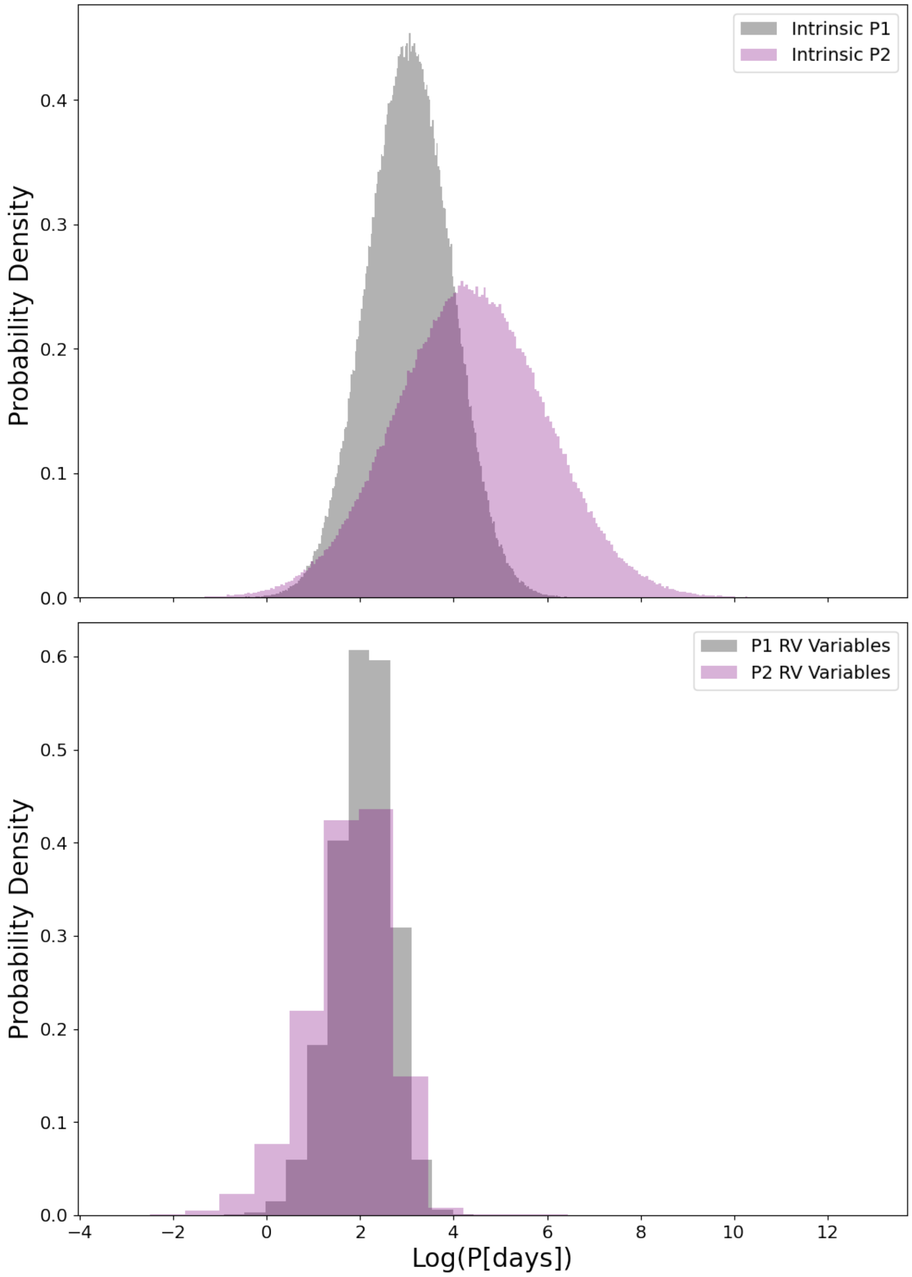
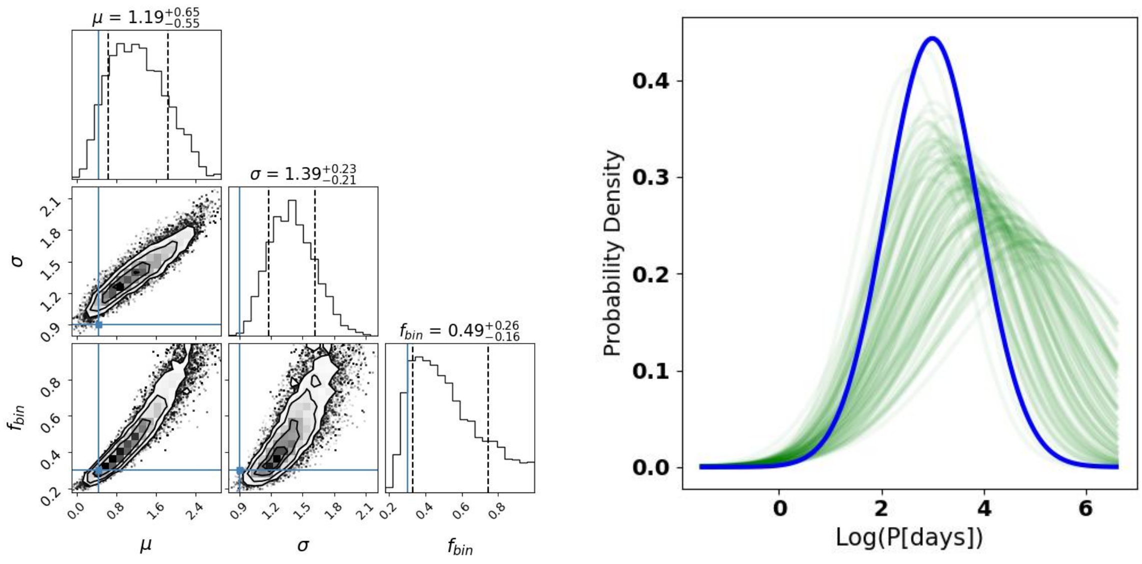


Figure 3.3: The comparison between two intrinsic and RV variable binary orbital period distribution. The top panel shows the two examples of the intrinsic distributions and the bottom panel shows the corresponding period distributions for RV variable stars. The grey shows an example of a log-normal distribution with a mean of 3 and a standard deviation of 0.9 and the purple shows an example of a log-normal distribution with a mean of 4.4 and a standard deviation of 1.6.

In this figure, we can see two examples of intrinsic binary orbital period distribution with one (grey) centered at a lower period ($\mu = \log(P[\text{days}]) = 3$) and with a smaller standard deviation ($\sigma = 0.9$) and another (purple) centered at a longer period ($\mu = \log(P[\text{days}]) = 4.4$) and with a larger standard deviation ($\sigma = 1.6$). These distributions are generated with a sample of stars comparable to what we see in a typical real dataset (our dataset for NGC 3201 has 3501 stars for example; see Chapter 4). While we can see clear differences between the two intrinsic distributions, it is difficult to distinguish between the two resulting distributions for the period of the binaries detected as RV variable, as we can see in the bottom panel of Figure 3.3. However, one thing that is noticeably different between these two cases is the fraction of binaries that are undetected. This is expected since the purple distribution has more long-period binaries that are not detected as RV variable, therefore increasing the fraction of binaries among the apparently single stars (f'_{bin}). In this example, 19.5% of the apparently single stars in the purple distribution are actually binaries and 9.3% in the grey distribution. This means that in order to better constrain the long-period regime of the period distribution, we also need to fit on the binary fraction of the cluster by taking into account the fraction of binaries that were detected or not detected as RV variable. When doing this, the prior we use on the binary fraction is an uniform prior between 0 and 0.5 as the binary fraction cannot be less than 0 by definition and we do not expect it to be larger than 0.5 for a globular cluster.

To do this, we can find the likelihood of the observed fraction of binaries given the expected fraction of all binaries in the cluster and the fraction of binaries that we would expect to observe as RV variable based on the velocity uncertainties and the time sampling, f_{det}. For each star in the sample, we find f_{det} by generating a large sample of binaries with the same time sampling and measurement errors as that star. We then determine what fraction of these would be detected as RV variable, and this fraction becomes f_{det}. This means that if we had a case where we detected all of the



(a) **(b)**

Figure 3.4: The same as Figure 3.2 but with the addition of the binary fraction ($\mathrm{f_{bin}}$) as a model parameter.

binaries as RV variable, then f_{det} would be 1. As we have both radial-velocity variable and non-variable stars in the sample, this likelihood is a binomial distribution with the probability of detecting a radial velocity variable for all stars deemed non-variable, $f_{\mathrm{bin}} f_{\mathrm{det}}$, and the probability of not detecting variability for all the stars detected as single, $(1 - f_{\mathrm{bin}} f_{\mathrm{det}})$. We can then take the product of $f_{\mathrm{bin}} f_{\mathrm{det}}$ for all stars detected as variable and the product of $(1 - f_{\mathrm{bin}} f_{\mathrm{det}})$ for all stars detected as non-variable. Combining these to form the total likelihood gives

$$\mathcal{L}_{f_{\mathrm{bin}}} = \prod_{\mathrm{variable}} (f_{\mathrm{bin}} f_{\mathrm{det}}) \prod_{\mathrm{non-variable}} (1 - f_{\mathrm{bin}} f_{\mathrm{det}}). \tag{3.6}$$

However, while this allows us to provide some constraints to the binary fraction and thus some constraints on the period distribution, it does not fully constrain either. In fact, a degeneracy is present as similar posterior probabilities are obtained for small binary fractions with period distributions skewed towards shorter periods, and higher binary fractions with period distributions skewed towards longer periods, as we can see in Figure 3.4 which shows results using the same mock data as in Figure 3.2 but with the binary fraction now a fitting parameter. This degeneracy is due to both cases resulting in similar numbers of radial velocity variables, meaning that they cannot be distinguished with just this likelihood term.

Therefore, as we discuss next, we also fit on the velocity distribution of the objects in the cluster, which provides additional constraints on the binary fraction and allows us to partly lift this degeneracy.

3.3 Fitting on the Velocity Distribution

Following the work done by Cottaar & Hénault-Brunet (2014), by fitting on the line-of-sight velocity distribution of the objects (both single stars and the centre-of-mass velocity for binary stars) in a cluster, we should be able to better constrain the binary fraction. This velocity distribution is assumed to be Gaussian and is described by the mean of the distribution (the systemic velocity of the cluster, v_{sys}) and the width of the Gaussian (the intrinsic velocity dispersion of the cluster, σ_v).

In Figure 3.5, we can see the effect of binaries on the velocity distribution. This figure was generated from mock data sets where the true velocity distribution (the true velocity for single stars and the barycentre for the binaries) is known (shown in blue) as well as the mean velocities for the stars classified as apparently single (shown in orange). If all binaries were detected, then these distributions would match. However, we can see that the distribution for the apparently single stars is more inflated in the wings. These differences are due to the undetected binaries as undetected binaries can have more extreme radial velocities due to their orbital motions that haven't been accounted for. By comparing the two panels (differing binary fractions), we can see that this effect is most prominent with higher binary fractions (left side) but that it is still present with lower binary fractions (right side). Therefore, by fitting simultaneously on the binary fraction and the velocity distribution, we should be able

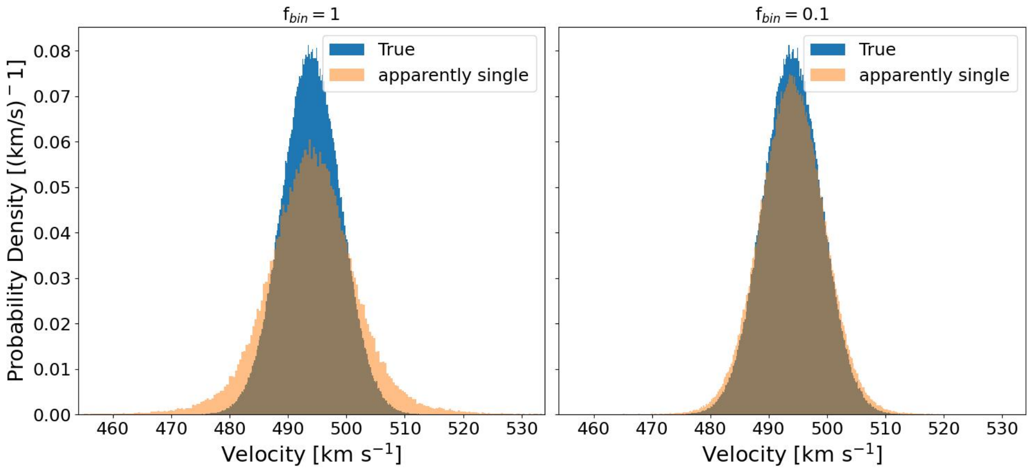


Figure 3.5: The true velocity distributions for all the stars in the cluster (blue) and the observed velocity distribution for stars detected as single (orange). The left panel shows an example of a cluster where all of the stars are in binaries and the right panel shows an example where the cluster has a binary fraction of 0.1.

to better constrain the binary fraction and thus the period distribution.

While the velocity distribution of stars in clusters has traditionally been fit just using the seemingly single stars in the sample (e.g. Cottaar & Hénault-Brunet, 2014), the posterior samples from our radial velocity curve fits provide information regarding the barycentric velocity of the binaries - allowing us to include them in the fit.

As was found in Equation 3.1, once we have the likelihood for every star in the sample, these individual likelihoods can be multiplied together to find the total likelihood. Thus,

$$\mathcal{L}_{\vec{v}} = \prod_{\mathrm{non-variable}} \mathcal{L}_{\vec{v},s} \prod_{\mathrm{variable}} \mathcal{L}_{\vec{v},b}, \tag{3.7}$$

where $\vec{v}$ are the parameters that describe the velocity distribution $(v_{sys}, \sigma_v, f_{\mathrm{bin}})$, $\mathcal{L}_{\vec{v},s}$ is the velocity likelihood for the apparently single stars, and $\mathcal{L}_{\vec{v},b}$ is the velocity likelihood for the stars detected as binaries (RV variable). The equations for these likelihood terms are from Cottaar & Hénault-Brunet (2014), and are summarized below.

For the apparently single stars, we need to account for the fraction of binaries within the apparently single stars, f'_{bin}. This is given by

$$f'_{\mathrm{bin}} = f_{\mathrm{bin}} \frac{1 - f_{\mathrm{det}}}{1 - f_{\mathrm{bin}} f_{\mathrm{det}}}, \tag{3.8}$$

with f_{det} defined in Section 3.2 and where f_{bin} is the binary fraction of the cluster.

Thus, the likelihood of the velocity for each individual star, i for the apparently single stars is given by

$$\mathcal{L}_{\vec{v},s}(v_i, \sigma_i) = f'_{\mathrm{bin}} \mathrm{pdf}_{\mathrm{binary}}(v_i, \sigma_i) + (1 - f'_{\mathrm{bin}}) \mathrm{pdf}_{\mathrm{single}}(v_i, \sigma_i), \tag{3.9}$$

where the probability of the observed velocities for a truly single star is $\mathrm{pdf}_{\mathrm{single}}$ and is given by a Gaussian with measurement uncertainties:

$$\mathrm{pdf}_{\mathrm{single}}(v_i, \sigma_i) = \frac{1}{\sqrt{2\pi(\sigma_v^2 + \sigma_i^2)}} e^{-\frac{(v_i - \mu)^2}{2(\sigma_v^2 + \sigma_i^2)}}, \tag{3.10}$$

where v_i is the weighted mean velocity of the star i and the uncertainty on the weighted mean velocity is σ_i. μ is the mean radial velocity (the cluster systemic velocity) and σ_v is the cluster velocity dispersion.

In Equation 3.9, $\mathrm{pdf}_{\mathrm{binary}}$ is the probability distribution for the velocity for an apparently single star that is actually an undetected binary and is given by:

$$\mathrm{pdf}_{\mathrm{binary}}(v_i, \sigma_i) = \int \mathrm{pdf}_{\mathrm{single}}(v', \sigma_i) \mathrm{pdf}_{\mathrm{offset}}(v_i - v') dv'. \tag{3.11}$$

This comes from the convolution of the pdf for the single stars (Equation 3.10) and the pdf for the radial velocity due to binary motions ($\mathrm{pdf}_{\mathrm{offset}}$). To get the pdf due to binary motions, a large number of N random binaries are generated from the adopted orbital period, mass ratio, and eccentricity distributions. Each with randomly oriented three-dimensional velocities, v_k. Thus, the projected velocity offset will be uniformly distributed between $-v_k$ and v_k, giving the following pdf for the radial velocity offsets:

$$\mathrm{pdf}_{\mathrm{offset}}(v_{\mathrm{bin}}) = \frac{1}{N}\sum_{k=1}^{N}\begin{cases} 0 & \text{if } v_k < |v_{\mathrm{bin}}| \\ \frac{1}{2v_k} & \text{if } v_k > |v_{\mathrm{bin}}| \end{cases}, \tag{3.12}$$

where v_{bin} is the velocity offset due to binary orbital motion projected along the line of sight.

To get the likelihood for the barycentric velocities for the detected binaries, the same likelihood that was used for fitting the period distribution (using posterior samples form the RV curve fitting procedure; Equation 3.5) can be used where $\vec{\alpha}$ is now the parameters governing the velocity distribution (cluster systemic velocity and the velocity dispersion). As the velocity distribution is assumed to be Gaussian, $f_{\vec{\alpha}}(P_{nk})$ is now replaced by $f_{\vec{\alpha}}(v_{nk})$ where f is the posterior probability distribution for a Gaussian and describes the probability of the barycentric velocity, v_{nk}, given $\vec{\alpha}$, where v_{nk} are the posterior samples for the barycentric velocity from the fits to the RV curves.

Putting this all together, the likelihood of observing the velocities (mean velocities for the non-variable stars and barycentric velocities for the binaries) in the dataset is given by:

$$\begin{aligned}\mathcal{L}_v &= \mathcal{L}_{v,\mathrm{single}} \times \mathcal{L}_{v,\mathrm{binary}} \\ &= \prod_{\text{non variable}} [f_{\mathrm{bin}}\mathrm{pdf}_{\mathrm{binary}}(v_i,\sigma_i) + (1-f_{\mathrm{bin}})\mathrm{pdf}_{\mathrm{single}}(v_i,\sigma_i)]\times \\ &\prod_{\text{variable}} \frac{1}{K}\sum_{k=1}^{K}\frac{f_\alpha(v_{nk})}{p_0(v_{nk})}.\end{aligned}, \tag{3.13}$$

By adding in this additional constraint from the velocity distribution, we are able to better constrain the binary fraction of the cluster, and therefore recover the intrinsic binary orbital period distribution. Examples showing this can be seen in the following section.

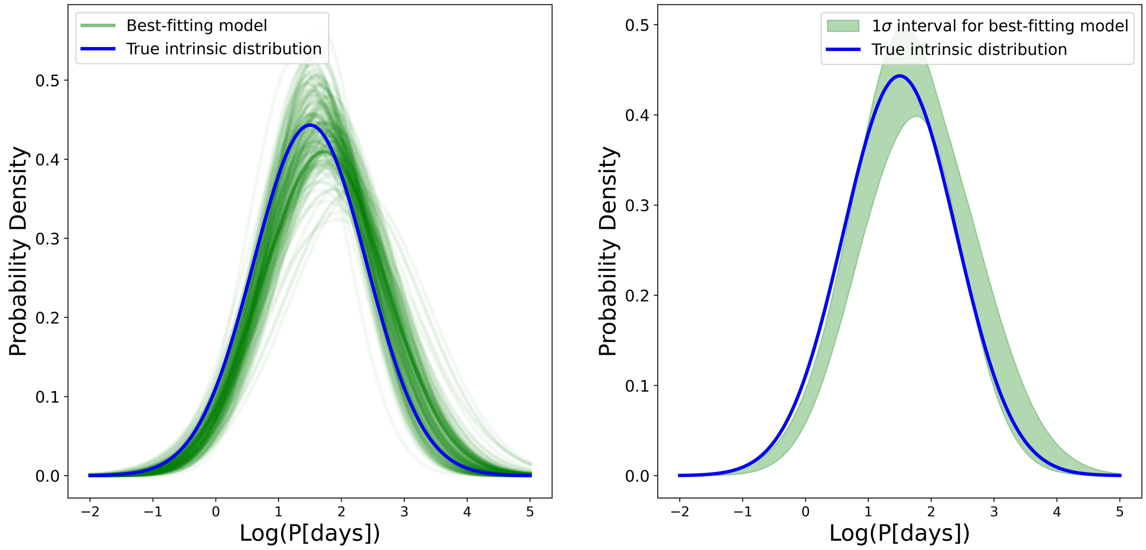
To summarize, when determining the intrinsic orbital period distribution, we fit on the parameters that describe the period distribution (Equation 3.5), the binary fraction (Equation 3.6), and the velocity distribution (Equation 3.13) by multiplying the likelihoods for each of these terms. With this total likelihood and the priors on each parameter (using uniform priors for each parameter), we can use a sampling method such as `emcee` or `dynesty` (see Sections 2.5 and 2.7) to sample the posterior probability distributions for the model parameters given the data. Unlike when fitting the radial velocity curves, we found that `emcee` was the more efficient sampling method for this case, due to the smaller number of parameters, and was used in each of the model fits.

Overall, this method allows us to take the multi-epoch radial velocities for stars in a cluster and use them to determine the intrinsic binary period distribution.

3.4 Testing the Method with Standard Distributions

Once we were confident that we were able to fit the radial velocity curves for stars detected as radial velocity variable and we had a method for fitting the intrinsic period distribution, this method was applied to mock datasets generated with a known standard analytic distribution for the orbital period of the binaries, such as a log-normal or a log-weibull. These mock datasets were generated using `velbin` (see Section 2.1) and were created to reflect the NGC 3201 dataset that will be analysed in Chapter 4 (by ensuring that the number of stars generated and radial velocity time sampling was similar to that in the real dataset).

The log-normal and log-weibull distributions were chosen as while we do not know the true shape of the intrinsic distribution, these distributions are expected to be reasonable approximations that will allow us to probe the accuracy of the method (e.g. Ivanova et al., 2005). As we will see in Section 3.5, we later expanded this method to allow free-form fitting of the distribution to minimize assumptions about the shape of the distribution.



(a)

(b)

Figure 3.6: The recovered binary orbital period distribution (green) fit to a mock dataset generated with a log-normal distribution for the binaries centered at $\mu = \log(30[\text{days}]) = 1.48$ and with a width of 0.9. The left panel shows the model results in green for 200 random samples from the converged MCMC chain and with the true distribution shown in blue. The right panel shows the 1σ confidence intervals in green with the the true distribution again shown in blue. See the text for more details on the mock data.

The first distribution considered was a log-normal which has the probability density function:

$$f(y) = \frac{e^{-y^2/2}}{\sqrt{2\pi}}, \tag{3.14}$$

where y in our case is $(\log P - \mu)/\sigma$. Thus, the mean, μ, and standard deviation, σ, of the distribution are the $\vec{\alpha}$ parameters in Equation 3.5. In Figure 3.6, we can see an example of a period distribution fit to mock data that was generated with a log-normal binary period distribution and then fit assuming a log-normal using `emcee`. In this example, the mock data was generated with $\mu = \log(30[days]) = 1.48$, $\sigma = 0.9$, $f_{\text{bin}} = 0.1$, $v_{\text{sys}} = 495$ km s^{-1}, and $v_{\text{disp}} = 5$ km s^{-1}. When generating this dataset,

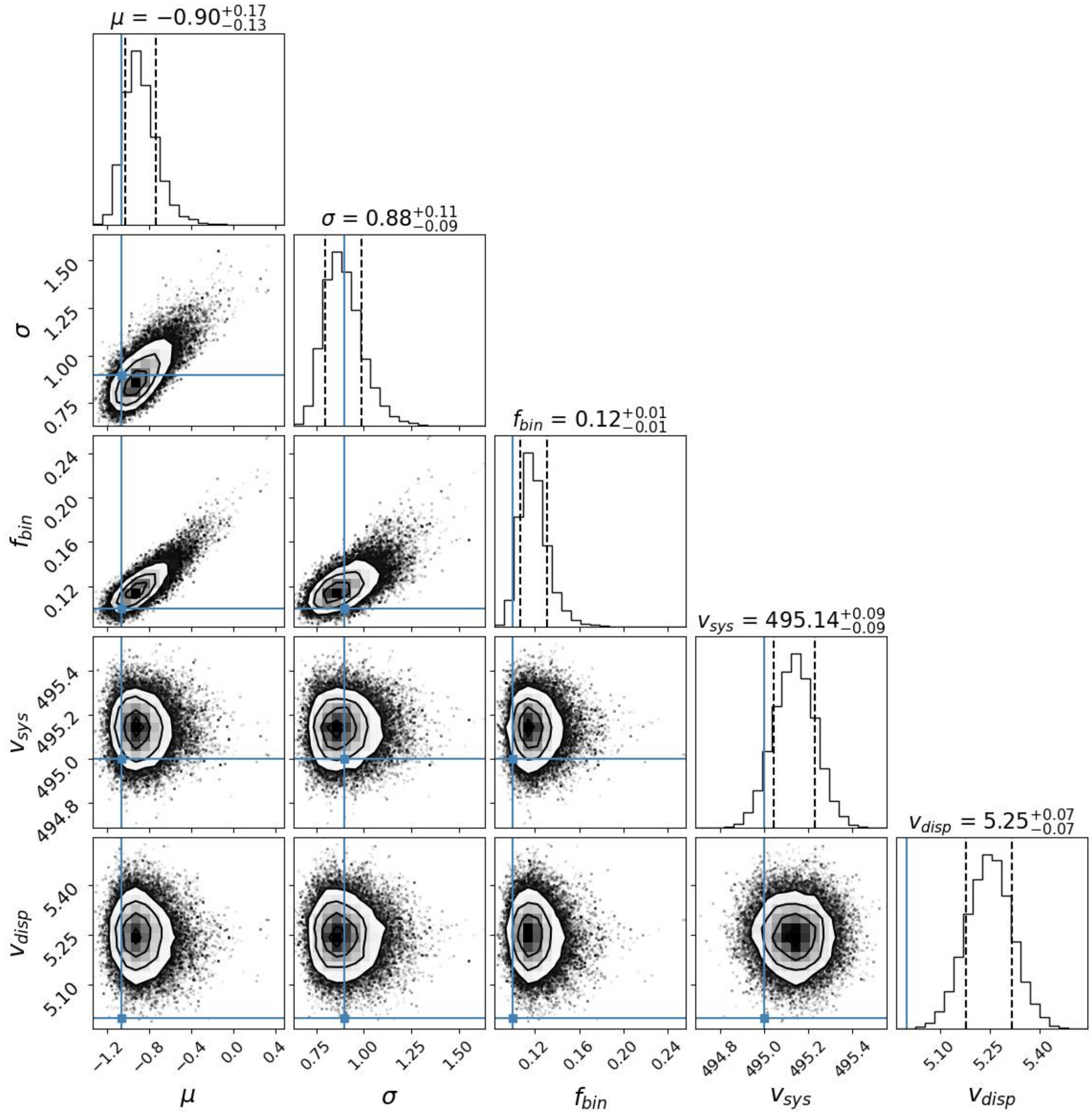
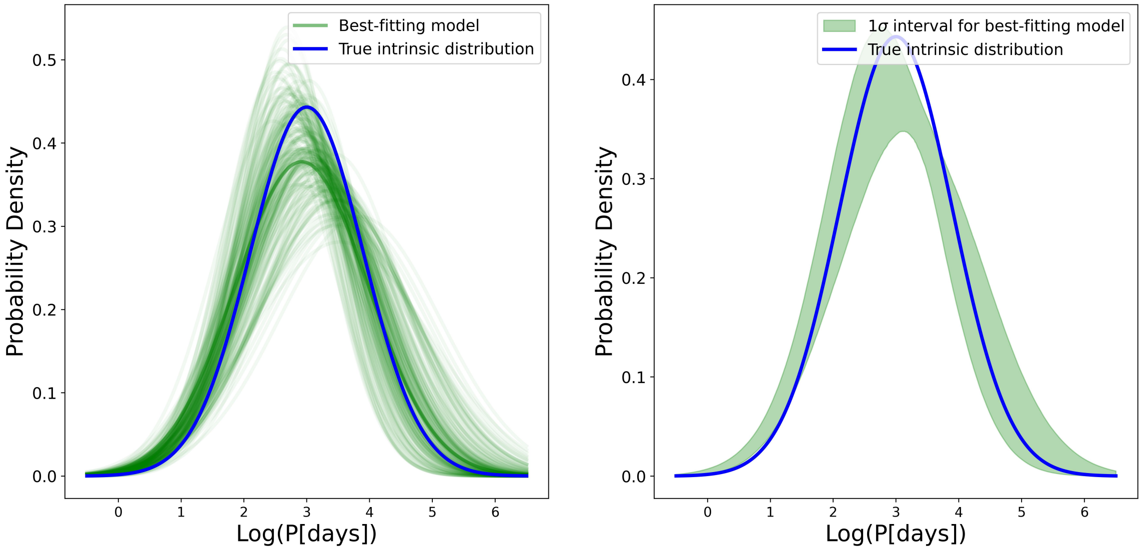


Figure 3.7: The corner plot for the fit to the population of binaries shown in Figure 3.6. The blue lines show the true values and the vertical dashed lines in the 1d marginalized posterior distributions show the 1σ interval.

3540 stars were sampled (same number of stars as in the NGC 3201 dataset), each with 20 epochs of data over approximately four years and with time baselines mimicking the NGC 3201 dataset. 267 of these stars were found to be radial velocity variable and classified as binaries using the χ^2 test, meaning that only 87 (25%) were undetected. Radial velocity curves were then fit to the data for each of these 267 radial velocity variable stars using the method outlined in Chapter 2. The posterior samples from these RV curve fits were then used in the method outlined in this chapter to fit on the population distribution for the binary orbital periods.

Both panels in Figure 3.6 show the results from the same fit. Figure 3.6a shows the distribution for 200 random samples from the converged MCMC chain (all taken after the MCMC burn-in phase) and Figure 3.6b shows the 1σ confidence intervals determined using all of the post MCMC burn-in chain elements. In both panels the blue line represents the true intrinsic distribution. We can see that the true distribution is well recovered. Looking at the corresponding corner plot in Figure 3.7, we can see that not only are the parameters describing the period distribution recovered, but the binary fraction and systemic velocity are as well, although uncertainties may be underestimated. However, the velocity dispersion is slightly overestimated.



(a) **(b)**

Figure 3.8: Same as Figure 3.6 for an example of a mock dataset generated with a log-normal distribution centered at $\mu = \log(1000[\text{days}]) = 3$ and with a width of $\sigma = 0.9$.

We can see another example of a fit using a log-normal period distribution for the orbital periods of the binaries in Figure 3.8. In this case the log-normal is centered at $\mu = \log(1000[\text{days}]) = 3$ and has a width of σ =0.9. This mock dataset was generated using a binary fraction of 0.3, a cluster systemic velocity of 250 km s^{-1} and a velocity

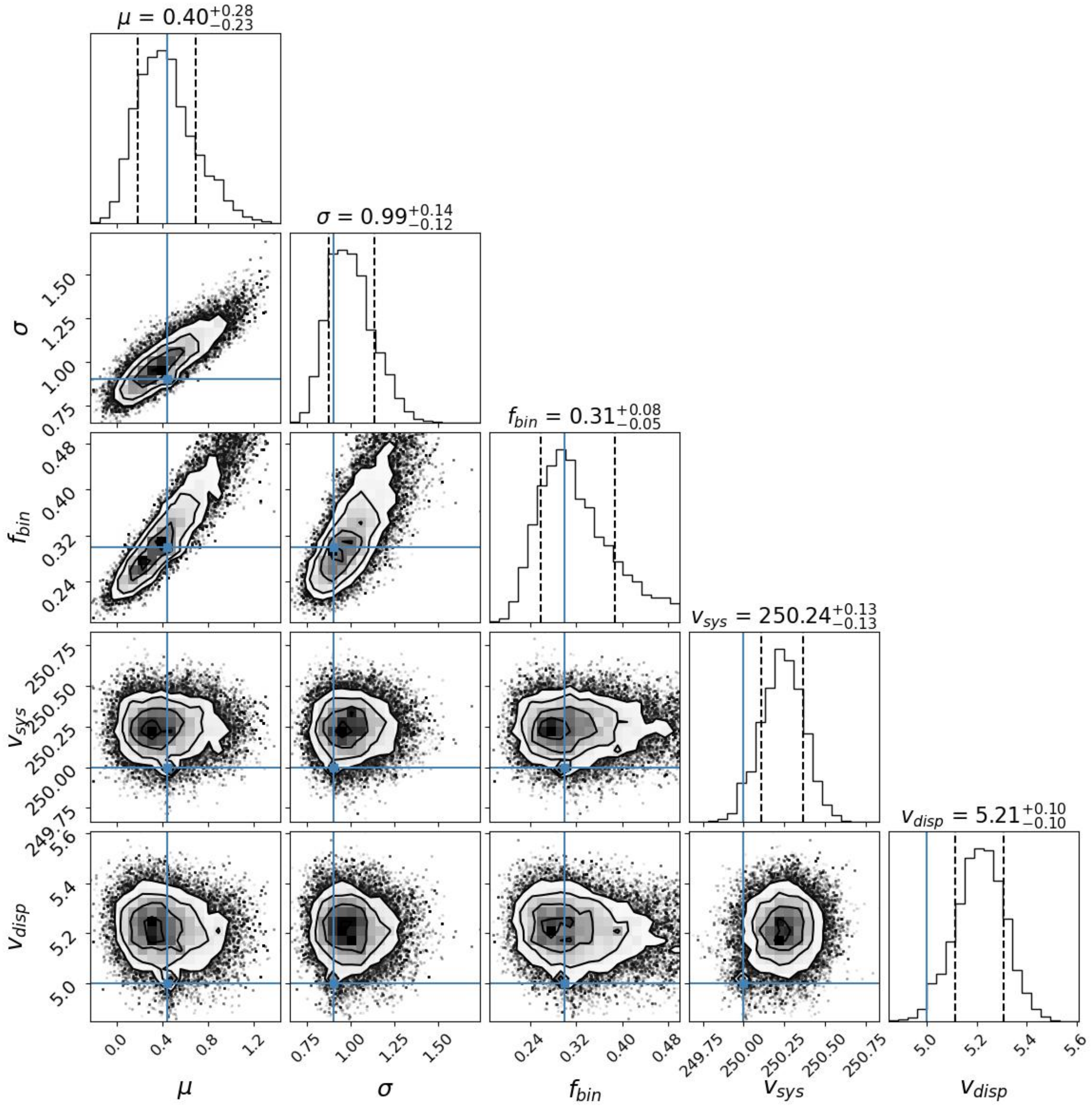


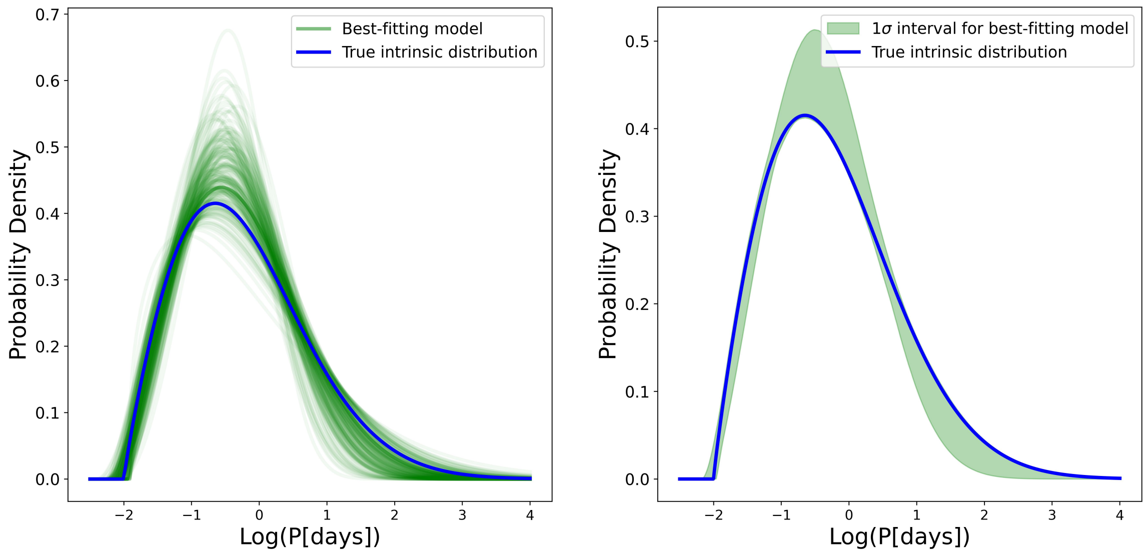
Figure 3.9: The corner plot for the fit to the population of binaries shown in Figure 3.8.

dispersion of 5 km s^{-1}. In this case, 2000 stars were generated and 241 were classified as binaries. Thus, 359 binaries (almost 60%) were undetected as there is a larger fraction of long-period binaries for the adopted period distribution in this second test case. As we can see in the resulting period distribution fit (Figure 3.8) and in the corner plot for this example (Figure 3.9), while there is more uncertainty in each of the model parameters (and this may be underestimated), we are still able to recover the true model parameters. This means that this method is still viable even in cases with fewer stars in our sample and a larger fraction of undetected binaries.

In order to ensure that the method works in general and not just for log-normal distributions, this method was also applied to an example using a log-weibull distribution. This distribution is skewed and may be more representative of the present-day period distribution of binaries in a globular cluster where the hard binaries have become harder over time, skewing the distribution towards shorter periods. The probability density function for this distribution is given by

$$f(y, c) = cy^{c-1}e^{-y^c}, \tag{3.15}$$

where c is the shape parameter of the distribution and in our case y is $\log P - L$ where L is the minimum period with non-zero probability density. Thus, in the case of the log-weibull distribution, the $\vec{\alpha}$ model parameters are c and L. An example of a fit to a mock dataset generated with a log-weibull distribution for the period of the binaries is shown in Figures 3.10 and 3.11. In this example, 3540 stars were generated with periods for the binaries drawn from a log-weibull distribution and the input parameters were $c = 2.0$, $L = -2.0$, $f_{bin} = 0.1$, $v_{sys} = 495$ km s^{-1} and $v_{disp} = 5$ km s^{-1}. In this case, 186 stars were found to be radial velocity variable; meaning that 168 (or 47%) binaries were not found to be variable. As was the case with the log-normal distributions, we can see that the true distribution and the other model parameters are well recovered in the case of the log-weibull distribution as well. Overall, this method appears to be able to recover the true period distribution in cases where the functional form for the true underlying intrinsic distribution is known.



(a) **(b)**

Figure 3.10: Same as Figure 3.6 for an example of a mock dataset generated with a log-weibull distribution with a slope of 2.0 and a minimum log-period (with the period in days) of -2.0.

3.5 Testing the Method with a Flexible Model for the Intrinsic Period Distribution

While we have shown that our method works in cases where the functional form for the period distribution is known, in practice this is not something that we know for a real dataset. Therefore, we need a method that allows the shape of the distribution to remain flexible. To achieve this, instead of fitting on parameters that describe a specific distribution, we instead fit on a larger number of parameters describing a flexible piecewise function. This piecewise function is defined such that the intervals are held fixed but the height of the function in each interval can change. As the distribution is still a probability distribution, the area under the distribution must always sum to one. This can also be thought of as fitting on a normalized probability density histogram where the free parameters are the bin heights of the histogram. By approaching the problem this way, we are able to minimize the assumptions about the underlying distribution and therefore gain a better understanding of the true

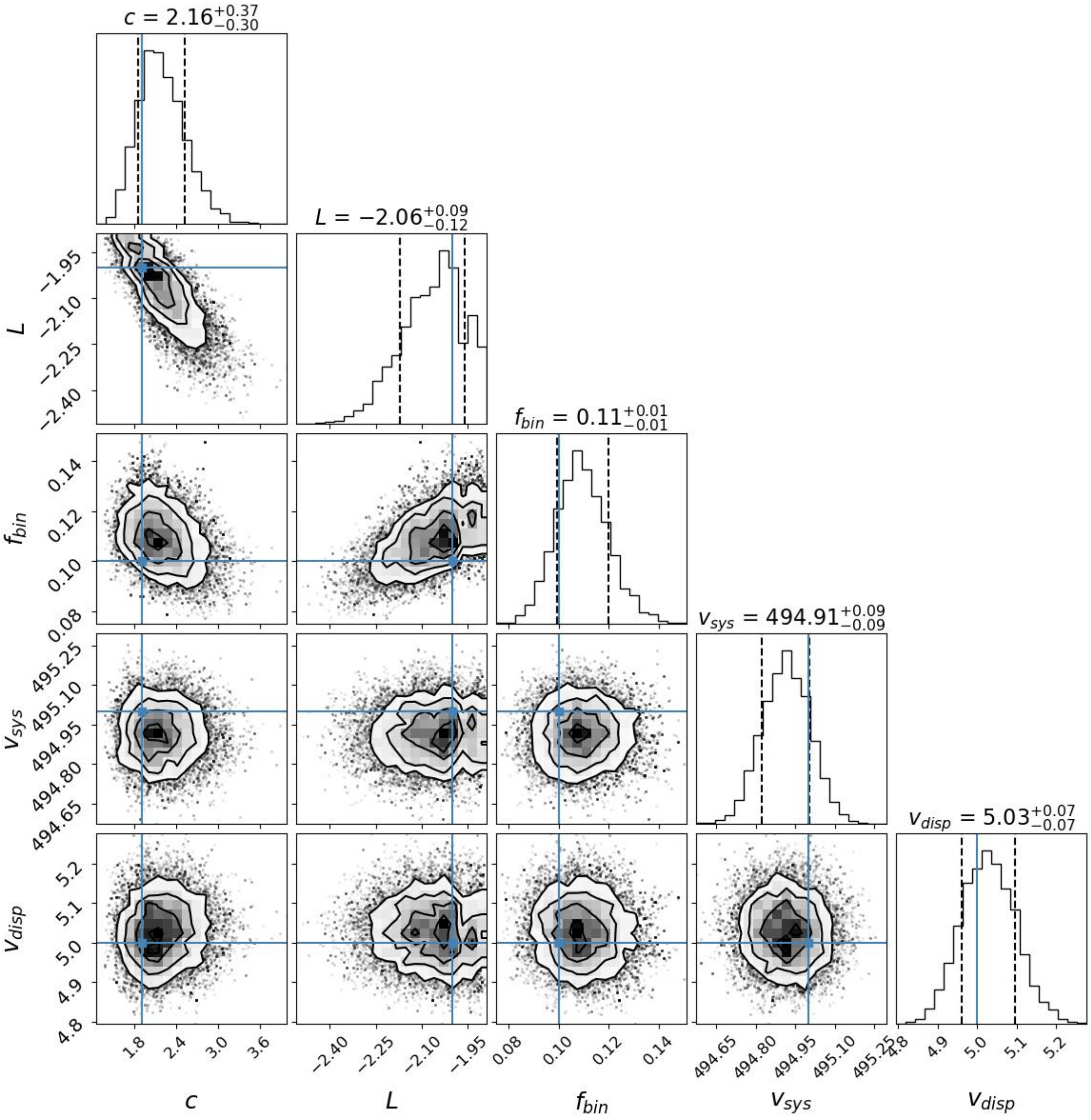


Figure 3.11: The corner plot for the fit to the population of binaries shown in Figure 3.10.

distribution and thus the hard/soft limit of the cluster binaries.

In Figure 3.12 we can see an example of this fitting procedure applied to the same dataset that was used in Figure 3.6. In this case, 15 equally spaced intervals were used to fit on the amplitudes of the period distribution piece-wise function spanning from the shortest expected period ($\log(P[\text{days}]) = -2$) to the longest expected period ($\log(P[\text{days}]) = 5$) for this dataset. In this figure, the distributions given by the parameters in the MCMC chain are shown with the green lines where darker areas

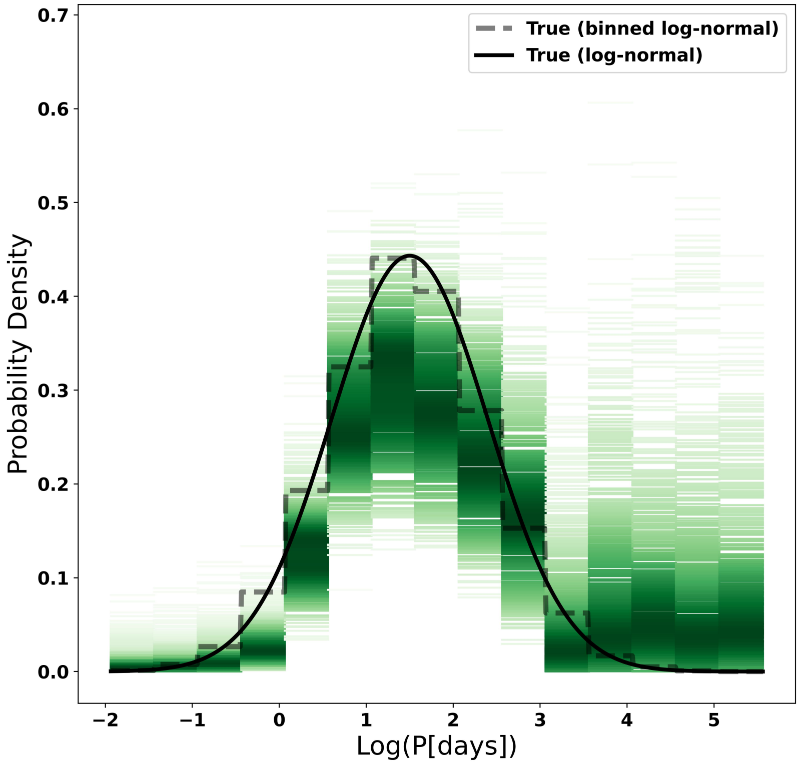


Figure 3.12: The recovered binary orbital period distribution for the binary population shown in Figure 3.6. The green shows the best-fitting model where the darker colors correspond to areas of higher probability. The solid black line shows the true intrinsic distribution (log-normal) and the grey dashed line shows the true intrinsic distribution when binned.

corresponding to areas of higher frequency (thus higher probability). The true log-normal distribution that the data was drawn from is shown with the solid black line and the corresponding true distribution when binned into the given intervals is shown with the dashed grey line. As we can see from this figure, there is technically room for non-zero probability in the long-period regime of the distribution, however, this is most likely due to the broad posterior probability distribution for the period of RV variable stars that are at the limit of what we can detect and thus cannot be well constrained. However, we do see that overall the true distribution is well recovered even though we have removed the model assumption of a log-normal distribution.

We can see that this again works for cases other than a log-normal distribution by looking at the example of the weibull distribution in Figure 3.13. This example is also now fit with 15 intervals spanning from $\log(P[\text{days}]) = -1$ to 7. As we can see, we are again able to recover the true distribution.

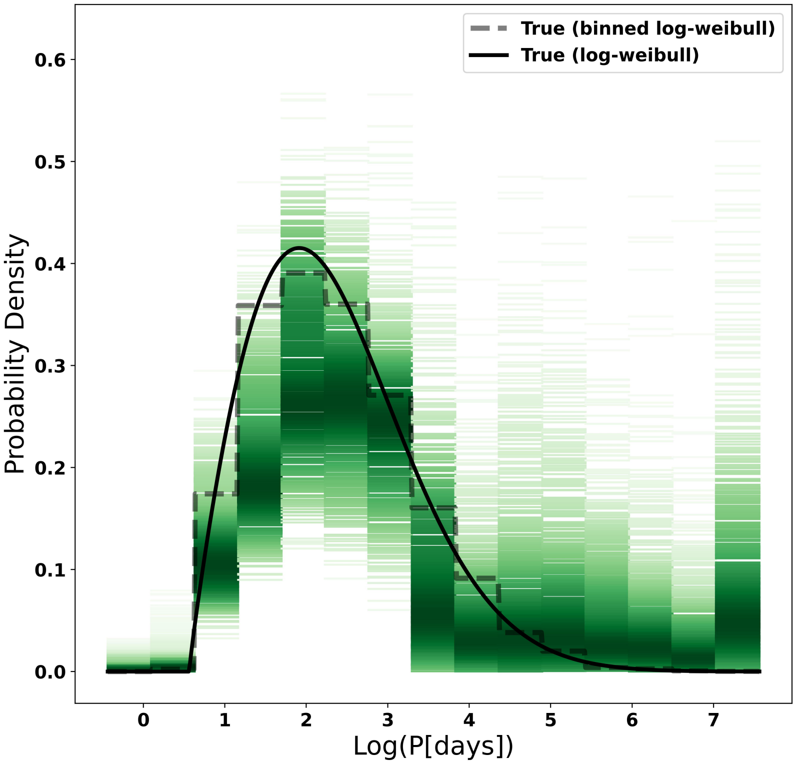


Figure 3.13: The recovered binary orbital period distribution for the binary population shown in Figure 3.10. The green shows the best-fitting model where the darker colors correspond to areas of higher probability. The solid black line shows the true intrinsic distribution (log-weibull) and the grey dashed line shows the true intrinsic distribution when binned.

Overall, we can see that by extending this method to this more flexible model for the intrinsic period distribution of binaries, we are still able to recover the true intrinsic binary orbital period distribution and remove assumptions about the shape of the distribution. Therefore, this method is now ready to be applied to real cluster data.

Chapter 4

The Orbital Period Distribution of Binaries in the Globular Cluster NGC 3201

Now that we have shown that our method works for mock datasets where the true underlying orbital period distribution of binaries is known, it can now be applied to a real globular cluster for which we want to infer the period distribution of binaries.

NGC 3201 is the ideal first cluster to apply this method to due to the extent of radial velocity data available. As stated in Section 1.4, detecting binaries and finding orbital solutions is generally biased towards shorter period binaries. However, with more radial velocity measurements, this bias can be mitigated. In the case of NGC 3201, the number of epochs of data for each star ranges from 3 observations per star to 63 observations per star over a time baseline of a few years, meaning that we will be able to probe more long-period binaries.

In this chapter, we discuss the data for NGC 3201 in Section 4.1, the radial velocity fits for the RV variables in Section 4.2, the constraints that we can place on the intrinsic orbital period distribution in Section 4.3, and finally we discuss the implications of this period distribution on the initial density of NGC 3201 in Section 4.4.

4.1 MUSE Data

As part of a survey of 25 Galactic globular clusters with the MUSE (Multi-Unit Spectroscopic Explorer) integral-field spectrograph (Kamann et al., 2018), Giesers et al. (2019) analyzed the binary content of the cluster NGC 3201 and it is the publicly available data from that study that is used in this work.

MUSE is a visible-light integral-field spectrograph, meaning that it produces a spectrum for each pixel in the image. Mounted on the Yepun UT4 8.2m telescope at the Very Large Telescope (VLT) of the European Southern Observatory (ESO), it has a good spatial resolution and a relatively wide field of view. Its field of view is 1x1 arcmin and its spatial resolution is about 0.4 arcsec, with a spatial sampling of 0.2 arcsec per pixel. This spatial resolution is important when studying globular clusters as globular clusters are dense structures, resulting in crowding in their central regions, therefore making it difficult to distinguish between stars. However, with MUSE, it is possible to get measurements for individual stars (Bacon et al., 2010). The wide field of view is useful, as it means that when MUSE is pointed at a globular cluster, it can collect data for thousands of stars at once (Giesers et al., 2019). MUSE also has a spectral resolving power of 1770 at 465nm to 3590 at 930nm and the radial velocity uncertainties from the MUSE spectra for the NGC 3201 radial velocity data range from around 1 km s^{-1} for brighter stars to a few km s^{-1} to over 10 km s^{-1} for fainter stars.

The pointing scheme used for the NGC 3201 data can be seen in Figure 4.1. As NGC 3201 is at a distance of 4.73 kpc (Baumgardt & Vasiliev, 2021), 1 arcmin corresponds to 1.37 pc. As the half-mass radius of NGC 3201 is 8.68pc (Baumgardt &

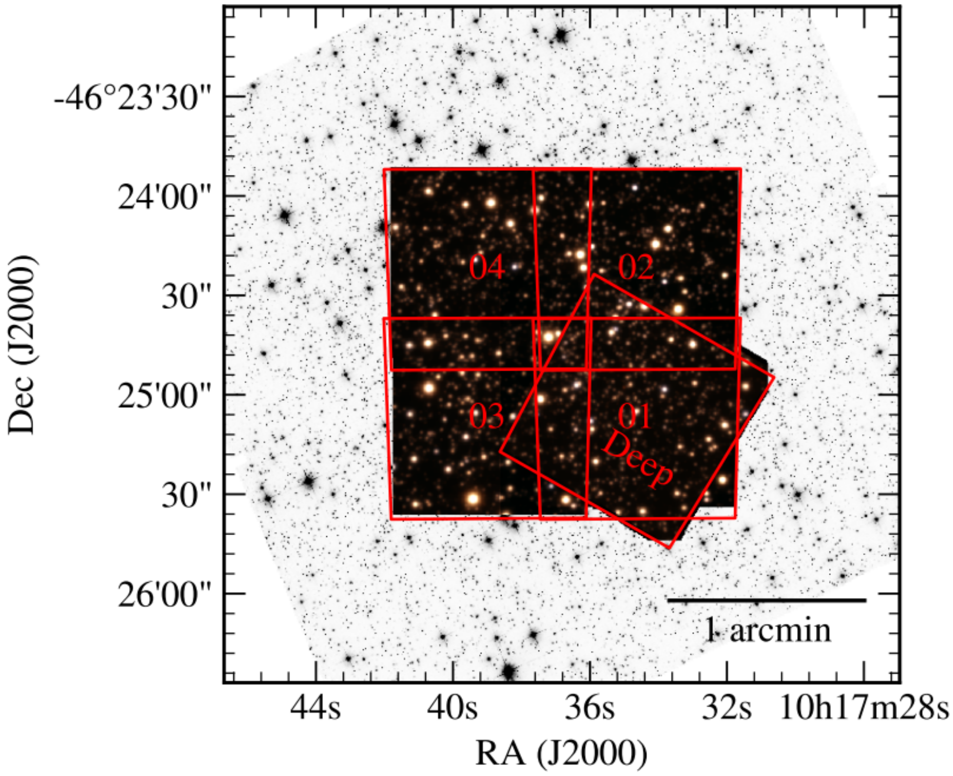


Figure 4.1: Pointing scheme for the NGC 3201 data where 1 arcmin corresponds to 0.137 pc. Figure from Giesers et al. (2019).

Hilker, 2018), the observations are within the central region of the cluster. Figure 4.1 shows the five pointings showing that the pointings partly overlap, allowing for more epochs of data in the overlapping regions. There are between four and 17 observations at different epochs for each pointing. These observations are unevenly spaced in time over 1500 days (over 4 years) and cover both short and long baselines with the shortest time between observations on the timescale of hours. This spacing is important as it means that we can probe more of the radial velocity curves on different timescales, making it easier to distinguish radial velocity variables from non-variables.

Overall, a varying number of observations were obtained per star (Figure 4.2) with the majority of the stars have around 15 observations. However, there are stars with up to 63 observations due to pointing overlaps. From these observations, radial velocity measurements were obtained for 3540 member stars where membership was

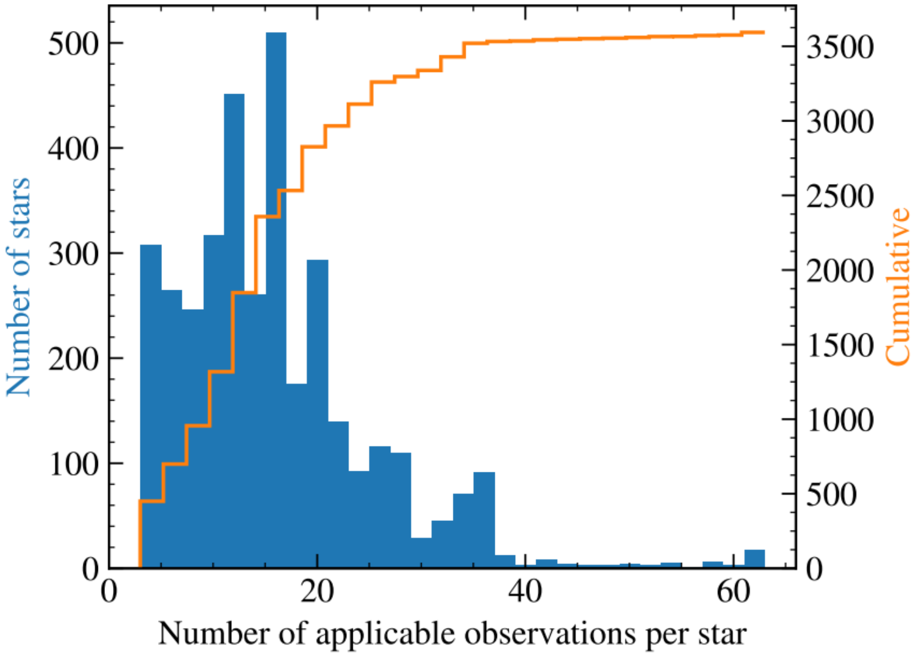


Figure 4.2: Number of observations per star in blue and cumulative distribution of the number of stars with a given number of observations in orange for the MUSE dataset of NGC 3201 used in this work. Figure from Giesers et al. (2019).

determined by Giesers et al. (2019) by removing stars from the sample that had radial velocities inconsistent with typical NGC 3201 radial velocities (taking into account the heliocentric radial velocity of NGC 3201, the central velocity dispersion, and an extra tolerance for typical binary amplitudes) . One important thing to note is that even with the spatial resolution of MUSE, the two stars in a binary are too close to be spatially resolved at the distance of globular clusters. Thus, binaries will appear as point sources and the spectra from MUSE will be dominated by the brighter star.

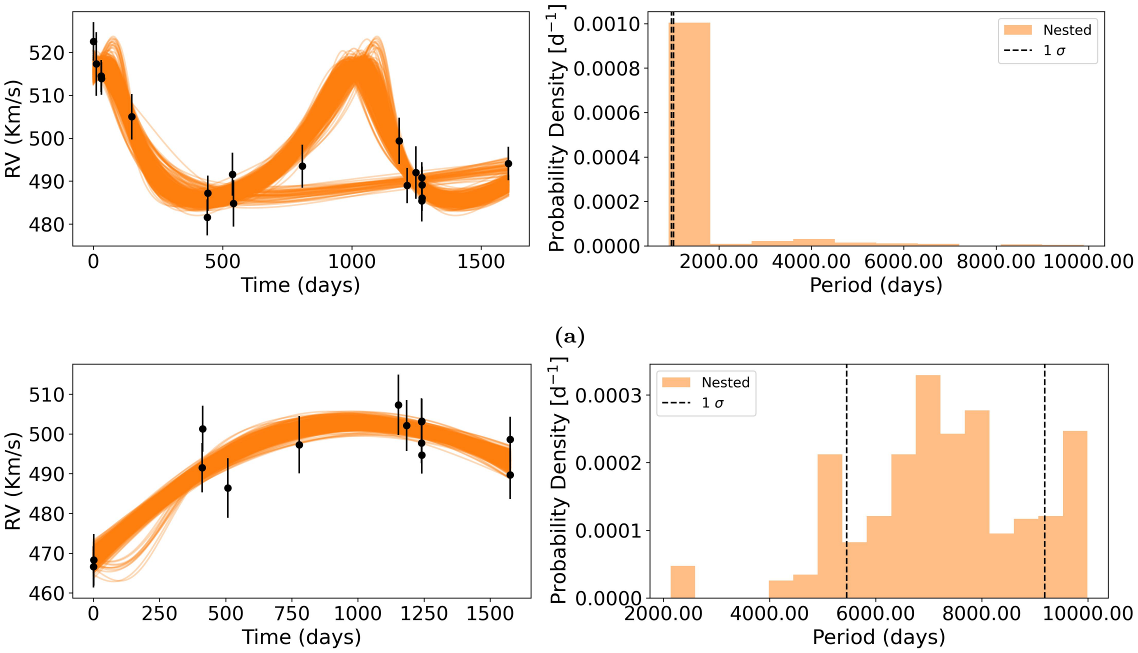
This multi-epoch radial velocity dataset can be used to fit radial velocity curves as described in Chapter 2 and derived in Section 1.2. This will provide posterior distributions for the binary orbital parameters including the period (P), eccentricity (e), argument of pericentre (ω), pericentre phase (ϕ_0), velocity semi-amplitude (K), and the barycenter velocity (v_0).

4.2 Fitting Radial Velocity Curves for NGC 3201 Binaries

Taking the data for NGC 3201 and applying the χ^2 test to each of the stars in the cluster, we found 224 stars to be radial velocity variable. Comparing this to the sample of binaries analyzed by Giesers et al. (2019), there were 94 in common (out of the 95 total in the Giesers et al. (2019) sample). These numbers differ as we adopted a different criterion for detection than these authors.

In the study by Giesers et al. (2019), they found the probability of a star to be variable using a reduced χ^2 and then found the probability that each star was variable by comparing the empirical cumulative distribution function for the measured χ^2 values to the expected cumulative distribution function for the null hypothesis (see Section 5.1 in Giesers et al. (2019) for more information). They then classified every star that had a probability over 50% as a binary, resulting in a first selection of 515 stars. With this preliminary sample, they then used `TheJoker` to fit the radial velocity curve for each of these potential radial velocity variables. However, in their final sample they only selected the stars that had well-fit radial velocity curves with clear uni-modal or bi-modal solutions, bringing their total number of binaries down to 95.

Our approach in this work differs as we do not currently have a way to handle false positives (single stars that would be classified as binaries) when fitting on the period distribution, so we impose a stricter requirement on what is RV variable. We also use `Dynesty` instead of `TheJoker` as it provided good fits to the radial velocity curves (see Section 2.7) in all cases, including cases where `TheJoker` provided an insufficient number of samples. This is crucial when trying to constrain the population distribution as we want to include as many RV variables as possible. This means that we can use our full sample of detected binaries to constrain the intrinsic period



(a)

(b)

Figure 4.3: Examples of RV curve fits for NGC 3201 binaries. The RV fits are shown on the left and the histogram of the period posteriors is shown on the right.

distribution instead of only selecting the ones with good fits. Therefore, while Giesers et al. (2019) performed a detailed study of a selected sub-sample of binaries (with a bias towards short-period binaries), we can take advantage of a larger sample of RV variable stars (including long-period binaries) to constrain the binary orbital period distribution, including in a longer period regime than that probed by the sample of binaries from Giesers et al. (2019). We can see an example of RV curve fits to binaries that Giesers et al. (2019) did not consider in Figure 4.3. These binaries are on the limit of what we can detect, but we are still able to fit their RV curves. As we can see in this figure, this provides RV curve solutions with longer periods. Thus, by using this method we are able to probe the binary period distribution at larger binary periods than was done by Giesers et al. (2019) by accounting for these additional binaries.

For the stars that we have in common with Giesers et al. (2019), we show a

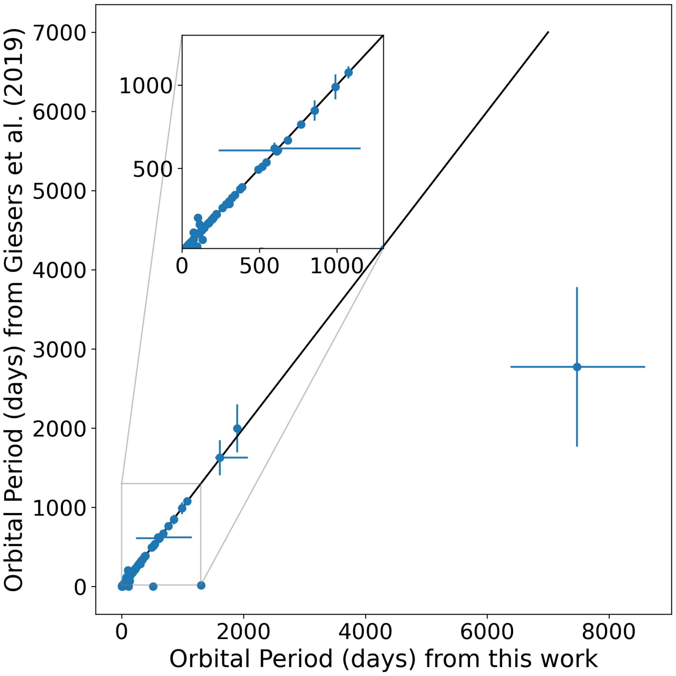
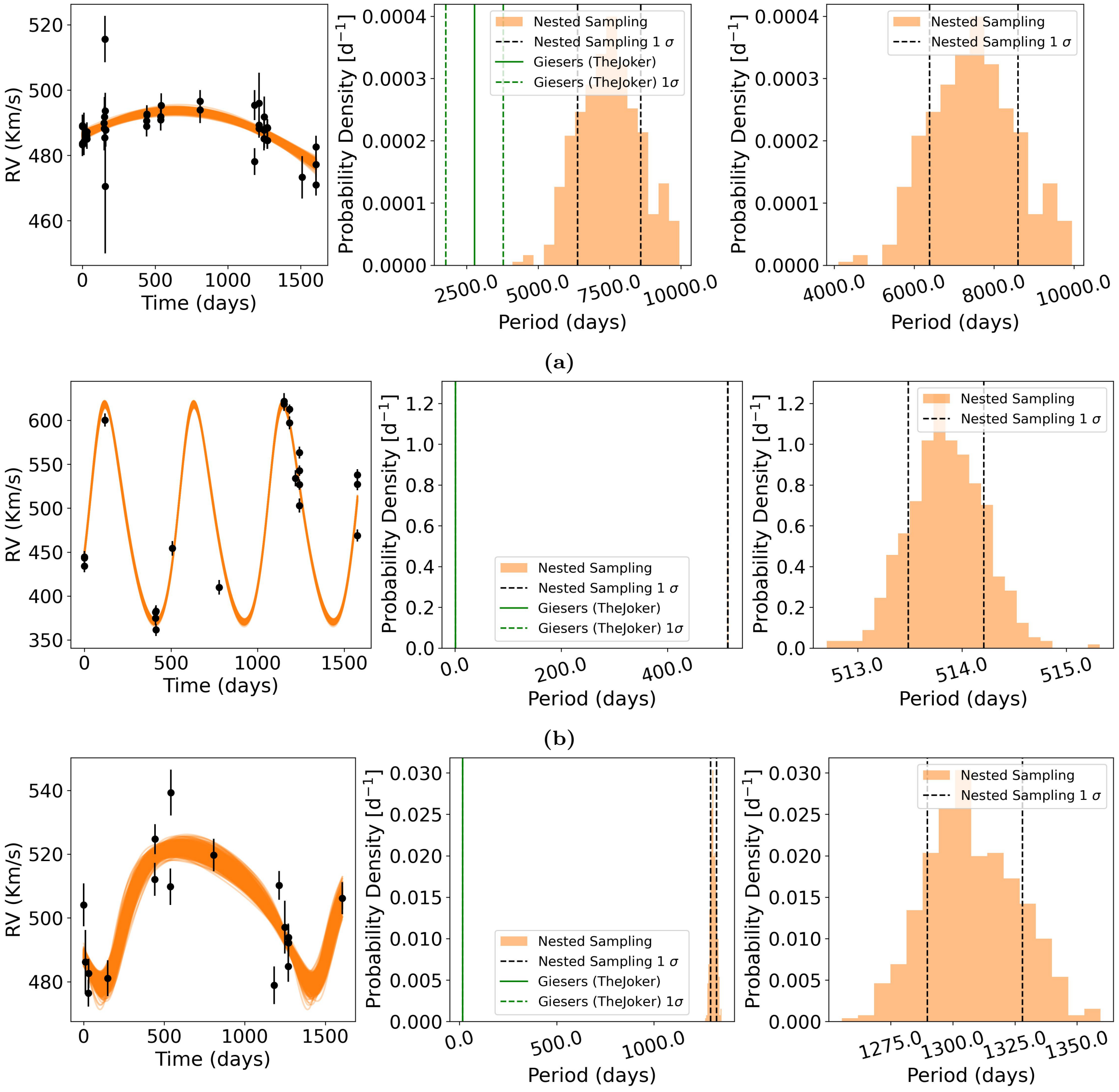


Figure 4.4: Comparison of the orbital binary periods found for individual binaries by Giesers et al. (2019) (y axis) and this work (x axis). The one-to-one line is shown in black.

comparison between the orbital period values found by these authors and the values found in this work in Figure 4.4. We can see that for three stars we find different values while the rest of the stars are in agreement. Looking closer at these three stars in Figure 4.5, we can see that while our values might not agree with the Giesers values, our best-fitting RV curves do match the data well. This suggests that Giesers et al. (2019) likely found a local maximum in the posterior period distribution due to sampling from a limited number of prior samples when using `TheJoker`. However, when we use `Dynesty`, we are able to recover all possible solutions and their relative probabilities.



(a)

(b)

(c)

Figure 4.5: The radial velocity curve fits for the three stars not in agreement with the results found by Giesers et al. (2019) using `TheJoker` in Figure 4.4. The left panels show the nested sampling fits to the radial velocity curve in orange with the data shown in black. The middle panels show the histogram of the orbital period posteriors in orange with the 1σ credibility intervals shown with the black dashed lines and the Giesers et al. (2019) best-fit value and 1σ uncertainties shown in a solid and dashed green lines, respectively. The right panels also show the orbital period posteriors but just for nested sampling.

4.3 The Orbital Period Distribution of Binaries in NGC 3201

The next step in this process is to take the posterior samples from these radial velocity fits and use the method outlined in Chapter 3. In particular, we use the method described in Section 3.5 as we do not know the true shape for the binary orbital period distribution for NGC 3201.

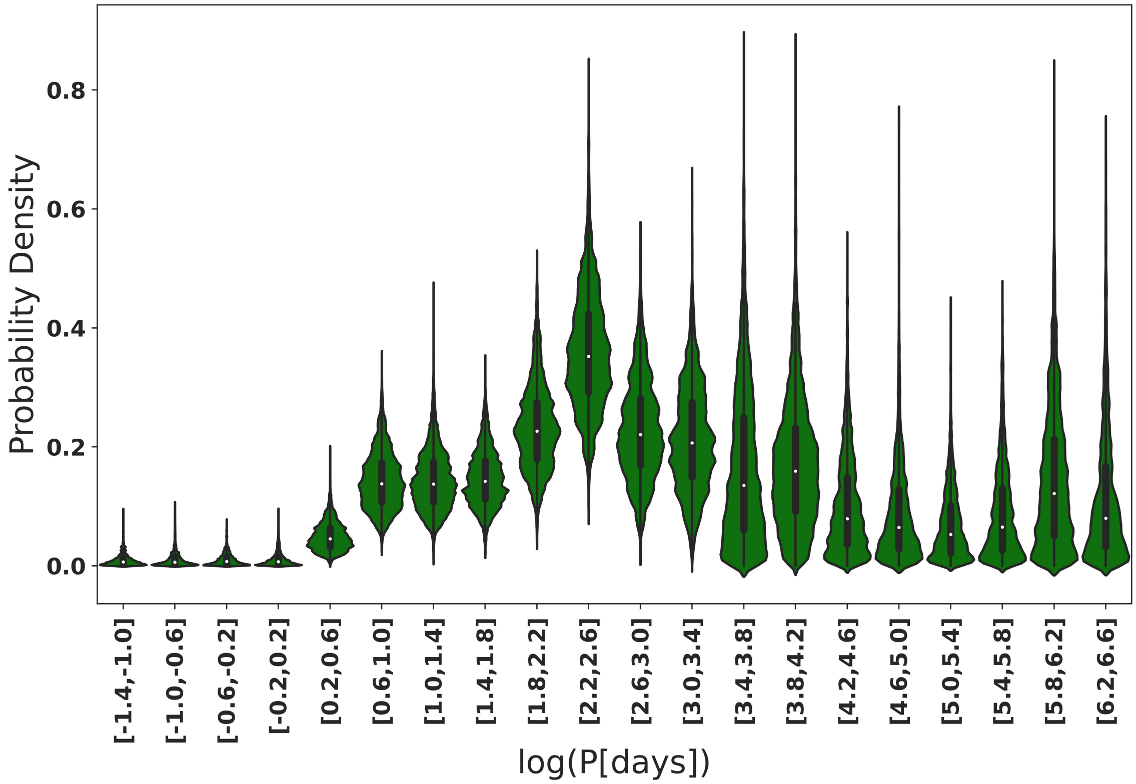


Figure 4.6: The period distribution for NGC 3201. For each interval in the piece-wise function, the smoothed posterior probability distributions for the interval height is shown with the 1σ credibility interval for the distribution in each period interval shown with a thick black line.

Applying this method, we can see the results for NGC 3201 shown in Figures 4.6 and 4.7. In this analysis, the range of periods considered is such that it includes the binaries with the shortest expected period (0.3 days) and the largest period binary that could survive today without being ionized (approximately 4×10^4 days - see Section 4.4) to ensure the full distribution is captured. In Figure 4.6, we can see a violin plot representing the smoothed posterior probability distributions for the probability density in each period interval (the 'heights' of the different bins in the piece-wise function representing the intrinsic period distribution). The 1σ credibility interval for the distribution in each period bin is shown with a thick black line. From this, we can see that the probability of a binary with an orbital period up to $\log(P[\mathrm{days}]) = 0.2$ is in agreement with zero, meaning that we do not expect to find many (or any) binaries with a period this low. The peak of the distribution is also clear in this figure and we can see that it occurs in the range of $\log(P[\mathrm{days}])$ =2.2 to 2.6 which corresponds to approximately 158 days to 398 days. At the other end, we can see that beyond $\log(P[\mathrm{days}]) = 4.2$ the probability density begins to tend towards zero. While we might expect this to be zero as we do not expect binaries with such long periods to survive (given the hard-soft limit of the cluster at the present day; see Section 4.4), this regime is difficult to fully constrain. This because we would likely not detect any binaries at these longer periods but they would still give rise to the same observed radial velocities. Therefore, while we are able to partially mitigate the effects of the detection bias, it is still present and is likely the cause of the non-zero probability in the longer-period end of the distribution.

The top panel of Figure 4.7 shows the same results as Figure 4.6, however, now each period distribution for the samples in the MCMC chain is shown. The darker regions correspond to areas of higher probability and the lighter regions are regions of lower probability. The vertical lines show the hard-soft limits that correspond to different cluster densities and will be discussed in Section 4.4. The bottom panel of this figure shows the histogram of the binary orbital periods as found by Giesers et al. (2019). One thing to note is that the bottom panel was found by taking the histogram of the reported binary periods, not using the method described in this work or taking into account the period errors. Thus, the bottom panel does not represent the full intrinsic binary period distribution, taking into account the detection bias.

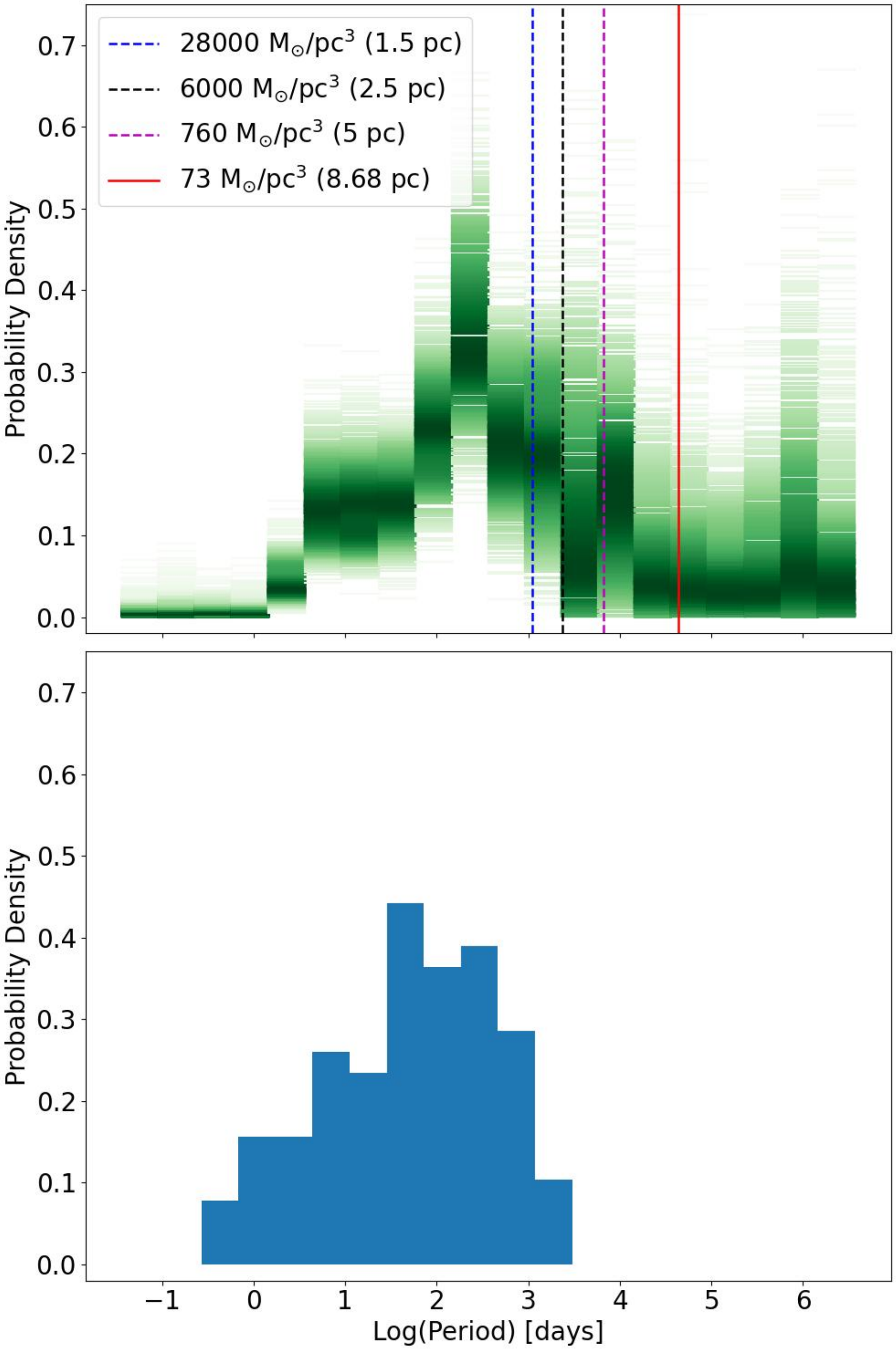


Figure 4.7: Results for the fit of the NGC 3201 binary orbital period distribution in the top panel. Each green line corresponds to one of the possible solutions and the darker regions show the areas of higher probability. The vertical lines represent the estimated hard-soft lines based on different cluster densities. The dashed lines represent the initial cluster densities (and corresponding initial half-mass radii) and the corresponding period at the initial hard-soft limit and the solid line represents the present day cluster density (and present day half-mass radii) and period corresponding to the present day hard-soft limit. The bottom panel shows the histogram of the binary periods found by Giesers et al. (2019).

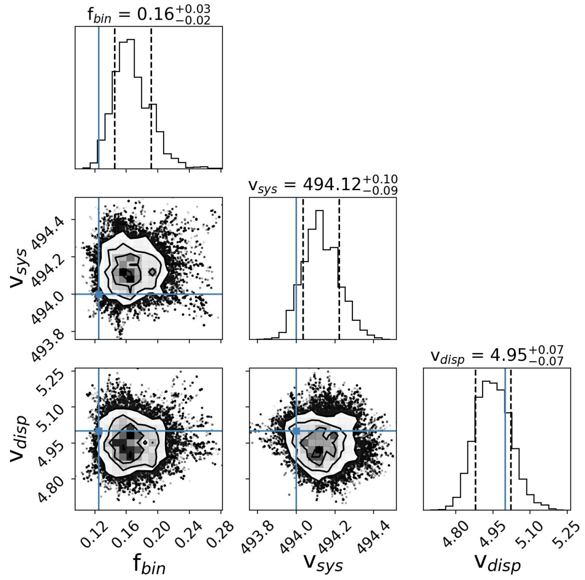


Figure 4.8: The corner plot for the binary fraction, cluster systemic velocity, and the velocity dispersion of NGC 3201 with the values reported by Giesers et al. (2019) shown in blue.

Nonetheless, by comparing the two panels, we can see that our results make sense in the context of their work and that we were able to probe the distribution at longer periods than Giesers et al. (2019) as they limited their analysis to binaries that had well-fit RV curves. However, it does appear that Giesers et al. (2019) found binaries with shorter periods than we did in this work. Upon closer inspection of these shorter period binaries this can be explained by the fact that these binaries consist of a system that Giesers et al. (2019) identified as a binary and we did not, and binary systems (15) for which Giesers et al. (2019) found a very short period solution when we found that a longer period solution had a higher probability (see Figure 4.5).

In Figure 4.8, we can see the corner plot for the binary fraction, f_{bin}, NGC 3201's systemic velocity, v_{sys}, and the velocity dispersion, v_{disp}. Comparing our results to the values reported by Giesers et al. (2019) (in blue), we can see that while our results are a bit overestimated compared to what they found, our results are consistent. One thing to note is that the method use by Giesers et al. (2019) to determine the binary fraction varies from the method used in this work and therefore we do not necessarily expect to recover the same value. As stated in Section 3.2, we constrained the binary fraction by using the velocity distribution and the knowledge that undetected binaries will inflate the wings of the velocity distribution. On the other hand, Giesers et al. (2019) took their fraction of detected binaries (stars with a binary probability above 0.5 in their case) and ran MOCCA (Monte Carlo Cluster Simulator) simulations to determine how this compared to the true binary fraction. From this, they found the percentage of binaries missing and then accounted for this when finding the binary fraction for the NGC 3201 dataset. Nonetheless, our binary fraction shows that our results make sense in the context of their work.

4.4 Discussion: the Hard-Soft Binary Limit for NGC 3201 and Constraints on its Initial Density

Finally, now that we have inferred the intrinsic binary orbital period distribution for NGC 3201, this can be used to place constraints on the initial density of this globular cluster. The relationship between the hard-soft limit and the cluster's initial density is born out of the idea that denser star clusters will have a larger velocity dispersion (at a fixed mass) and thus there are more binaries at shorter orbital periods that can be disrupted. This relates back to the initial density as clusters are the densest when they are first formed and then expand over time due to mass loss. Thus, the hard-soft limit will move to longer periods as the cluster evolves, making the initial cluster environment the most stringent binary period constraint. This was discussed in detail in Section 1.3.

In general, this hard-soft limit is the point where the kinetic energy of a typical star in the cluster is equal to the binding energy of the binary. Thus,

$$E_b = E_K \rightarrow \frac{Gm_1m_2}{2a} = \frac{1}{2}\langle m\rangle v_{disp}^2, \tag{4.1}$$

where $\langle m\rangle$ is the typical mass of a star in the cluster and m_1 and m_2 are the masses of the binary components, a is the semi-major axis of the binary, and in this context, v_{disp} is the initial velocity dispersion of the cluster. However, in this work we are considering the binary orbital period distribution, not the semi-major axis distribution. To rewrite this equation in terms of the period, we can use Kepler's third law: $P^2 = \frac{4\pi^2}{G(m_1+m_2)}a^3$. Thus, rewriting Equation 4.1 in terms of the period and solving for the period of a binary right at the hard-soft limit gives:

$$P = \left(\frac{4\pi^2}{G(m_1+m_2)}\right)^{1/2}\left(\frac{Gm_1m_2}{\langle m\rangle v_{disp}^2}\right)^{3/2}. \tag{4.2}$$

While the end goal is to constrain the initial cluster density, the initial mass of Milky Way globular clusters is generally better understood (e.g. Baumgardt et al., 2019). Therefore, the parameter to vary in these calculations to find the period of a binary at the hard-soft limit when the cluster formed is the initial half-mass radius of the cluster, which can then be used to determine the initial density of the cluster within the half-mass radius.

To express the initial velocity dispersion in terms in terms of the half-mass radius, we assume a Plummer model (Plummer, 1911), for which the one-dimensional velocity dispersion is given by

$$v_{\mathrm{disp}}^2(r) = -\frac{1}{6}\phi(r), \tag{4.3}$$

where $\phi(r)$ is the potential and is given by

$$\phi(r) = -\frac{GM}{a}\left(1+\frac{r^2}{a^2}\right)^{-\frac{1}{2}}, \tag{4.4}$$

where M is the total mass of the cluster and a is a scale factor that is approximately $R_h/1.305$ where R_h is the half-mass radius. The Plummer model was chosen as it is a convenient and reasonable way to approximate the mass profile and kinematics of a globular cluster in this case. It is a self-consistent dynamical model that can be written in terms of simple functions and solved analytically. It has historically been used as initial conditions for evolutionary calculations of globular clusters (e.g. Spitzer & Shull, 1975).

As we are interested in the typical velocity dispersion within the half-mass radius and not just the velocity dispersion at the half-mass radius, we need to integrate to compute the average of the velocity dispersion from the centre of the cluster out to R_h. However, the density changes with radius so needs to be accounted for in this integral. Thus,

$$\langle v^2_{\mathrm{disp}} \rangle = \frac{\int_0^{R_h} v^2_{\mathrm{disp}}(r) 4\pi r^2 \rho(r) dr}{\int_0^{R_h} 4\pi r^2 \rho(r) dr} \tag{4.5}$$

where $\langle v^2_{disp} \rangle$ is the average velocity dispersion within the region, R_h is the half-mass radius, $v_{\mathrm{disp}}(r)$ is the velocity dispersion at distance r and $\rho(r)$ is the mass density at distance r.

For the Plummer model, the density profile is given by

$$\rho(r) = \frac{3M}{4\pi a^3} \left(1 + \frac{r^2}{a^2}\right)^{-\frac{5}{2}} . \tag{4.6}$$

Thus, by solving Equation 4.5 and using $\langle v^2_{disp} \rangle$ in Equation 4.2, we can find the largest orbital period that a binary could have (without becoming ionized) given a specified initial half-mass radius.

Using the initial half-mass radii, the initial cluster densities and their corresponding hard-soft limit periods can be seen in the dashed lines in Figure 4.7. These values were determined by assuming that the present-day cluster mass of NGC 3201 is 1.93×10^5 $M_\odot$, that the initial cluster mass is 4.1×10^5 $M_\odot$ (Baumgardt, 2017) (about 50% of the mass has been lost due to stellar evolution), that the typical mass of a star in the cluster is $\langle m \rangle = 0.8$ $M_\odot$ and that the mass components of the binary are $m_1 = 0.8$ $M_\odot$ and $m_2 = 0.5$ $M_\odot$. These values for the binary mass components were taken from the typical values found for the binary visible and invisible mass components by Giesers et al. (2019). For the invisible mass component, this involved taking into account the average $\sin(i)$ value (0.78 when assuming random orientations of the orbital plane and inclination) as the reported masses are the lower limits due to unknown inclinations.

In Figure 4.7, the first three vertical lines correspond to the period of a binary at the hard-soft limit for different possible initial cluster densities (found using initial half-mass radii of 1.5 pc, 2.5pc, and 5.0 pc from left to right) and the last vertical line to the right (in red) corresponds to the period of a binary at the hard-soft limit given the present-day cluster density (found by taking the present day half-mass radii to be 8.68 pc). This last limit shows us that our results align with our expectations in the very long period regime as we certainly do not expect any binaries to survive with periods larger than the present-day hard-soft limit. The hard-soft limits corresponding to different initial half-mass radii are interesting as they suggest that the initial half-mass radius of NGC 3201 was likely between approximately 2.5 and 5 pc. 1.5 pc (and other smaller half-mass radii) was rejected as a potential initial half-mass radii as the probability density is significantly different from zero at the period corresponding to this hard-soft limit.

This is interesting as in the literature, while the initial density and the initial half-mass radius are poorly constrained, Monte Carlo cluster models that are chosen to be the best match for NGC 3201 have smaller initial half-mass radii. These models are usually selected based on their general agreement with the structure (surface brightness profile) and kinematics (velocity dispersion profile) of the cluster at present day. For example, Giesers et al. (2019) used a MOCCA model with an initial half-mass radius of 1.65 pc to represent NGC 3201 and Rui et al. (2021) used a CMC Cluster Catalog model with an initial virial radius of 2 pc (assuming a Plummer model, this corresponds to a half-mass radius of approximately 1.54 pc). However, these values were chosen alongside other parameters to best reproduce the present-day cluster and were not well-constrained, partly because the grids of Monte Carlo models from which they are selected are quite sparse. Our approach provides a better and completely independent constraint on this initial half-mass radius, which can be used in the future to find Monte Carlo models that better replicate NGC 3201, and more generally which can be used in combination with other information to constrain the initial conditions and evolution of globular clusters.

Taking the initial half-mass radius of NGC 3201 to be in the range of $2.5-5$ pc, this corresponds to an initial mass density within the half-mass radius of approximately 760 - 6100 $\mathrm{M}_{\odot}/\mathrm{pc}^3$. While the initial density of globular clusters is not currently well constrained, this aligns with recent work done by Gieles & Gnedin (2023) who showed using N-Body models that for globular clusters that survive to the present day, and to reproduce the global properties of the population of Milky Way globular clusters, the initial density of globular clusters is expected to be in the range of approximately 10 to 10^5 $\mathrm{M}_{\odot}/\mathrm{pc}^3$. Thus, our estimate for NGC 3201 nicely fits within this range, reaffirming that our results are reasonable. This constraint on the initial cluster density is also important for things such as the expected black hole retention of the cluster as denser clusters will have ejected more black holes than less dense clusters due to more frequent interactions between black holes in the cluster core (e.g. Gieles et al., 2021).

Overall, by implementing the new method developed in this work, we are able to fit the radial velocity curves of individual binaries, better constrain the current day intrinsic binary population orbital period distribution, and use this period distribution to place new constraints on the required initial cluster density that could give rise to this distribution in the globular cluster NGC 3201.

Chapter 5

Conclusions

In this work we have developed a hierarchical Bayesian method that takes in multi-epoch radial velocity measurements for stars in globular clusters, fits the radial velocity curve for stars detected as RV variable, and then uses the posterior samples from these fits to construct the intrinsic binary orbital period distribution, while accounting for observational biases.

In Chapter 2, we first determined which stars where RV variable by using a χ^2 test. For each of the RV variables, we then needed a method for fitting the radial velocity curves. This involved assessing which sampling method was most suitable for what we required. The methods considered were MCMC (using `emcee`), rejection sampling (using `TheJoker`), and nested sampling (using `Dynesty`). From this, we found that while `TheJoker` is what is most commonly used in the literature, specifically for cases with sparse radial velocity data, it does not perform as well in cases with well-sampled radial velocity curves. This leads to stars with few posterior samples, meaning that they would not be usable in the fit for the binary orbital period distribution. Although `emcee` was able to recover enough posterior samples, it can get stuck in a local maximum of the posterior probability distribution and miss other peaks in the distribution. On the other hand, `Dynesty` was found to reliably recover the true solution for radial velocities from mock datasets and ensured a sufficient number of posterior samples. Therefore, from this point on we took `Dynesty` as the method for fitting radial velocity curves in this work.

With the posterior samples from these radial velocity curve fits, we then developed a method to take these in and then fit the period probability distribution (Chapter 3). This was not straightforward as we needed to account for the detection bias and in order to fully constrain the distribution (particularly the long-period end of the distribution), we needed to add in the cluster binary fraction and the parameters characterizing the line-of sight velocity distribution with the cluster as free parameters. By adding in these additional terms, we first showed that we were able to recover the true distribution in cases with mock data where the true distribution was known and we were fitting with a standard parameterized distribution function (log-normal and log-weibull). However, this could not be applied to a real dataset at this point as the intrinsic globular cluster binary period distribution does not have a known functional form and it is not expected to be universal across all globular clusters. Therefore, we needed a flexible method that could be applied to any cluster with minimal assumptions about the shape of the distribution. To achieve this, we fit the period distribution using a piece-wise function where the intervals remained constant but the height of the distribution in each interval was allowed to vary. By doing this, we showed that we were still able to recover the true period distributions.

After showing that this method works for mock data, the next step was to apply it to the real data for NGC 3201 (Chapter 4). NGC 3201 was chosen as the first cluster to apply this method to due to the extent of radial velocity data available, meaning that more long-period binaries can be identified, thus helping us constrain the intrinsic binary period distribution. First, we fit the radial velocity curves for every star detected as RV variable and, for the stars that we had in common with Giesers et al. (2019), we compared our results. For the most part our results are in agreement. In the cases where we differ, our results are a good fit for the radial velocity data, giving us confidence in our results. With the posterior samples from these fits, we then fit the orbital period distribution for NGC 3201 using our flexible fitting procedure. Comparing our results for the binary periods, binary fraction (0.16±0.03) , the cluster systemic velocity (494.12 km s± 0.1 km s^{-1}), and the velocity dispersion (4.95 ±0.07 km s^{-1}) to those found by Giesers et al. (2019), we showed that our results make sense in the context of their work and that our method has provided us with more information about the intrinsic binary period distribution.

As the majority of binary disruptions occur when a cluster is at its densest and clusters are the densest at the early stages of their evolution, we are able to use this present-day intrinsic binary period distribution to infer the initial cluster conditions. From this, we were able to show that our results suggest an initial half-mass radius of around 2.5-5 pc which corresponds to an initial cluster density of 760-6100 $\mathrm{M}_{\odot}/\mathrm{pc}^3$. This represents a new, independent, and more stringent constraint on the initial density of this cluster. This will further inform dynamical modelling of the evolution of NGC 3201 and its black hole population, including implications for the retention of black holes over time, as a cluster with a higher initial density will result in the ejection of more black holes. This black hole retention has implications for the formation of possible black hole binary mergers that produce gravitational waves.

5.1 Future Work

In future work, this method could be improved by accounting for the radial dependence for parameters such as the binary fraction and velocity dispersion of the cluster. Currently, we assume no radial dependence for any parameters within the field of view of the MUSE pointings. However, in reality the binary fraction and velocity dispersion are expected to increase towards the core of the cluster due to mass segregation and the mass distribution within the cluster. Therefore, by accounting for this, we could improve the accuracy of the method.

Additionally, now that we have a method that works, it can be applied to other globular clusters with multi-epoch radial velocity data. For example, Kamann et al. (2018) obtained MUSE radial velocity measurements for 25 globular clusters so this method could be applied to these clusters in the future. In particular, 47 Tucanae and Omega Centarui are well-studied clusters and have been observed at a large number of epochs (comparable to the NGC 3201 dataset used in this work) and would be prime targets for applying this method.

As well, now that we are able to place constraints on the initial density of globular clusters using the imprint of the initial density on the period distribution of binaries, this can be paired with N-Body models to look at the implications on aspects such as the evolution of the cluster, the retention of black holes, and the resulting gravitational waves from black hole binary mergers.

Bibliography

Antonini, F., & Gieles, M. 2020, Phys. Rev. D, 102, 123016, doi: 10.1103/PhysRevD.102.123016 21

Ashman, K. M., & Zepf, S. E. 1998, Globular Cluster Systems (Cambridge University Press) 2

Bacon, R., Accardo, M., Adjali, L., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III, ed. I. S. McLean, S. K. Ramsay, & H. Takami, 773508, doi: 10.1117/12.856027 79

Bahcall, J. N., Hut, P., & Tremaine, S. 1985, ApJ, 290, 15, doi: 10.1086/162953 15

Baumgardt, H. 2017, MNRAS, 464, 2174, doi: 10.1093/mnras/stw2488 93

Baumgardt, H., & Hilker, M. 2018, MNRAS, 478, 1520, doi: 10.1093/mnras/sty1057 79

Baumgardt, H., Hilker, M., Sollima, A., & Bellini, A. 2019, MNRAS, 482, 5138, doi: 10.1093/mnras/sty2997 91

Baumgardt, H., & Vasiliev, E. 2021, MNRAS, 505, 5957, doi: 10.1093/mnras/stab1474 79

Binney, J., & Tremaine, S. 2008, Galactic Dynamics (Princeton University Press) 3, 4

Brodie, J. P., & Strader, J. 2006, ARA&A, 44, 193, doi: 10.1146/annurev.astro.44.051905.092441 5

Cottaar, M., & Hénault-Brunet, V. 2014, A&A, 562, A20, doi: 10.1051/0004-6361/201322445 24, 26, 64, 65

Duquennoy, A., & Mayor, M. 1991, A&A, 248, 485 17

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: 10.1086/670067 33

Gieles, M., Baumgardt, H., Heggie, D. C., & Lamers, H. J. G. L. M. 2010, MNRAS, 408, L16, doi: 10.1111/j.1745-3933.2010.00919.x 5

Gieles, M., Erkal, D., Antonini, F., Balbinot, E., & Peñarrubia, J. 2021, Nature Astronomy, 5, 957, doi: 10.1038/s41550-021-01392-2 21, 94

Gieles, M., & Gnedin, O. Y. 2023, MNRAS, 522, 5340, doi: 10.1093/mnras/stad1287 94

Giesers, B., Kamann, S., Dreizler, S., et al. 2019, A&A, 632, A3, doi: 10.1051/0004-6361/201936203 vii, viii, 6, 30, 32, 39, 79, 80, 81, 82, 83, 84, 85, 87, 88, 89, 90, 93, 94, 97, 98

Goodman, J., & Hut, P. 1989, Nature, 339, 40, doi: 10.1038/339040a0 5

Goodman, J., & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, 5, 65, doi: 10.2140/camcos.2010.5.65 33

Gratton, R., Bragaglia, A., Carretta, E., et al. 2019, A&A Rev., 27, 8, doi: 10.1007/s00159-019-0119-3 16

Harris, W. E. 2010, arXiv e-prints, arXiv:1012.3224, doi: 10.48550/arXiv.1012.3224 2

Heggie, D., & Hut, P. 2003, The Gravitational Million-Body Problem: A Multidisciplinary Approach to Star Cluster Dynamics (Cambridge Univiversity Press) 2, 4, 12, 14, 16

Heggie, D. C. 1975, MNRAS, 173, 729, doi: 10.1093/mnras/173.3.729 13, 14

Heggie, D. C., Hut, P., & McMillan, S. L. W. 1996, ApJ, 467, 359, doi: 10.1086/177611 13

Hogg, D. W., Myers, A. D., & Bovy, J. 2010, ApJ, 725, 2166, doi: 10.1088/0004-637X/725/2/2166 57

Hut, P., McMillan, S., Goodman, J., et al. 1992, PASP, 104, 981, doi: 10.1086/133085 8

Ivanova, N., Belczynski, K., Fregeau, J. M., & Rasio, F. A. 2005, MNRAS, 358, 572, doi: 10.1111/j.1365-2966.2005.08804.x 12, 16, 18, 19, 69

Ivezić, Ž., Connolly, A. J., VanderPlas, J. T., & Gray, A. 2020, Statistics, Data Mining, and Machine Learning in Astronomy. A Practical Python Guide for the Analysis of Survey Data, Updated Edition 34

Kamann, S., Husser, T. O., Dreizler, S., et al. 2018, MNRAS, 473, 5591, doi: 10.1093/mnras/stx2719 79, 99

Kipping, D. M. 2013, MNRAS, 434, L51, doi: 10.1093/mnrasl/slt075 31

Kruijssen, J. M. D., Pfeffer, J. L., Reina-Campos, M., Crain, R. A., & Bastian, N. 2019, MNRAS, 486, 3180, doi: 10.1093/mnras/sty1609 5

Marks, M., Kroupa, P., & Oh, S. 2011, MNRAS, 417, 1684, doi: 10.1111/j.1365-2966.2011.19257.x 19

McVean, J. R., Milone, E. F., Mateo, M., & Yan, L. 1997, ApJ, 481, 782, doi: 10.1086/304067 6, 8

Milone, A. P., Piotto, G., Bedin, L. R., et al. 2012, A&A, 540, A16, doi: 10.1051/0004-6361/201016384 7, 8

Perryman, M. 2011, The Exoplanet Handbook (Cambridge University Press) 6, 9, 10

Plummer, H. C. 1911, MNRAS, 71, 460, doi: 10.1093/mnras/71.5.460 91

Price-Whelan, A. M., Hogg, D. W., Foreman-Mackey, D., & Rix, H.-W. 2017, ApJ, 837, 20, doi: 10.3847/1538-4357/aa5e50 32, 38

Price-Whelan, A. M., Hogg, D. W., Rix, H.-W., et al. 2020, ApJ, 895, 2, doi: 10.3847/1538-4357/ab8acc 31, 32, 39

Raghavan, D., McAlister, H. A., Henry, T. J., et al. 2010, ApJS, 190, 1, doi: 10.1088/0067-0049/190/1/1 17

Rui, N. Z., Kremer, K., Weatherford, N. C., et al. 2021, ApJ, 912, 102, doi: 10.3847/1538-4357/abed49 94

Sana, H., de Mink, S. E., de Koter, A., et al. 2012, Science, 337, 444, doi: 10.1126/science.1223344 19

Skilling, J. 2004, in American Institute of Physics Conference Series, Vol. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. R. Fischer, R. Preuss, & U. V. Toussaint, 395–405, doi: 10.1063/1.1835238 43

Sollima, A. 2008, MNRAS, 388, 307, doi: 10.1111/j.1365-2966.2008.13387.x 13

Speagle, J. S. 2020, MNRAS, 493, 3132, doi: 10.1093/mnras/staa278 vi, 43, 44

Spitzer, L., J., & Shull, J. M. 1975, ApJ, 200, 339, doi: 10.1086/153793 92

Spitzer, L. S. 1987, Dynamical Evolution of Globular Clusters (Princeton University Press) 13, 14, 16