An Analysis of Variable Objects in the Globular Cluster M4 Using Observations from the NASA K2 Mission

Joshua John Wallace

A Dissertation
Presented to the Faculty of Princeton University in Candidacy for the Degree of Doctor of Philosophy

Recommended for Acceptance by the Department of Astrophysical Sciences

Adviser: Gáspár Áron Bakos

September 2019
Abstract

This dissertation presents an analysis of data from the NASA K2 mission for the globular cluster (GC) M4, representing the longest continuous observation of a GC. The image reduction and variable object identification pipeline is described. Light curves for 4554 objects are extracted and made publicly available. Among cluster-member stars, 66 variables are found, including 52 new discoveries. Among field stars, 24 variables are found, including 20 new discoveries. Additionally, 57 cluster-member suspected variables, 10 cluster-nonmember suspected variables, and four variables with ambiguous cluster membership are discovered. Two objects appear to be representative of a new variable subclass, which are dubbed “millimagnitude RR Lyrae” (mmRR). Asteroseismic activity is detected in seven horizontal branch stars, as well as 19 stars along the red giant branch. A previously known cluster-member RR Lyrae variable is identified as a candidate Blazhko variable, with a Blazhko period in excess of 78 days. A search for transiting planets in M4 is also performed. Previous surveys have at best been sensitive to planets with periods $P \lesssim 16$ days and, at the shortest periods, planets with radii $R_p \gtrsim 0.8$ $R_J$—this search is sensitive to larger periods ($P \lesssim 35$ days) and, at short periods, smaller planet radii ($R_p \gtrsim 0.3$ $R_J$) than any previous survey. Seven planet candidates are presented which, if any are confirmed, would be the first transiting planet known in a GC. An analysis of the systematic noise reveals that the false alarm probabilities for these candidates are high. Upper limits are placed on planet occurrence and compared to previous results. In the new GC parameter space explored by the results, I calculate $3\sigma$ occurrence rate upper limits of 6.6% for 0.71–2 $R_J$ planets with 1–36 day periods and 17% for 0.36–0.71 $R_J$ planets with 1–10 day periods. A theoretical motivation for the planet search, examining collisional fragmentation during the formation of short-period rocky planets, is presented, as is the M4 proper motion membership catalog I compiled, which is the most complete assembled for M4 to date.
Acknowledgements

There have been hundreds of people I have interacted with over the last five years and who have had an influence during my five years preparing this dissertation, but of necessity I will focus on just a few.

I would first like to thank my advisors, Gáspár Bákos and Joel Hartman, under whose mentorship my skills as a scientist have significantly grown. Among other things, I would like to particularly thank Gáspár for his kindness and understanding as I have worked to balance my various responsibilities, and Joel for always being willing to answer my questions even when I stopped by his office unannounced, and also for taking the time to travel to Arizona with me to give me observing experience. I must also thank my semester project advisors, Adam Burrows and Scott Tremaine, who gave me beneficial mentorship at the beginning of my graduate research experience and have helped me develop breadth as a scientist. I also want to thank all of the above plus Josh Winn for being on my thesis committee these three years and the valuable ideas and feedback they have given me in that capacity. And last but not least, I thank Waqas Bhatti for the incredibly valuable help and mentorship he gave me as I have engaged in this research.

Moving on to other Peytonites, I would also like to thank the many graduate and post-baccalaureate students who have valued me as a colleague and as a friend. Their support and friendship have made the basement brighter than any windows could have. I want to give particular thanks to my cohort: Kris Pardo, Melinda Soares-Furtado, and David Vartanyan. Thanks for the laughs, the love, and the sushi. The faculty and staff of the department have been helpful and supportive, and in particular, I would like to thank Polly Strauss, whose able and happy administration of the grad student program has so positively transformed the experience for all of us.
Only a relatively small fraction of Princeton graduate students have children, so I am grateful for the group Tigers with Cubs to bring us together. Thank you to all of the student–parents out there who make their campuses a better and richer place by bringing their children along for the difficult ride. I have had the pleasure of serving first as Advocacy Liaison and then President for Tigers with Cubs, and I am extremely grateful for all the work Tigers with Cubs members and the Graduate Student Government put in to various advocacy efforts, which have had such great success. In particular, I would like to thank the Graduate Student Government presidents who have helped me find a voice on campus: Sean Edington, Akshay Mehra, Mircea Davidescu, Jonathan Balkind, and Mai Nguyen.

There are a variety of university-level administrators whom I would like to thank for their work with me and with others in our campus advocacy. From the graduate school, Sanjeev Kulkarni and Sarah-Jane Leslie as deans and Cole Crittenden as deputy dean, and associate and assistant deans Nicole Barkley, Mary Bechler, Lisa Schreyer, and Lily Secora. I would also like to thank the housing office, and Michael Hebditch and Jim Poole in particular, who helped me with many questions related to my housing and made an inconvenient situation more bearable.

In my career considerations and job search, Susanne Killian was a vital source of help and counsel whom I recommend to all graduate students, and Amy Pszczolkowski and Susanne provided much programming that was very helpful in my career search and preparations.

I wish also to thank my undergraduate research advisor, Zheng Zheng, who was the first to show me the full breadth and depth of the wonder of astrophysics. My work with Zheng is the entire reason I became interested in the field in the first place.

I must also thank the Princeton 1st Ward of The Church of Jesus Christ of Latter-day Saints for being my family’s home away from home. They have taken such good care of my family, and I have learned so much from so many people in the ward.
In particular I must thank Bryan Law whose truly inspired leadership as a bishop helped turn around some of my darkest hours here; Renaud Couillens, Caleb Cluff, Mike Tolley, and Fernando Balderrama for the great experiences leading an elders quorum together; and those grad student families in the ward whose time here has overlapped with most of ours: Matt and Michele Nielsen, David and Nana Romney, Jane and Jared Balaich, and Amanda and Patrick Monson. Each of these families has answered my family’s prayers in some way. And in that vein, I also want to thank the God of the heavens I study for the sustaining power I have received throughout my time here.

And finally, I must thank my family. First my parents, who were willing to have their first two grandchildren born far away from them so that I could pursue this degree. Next my children, Simon and Felicity, who were both born here: thanks for the excitement and purpose of life you brought me. And last mentioned but first in importance, my dear wife and partner Jordan. She was the first person to encourage me to pursue a graduate degree and has always been my number one support, believing in me even when I did not believe in myself. There is much she has done and sacrificed to allow me to achieve what I have. Thank you dear. Here’s to our next chapter together.
Preface

This dissertation focuses on the analysis and interpretation of data acquired by the Kepler space telescope during the K2 portion of its mission of a certain star cluster, a globular cluster with the catalog designation M4. It consists of five chapters and one appendix. The first chapter provides an introduction to globular cluster time series observations and the Kepler telescope and K2 mission. The second through fifth chapters are framed as self-contained, stand-alone works: Chapter 2 presents observations and analysis of the first published discovery from my work, that of the millimagnitude RR Lyrae; Chapter 3 provides a detailed description of my data processing pipeline as well as a complete catalog of the variable objects discovered in the data; Chapter 4 presents a theoretical analysis of collisional fragmentation of rocky planets during their formation in very short-period orbits; and Chapter 5 describes planet candidates found in the data from M4 and a planet occurrence rate calculation based on these data. The Appendix (Appendix A) discusses my proper motion cluster membership catalog calculated from Gaia data, which informs the previous studies.

Portions of this dissertation have been published or submitted to peer-reviewed or editor-reviewed journals, and have been subjects of oral presentations, as detailed below. The corresponding chapters have each benefitted from collaboration with and mentorship from co-authors listed for each publication and their assistance is gratefully acknowledged. Additionally, Chapter 5 is planned to be submitted to The Astrophysical Journal, with Joel Hartman and Gáspár Bakos as co-authors.

Publications

1. Collisional Fragmentation Is Not a Barrier to Close-in Planet Formation
2. Ultralow-amplitude RR Lyrae Stars in M4

3. M4 Membership Catalog from Gaia Proper Motions
Joshua J. Wallace 2019, RNAAS, 2, 213 (cf. Appendix A)

4. A Search for Variable Stars in the Globular Cluster M4 with K2

Presentations

1. Detecting Time Variability of Stars in Crowded Images

2. Astrophysical Investigations
   Invited Talk at Intel Co., December 2018, Hillsboro, OR

3. Time Series Analysis of M4 by the Kepler Telescope: A Search for Variable Objects
   Invited Talk at Edgestream Partners L.P., August 2018, Princeton, NJ

4. Collisional Fragmentation Is Not a Barrier to Close-in Planet Formation
   Submitted Talk at DPS 49, October 2017, Provo, UT

5. Is Collisional Fragmentation a Barrier to the Formation of Short-Period Planets?
   Submitted Talk at AAS 229, January 2017, Grapevine, TX

---

1 The Research Notes of the AAS is an editor-reviewed journal of short publications; in contrast, the other journals listed here are all peer reviewed.
For Jordan,

who was the first to encourage me to take this journey

and has walked every step with me, rain or shine.
# Contents

Abstract ................................................................. iii
Acknowledgements ..................................................... iv
Preface ................................................................. vii

1 Introduction .................................................................. 1
   1.1 Globular Clusters .................................................... 1
   1.2 Variable Stars ....................................................... 4
   1.3 The *Kepler* Space Telescope .................................. 8

2 Ultralow-amplitude RR Lyrae Stars in M4 ....................... 10
   2.1 Abstract .............................................................. 10
   2.2 Introduction ........................................................ 11
   2.3 Observations and Analysis ....................................... 12
      2.3.1 K2 Image Subtraction, Reduction, and Variable Search 12
      2.3.2 The Horizontal Branch Stars .............................. 13
      2.3.3 Blend Scenarios .............................................. 16
   2.4 Discussion .......................................................... 21

3 A Search for Variable Stars in the Globular Cluster M4 with K2 24
   3.1 Abstract .............................................................. 24
   3.2 Introduction ........................................................ 25
   3.3 Method ............................................................... 27
4.4.2 Equal-mass Case ........................................... 109
4.4.3 Unequal-mass Case ....................................... 114
4.5 Numerical Methods ........................................... 118
  4.5.1 Collision Algorithm ..................................... 118
  4.5.2 Initial Conditions ....................................... 121
4.6 Results ....................................................... 126
  4.6.1 Fiducial runs ............................................. 126
  4.6.2 Other scenarios ......................................... 137
4.7 Discussion ................................................... 139
4.8 Conclusion ................................................... 142

5 A Search for Transiting Planets in the Globular Cluster M4 with K2: Candidates and Occurrence Limits 143
  5.1 Abstract .................................................... 143
  5.2 Introduction ............................................... 144
  5.3 Method ...................................................... 148
    5.3.1 Photometric Processing ................................. 148
    5.3.2 Transit Search ......................................... 149
    5.3.3 Signal-to-pink Noise Threshold Determination ....... 151
    5.3.4 Occurrence Rate Calculation .......................... 153
  5.4 Transit Search Results ...................................... 159
  5.5 Occurrence Rate Results .................................... 164
    5.5.1 Transit Recovery Results .............................. 164
    5.5.2 Planet Occurrence Rates and Limits .................. 168
  5.6 Discussion .................................................. 174
  5.7 Conclusion .................................................. 179

Appendix 181
Chapter 1

Introduction

1.1 Globular Clusters

Globular clusters (GCs) are curious astronomical objects. They are dense clusters of up to millions of stars bound together by their own self-gravity (which leads to many interesting astrophysical phenomena such as gravothermal collapse and evaporation, which are outside the scope of this work but the interested reader is referred to Binney & Tremaine 2008). Many, including the ones discussed in this work, are almost as old as the universe at some \(\sim 12\) billion years in age. The Milky Way has approximately 150 GCs, at least of those discovered to date (Harris 1996, 2010 edition). In this enumeration, though, the Milky Way appears to be on the low end among large galaxies, with M31 (our closest large neighbor, the Andromeda galaxy) having \(460 \pm 70\) GCs (Barmby & Huchra, 2001) and the massive galaxies M87 and M89 having, respectively, \(12000 \pm 800\) and \(1400 \pm 170\) GCs (Tamura et al., 2006).

The existence of GCs has been known since the 17th century, but initial observers could only see them as smudges of light in the sky and thus they were initially called “nebulae.” They were not realized as collections of stars until the observations of Charles Messier (Messier, 1771) were able to resolve individual stars in M4, the GC
Figure 1.1: The globular cluster M4. The appearance of M4 is characteristic of globular clusters: a dense central core of stars, with the stellar density tapering off away from the center in a fairly symmetric fashion. Most of the brightest stars visible in this image are foreground stars and not part of the cluster. Image credit: ESO Imaging Survey, distributed under a Creative Commons Attribution 4.0 International License.
that is also the main subject of this thesis. In fact, the “M” in M4 stands for “Messier” as it is the fourth object listed in Messier’s catalog. Figure 1.1 shows M4 and is included here as an introductory visual reference for the appearance and structure of GCs. The term “globular cluster” was first coined by William Herschel (Herschel, 1789), who also realized in the same work that the increasing brightness towards the center of GCs could not be explained alone by the spherical shape of the clusters but must also be accompanied with an increasing concentration of stars towards the center. Interestingly, though not of consequence for this work, the asymmetric distribution of GCs on the sky as seen from Earth was what led Harlow Shapley to conclude that Earth was not the center of the Milky Way, the first evidence of this fact (Shapley, 1918). He supposed that the GCs, which when observed from Earth are highly clustered in one part of the sky, might actually be evenly distributed about the center of the Milky Way. By measuring distances to many of the GCs, he discovered that from a certain distance away from Earth the GCs did indeed arrange themselves into a symmetric distribution, which Shapley correctly concluded was the center of our Galaxy.

Broadly around the same time as Shapley’s work, variable stars in GCs were a very active avenue of research. The term “variable star” means a star that varies in brightness over time. A selection of works detailing and cataloguing such discoveries includes Pickering & Bailey (1895), Barnard (1900), Leavitt & Pickering (1904), Cannon (1907), Davis (1917), Woods & Bailey (1919), and Bailey et al. (1919). In particular, Leavitt & Pickering (1904) discovered some of the largest amplitude variables examined in this work. It was realized that many of these variable stars had similar properties, and these were initially called the “cluster variables” owing to their concentrated presence in GCs. Around this time, Williamina Fleming discovered the field star RR Lyr to be variable in similar fashion as the cluster variables (Pickering et al., 1901). It was soon realized that the cluster variables actually belong to
a broader class of variable stars that were not just found in clusters; this class of variable star was named “RR Lyrae” after that first field star realized to possess this variability.

1.2 Variable Stars

Despite how static and fixed things look in the night sky to the human eye, the universe is full of movement and variation. In particular, there are many types of stars that change in brightness over time. In reality, all stars vary in brightness to some degree, but for most stars their levels of brightness variability are too small to detect. The term variable star typically refers to just those stars whose brightness variations are large enough for us to detect. Describing the full zoo of such variable objects is outside the scope of this work, and instead I will only focus on a few classes of stellar variability that are relevant for the data and discussion on hand.

The RR Lyrae variables previously mentioned are “pulsating variables.” These stars have brightness variations produced by pulsations of the star—the star gets larger and smaller over time. RR Lyrae variables pulsate in either their fundamental or first overtone radial modes or sometimes a mix of the two modes. Each star pulsates at a period that is more or less constant over short intervals of time, though period changing has been observed (e.g., Netzel et al. 2018 and Arellano Ferro et al. 2018 as recent examples, and Clementini et al. 1994 for V15, one of the stars in this work), and RR Lyrae variables are expected to have their variability periods change as they evolve, primarily due to their changing densities as their cores fuse helium into carbon. A typical period range for RR Lyrae variables is 0.1–1 days, with the first-overtone variables having shorter periods than the fundamental mode pulsators.

The mechanism causing the RR Lyrae variability is as follows (based on the discussions in Carroll & Ostlie 2007, Aerts et al. 2010, and Catelan & Smith 2015).
The opacity—a measure of how opaque a material is to electromagnetic radiation—of material in a star approximately obeys the Kramers opacity law throughout much of the star, which is commonly expressed in proportional form as

\[ \kappa \propto \rho \frac{T^3}{\rho^{3.5}}, \tag{1.1} \]

where \( \kappa \) is the opacity, \( \rho \) is the density, and \( T \) is the temperature of a given region of the star. Typically in a star, when material gets compressed, both density and temperature increase, and the associated change in opacity with the compression depends on which of these values increases more in Equation (1.1). Since it has the larger exponent, temperature usually wins out, and the opacity decreases with compression. This makes the material more transparent, allowing any radiation bottled up with the compression to more easily escape after the compression than before. Similarly, when material in a star expands, the decrease in temperature wins out over the decrease in density in Equation (1.1), and the opacity increases. This leads to the material becoming opaque, holding in the electromagnetic energy it has. The interplay of these compressions, expansions, and opacity as described here leads to a system that is stable and relatively static. However, as mentioned, RR Lyrae variables are pulsating variables. For pulsations to occur, there must be a change to the picture painted above: there must be a way to store heat in the material during compression, rather than having it radiate away. For RR Lyrae variables, this is accomplished in partial ionization zones, specifically the “He II partial ionization zone,” or the region where the helium is all singly ionized and partially doubly ionized. When these partial ionization zones are compressed, part of the work performed in the compression goes into ionizing the material rather than increasing its temperature. This smaller increase in temperature with compression causes the density term in Equation (1.1) to dominate and opacity increases as a result of the compression, with energy being.
stored in the additional ionization of the material. Eventually, with the increased opacity, the pressure interior to the partial ionization zone builds up enough to push the zone out, causing the star to expand. This expansion is powered in part by the energy stored in the ionized electrons, which release their energy as they recombine with the helium ions as the star expands and cools. The decrease in the material’s temperature during the expansion is partially mitigated by this released energy, and density again dominates in Equation (1.1). This leads to decreasing opacity in the expansion, allowing the built-up energy to more easily escape. The pressure support for the expansion is lost, and the outer layers of the star collapse to begin the cycle all over again. Owing to the role that opacity plays in this pulsation mechanism, it is called the \( \kappa \) mechanism.

The \( \kappa \) mechanism is enhanced by the fact that the partial ionization zone—being cooler than the surrounding layers during compression—pulls in energy from the neighboring layers. This increases the amount of energy stored in the ionization beyond that simply introduced by the compression work. Because of this mechanism’s relation to the adiabatic exponent \( \gamma \) (because of the differing heat capacities in the partial ionization zones relative to other layers of the star), it is called the \( \gamma \) mechanism. Both the \( \kappa \) and \( \gamma \) mechanisms operate in RR Lyrae variables and are the reason for the pulsations we observe.

A second type of stellar variability relevant to the present analysis is asteroseismic variability. A detailed exposition of this kind of variability is beyond the scope of this work, and readers are referred to Aerts et al. 2010 for a general reference. This variability is due to the (typically small) oscillations of stars in the various natural frequencies and modes related to their internal structures. These modes come in two dominant flavors that are most relevant for the stars in this work: \( p \) modes and \( g \) modes. The former, \( p \) modes (“\( p \)” standing for “pressure”) are simply pressure-driven sound waves that occur in the star, the same as the sound waves we can hear. The
latter, g modes (“g” standing for “gravity”) are waves driven by buoyancy forces, the same as waves on the ocean. These asteroseismic oscillations lead to observable periodic brightness changes in the stars as they expand and contract. The oscillations of stars in these modes are directly tied to the density, temperature, and other physical aspects of the star, so measurements of asteroseismic variability permit probing of the internal structure of the stars. As the interiors of stars are otherwise obscured from view, such measurements are extremely valuable in understanding stellar interiors and calibrating stellar models. For the *Kepler* space telescope, discussed in the next section, the observation timescales are conducive for observing asteroseismology of giant stars, particularly red giants. Because of their densities, the relatively smaller hydrogen-burning stars that are a majority of the stars in the universe, including the Sun, have oscillation periods that are typically too short to be observed with *Kepler’s* observation averaging time of 30 minutes, but the lower densities of giant stars put them in the right period range. The oscillations of red giants stars are driven by the convection that occurs near the surface of such stars.

A third type of stellar variability is that due to eclipses between a star and another object. A common example of this is a binary star system that is oriented such that we see one or both of the stars periodically cross in front of each other (see Kallrath & Milone 2009 for a general reference). These stars are too distant for us to resolve them as separate light sources. Instead, what we can see are periodic dips in the combined light coming from these stars as they cross in front of each other and take turns blocking (some or all of) each other’s light. These eclipsing binary star systems are known simply as “eclipsing binaries.” There are different kinds of eclipsing binaries and I focus on two here. The first, detached eclipsing binaries, are systems where the stars orbit far enough apart that they are not touching and the eclipses are seen as distinct dimmings from the combined brightness level of the two stars. The second, contact eclipsing binaries, are binary stars that are so close together that not only do
their tidal forces create large bulges on each other, which create their own brightness variations as the bulges spin in and out of view, but the bulges are so large and the orbital separations so small that the stars are actually in contact with each other. These produce quasi-sinusoidal brightness variations in time.

1.3 The *Kepler* Space Telescope

The *Kepler* Space Telescope (Borucki et al. 2010) and associated mission was a NASA Discovery-class mission launched in March of 2009. Its primary mission was to stare at a 115 square degree portion of the sky (an area around one of the wings of the constellation Cygnus, the swan) and take a extensive series of images of that part of the sky, recording an image about every 30 minutes. The purpose of these images was to make a record of the brightnesses of the stars over time, with the primary intent to look for periodic small dips in the stellar brightnesses caused by planets orbiting those stars. As planets orbit distant stars, if the orbit is oriented correctly, the planet will partially eclipse the star and block some of the light, similar to the eclipsing binaries described above. This method of planet discovery was not new at the time of *Kepler*’s launch. HD 209458 b, initially discovered through the radial velocity technique, provided the first example of a detectable exoplanet transit (Charbonneau et al. 2000; Henry et al. 2000), and OGLE-TR-56b (Udalski et al. 2002) was the first planet discovered by its transits.

The *Kepler* mission has been a huge success, with 2345 of the 4016 planets recorded at the NASA Exoplanet Archive as of the time of this writing (2019 July 23) listing *Kepler* as their discovery facility and 382 listing K2 (Howell et al., 2014), the second phase of the *Kepler* mission. *Kepler* has given us our first handle on the occurrence frequency of Earth-like planets in Earth-like orbits. Petigura et al. (2013) used results

---

1This technique measures the velocity of a star to look for periodic changes that are due to a planet’s gravitational tug on its host star, which changes the star’s velocity.
from *Kepler* to conclude that 3.5–7.4% of Sun-like stars host an Earth-like planet (defined as having a radius between 1 and 2 times the radius of Earth) with an orbital period between 200 and 400 days.

Unfortunately, about four years into the mission, *Kepler* broke. Of the four reaction wheels it had to help it maintain its fine pointing stability, it needed three to maintain the stability. One broke about three years into the mission and a second broke four years into the mission. A clever engineering solution was determined as a workaround for this pointing problem, and the mission was reborn as K2. The two remaining reaction wheels maintained pointing about the pitch and yaw axes, and roll stability was maintained via periodic thruster firings, though not to as fine a degree as with a working reaction wheel. To help maintain roll stability, the telescope was pointed such that the net torque from the solar pressure was minimized to reduce the amount of roll the spacecraft experienced. This required pointing the telescope at a new field of view approximately every three months to keep sunlight from entering the telescope. The reduced pointing stability and the need to repoint the telescope every three months reduced the telescope’s planet discovery rate but opened up many other avenues of interesting science. The K2 mission ended on 2018 October 30 after the spacecraft ran out of fuel.

The second campaign of the K2 mission, occurring between 2014 August 23 and 2014 November 10, included as part of its observations the GC M4. These observations represent the longest continuous observation of a GC to date. This dissertation presents the first detailed look at these data and several discoveries that have arisen from the analysis. Of note is a new class of variable star that is likely related to RR Lyrae variability and the presentation of transiting planet candidates, potentially the first transiting exoplanets known in a GC.
Chapter 2

Ultralow-amplitude RR Lyrae Stars in M4

2.1 Abstract

We report evidence for a new class of variable star, which we dub millimagnitude RR Lyrae (mmRR). From K2 observations of the globular cluster M4, we find that out of 24 horizontal branch (HB) stars not previously known to be RR Lyrae variables, two show photometric variability with periods and shapes consistent with those of first overtone RR Lyrae variables. The variability of these two stars, however, have amplitudes of only one part in a thousand, which is $\sim$200 times smaller than for any RR Lyrae variable in the cluster, and much smaller than any known RR Lyrae variable generally. The periods and amplitudes are: 0.33190704 d with 1.0 mmag amplitude and 0.31673414 d with 0.3 mmag amplitude. The strange RR Lyrae predicted by Buchler & Kolláth (2001) match these variables in amplitude but not in period. The stars lie just outside the instability strip, one blueward and one redward. The star redward of the instability strip also exhibits significant multi-periodic variability at lower frequencies. We examine potential blend scenarios and argue that they are all
either physically implausible or highly improbable. Stars such as these are likely to shed valuable light on many aspects of stellar physics, including the mechanism(s) that set amplitudes of RR Lyrae variables.

2.2 Introduction

RR Lyrae stars are valuable astronomical tools. They are used as standard candles, and to measure the helium abundance of stars in GCs. Space-based monitoring of RR Lyrae variables by missions such as MOST (Walker et al., 2003), CoRoT (Baglin & COROT Team, 1998), and Kepler/K2 (Howell et al., 2014) has revealed new information on these objects. For example, Kepler has revealed additional, low-amplitude oscillation modes in fundamental mode (RRAB) RR Lyrae variables (Molnár et al., 2012), including RR Lyr itself (Benkó et al., 2010). See Molnár (2018) for a more complete list of these discoveries.

As part of continuing efforts to observe RR Lyrae stars, the GC M4 (NGC 6121) was observed by Kepler/K2\(^1\) in 2014 during its Campaign 2 using a large superstamp that contained thousands of stars. This and other K2 observations of GCs are the longest continuous photometric surveys of populations of GC stars, monitored at the high precision that has been Kepler’s hallmark. As part of our analysis of these data, we have discovered two HB stars just outside the instability strip that have photometric variations similar to first overtone RR Lyrae (RRC) pulsators but with an amplitude \(\sim\)200 times lower than the typical lowest amplitude RRCs. We tentatively give these stars the name “millimagnitude RR Lyrae”, or “mmRR” for short. The two stars are Gaia DR2 6045466571386703360 (mmRR 1) and Gaia DR2 6045478558624847488 (mmRR 2). There is no previously identified variable class that matches the properties of these stars, and if their variability is associated with RRC

\(^1\)Funding for the Kepler mission is provided by the NASA Science Mission Directorate.
variability, then they would be by far the lowest amplitude RR Lyrae variables yet
discovered. Previous RR Lyrae searches would likely have been unable to find such
low-amplitude objects, so it is not surprising that they are only now being discovered
by K2.

2.3 Observations and Analysis

2.3.1 K2 Image Subtraction, Reduction, and Variable Search

Our light curve extraction pipeline is very similar to the image subtraction pipeline
of Soares-Furtado et al. (2017). The specific pipeline, briefly described here, will
receive a full description in Chapter 3. We downloaded the 16 target pixel files
(K2 IDs 200004370–200004385) that make up the M4 superstamp from the Mikulski Archive for Space Telescopes\(^2\) and stitched them together using \texttt{k2mosaic} (Barentsen, 2016), producing a total of 3856 images. We removed images that were blank or that otherwise would produce low quality photometry (usually due to excessive drift) and were left with 3724 images covering \(\sim78\) days. We reduced these images to a set of registered, subtracted images using tools from the \texttt{FITSH} software package (Pál, 2012).

We used the \textit{Gaia}\(^3\) DR1 source catalog (Gaia Collaboration et al., 2016b,a) as both
an astrometric (Lindegren et al., 2016) and photometric (van Leeuwen et al., 2017)

\(^2\)The Mikulski Archive for Space Telescopes (MAST) data archive is at the Space Telescope Science Institute (STScI). Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NAG5-7584 and by other grants and contracts. STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 526555.

\(^3\)Data from the European Space Agency (ESA) mission \textit{Gaia} (https://www.cosmos.esa.int/gaia) is processed by the \textit{Gaia} Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the \textit{Gaia} Multilateral Agreement.
reference catalog. DR1 was used instead of DR2 because our analysis began prior to DR2’s release. A conversion between Gaia magnitude $G$ and Kepler magnitude $K_p$ was determined, which, owing to the similar bandpasses of the two telescopes, was purely linear. The converted $G$ magnitudes were used as reference magnitudes for performing image subtraction photometry on the subtracted images, using fphot from FITSH and a series of aperture sizes. The aperture used for a given magnitude was determined by calculating the RMS scatter of the final light curves and finding the aperture that had the lowest median RMS value in half-magnitude bins.

The light curves suffered from residual systematic variations due to the roll of the spacecraft. We performed a decorrelation of the measured photometry against the telescope roll using the process described by Vanderburg & Johnson (2014) and Vanderburg et al. (2016). As part of the decorrelation, a B-spline was also fit to the data with breakpoints set every 1.5 days and removed from the data. The VARTOOLS implementation (Hartman & Bakos, 2016) of the trend filtering algorithm (TFA; Kovács et al., 2005) was then used to further clean up global trends in the final photometry.

Light curves were obtained for 4600 Gaia DR1 sources, which were searched for variability using the Generalized Lomb–Scargle (GLS; Lomb, 1976; Scargle, 1982; Zechmeister & Kürster, 2009), phase dispersion minimization (Stellingwerf, 1978), box least squares (Kovács et al., 2002), and auto-correlation function (McQuillan et al., 2013) methods as implemented in astrobase (Bhatti et al., 2017). The results from these methods were searched by eye for significant variability.

### 2.3.2 The Horizontal Branch Stars

To determine cluster membership, we used Gaia DR2 proper motion measurements (Lindegren et al., 2018a) to determine cluster membership. The proper motion of M4 ($\mu_\alpha = -12.5$ mas/yr, $\mu_\delta = -19.0$ mas/yr) is well separated from that of the field
Table 2.1. Data on mmRRs

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>W2015</td>
<td>13.212</td>
<td>16:23:32.30</td>
<td>$-26:28:53.5$</td>
<td>0.33190704</td>
<td>1.0</td>
<td>2059.57</td>
</tr>
<tr>
<td>W2386</td>
<td>13.047</td>
<td>16:23:35.93</td>
<td>$-26:26:20.9$</td>
<td>0.31673414</td>
<td>0.3</td>
<td>2059.47</td>
</tr>
</tbody>
</table>

Note. — ID is the identifier for the object as used in Table 3.1. W2015 is Gaia DR2 6045466571386703360 and W2386 is Gaia DR2 6045478558624847488. W2015 is also referred to in this work as mmRR 1 and W2386 is also referred to as mmRR2. Magnitude and position information from Gaia DR2. Epoch is the time of maximum brightness, expressed in Kepler BJD, which is BJD-2454833.0.

population. We used scikit-learn (Pedregosa et al., 2011) to fit a two-component Gaussian mixture model (GMM) to the proper motion measurements of all Gaia DR2 sources within 30′ of the cluster center with reported proper motions (for full details, see Wallace 2018b or Appendix A).

This chapter presents results of a variability search among the 34 HB stars for which we had light curves. HB stars were selected to be those with $14.3 < G_{BP} < 13.0$ and $G_{BP} - G_{RP} < 1.5$ and a >95% cluster membership probability. Of these, 10 were previously identified as RR Lyrae variables (Clement et al., 2001). Of the other 24 HB stars, we identified two low-amplitude variables with $\lesssim 1$ mmag amplitude sinusoidal variability and periods of $\sim 0.3$ day that fell outside the locus of identified RR Lyrae stars. Table 2.1 contains some information on these objects and Figure 2.1 shows their light curves (full and phase-folded) and associated GLS periodograms. The light curves are published online4. The periods are consistent with RR Lyrae variability, and the light curve shapes—in particular the possible notches just before maximum brightness for both stars—are similar to RRCs. The amplitudes, however, are much smaller than any known RRC, which have amplitudes of $\sim 200–350$ mmag. The

Figure 2.1: Light curves and periodograms for the two variable stars. Panels A, B, and C correspond to mmRR 1 and panels D, E, and F to mmRR 2. Panels A and D show the GLS normalized power (N.P.) spectra with the three highest peaks in each case labeled. Panels B and E show the phased light curves of each star, each folded at the GLS period with the highest peak. Panels C and F show the full light curves for each star. Gray points show individual measurements and blue points show binned-median values. All light curves have their median magnitudes subtracted off (mmRR 1: 13.14153, mmRR 2: 12.91269). We note possible notches in both phase-folded light curves just prior to maximum brightness.
variability search for mmRR 1 detects only this sinusoid variability and its harmonics and aliases, while mmRR 2 shows low-amplitude variability at a number of longer periods as well. The positions of these stars in the color–magnitude diagram (CMD) are shown in Figure 2.2. Star mmRR 1 is blueward of the locus of RR Lyrae stars and mmRR 2 is redward. Several of the other stars redward of the full-amplitude RR Lyrae variables show low-amplitude variability at multiple periods in the approximate range 0.3–5 day.

We note a possible third star of interest, *Gaia* DR2 6045489283174903168 (G3168), which has a $\sim 0.64$ day sinusoidal period and $\sim 0.5$ mmag amplitude and is in the locus of RR Lyrae variables (marked in Figure 2.2 with a black cross). We do not include it as an mmRR because it has stronger variability than mmRR 2 at other periods (for example, a sinusoid variability at 1.67 day period of slightly smaller amplitude than the 0.64 day signal). We will further discuss this and the other M4 variables in Chapter 3.

### 2.3.3 Blend Scenarios

Figure 2.3 shows images of the two mmRRs, with nearby *Gaia* DR2 sources marked. The aperture used for photometry extraction is indicated, which has a radius of 2.25 Kepler pixels, or $\sim 9\,''$. Both the apertures used and the individual K2 pixels that these stars lie on are significantly blended. We focused our blend analysis on mmRR 1 due to its higher signal-to-noise ratio, but many of our conclusions extend to mmRR 2.

We searched for variability among the blended sources by using an array of 0.51 pixel radius apertures on and around mmRR 1 to obtain focused photometry of the blended objects from the K2 data. This photometry underwent the same roll decorrelation previously described but not TFA cleaning. We then searched for variability at a period and flux amplitude matching the aperture centered on mmRR 1. Three apertures had a corresponding variability: the one centered on mmRR 1, and the two
Figure 2.2: *Gaia* DR2 CMDs for M4, with $G_{RP}$ and $G_{BP}$ data taken from *Gaia* DR2 (Riello et al., 2018). Only objects with membership probabilities greater than 95% are included. Left panel: red x’s mark stars previously identified as RR Lyrae in the catalog of Clement et al. (2001), June 2016 edition. The blue star marks mmRR 1 and the green mmRR 2. The blue cross marks a particular star blended with mmRR 1 (Blend 1). Right panel: zoom-in of the portion of the left panel delineated by the dashed lines. RR Lyrae variables are differentiated by subclass: light red for RRABs and dark red for RRCs, as indicated. For this panel, sources for which we have a K2 light curve are marked in black instead of gray. Star G3168 is marked with a black cross. Our $G_{BP} - G_{RP} < 1.5$ cut for investigated objects is shown with a vertical dashed line.
Figure 2.3: Images of the two variable stars. Crosses mark Gaia DR2 source positions. Blend 1 is additionally marked with a red x. The blue circles mark the size of aperture used for photometry extraction from the K2 data. The image for mmRR 1 is from the M4 reference image of Kaluzny et al. (2013a) and is rotated slightly relative to the Gaia source positions. The image for mmRR 2 (saturated in this image) was taken with a Sinistro detector on an LCOGT 1 m telescope operated by Las Cumbres Observatory. The pixels in all three images are expressed in a logarithmic scale.
apertures located 0.51 pixels left and right roughly along the x-axis of the image in Figure 2.3. We concluded the variability source could only be mmRR 1, *Gaia* DR2 6045466571377393792 (Blend 1, marked in Figure 2.2 with a red x), or an unresolved blended source.

Blend 1’s location (blue cross in Figure 2.2) in the CMD is unusual, particularly given its >99% probability of cluster membership. The *Gaia* detector windows to measure $G_{\text{BP}}$ and $G_{\text{RP}}$ are $2''.1 \times 3''.5$ (Arenou et al., 2018), so it is possible that the color measurements are significantly blended with mmRR 1, perhaps inhomogeneously between the two filters for it to appear bluer than mmRR 1. It is also possible that Blend 1 is a subdwarf B (sdB) star or a white dwarf (WD) blended with a main sequence (MS) star. We were unable to determine any physical MS–WD combination that matched the measured color and magnitude for this object. Many sdB stars are variable, but none in a way to match the variability seen for this object (which, given the $G\approx18$ magnitude of this object, would need to have a $\sim0.1$ magnitude amplitude). All known types of sdB variables are some combination of too short of period, too small of amplitude, or too incoherent of pulsations to explain the variability (Catelan & Smith, 2015, chapter 12). We also were unable to find any ellipsoidal variability of an sdB–MS binary that provided the necessary variability amplitude (the highest unblended amplitudes obtained were $\sim0.01$ mag).

If the color measurements are in error and this is an MS star, the only variability scenarios that could match the observed shape and period are rapid rotation of a heavily spotted star or ellipsoidal variability. We were unable to find any physically plausible ellipsoidal variability scenarios that matched the observed variability and $G$-band magnitude. To estimate the probability of blending with a heavily spotted fast rotator, we looked through the light curves for all objects with $G>15$ and found five objects with periods less than one day and sinusoidal variability of roughly appropriate amplitude when blended with an HB object. The search field was $\sim149$ square
arcminutes. Since, from the aperture analysis, we know the blend must be within about a Kepler pixel radius (≈4′′), the probability of one of these objects blending with mmRR 1 is ∼5×10⁻⁴. The probability of finding two chance alignments out of 24 targets is very small at ∼7×10⁻⁵.

Returning to mmRR 1, the orbital separation needed for a ∼0.66 day binary orbit including mmRR 1 is ∼3–4 R⊙. Gaia DR2 (Andrae et al., 2018) measures the radius of mmRR 1 to be 2.8–4.1 R⊙ (16th–84th percentiles). We used PHOEBE (Prša & Zwitter, 2005) to examine contact binary scenarios and could find no physical scenario with the radius of mmRR 1 being larger than ∼2.4 R⊙, the Roche limit. If the radius of mmRR 1 is indeed exceptionally small to allow a contact binary scenario, only companions with masses between 0.08 M⊙ (with a face-on orbit) and ∼1 Jupiter mass (with inclination ≲45°) could produce millimagnitude amplitudes. Given the even larger radius that mmRR 1 would have had when on the red giant branch, such a system would be a post-common-envelope-binary. Approximately one third of WDs are known to have short-period post-common-envelope binary companions, with the majority having secondary stars of mass less than 0.25 M⊙ (Schreiber et al., 2010). While we are unaware of estimates for the occurrence rate of such systems with HB primary stars, the possibility of an HB star having a low-mass contact binary companion cannot be dismissed out of hand. However, finding two of these systems on low inclination orbits (which are less likely than higher inclinations assuming random orientations), without also finding systems on higher inclination (and thus higher photometric amplitude) orbits is unlikely. Moreover, the inconsistency between the measured stellar radius from Gaia and the upper limit on the radius for a contact binary is strong evidence that this scenario does not explain the observations.

Finally, we consider an undetected background RRC or short-period Cepheid variable. An RRC would need to be ∼200–350 times dimmer than mmRR 1 to get a millimagnitude blended amplitude. With mmRR 1 having G=13.23, the background
RRC would need to have $G \approx 18.9$–19.6. We used the Gaia DR2 RR Lyrae variable catalog (Holl et al., 2018; Clementini et al., 2019) to determine the surface density of RR Lyrae variables with $G$ magnitudes in the appropriate range in the field near M4, finding $\sim 2$ RR Lyrae variables per 0.7 square degrees. Mirroring our estimation of rapidly rotating spotted star blending, we get a blend probability of $\sim 6 \times 10^{-6}$. The probability of finding two chance alignments out of 24 targets is vanishingly small at $\sim 1 \times 10^{-8}$. Even if the catalog of RR Lyrae we are using has a completeness as low as 15% as it does in the Galactic Bulge (Holl et al., 2018, Table 3), the probability of chance alignment is still minuscule. There are even fewer background Cepheid variables (Gaia detected none the areas we searched for RR Lyrae variables) and they typically have much longer periods, so the probability of blending with a background Cepheid is even smaller.

The signal-to-noise ratio for mmRR 2 was not high enough for our small aperture array to disentangle specific possible sources of the variability. We note, however, that all of the Gaia DR2 sources within 5″ of mmRR 2 are proper motion members of the cluster. Because of this, arguments similar to those for the possible blend scenarios of mmRR 1 and Blend 1 prevail. We note that mmRR 2 has a relatively large radius in Gaia DR2 (7.8–8.3 R$_\odot$), which would make it impossible to host a binary object at a $\sim 0.63$ day period orbit. We also checked that the periods of the three variables do not match any previously identified RR Lyrae star in the cluster, nor do they match any other variable signal found in our light curves from the M4 superstamp. Gaia detects no variables within 10″ of the two mmRRs.

### 2.4 Discussion

From the evidence presented, we conclude that the most likely explanation for the observed variability is a previously unreported kind of stellar variability that, based
Figure 2.4: Periods and amplitudes of the two mmRRs and the RR Lyrae stars in M4. As in Figure 2.2, the blue star marks mmRR 1 and the green marks mmRR 2. RRABs are shown as black triangles and RRCs are shown as gray squares. The data for RR Lyrae variables are from Clement et al. (2001). The two new variables have much lower amplitudes than any RR Lyrae star in the cluster.

on the locations in the CMD and variability periods and shapes, is possibly related to RR Lyrae variability. Figure 2.4 shows the periods and amplitudes of the mmRRs relative to the RR Lyrae variables in M4. Their amplitudes (mmRR 1: 1.0 mmag, mmRR 2: 0.3 mmag) are much lower than any previously observed RR Lyrae star, which have amplitudes of ∼200 mmag and greater.

We note here the works of Buchler et al. (2005) and Buchler et al. (2009), who used data from the MACHO and OGLE databases to find ∼30 objects near the Cepheid instability strip of the LMC with amplitudes ≲0.01 mag. At least ∼20 of these objects are members of the LMC. These match the predicted strange Cepheids of Buchler et al. (1997). These mmRRs may be the very similar strange RR Lyrae predicted by Buchler & Kolláth (2001). The strange modes of Cepheids and RR Lyrae are due their hydrogen partial ionization zones, which create a large potential barrier.
close to the surface of the star and can cause pulsation modes with nodes located at this potential barrier to pulsate in the outer portion of the star. The amplitudes, shapes, and CMD locations of the mmRRs match the predictions, but the periods (which would be coming from the 8th-10th radial overtones based on the calculations of Buchler & Kolláth 2001) are longer than predicted.

We also note once again G3168, the possible third mmRR we found, as well as the other HB stars redward of the known RR Lyrae stars that had multi-periodic photometric variability of periods of approximately 0.3–5 days. These stars are perhaps connected to the mmRRs and will be described more completely in Chapter 3.

If these objects do represent a new class of variability, why have no similar objects been discovered previously? As mentioned in Section 2.2, Kepler/K2 has enabled discovery of very small amplitude modes in RR Lyrae variables, seemingly commonplace yet undetected in over a century of observations of these stars. The mmRRs appear to share a similar story. We make particular mention of RR Lyr, an RRAB, which has been shown by Kepler to have small amplitude first overtone pulsations (Molnár et al., 2012), a phenomenon perhaps connected to these mmRRs. Finally, theoretical work indicates that convection and viscous damping are the likely physical process that set the amplitudes of RR Lyrae variables (Kolláth et al., 1998; Smolec & Moskalik, 2008; Geroux & Deupree, 2013); mmRRs could be valuable in further developing this understanding.

The work in this chapter utilized the following software: astrobase (Bhatti et al., 2017), astropy (Astropy Collaboration et al., 2018), FITSH (Pál, 2012), k2mosaic (Barentsen, 2016), matplotlib (Hunter, 2007), numpy (Oliphant, 2006), PHOEBE 1.0 (Prša & Zwitter, 2005), scikit-learn (Pedregosa et al., 2011), scipy (Jones et al., 2001), VARTOOLS (Hartman & Bakos, 2016).
Chapter 3

A Search for Variable Stars in the Globular Cluster M4 with K2

3.1 Abstract

We extract light curves for 4554 objects with $9 < G$ (Gaia) $< 19$ in the K2 superstamp observations of the globular cluster M4, including 3784 cluster members, and search for variability. Among cluster-member objects, we detect 66 variables, of which 52 are new discoveries. Among objects not belonging to the cluster, we detect 24 variables, of which 20 are new discoveries. We additionally discover 57 cluster-member suspected variables, 10 cluster-nonmember suspected variables, and four variables with ambiguous cluster membership. Our light curves reach sub-millimagnitude per-point precision (at a cadence of 1765.5 s) for the cluster horizontal branch, permitting us to detect asteroseismic activity in six horizontal branch stars outside the instability strip and one inside the strip but with only $\sim$1 mmag amplitude variability. 19 additional stars along the red giant branch also have detected asteroseismic variability. Several eclipsing binaries are found in the cluster, including a 4.6 day detached eclipsing binary and a W Ursae Majoris eclipsing binary (EW), as well as an EW with uncertain
cluster membership and three other candidate EWs. A 22 day detached eclipsing binary is also found outside the cluster. We identify a candidate X-ray binary that is a cluster member with quiescent and periodic \( \sim 20 \) mmag optical variability. We also obtain high-precision light curves for ten of the previously known RR Lyrae variables in the cluster and identify one as a candidate Blazhko variable with a Blazhko period in excess of 78 days. We make our light curves publicly available.

### 3.2 Introduction

The GC M4 (NGC 6121), located in the constellation Scorpius, is the closest GC to Earth at a distance of \( \sim 1.8 \) kpc (Kaluzny et al. 2013b; Braga et al. 2015; Neeley et al. 2015). M4 is an old GC, with recent age measurements falling between \( \sim 11–12 \) Gyr (Bedin et al. 2009; Kaluzny et al. 2013b; VandenBerg et al. 2013), and it has a metallicity of \([\text{Fe/H}] \approx -1.2\) (Harris 1996, 2010 edition). Given its relative proximity to us and also the relative sparseness of its core, M4 is a prime target for the detailed study of individual GC member stars.

M4 is rich in variable objects—90 in the current count of Clement et al. (2001), 2016 June edition—such as pulsating variables (including dozens of RR Lyrae variables), eclipsing binaries, and cataclysmic variables (Clement et al. 2001; Bassa et al. 2004; Kaluzny et al. 2013a,b; Stetson et al. 2014; Samus et al. 2017; Watson et al. 2017 and references therein). Some recent examples of the scientific utility of these variables include using RR Lyrae variables for an M4 distance determination (e.g., Braga et al. 2015) and using M4 eclipsing binaries to provide constraints on the mechanism of formation of close binaries in GCs (Kaluzny et al. 2013a). Given the large number of variable objects already known in M4 and the scientific impact of both better understanding known variables and discovering new ones, any data that allows variable discovery and characterization are of great value.
M4 was in the field of view of the *Kepler* telescope during Campaign 2 (running from 2014 August 23 to 2014 November 10) of the K2 mission (Howell et al., 2014), and continuous observations of a portion of this cluster in the form of a “superstamp” were included in the data downloaded from the observatory. These and other K2 observations of GCs represent, by far, the longest continuous observations of GCs to date, and in the case of M4, the longest continuous observation of what happens to be the closest GC. Additionally, these observations were taken by a space-based observatory designed and built with high-precision photometry as its goal. This is a prime data set for an object of great scientific interest and will likely be the best time series data we have for a GC for a while to come.

Unfortunately, *Kepler’s* design was not optimized for observing GCs. Its 3′′.98/pixel spatial scale leads to significant blending in the images, particularly close to the core. Fortunately, techniques exist to partially mitigate the effects of the blending, and given the expected richness and value of the derived light curves, the effort to work through these issues is still worthwhile. This chapter uses image subtraction (Alard & Lupton, 1998) among other techniques to deal with the blending, and, building off of Chapter 2 (published as Wallace et al. 2019b), it is, as far as we are aware, the first general analysis of the K2 observations of a GC. Previous work on these images is limited in number and scope: Miglio et al. (2016) looked at asteroseismic oscillations in K giants, and Kuehn et al. (2017) looked at the RR Lyrae variables. The results from these limited searches demonstrate the incredible potential of the M4 K2 superstamp data. This work is focused more on breadth (production of quality light curves and identification of variables) rather than depth (full characterization of individual variable objects) and is a starting point for analysis of these data. We describe our methods to extract and analyze data from the images in Section 3.3, and in Section 3.4 we present the results of our variability search. A discussion is presented in Section 3.5, and we conclude in Section 3.6.
Figure 3.1: The astrometric reference image of the K2 superstamp of M4. The image is 300 pixels by 150 pixels, or approximately 20′ by 10′, and is displayed with arbitrary z-scale and colors inverted. The white regions in the upper-left and -right corners are regions that were not included in the superstamp. The core of the cluster is ∼1′ off of the bottom edge of the image.

3.3 Method

We present here a detailed description of our data reduction and variable identification pipeline.

3.3.1 Image Preparation

The images we used are the 16 target pixel files (TPFs) that make up the M4 superstamp from the Mikulski Archive for Space Telescopes. Each is 50 pixels by 50 pixels in dimension. These files had the K2 EPIC ID numbers 200004370 – 200004385. We stitched the TPFs together using k2mosaic (Barentsen, 2016), producing a series of images with dimensions of 150 pixels by 300 pixels, each missing two 50-pixel-by-50-pixel notches. These images were ∼10′ by ∼20′ on the sky. One of the images is shown in Figure 3.1. The superstamp is not centered on the cluster but rather avoids the cluster center and is focused more on the cluster outskirts on one side of
the cluster. A total of 3856 superstamp images are produced, one for each cadence. By mission design, 39 of the images had no data recorded as they took place during resaturation events (major thruster fires used to spin up the reaction wheels) that occurred every 96 cadences and were thus not usable in our analysis.

Our data extraction and reduction pipeline is very similar to that of Soares-Furtado et al. (2017). After assembling the superstamp images, we used the fistar tool from the open-source FITSH software package (Pál 2012) for source detection in preparation for image registration. We used an asymmetric Gaussian model for the point spread function (PSF), a detection threshold of 400 ADUs, the default uplink candidate extraction algorithm, and two symmetric and one general iterations. From this, we generated a list of source positions, fluxes, and PSF shape and width parameters for each detected source. The image with the smallest median PSF full width at half maximum (FWHM) across all the detected sources was chosen as the astrometric reference image. This smallest median FWHM was 1.457 pixels, and the collection of median FWHM values across the images had a mean of 1.503 pixels and a standard deviation of 0.018 pixels. The selected astrometric reference frame image—the 1197th cadence in the campaign, which is shown in Figure 3.1—also had one of the most symmetric FWHMs of all the images.

The grmatch tool from FITSH was then used to match the detected sources in each image to the similar list of sources on the selected astrometric reference image, and thus find the spatial transformation between the individual frames and the astrometric reference image. To determine the best parameters for the match, a grid was employed consisting of two different transformation orders (1 and 2) and many different values (170–500) for the maximum number of sources to select from the reference and image source lists (ordered by greatest flux to least) to use for the triangle matching. We ran the grmatch code for each image for all the parameters on this grid. For each image, we adopted the set of parameters that maximized the number of matched objects
normalized by the square of the weighted residual, subject to the restriction that at least 100 objects were matched, and that the match was accurate (i.e., the weighted residual reported by `grmatch` was less than 0.001 and the reported unitarity was greater than 0.015). The `FITSH` tool `fitrans` was then used to register each image to the frame of the astrometric reference image using the selected transformation calculated by `grmatch`.

After registering the images, the next task was to create a photometric reference image to use for image subtraction. For each image, the Euclidean distance (in pixels) of the transformation of a point at the center of the image to the astrometric reference image was calculated. The closeness of the PSF size and shape (as measured by the median S, D, and K parameters) of the image to the astrometric reference image was also calculated. Cutoff values for the transformation distance and the SDK closeness (respectively, 0.0998 pixels and 0.1) were selected such that there were 100 images chosen to be used in the creation of a photometric reference image. The chosen images were taken mostly during the first half of the campaign, which is unsurprising, considering the much larger drift in the second half of the campaign. These 100 images were then median combined using `ficombine` from `FITSH` to create the master photometric reference image.

### 3.3.2 Image Subtraction and Photometry Extraction

`FITSH`’s `ficonv` tool was then used to subtract the master photometric reference image from each of the K2 images. A first-order polynomial was fit to the background and also subtracted. A constant discrete convolution kernel with a half-size of 4 pixels was used to match the PSF and flux scale of the reference image to that of each individual K2 image. This unfortunately meant that objects that were within 4 pixels of the edge of the image (a little less than 1% of the image, referred to in this work as “the edge region”) were not included in the image subtraction calculation, and objects near
to the edge region with parts of their images were not included in the photometry
calculation. Nine isolated, relatively bright stars across the least crowded portions
of the superstamp (left, right, and upper portions) were selected by eye and used
to optimize the parameters of the background transformation and the convolution
kernel.

Ideally, what remains after the image subtraction (barring any uncorrected sys-
tematics and/or an incorrect background fit) is an image free of any non-variable
sources with random scatter about a statistical average of zero. Stars leave behind
larger magnitudes of scatter than the source-less background due to the intrinsic Pois-
son noise of these sources and the value of that noise being larger for brighter sources.
Saturated stars leave behind visible artifacts, which are due in part to the column
bleeding that occurs upon pixel saturation and the periodic drift of the telescope
moving the bleed columns around on the detector from cadence to cadence. Fig-
ure 3.2 shows the same image as Figure 3.1 after subtracting the master photometric
reference image as described above.

Extracting photometry from the subtracted images requires a catalog of source
positions as well as reference fluxes/magnitudes for each source to properly calibrate
the amplitude of the variable signals found in the subtracted images. We used the
Gaia first data release (DR1) source catalog (Gaia Collaboration et al. 2016a,b) as
both our astrometric (Lindegren et al. 2016) and photometric (van Leeuwen et al.
2017) reference catalog. At the time of the Gaia second data release (DR2; Brown
et al. 2018), our work had sufficiently progressed that we chose to stick with the Gaia
DR1 data despite DR2’s superior quality. That being said, data from Gaia DR2 were
used as part of our analysis (for example, its identification of duplicate DR1 sources).

The Gaia DR1 source catalog is virtually complete at the magnitude range of the
main-sequence turnoff stars in M4 (G≈16–17), and its excellent astrometry allows
for precise source position determination, and it aids in identifying and disentangling
Figure 3.2: Subtracted image for the image in Figure 3.1, with arbitrary z-scale of the same dynamic range as Figure 3.1 and no color inversion. The white regions in the upper-left and -right corners are the same as the white regions in Figure 3.1, regions that were not included in the superstamp. The RR Lyrae variables are of sufficiently large amplitude to be visible to the naked eye in the subtracted image: the black “holes” in the middle and in the upper right of the image are two RR Lyrae variables, as are the bright spots (i.e., no dark pixels in the star’s image) left and slightly down as well as right and down from the middle hole. Residual noise and saturation artifacts are visible. The 4 pixel border of zero-value pixels filling the edge region, as described in the text, is also present, and can be made out at the bottom of the image.
close neighbors that are impossible to differentiate in the K2 images. That being said, crowded regions limited Gaia’s completeness in both DR1 and DR2 (e.g., Gaia Collaboration et al. 2016a). These limitations in completeness correlate with crowdedness and are located in the most crowded regions of our images. Because of this, any star missing from our astrometric reference catalog is likely to appear in some other star’s photometric processing aperture, and so we decided to proceed despite the potential completeness issues. Kepler’s and Gaia’s bandpasses are also similar, which we found eliminated any need to derive more than an additive conversion from our instrumental magnitudes to Gaia G magnitudes.

From the Gaia DR1 archive, we extracted those sources that fell inside or near to the region of the M4 superstamp and had a $G$ magnitude brighter than 19. This cutoff does not go deep enough to cover all the stars in the cluster, nor does it go deep enough to cover the possible variable stars in the background, many of which may be sufficiently unblended in the images to detect variability. The choice of this magnitude cutoff was based on the photometric performance achieved by Soares-Furtado et al. (2017) and our initial goal to primarily search for transiting exoplanets rather than larger-amplitude variables. The right ascension and declination values obtained for the Gaia DR1 sources were projected onto a pixel-based image coordinate system and then matched using grmatch with the extracted sources of the selected astrometric reference image. The matching, similar to before, was performed over a grid of spatial orders (1 and 2) and number of objects (200–400), ordered by decreasing brightness, to include in the triangle matching. The Gaia DR1 sources were weighted as the second power of the flux. The best transformation was then chosen as the match with at least 100 matched objects, weighted residual less than 0.001, and unitarity greater than 0.015 that had the largest number of matched objects normalized by the square of the weighted residual. We found that a second order spatial fit performed the best. We then transformed the coordinates of the Gaia DR1 sources to the
astrometric reference frame using \texttt{grmatch} based on the transformation calculated above. After removing those sources with transformed coordinates that fell outside the astrometric reference image, there were 5914 sources. We refer to this as our source position catalog.

The next step was to calculate the photometry for each of the 5914 sources from the subtracted images. This required first deriving a conversion from the $G$ magnitudes of the photometric reference catalog to the instrumental magnitudes of the K2 images. To accomplish this, we first used the \texttt{FITSH} tool \texttt{fiphot} to obtain photometry from the master photometric reference image for a set of circular apertures, with 15 apertures ranging from 1.15 to 2.55 pixels. These radii were selected to obtain a good measure of how changing the aperture size affected the amount of flux measured. The median FWHM of the PSF across the images was $\sim$1.5 pixels, with the range 1.45–1.55 pixels covering nearly all the median PSF widths. The apertures were centered at each of the positions of the 1024 objects that had been directly matched between the \textit{Gaia} DR1 source catalog and the astrometric reference image. (Since \texttt{fiphot} found only 1073 sources in the images, probably due to the degree of blending, there were far fewer matched sources than the total available from the \textit{Gaia} DR1 source catalog.) For this calculation, the sky was subtracted based on the mode of pixel values in an annulus with inner radius of 17 pixels and outer radius of 30 pixels. A radius of 3 pixels around any source in the set of 1024 matched sources was excluded from this background calculation, and the pixel values were sigma clipped ($3\sigma$, two iterations) prior to the calculation.

After performing this reconnaissance photometry, we determined a transform from \textit{Gaia} $G$ to \textit{Kepler} instrumental magnitudes. As mentioned previously, we found that an additive transform was all that was needed for this conversion, likely because of the very similar bandpasses of the two instruments. Since there is significant blending of the sources in the K2 images, we first selected out those K2 sources for which
we thought there were negligible contributions from neighbors. Several unblended sources, as well as a few unsaturated bright sources for which any blending from neighbors would be small, were selected from the astrometric reference image by eye and were verified to be negligibly blended by using the Gaia DR1 source catalog. After this, sources with instrumental magnitudes in a narrow range around the transformed $G$ magnitudes (and thus presumably negligibly blended on the images) were selected and then fit to determine a more precise value for the additive constant. For all this, we used a 2.5 pixel radius aperture to calculate the instrumental magnitudes. Next, we determined the effect that changing the aperture size had on this conversion factor. For the brightest unblended and unsaturated stars, we normalized the fluxes calculated over a range of aperture sizes to the flux in the 2.5 pixel aperture and then determined the median-normalized flux for each aperture size across the selected stars. We then fit the integral of a Gaussian function to the median-normalized fluxes to determine a conversion from the flux at a given aperture size to that of the 2.5 pixel aperture.

Aperture photometry was then performed on the master photometric reference image for all the positions in the astrometric source catalog. As before, the sky background was calculated as the mode of pixels values in an annulus with inner radius of 17 pixels and outer radius of 30 pixels, with the same sigma clipping and source exclusion as before. The background was then subtracted. We performed the photometry calculation with apertures 1.5, 1.75, 2.0, 2.25, 2.5, 2.75, and 3.0 pixels in radius. Then, using the $G$ magnitudes from the photometric reference catalog, we substituted the reference fluxes for each object with the values determined from the converted $G$ magnitudes, additionally modified based on the aperture size. This provided reasonably accurate and unblended reference fluxes for each of the objects.

We then calculated the image subtraction photometry using fphot and the derived reference fluxes. The sky background, having been previously subtracted when
the subtracted images were created, was not fit in this step. We also used the same convolution kernels calculated for the creation of the subtracted images.

At this point, Kepler BJDs (KBBDs; BJD = 2454833.0) were assigned to each cadence for each object. Each of the original 16 TPFs was assigned only a single KBJD for each cadence, calculated along the center of the TPF. We assigned to each object the KBJDs from the TPF image in which it was found, corresponding to the location in the superstamp that the source position catalog (i.e., position derived from a transformation of the Gaia DR1 source catalog position to the astrometric reference frame) indicates. In reality, the KBJD will differ slightly for objects across each TPF from the value at the TPF center. We do not account for this. The greatest difference in KBJDs across the TPFs is only \( \sim 3 \) s (compared to the cadence of 29.4 minutes), with differences across individual TPFs being even smaller than this, so the timing error introduced this way for individual objects is negligibly small.

After the photometry calculation from the image subtraction, we obtained light curves for 4601 objects. The reason for the reduction from the original 5914 objects was that some were excessively blended with much brighter neighbors and were unable to have photometry measured, and that some of the objects fell in or excessively overlapped with the excluded edge region. The brightest stars (for cluster members, this corresponds to many of the giant stars) were saturated. We did not perform any special treatment of saturated stars, though, because they were so bright that the largest apertures employed in our processing (3 pixel radius) were used. Additionally, there was one previously known RR Lyrae variable, V27 of Clement et al. (2001), that was not a Gaia DR1 source and thus did not have a light curve calculated. We separately extracted a light curve for this star following the procedure described above and based on the transformed Gaia DR2 position for this object. The light curve for V27 did not undergo any of the following post-processing procedures since large-amplitude variables were not served well by the roll decorrelation, described in
Section 3.3.3. Including V27, we produced 4602 light curves in total. The light curves at this stage are what we refer to as the “raw light curves” throughout the rest of this work. All of our raw and processed light curves are published and publicly available online\(^1\), with a citation of Wallace et al. (2019a).

### 3.3.3 Photometry Post-processing

The roll of the telescope during the K2 mission introduced systematic variations to the brightness of objects as they moved across the detector (Howell et al., 2014). This is due to differences in pixel sensitivity unaccounted for in the K2 data reduction. These brightness variations are correlated with the object position on the detector and are not fully corrected by the image subtraction photometry. The remaining systematic variations can be decreased by performing a decorrelation of flux variations against object position with a procedure based on Vanderburg & Johnson (2014) and Vanderburg et al. (2016). We divided the light curves, normalized to their median values, into the same eight time chunks as Vanderburg et al. (2016) did for Campaign 2 (A. Vanderburg 2017, private communication). To determine the drift position of each object, the positions in the source position catalog were transformed for each cadence using the inverse of the transformation originally used to register each cadence’s image to the astrometric reference frame. Since the drift of objects across the detector was primarily in one direction, for each object we used a principal components analysis (PCA) to determine this primary direction of drift. The object positions for each cadence were transformed to the axes defined by the PCA, and then a fifth-order polynomial was fit to the positions. Each object’s drift’s arc length along the polynomial at each cadence was calculated and stored for later decorrelation.

\(^1\)Published at Princeton University’s DataSpace and licensed under a Creative Commons Attribution 4.0 International License, accessible via the permanent URL http://arks.princeton.edu/ark:/88435/dsp01h415pd368
For each time chunk, we iterated over fitting long-term trends with a B-spline fit and decorrelating against the roll. For the B-spline, we had breakpoints set nominally every 1.5 days. The 1.5 day breakpoint spacing was adjusted to allow for knots to be distributed evenly across the time chunk. Also, where possible, 0.75 days from adjacent time chunks were included to improve the smoothness and accuracy of the spline fit across time-chunk boundaries. We then excluded $3\sigma$ outliers to the B-spline fit, refit the spline, and repeated this until no outliers remained to be removed. The median-normalized light curve was then divided by the spline fit. The spline fit is not ever reintroduced into our light curves, so smoothly varying signals with timescales longer than the 1.5 day knot placement are likely to either be altered or removed. Objects with such signals are best studied from our data using the raw light curves.

After this, the fluxes of each chunk of the light curve were binned into 15 bins in arc length. $3\sigma$ in flux outliers were excluded in each bin, and then a linear interpolation was made using the mean flux values of each non-empty bin. In cases where bins had only a single point, an interpolation between adjacent bins was made. If there was a single point in the last bin (usually corresponding to outliers in the pointing), no fit was made for that point. The light curve was then divided by this interpolation. This process of fitting a spline to the longer trends and decorrelating against position was repeated eight times or until convergence, whichever came first. We selected eight to be the maximum number of times because we found those light curves that required more than eight iterations were usually oscillating between two close fits to the data that were not quite close enough to be counted as converging. If less than 10 points were in a time chunk, the decorrelation against drift position was not performed.

The trend filtering algorithm (TFA; Kovács et al. 2005) as implemented in VARTOOLS (Hartman & Bakos 2016) was then used to clean up systematics common across the light curves. For each aperture, 250 light curves with at least 97% of the maximum number of light curves points were selected from uniform bins of source position and
magnitude to be used as the trend light curves. For light curves with less than 2500 points, a subset of the selected 250 trend light curves was used in the detrending, with the number of selected trend light curves being close to but less than 10% the number of light-curve points. Since the KBJDs for a given observation differed slightly depending on which TPF an object was located (see Section 3.3.2) and common instrumental effects were likely correlated based on actual observation time than KBJD, detrending was performed based on cadence number rather than KBJD. Light curves from stars that were known to be RR Lyrae variables or saturated were not included as potential TFA template trend light curves. All light curves were then detrended against the trend light curves for the given aperture size with trend light curves excluded from the detrending if they were closer than 6 pixels, which is ∼4 FWHMs. The light curves that resulted were the ones used in our variability search and are referred to in this work as “final light curves.” Figure 3.3 shows the root-mean-square (RMS) scatter of the sigma-clipped (3σ clipping, iterated three times) final light curves for those objects included in our variability search. Owing to significant outlier points in our final light curves, outlier removal was necessary for our subsequent period search. These outliers seem to be due to still-uncorrected systematics, the worst of which occurred when the telescope changed its roll direction about halfway through the campaign.

The photometric performance displayed in Figure 3.3 shows that our sigma-clipped light curves are able to reach millimagnitude RMS scatter down to $G\approx15$, and 0.01 mag RMS scatter down to $G\approx18$. There is a large envelope of points with significantly larger scatter than is typical for objects of their magnitude. Some of these are variable stars, while the rest have excessive scatter due to the amount of blending present in the images or also possibly due to breakdowns of the photometric processing for individual objects. We also note that our saturated giant/bright foreground stars do not have significantly larger scatter than, e.g., our HB stars at $G\approx13$. The point at
Figure 3.3: RMS scatter as a function of Gaia DR1 G magnitude for our final light curves. The RMS is calculated from our magnitude light curves, which have been sigma clipped with 3σ clipping iterated three times. All of our final 4554 objects under consideration except V27 are plotted here; V27’s light curve did not undergo the same processing as those of the other objects (see text for details). The solid line shows a calculation of our expected RMS scatter and the dashed line shows that same calculation reduced by a factor of three (see text for a discussion). The collection of objects with excessive RMS values at \(G\approx13\) are the RR Lyrae variables, though we note that our light-curve processing pipeline impacted the amplitudes of large-amplitude variables.
$G \approx 9.5$ is a star that is an intrinsic variable, hence the larger scatter. The clump of points with high RMS scatter at $G \approx 13$ are the RR Lyrae variables.

The solid line in Figure 3.3 shows our expected RMS performance based on source Poisson noise and the background sky flux as seen in our photometric reference image, and the dotted line shows the same expected RMS performance reduced by a factor of three. We have not entirely determined the reasons for our photometric performance to fall as far below our expected performance as it does, but it is perhaps attributable to some combination of an incorrect gain value, an incorrect sky background characterization, an incorrect magnitude zero-point determination, or outliers being excessively clipped due to large, non-Gaussian errors. We note that our roll decorrelation and TFA calculations have some free parameters, but this at most could account for only a few percent decrease of the scatter relative to the expected.

### 3.3.4 Skipped Images

Now that the photometric processing pipeline has been explained, sufficient context is available to discuss why certain cadences were not used in our analysis. In what follows, the cadence numbering starts at 1 for the first cadence in the campaign (which corresponds to the *Kepler* long cadence number of 95497). Of the 3856 cadences in Campaign 2, 39 were blank due to resaturation events, an additional six were blank due to other reasons (cadences 216–218 and 2856–2858), 12 were excluded due to our noticing excessive telescope slew during the exposure (cadences 50, 191, 202, 203, 205–207, 209, 383, 863, 1535, and 1823), 68 were excluded due to being excessive pointing outliers (1–49, 51–57, 192–201, 204, and 727), one was excluded due to a hot pixel column we noticed (208), and six were excluded due a majority of the light curves having large outliers (at least 50% off) in flux measurements relative to the median flux value across the whole light curve (2150, 2151, and 2153–2156). Those six cadences occurred around the point in the observations when the telescope roll
direction switched. For the pointing outliers, cadences 1–49 were all pointed in a locus several pixels away from the main group of pointings, and this was an insufficient number to perform our roll decorrelation just on these points; cadences 192–201 and 204 were similarly pointed in a different locus several pixels away from the main; cadences 51–57 were pointed in a locus close to the main locus of pointing but not close to the pointings of its time chunk; and similarly, the pointing of cadence 727 was quite disparate from any in its time chunk. This is a total of 132 cadences that were entirely removed from or not available for our consideration, leaving 3724 (96.6%) of the cadences for the final analysis. We note that most of these cadences were removed from both our raw and final light curves, but that cadences 1–49 are still present in the raw light curves.

### 3.3.5 Removal of Objects

We removed from consideration objects with light curves with less than 800 points (out of a maximum number of 3724 for the final light curves). There were 32 such objects in total, leaving 4570 objects. These removed objects tended to be highly blended with a much brighter object, and this caused many of the photometric calculations to fail. In practice, we found that such light curves were not productive to search for variability. The selected cutoff of 800 was rather conservative and still permitted other relatively sparse and blended light curves that were not useful, so the removal of these objects is not likely to remove anything that might be detected as a variable.

### 3.3.6 Additional Data Used for Analysis

We used the \textit{Gaia} DR2 \texttt{gaia\_dr2\_dr1\_neighbourhood} crossmatch catalog to inform us which of the examined \textit{Gaia} DR1 sources were duplicates. There were 16 DR2 sources matched to two entries in the DR1 source catalog. So that the photometric
aperture used corresponded as closely as possible to the DR2 source position, in each case we kept whichever of the two DR1 sources was closest in position to the corresponding DR2 source. This also happened to correspond in each case with the DR1 source with the best “RANK” value—a calibrated measure of how close a DR1 source is to a DR2 source in both position and magnitude—between the two DR1 sources. We removed the 16 extraneous DR1 sources from the analysis and were left with a final set of 4554 objects with usable light curves. Information on these objects and their light curves is presented in Table 3.1.

As part of our analysis, knowledge of the cluster membership of each of the stars was necessary. We used the membership catalog described in Appendix A or in Wallace (2018a), with the catalog itself available on GitHub2 or in Wallace (2018b). This catalog fitted a two-component Gaussian mixture model to Gaia DR2 proper motions (Lindegren et al., 2018a) to calculate a membership probability for all Gaia DR2 sources with reported proper motions. A very large majority of the calculated membership probabilities were < 1% or > 99%, essentially allowing the catalog to function as a binary classification in all but a few cases. Of the 4554 objects with usable light curves, 4469 of them—98.1%—were matched (again, using the `gaia_dr2.dr1.neighbourhood` crossmatch catalog) to a single DR2 source with reported proper motions and thus were able to be assigned a cluster membership probability. Of the remaining 85 objects, 74 were matched to DR2 sources that lacked reported proper motions, six were matched to more than one DR2 source, and five were not matched to any DR2 sources. Membership probabilities for these 85 objects were not calculated. Of the 4469 objects with reported proper motions, 3784 of them had calculated membership probabilities of ≥99%.

2https://github.com/joshuawallace/M4_pm_membership
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(hh:mm:ss)</td>
<td>(dd:mm:ss)</td>
<td>(mag)</td>
<td>(mag)</td>
<td>(mag)</td>
<td></td>
<td>(mmag)</td>
<td></td>
</tr>
<tr>
<td>V6</td>
<td>DR2 6045478696063803648</td>
<td>16:23:25.76</td>
<td>−26:26:16.7</td>
<td>13.25</td>
<td>13.68</td>
<td>12.63</td>
<td>3773</td>
<td>120.82</td>
<td>1.000</td>
</tr>
<tr>
<td>V7</td>
<td>DR2 6045478391137284224</td>
<td>16:23:25.92</td>
<td>−26:27:42.3</td>
<td>13.28</td>
<td>13.80</td>
<td>12.61</td>
<td>3762</td>
<td>285.46</td>
<td>1.000</td>
</tr>
<tr>
<td>V8</td>
<td>DR2 6045477910100852736</td>
<td>16:23:26.12</td>
<td>−26:26:12.0</td>
<td>13.23</td>
<td>13.64</td>
<td>12.49</td>
<td>3773</td>
<td>255.69</td>
<td>1.000</td>
</tr>
<tr>
<td>V9</td>
<td>DR2 6045477910100361600</td>
<td>16:23:26.76</td>
<td>−26:29:42.0</td>
<td>13.23</td>
<td>13.64</td>
<td>12.49</td>
<td>3773</td>
<td>256.94</td>
<td>1.000</td>
</tr>
<tr>
<td>V10</td>
<td>DR2 6045478322417726848</td>
<td>16:23:29.17</td>
<td>−26:28:54.7</td>
<td>13.19</td>
<td>13.64</td>
<td>12.53</td>
<td>3087</td>
<td>155.06</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note. — There is no W1873 in this table. The identifiers beginning with “W” are sequential otherwise. Light curves for all of these sources are available at Wallace et al. (2019a). Table 3.1 is published in its entirety as a plain text table, object_information.txt, at Princeton University’s DataSpace, accessible via the permanent URL http://arks.princeton.edu/ark:/88435/dsp01h415pd368. A portion is shown here for guidance regarding its form and content.

aThe identifier by which this object is known in this work. Those prepended with “V” are previously identified variables from the catalog of Clement et al. (2001), 2016 June edition, not marked as constant. Those prepended with “SC” are candidate variables from Stetson et al. (2014), and those prepended with “W” are additional Gaia DR1 sources examined in this work.

bGaia source ID, taken from DR1 or DR2 as indicated. The DR2 ID was preferentially used, and only 11 objects in this table have their DR1 IDs quoted.

cJ2000.0; data taken from Gaia DR1 (Lindegren et al., 2016) or DR2 (Lindegren et al., 2018a) as indicated in the “Gaia ID” column (see table note b).

dGaia G magnitude taken from either Gaia DR1 (van Leeuwen et al., 2017) or DR2 (Riello et al., 2018) as indicated in the “Gaia ID” column (see table note b). Please note that G had a different definition between DR1 and DR2 (Evans et al., 2018). G_Bp and G_Rp are taken only from Gaia DR2 and were not included in Gaia DR1, nor are they available for all Gaia DR2 sources.

eNumber of points in the light curve. Raw light curves are used for objects with identifiers beginning with “V” and final light curves for all others. Raw light curves can include data from cadences 1–49 and so may have more points than the maximum of 3724 for the final light curves.

fRMS of the light curve, with sigma clipping (3σ, iterated three times). Raw light curves are used for objects with identifiers beginning with “V” and final light curves for all others.

gMembership probability as calculated in Appendix A. “N. DR2” means this object was not matched to a Gaia DR2 source. “N. D.” means this object lacked proper motion data in Gaia DR2 and its membership probability could not be calculated. “Dup.” means this DR1 source was matched to multiple DR2 sources.
3.3.7 Search for Variability

We used three algorithms for finding periodic signals in our data: the Generalized Lomb–Scargle (GLS; Lomb, 1976; Scargle, 1982; Zechmeister & Kürster, 2009), phase dispersion minimization (PDM; Stellingwerf, 1978), and box-fitting least squares (BLS; Kovács et al., 2002) algorithms. The *astrobase* (Bhatti et al., 2017) implementations of these algorithms were used. With the amount of signal blending in the data, we incorporated a blend detection and removal procedure with the period search. It is worth noting that this blend search incorporated only data that was available from the section of the superstamp we examined. Any blending or systematics due to objects that were in the edge region of the superstamp or beyond could not be readily identified. Additionally, with the amount of systematic noise remaining in the data, it was necessary for us to employ a custom and period-dependent signal-to-noise ratio (SNR) threshold, determined from our examination of the data. The code written to perform both of these tasks, *simple_deblend*, is available at Wallace & Hoffman (2019) or on GitHub\(^3\).

The basic framework of the algorithm used by *simple_deblend* is as follows. For a given period search method (GLS, PDM, BLS) and star, the code:

- Determines the best period based on the period search.
- Checks the periodogram SNR of this period against the threshold; if below the threshold, then it quits the period search.
- Phase-folds neighbor light curves at the given period and figures out which of all the objects has the highest flux amplitude of variability.
- Records the star as the source of that variability if the star has the highest flux amplitude of variability.

\(^3\)https://github.com/simpledeblendorganization/simple_deblend
- Fits out the found period using a Fourier series fit to the data, then repeats.

This is repeated for the desired number of periods—three for our analysis—or until no more robust signals are found.

We now provide a more detailed description of the above method. For a given period search method and star, the code runs the `astrobase` implementation of the period search algorithm. In each search, working in magnitudes (and not fluxes), the minimum period searched was 0.06 days and the maximum period search was 78 days for GLS and PDM—about as long as the maximum duration of the final light curves—or, for BLS, half the observation duration of the light curve. A frequency grid for the search was selected automatically with the `autofreq` parameter set to true. For GLS and PDM, this produced a frequency grid with frequency spacing $\Delta f = 1/(5 \times L)$, with $L$ being the duration of the observations. For BLS, this produced a frequency grid with $\Delta f = 0.25 \times q_{\text{min}}/L$, with $q_{\text{min}}$ being the minimum transit duration in units of fractional phase. This was set to 0.02 and the maximum transit duration was set to 0.55. For BLS, the number of phase bins also needed to be set, and was set to 200.

After running a period search, the resultant periodogram was median filtered to correct for trends that were presumably due to non-white noise. For each point in the periodogram, either 40 (for GLS and PDM) or 100 (for BLS; larger due to its smaller $\Delta f$) of the periodogram values on each side, outside of an exclusion area that was equal to $4/L$ on each side, were collected and were $3\sigma$ sigma clipped before calculating their median, which was then subtracted to produce the filtered periodogram. For PDM, which has periodogram values of one for frequencies with no power, the filtered periodogram values had one added back on. The peak with the highest power was then found, and the robustness of this peak was determined using an SNR calculation on the median-filtered periodogram values. The noise for the ratio was calculated using the standard deviation $\sigma_{\text{per}}$ of nearby periodogram values collected in the exact same way as described above for determining the median filter. The SNR value was

45
then simply the ratio of the periodogram value $p$ with this standard deviation, $p/\sigma_{\text{per}}$, or, for PDM, $(1 - p)/\sigma_{\text{per}}$. Appropriate thresholds for this SNR were determined as a function of period by comparing the SNR values for objects and periods with previously determined variability and (for BLS) injected transits with the rest of the detected periods. This and the selected thresholds are shown in Figure 3.4. If the SNR did not exceed the threshold, the period is marked as not robust and the period search for this object was done.

If the period was determined to be robust using the SNR threshold described, the next step was to check for blends. The light curve was fit with a seven-harmonic Fourier series, which was then evaluated at 200 evenly spaced points. A flux amplitude was then calculated using the minimum and maximum of these Fourier series evaluations, converted from magnitudes. Subsequently, all neighbors within 12 pixels had their flux amplitudes at the same period determined in the same fashion. The choice of 12 pixels was determined by choosing two RR Lyrae variables and looking at all the light curves for surrounding objects to see how far their influence extended. If the object was determined to have the largest flux amplitude, then the period was considered a valid detection, and an 11-harmonic Fourier series fit at the period was subtracted, except for the offset term, from the light curve for subsequent period determination.

We noticed two cases where known low-amplitude variables—specifically, the mmRR variables of Chapter 2—were marked as blends. This was because their periods were $\sim 2/3$ that of some large-amplitude-variable neighbors. Although folding these neighbors’ light curves on the mmRR variability period did not produce the ideal folding for these neighbors’ variability, the folded neighbor light curves still had a large enough amplitude to be larger than the mmRRs’ $\sim$mmag variability. Because of this, if the object was determined as not having the largest flux amplitude, then the neighbor with the largest flux amplitude at the given period was checked to make sure
Figure 3.4: Thresholds for the periodogram SNR for the three period search methods. The corresponding period search method is shown in the upper right of each panel. The blue points show the values calculated from the best eight periods found for each object. For BLS, the orange dots show the values for light curves with injected transits from transiting objects with radii between 0.3 and 3.5 R$_J$. For all panels, the red dots show the periodogram SNR values for objects and periods we identified as being variables during some initial reconnaissance of the data. Not all variables are identified by all the methods, so there are red dots missing between the panels. The thresholds used in our analysis are plotted with a black line in each panel.
that period corresponded to a “real” period of the object. This was determined by
running the given period search method on the neighbor’s light curve and checking
whether the found period matched any of the neighbor’s top eight periods. If the
period matched any of the neighbor’s top eight found periods, then the period was
marked as a blend, and, as for the valid period, the light curve with an 11-harmonic
Fourier series fit removed (except for the offset term) was then used for a subsequent
period search. This recursed until either a valid period was found, a period was de-
determined to not be sufficiently robust, or, in the case of sequential finds of blending,
a recursion limit was hit. This recursion limit was set to be four for GLS and PDM,
and three for BLS. Additionally, if a particular object and period’s flux amplitude
was not the greatest but was greater than 90% the maximum flux amplitude of its
highest-amplitude neighbor, it was marked as a possible source of the variability.

The 1310 objects thus determined to have robust periods were then searched by
eye for classification and to weed out false positives. For this by-eye evaluation, we
used the checkplot submodule of astrobase. After variables and suspected variables
were identified, those with similar periods were checked against each other to look
for blends by evaluating the similar shapes and phasing of the variability. In many
cases, nearby stars were blended with each other, but in some cases, the identified
blends were quite spatially disparate and may have arisen from some effect of our
photometric processing. Appendix 3.A provides specific details on these manually
determined blends. We had 161 variables or suspected variables remaining after this
manual step.

The periodogram SNR selection criterion as we implemented it was not robust to
detect objects with strong variability at a variety of fairly close periods, such as giant
stars with solar-like oscillations. This is due to the calculated noise being artificially
high from the variability at these other periods. In fact, in Figure 3.4, most of the
red points that fall below to the thresholds belong to such asteroseismically active
objects. For simplicity and given the breadth-focused nature of this chapter, we did not make a special search for such variability in those stars for which we may have had a priori reasons for suspecting such variability, and we know our accounting of such variables in this work is incomplete. Readers interested in such variability are encouraged to download the light curves and perform their own searches.

3.3.8 Amplitude, Epoch, and Final Period and Period Uncertainty Determination

For each object determined to be a variable or a suspected variable, a final period search was made using one of our three period search methods with a fine frequency grid ($\Delta f = 10^{-6}$) in a restricted region of frequencies. These frequencies corresponded to possible periods based on the observation duration and the period originally detected in our variability search. The period with the strongest power in this finer search was selected as the final period for the object. For objects with narrow eclipses, a trapezoid model was instead fitted to determine the period, amplitude (trapezoid depth; quoted as a negative number in the case of inverse transits), epoch (center point of transit), and period uncertainty. For all other objects, the amplitude and epoch were derived from a multiharmonic fit to the phase-folded light curve, with amplitude being derived from the difference between the minimum and maximum values of the fit, and epoch being the KBJD of the minimum of the fit. The number of harmonics used varied from object to object, with the most being 11 (for the RRABs) and the least being one, and most objects having between one and five harmonics for their fits. Epochs were always adjusted to be within one period of the KBJD of the earliest observations of our final light curves, $\text{KBJD} = 2060.284181$. Period uncertainties were derived from bootstrap resampling, with 100 resamplings, and with the fine-grid search described above being performed on each resampling and the quoted uncertainty being the difference between the 15.865 and 84.135 per-
centiles of the calculated periods. Such values are more of a confidence interval than a formal uncertainty, but we still quote them as our period uncertainties. Uncertainties on epochs and amplitudes were not determined.

### 3.4 Variability Search Results

The presentation of the results is organized based on the cluster membership probability of the star, whether it is a horizontal branch (HB) star, and whether a given variability signal is certain, suspected, or indeterminably blended. As far as possible, we adopt the same variability classification scheme, including abbreviations, as used in the General Catalog of Variable Stars (GCVS), 2017 March edition (Samus et al., 2017), with additional designations to describe variability not described in this classification scheme. Other than W1189, W3756, and the variables in the Clement et al. (2001) catalog, none of the variables or suspected variables presented here are listed in the GCVS. As part of our breadth versus depth approach, most of our variables go unclassified.

#### 3.4.1 Summary Figures

We first present some figures showing general results from the variability search. Figure 3.5 shows the positions of the variables in the superstamp images, differentiated by cluster members, nonmembers, blended variables, and suspected variables. Figure 3.6 shows a color-magnitude diagram (CMD) for the examined stars, with the identified variables and suspected variables marked. The HB is visible at $G \approx 13$ and $0.5 \lesssim G_{BP} - G_{RP} \lesssim 1.6$, and the main-sequence turnoff is visible at $G \approx 16.5$ and $G_{BP} - G_{RP} \approx 1.2$. We note two stars that are proper motion cluster members and are well off the expected photometric track. The magenta triangle at $G_{BP} - G_{RP} \approx 0.0$ is W1136 and is blended with several other stars (Gaia DR2 source catalog has four other
Figure 3.5: Locations in the superstamp images of our detected variables and suspected variables. This is the same image as in Figure 3.1. Red circles mark the positions of the cluster-member variables, magenta circles mark the positions of variables that are not cluster members or with ambiguous cluster membership, gold circles mark the positions of variables that are indistinguishably blended (only one circle per set of blended stars), and blue circles mark the positions of suspected variables irrespective of cluster membership. Light curves were not obtained for stars in the edges of the images and so no variables were found in those areas (see text for details).
Figure 3.6: Color–magnitude diagram for the stars we obtained light curves for, with variables marked. The photometric data are taken from Gaia DR2 (Riello et al., 2018). Of the 4554 objects we obtained light curves for, 11 objects did not have any DR2 data, and 92 objects were missing $G_{BP}$ and/or $G_{RP}$ data and are not included here. None of the variables or suspected variables were missing these data. The gray points show the data for all the objects. Red points show the data for the RR Lyrae variables, blue points show the data for those objects classified as multiharmonic or millimagnitude RR Lyrae variables, gold points show the data for objects classified as some type of eclipsing binary (EA, EB, or EW), and magenta points show the data for other types of variables. Those variables that are cluster members are marked with closed symbols, and those that are not cluster members or have ambiguous cluster membership are marked with open symbols. Circle symbols represent those for which one object is identified as the variable, while triangle symbols mark variables that are indistinguishably blended. The inset shows the same data but with the suspected variables marked in black, and with the same open/closed symbol membership convention as the main panel. Note the differing scales between the main panel and the inset.
stars within 5″). However, the Gaia DR2 data does not indicate any potential errors in the photometric measurements: its $G_{\text{BP}}$ flux error over mean flux is $3.7 \times 10^{-3}$, and its $G_{\text{RP}}$ flux error over mean flux is $2.2 \times 10^{-3}$, with a $\text{phot\_bp\_rp\_excess\_factor}$ of 1.24. The magenta circle at $G_{\text{BP}} - G_{\text{RP}} \approx 2.5$ is W4490 and has no Gaia DR2 sources within 5″. Its $G_{\text{BP}}$ flux error over mean flux is $7.6 \times 10^{-3}$, and its $G_{\text{RP}}$ flux error over mean flux is $2.0 \times 10^{-3}$, while the $\text{phot\_bp\_rp\_excess\_factor}$ is 1.46. However, W4490 is a unique object (likely an X-ray binary) that we discuss further in Section 3.4.4. Figure 3.7 shows photometric data and variability amplitudes versus periods for all of the variables. Of particular note is the period-luminosity relationship seen in the upper-left panel for objects with multiharmonic variability that mirrors that seen for RR Lyrae variables. This will be further discussed in Section 3.4.4.

3.4.2 Clement et al. (2001) and Stetson et al. (2014) Variables

This subsection focuses exclusively on the previously known variables found in the catalog of Clement et al. (2001), 2016 June edition, with additions from Stetson et al. (2014). This does not include the other previously known variables of W1189, reported as a delta Scuti (DSCUT) variable by Yao & Tong (1989), W3756, reported as a gamma Doradus (GDOR) variable by Yao et al. (2006a), or the asteroseismic giant stars of Miglio et al. (2016); these are discussed later. We also note that none of the new variables of Safonova et al. (2016), which are not in the Clement et al. catalog, fell on the superstamp. A summary of the results for sources not marked “CST” (constant) in the Clement et al. catalog is found in Table 3.2, and the associated light curves are found in Figure 3.8. There are 12 variables from the Clement et al. (2001) and two from Stetson et al. (2014) that fell into our observable region. The 12 Clement et al. variables were first discovered by Leavitt & Pickering (1904) (V6–V10, V15, V19, V27, and V29), Yao et al. (1988) (V61), and Kaluzny
Figure 3.7: Relationships of photometric properties and variability amplitudes with variability periods. The left two panels are for variables that are cluster members, and the right two are for variables that are not cluster members or have ambiguous cluster membership (indicated by, respectively, “Mem.” and “Non.” in the lower-right corners of each panel). The top-left panel shows $G$ vs. period, and the top-right panel shows $G_{BP} - G_{RP}$ vs. period. Both the bottom-left and bottom-right panels show variability amplitude (see Section 3.3.8) vs. period, with amplitude converted to a positive value for those few cases with negative amplitudes as we have defined it. The x-axis scales are the same for panels in the same column, and the y-axis scales for the bottom-left and bottom-right panels are the same. The legend of the bottom-left panel applies to all panels: red diamonds are RR Lyrae variables, blue stars are multiharmonic and mmRR variables, gold “X”s are eclipsing variables, and magenta squares are all other variables. Solid symbols are for variables identified to single objects, and hollow symbols are for variables indistinguishably blended with others. In the x-axis labels, “d” stands for “day.” The period–luminosity relation of the RR Lyrae variables is seen, and the multiharmonic variables also appear to continue this relation to longer and shorter periods.
Figure 3.8: Light curves for 14 previously identified variables from Clement et al. (2001), 2016 June edition, and Stetson et al. (2014) that were in the K2 superstamp and for which we have light curves. Phase-folded light curves are shown in the top 12 panels, while the bottom two light curves show unphased light curves for V13 and SC3. The top left of each panel shows the identifier for the associated star, and the top right shows the period (or “Unphased” for V13 and SC3) at which the light curve is folded, with “d” standing for “day.” Gray points show the individual magnitude measurements, while the black points are binned-median values. For all light curves except that of SC4, the raw light curve output from our image subtraction is used. SC4, along with almost all the other light curves presented in this work, has the additional roll decorrelation and TFA post-processing as described in Section 3.3.3.
Table 3.2. Results for Variables from Clement et al. (2001) and Stetson et al. (2014)

<table>
<thead>
<tr>
<th>ID</th>
<th>Right Ascension (hh:mm:ss)</th>
<th>Declination (dd:mm:ss)</th>
<th>Gaia G magnitude</th>
<th>Period (day)</th>
<th>Period Uncertainty (10^{-5} day)</th>
<th>Amplitude (mag)</th>
<th>Epoch (KBJD)</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>V6</td>
<td>16:23:25.76</td>
<td>−26:26:16.7</td>
<td>13.25</td>
<td>0.320500</td>
<td>0.6</td>
<td>0.33</td>
<td>2060.58</td>
<td>RRC</td>
</tr>
<tr>
<td>V7</td>
<td>16:23:25.92</td>
<td>−26:27:42.3</td>
<td>13.28</td>
<td>0.498787</td>
<td>0.7</td>
<td>0.99</td>
<td>2060.55</td>
<td>RRAB</td>
</tr>
<tr>
<td>V8</td>
<td>16:23:26.12</td>
<td>−26:29:42.0</td>
<td>13.23</td>
<td>0.50822</td>
<td>1</td>
<td>0.87</td>
<td>2060.45</td>
<td>RRAB</td>
</tr>
<tr>
<td>V9</td>
<td>16:23:26.76</td>
<td>−26:29:48.4</td>
<td>13.10</td>
<td>0.57192</td>
<td>2</td>
<td>0.87</td>
<td>2060.36</td>
<td>RRAB</td>
</tr>
<tr>
<td>V10</td>
<td>16:23:29.17</td>
<td>−26:28:54.7</td>
<td>13.19</td>
<td>0.490723</td>
<td>0.4</td>
<td>0.87</td>
<td>2060.70</td>
<td>RRAB</td>
</tr>
<tr>
<td>V13</td>
<td>16:23:30.88</td>
<td>−26:27:04.4</td>
<td>10.04</td>
<td>∼20–30</td>
<td>...</td>
<td>∼0.1</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>V15</td>
<td>16:23:31.93</td>
<td>−26:24:18.5</td>
<td>13.38</td>
<td>0.443795</td>
<td>0.4</td>
<td>1.03</td>
<td>2060.57</td>
<td>RRAB</td>
</tr>
<tr>
<td>V19</td>
<td>16:23:35.02</td>
<td>−26:25:36.8</td>
<td>13.21</td>
<td>0.467809</td>
<td>0.4</td>
<td>0.99</td>
<td>2060.38</td>
<td>RRAB</td>
</tr>
<tr>
<td>V27</td>
<td>16:23:43.14</td>
<td>−26:27:16.7</td>
<td>12.96</td>
<td>0.612027</td>
<td>0.8</td>
<td>0.76</td>
<td>2060.74</td>
<td>RRAB</td>
</tr>
<tr>
<td>V29</td>
<td>16:23:58.22</td>
<td>−26:21:35.4</td>
<td>13.05</td>
<td>0.52250</td>
<td>1</td>
<td>0.75</td>
<td>2060.69</td>
<td>RRAB</td>
</tr>
<tr>
<td>V61</td>
<td>16:23:29.72</td>
<td>−26:29:50.7</td>
<td>13.08</td>
<td>0.265293</td>
<td>0.7</td>
<td>0.13</td>
<td>2060.49</td>
<td>RRC</td>
</tr>
<tr>
<td>V66</td>
<td>16:23:25.53</td>
<td>−26:29:12.1</td>
<td>16.59</td>
<td>0.269889</td>
<td>0.4</td>
<td>0.22</td>
<td>2060.29</td>
<td>EW</td>
</tr>
<tr>
<td>SC3</td>
<td>16:23:35.57</td>
<td>−26:27:08.3</td>
<td>16.32</td>
<td>∼19</td>
<td>...</td>
<td>∼0.1</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>SC4</td>
<td>16:23:44.77</td>
<td>−26:24:29.4</td>
<td>14.88</td>
<td>0.45963</td>
<td>2</td>
<td>0.033</td>
<td>2060.62</td>
<td></td>
</tr>
<tr>
<td>SC9</td>
<td>16:23:34.58</td>
<td>−26:25:41.6</td>
<td>18.73</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

a The identifier by which this object is known in this work, see Table 3.1. Those prepended with “V” are previously identified variables from the catalog of Clement et al. (2001), 2016 June edition, not marked as constant, and those prepended with “SC” are candidate variables from Stetson et al. (2014).

b J2000.0; data taken from Gaia DR2 (Lindegren et al., 2018a).

c Gaia G magnitude taken from Gaia DR2 (Riello et al., 2018).

d The period of the variability in days.

The uncertainty of the period of the variability (see Section 3.3.8 for details on how this is measured.)

f The amplitude of the variability in magnitudes (see Section 3.3.8 for details on how this is measured.)

h The epoch of the minimum of the variability, expressed in KBJD (BJD−2454833.0). See Section 3.3.8 for details on how this is measured.

i Classification based on the GCVS Variability Types, fourth edition (Samus et al., 2017).

j Not a cluster member.

k Not a cluster member; significantly blended with V19 and unable to determine its own variability.
et al. (1997) (V66; called V47 in the discovery work). Given the variability amplitudes for the Clement et al. variables, for Figure 3.8, the raw light curves were used, as our implementation of the Vanderburg-style roll decorrelation did not perform well for objects with large-amplitude variability at timescales shorter than our spline fit. As a note, we count 17 Clement et al. variables in the edge regions for which we did not obtain image subtraction photometry. We mention this here to show that there is still more that can be done with the superstamp data than what is presented in this work. For example, simple aperture photometry could be used on those stars in the less crowded portions of the edge region.

V6, V7, V8, V9, V10, V15, V19, V27, V29, and V61 are all RR Lyrae variables. V6 and V61 are RRCs, while the others are all RRABs. Our period-search method did not detect any significant variability at periods other than (sub)harmonics of the main period, but we wish to stress that our method was focused more on deblending and primary period finding than on a detailed analysis of small-scale variability in these RR Lyrae variables. Kuehn et al. (2017) performed such an analysis for the RR Lyrae variables in the M4 K2 superstamp.

V8, V9, and V61 are in fairly close proximity to each other and to a few other HB stars. In particular, V8 and V9 are blended, and we observed a beating effect between their two periods that created the increased scatter of their light curves seen in Figure 3.8. We did not correct for the blending between these two stars, though in principle, it should be possible. We do not know if V61’s relatively larger scatter is due to blending with V8 and V9 (it is further from them than they are from each other) or just generally higher noise in that part of the image due to the concentration of HB stars, or perhaps something else.

We checked for Blazhko variations among the RR Lyrae variables by searching plots of the (unphased) light curves by eye. Stetson et al. (2014) reported V15 and V29 as candidate Blazhko variables. Kuehn et al. (2017), who used the same K2
Figure 3.9: Light curves for V7 (candidate Blazhko variable) and V29 (Blazhko variable). The identifier is given in the upper-left corner of each panel. The x-axis scale applies to both light curves. The horizontal dashed lines in V7’s panel are arbitrary lines added to help highlight the suspected Blazhko variation.

superstamp data as us, reported V19 and V29 as Blazhko variables as detected via sidepeaks in the amplitude spectra. They also reported the V35 of Clement et al. (2001) as a Blazhko variable, but this star appeared in our edge region and so we did not extract a light curve for it. Here is what we note from our analysis, with Figure 3.9 showing the associated light curves for V7 and V29:

- V7: suspected Blazhko variable, with a period longer than the duration of the observation (most of a cycle is seen).

- V15: Our manual vetting did not find any Blazhko variability. As noted above, Stetson et al. (2014) marked this as a candidate Blazhko (though they did not record a period), while Kuehn et al. (2017) did not. V15 is itself a very peculiar object, as noted by Clementini et al. (1994) and we refer interested readers to that work and its references for full details. In short, the star has peculiarities in its light and radial velocity curves, which could be due either to this star
being in the process of transitioning from an RRAB to an RRC or a strong Blazhko variability.

- V19: Our manual vetting did not find any Blazhko variability. The sidepeak analysis of Kuehn et al. (2017) found a Blazhko period of 16.554 days.

- V29: Blazhko variable, as also noted by Kuehn et al. (2017) and listed as a candidate in Stetson et al. (2014). Kuehn et al. (2017) report a 22.419 day period, which is consistent with what we see.

V13 was first reported as a variable star in Leavitt & Pickering (1904) and is presently reported as being a semi-regular variable (SR). Eggen (1972) observed a \( \sim 40 \) day variability and an amplitude of \( \Delta V = 0.5 \) mag. In our raw light curve, we see low-amplitude variability of \( \sim 0.1 \) mag, quasiperiodic with a period range of \( \sim 20–30 \) days, as can be seen in Figure 3.8. The star is saturated in the images, so it is possible that systematics remain in our light curve. We also note that our final light curve for this object did not have any variability detected for this object, possibly due to the spline fit removing the long-term variability. We mention this as an example of long-term variability that can go undetected by the method employed in this work.

V66 is a \( \sim 0.26 \) day contact eclipsing binary of the W Ursae Majoris type (EW by the GCVS classification). From our analysis, it was not immediately clear which of four blended stars (V66, as well as W1347, W1380, and W1426) was the source of the variability, as all four had approximately the same flux amplitude in our light curves. However, the discovery observations (Kaluzny et al., 1997) were taken at much higher resolution (median seeing FWHM \( \sim 1''0–1''1 \) for five of the six nights of observation) than the separations of these four stars—which were comparable to but slightly greater than Kepler’s \( \sim 4''/ \) pixel image scale. We thus show the light curve only for V66 and not any of its blends.
SC3 is not a cluster member. Similar to V13, it did not have variability detected by our pipeline in its final light curve, again likely due to the long-term and smooth nature of the variability being fitted out by our spline fit. In the raw light curve, we observe approximately the same period and amplitude of variability as Stetson et al. (2014).

SC4, not a cluster member, was identified as a variable by Stetson et al. (2014). However, Gaia DR2 has a phot_variable_flag triggered on the nearby W3152, which is a cluster member, and not SC4. Our pipeline marked SC4 as the true variable and W3152 as blended with SC4, though the flux amplitudes are within ~15% of each other. The resolution of the images used by SC4 was sufficient to resolve these objects, which had 2.7′′ separation, so we stick with Stetson et al. (2014) in calling SC4 and not W3152 the variable.

SC5 is reported as a 0.4197 day period object with ~0.5 mag amplitude, and it should have easily been detected with our data and pipeline. However, it is separated from V19—itself having a 0.4678 day period—by 7′′6 and is quite blended with it. Our pipeline did not identify any variability for SC5 at the reported period. More careful removal of V19’s signal from the data may prove fruitful for this object, but we do not perform such an analysis here.

Our pipeline also produced light curves for V54 (this work: W3012), V55 (this work: W3267), and V80 (this work: W3471), all of which are marked “CST” in the Clement et al. (2001) catalog, meaning that there is uncertainty about whether they are actually variable. Our pipeline did not flag any significant variability for any of these objects, but that does not mean they are not variable. Given the caveats of our variable-search method and the relatively low noise levels our light curves were able to reach, we decided to take a closer look at these stars, particularly their raw light curves.
V54 was marked “CST” from the time of its initial listing in the Clement et al. (2001) catalog because the first report of its variability (Yao et al. 1981a; see also Yao et al. 1981b for an English translation) reported such a small amplitude for the star and it was observed over only a \( \sim 2\)-hr time window total. V54 is a giant star and a proper motion member of the cluster. It exhibits multiharmonic variability, with the strongest GLS power at \( \sim 1.02 \) day period, with a \( \sim 1 \) mmag variability. The reason this was not detected by our method is likely that the rich structure of the periodogram boosted the noise value used in the periodogram SNR calculation, thus leading to an SNR value that fell below the threshold. This variability, though, is of \( \sim 1 \) mmag amplitude, much smaller than the \( \sim 0.1–0.2 \) mag seen for this star in Yao et al. (1981a) and is probably unrelated to what they reported.

V55 was also first reported by Yao et al. (1981a,b) and was also marked “CST” from its initial entry into the Clement et al. (2001) catalog for the same reasons as V54. V55 is an HB star and a proper motion member of the cluster. The variability amplitude reported by Yao et al. (1981a) for V55 (\( \sim 0.1–0.2 \) mag) is larger than the \( \sim 3 \) mmag RMS value we obtain for the raw light curve or the \( \sim 0.3 \) mmag RMS noise value we obtain for the final light curve. The strongest GLS period is \( \sim 3.10 \) day, but this is somewhat weak and the periodogram overall is fairly noisy.

V80 is a subgiant member of the cluster. Variability was reported by Yao et al. (2007) (see Yao et al. 2006b for an English translation) as variable with a period of about a day and with amplitude of 0.05 mag in V. Despite our obtaining an RMS noise level of \( \sim 0.01 \) mag in its raw light curve and \( \sim 3 \) mmag in its final light curve, no significant variability is seen.

Thus, from our work, we think V54 should be marked as a low-amplitude astero-seismic variable and V55 and V80 should retain their “CST” designations, though it would seem the variability we observe for V54 is not the same variability, or at least significantly changed from, what was reported by Yao et al. (1981a).
3.4.3 Millimagnitude RR Lyrae and the Other Horizontal Branch Stars

Two of the HB stars—W2015 and W2386—have been more fully examined in Chapter 2 as potential low-amplitude RRC pulsators (millimagnitude RR Lyrae variables, or mmRRs as coined in that work). W2015 is mmRR 1 from that chapter, W2386 is mmRR2, and W4081 is G3168 briefly mentioned in that chapter. We define the HB in a similar fashion as in Chapter 2: stars with $14.3 < G_{BP} < 13.0$ and $G_{BP} - G_{RP} < 1.5$ and a $>95\%$ cluster membership probability (though the membership probabilities for all these stars are so high that a $99\%$ cutoff could be used with no loss). Excluding the 10 stars previously identified as RR Lyrae variables (see Table 3.2), we have light curves for 24 HB stars, eight of which we detected as significantly variable. Information on these HB variables is found in Tables 3.3 and 3.4, and Figure 3.10 shows the phase-folded light curves and GLS periodograms for these objects. The results are split across two tables for formatting reasons. We stress once again, though, that our periodogram SNR cutoff can sometimes exclude stars with significant variability at other periods close to the peak period, so it is entirely possible that multiharmonic variability is to be found among many of the other 16 non-RR-Lyrae HB stars. Indeed, a quick search that we performed revealed many of them—though not all—to possess multiharmonic variability. To maintain internal consistency with our search method, we do not report them in detail here but do note again that our light curves are available for download and analysis in Wallace et al. (2019a). Several of these objects are blended with other bright stars, so we advise appropriate caution in using them. Two particularly notable blends we noticed were W818, which is likely a blend with W1189, and W1607, which is either blended or otherwise left with a photometric footprint of the somewhat distant V10. W1607 has some power in its periodogram outside the blend period and may possess
Figure 3.10: Phase-folded light curves and periodograms for the eight stars identified as variable or suspected variable non-RR-Lyrae HB stars. Gray points show the individual magnitude measurements, and the black points are binned-median values. The y-axis shows $K_p$, in units of millimagnitude, relative to the median $K_p$ magnitude. In the top-right corner of each panel is shown (in order from top to bottom) the object identifier, the period used for the phase folding, the median magnitude subtracted off, and, for W3125, “SUSP” indicating that this is a suspected variable. For the inset periodogram in each panel, “GLS N.P.” stands for “Generalized Lomb–Scargle Normalized Power” and the red arrow points to the location in the periodogram of the phase-folding period. For W3125, the arrow points slightly off the maximum value of the peak as the period used was taken from a BLS determination of the period rather than a GLS determination. For W4081, the arrow is pointing at a period twice that of the periodogram peak, since upon visual inspection of the light curve we chose a period twice that found by GLS. In the inset x-axis and the listed period, “d” stands for “day.” W2015 and W2386 are the mmRRs from Chapter 2.
<table>
<thead>
<tr>
<th>ID</th>
<th>R.A.</th>
<th>Decl.</th>
<th>$G^c$</th>
<th>Period$^d$</th>
<th>Per. Unc.$^a$</th>
<th>Amp.$^f$</th>
<th>Epoch$^b$</th>
<th>Method$^b$</th>
<th>Type$^i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>W491</td>
<td>16:23:11.52</td>
<td>–26:26:41.1</td>
<td>13.37</td>
<td>0.27941</td>
<td>0.6</td>
<td>0.3</td>
<td>2060.51</td>
<td>Harm.</td>
<td>mmRR/mh</td>
</tr>
<tr>
<td>W508</td>
<td>16:23:12.04</td>
<td>–26:29:44.0</td>
<td>13.86</td>
<td>0.16280</td>
<td>0.3</td>
<td>0.2</td>
<td>2060.40</td>
<td>Harm.</td>
<td>mh</td>
</tr>
<tr>
<td>W521</td>
<td>16:23:12.36</td>
<td>–26:21:58.8</td>
<td>13.25</td>
<td>0.871</td>
<td>10</td>
<td>0.4</td>
<td>2060.45</td>
<td>Harm.</td>
<td>mh</td>
</tr>
<tr>
<td>W566</td>
<td>16:23:13.39</td>
<td>–26:29:15.7</td>
<td>17.82</td>
<td>0.28870</td>
<td>0.9</td>
<td>2</td>
<td>2060.50</td>
<td>Harm.</td>
<td>EW?</td>
</tr>
<tr>
<td>W689k</td>
<td>16:23:15.73</td>
<td>–26:25:58.0</td>
<td>15.47</td>
<td>0.992</td>
<td>10</td>
<td>0.7</td>
<td>2061.21</td>
<td>Harm.</td>
<td>EA?</td>
</tr>
<tr>
<td>W799</td>
<td>16:23:17.63</td>
<td>–26:27:10.6</td>
<td>13.01</td>
<td>0.4054</td>
<td>7</td>
<td>0.2</td>
<td>2060.41</td>
<td>Harm.</td>
<td>mh</td>
</tr>
<tr>
<td>W837</td>
<td>16:23:18.22</td>
<td>–26:29:07.6</td>
<td>14.39</td>
<td>0.09472</td>
<td>0.1</td>
<td>0.1</td>
<td>2060.37</td>
<td>Harm.</td>
<td>shortperiod</td>
</tr>
<tr>
<td>W869</td>
<td>16:23:18.68</td>
<td>–26:23:43.6</td>
<td>10.76</td>
<td>2.177</td>
<td>20</td>
<td>0.3</td>
<td>2061.98</td>
<td>Harm.</td>
<td>mh</td>
</tr>
<tr>
<td>W1091</td>
<td>16:23:21.68</td>
<td>–26:24:7.2</td>
<td>13.05</td>
<td>0.3583</td>
<td>1</td>
<td>0.3</td>
<td>2060.44</td>
<td>Harm.</td>
<td>mh</td>
</tr>
<tr>
<td>W1154k</td>
<td>16:23:22.5</td>
<td>–26:24:59.4</td>
<td>16.15</td>
<td>0.991</td>
<td>20</td>
<td>0.6</td>
<td>2060.66</td>
<td>Harm.</td>
<td>?</td>
</tr>
<tr>
<td>W1165</td>
<td>16:23:22.64</td>
<td>–26:26:22.5</td>
<td>12.08</td>
<td>1.3338</td>
<td>5</td>
<td>0.8</td>
<td>2061.10</td>
<td>Harm.</td>
<td>mh</td>
</tr>
<tr>
<td>W1349j</td>
<td>16:23:24.98</td>
<td>–26:29:25.3</td>
<td>13.23</td>
<td>0.3884</td>
<td>2</td>
<td>0.2</td>
<td>2060.43</td>
<td>Harm.</td>
<td>mh</td>
</tr>
<tr>
<td>W1582m</td>
<td>16:23:27.84</td>
<td>–26:29:11.9</td>
<td>13.78</td>
<td>0.17275</td>
<td>0.4</td>
<td>0.2</td>
<td>2060.36</td>
<td>Harm.</td>
<td>mh</td>
</tr>
<tr>
<td>W1608</td>
<td>16:23:28.13</td>
<td>–26:26:09.9</td>
<td>12.90</td>
<td>1.319</td>
<td>10</td>
<td>0.2</td>
<td>2061.60</td>
<td>Harm.</td>
<td>mh</td>
</tr>
<tr>
<td>W1735</td>
<td>16:23:29.15</td>
<td>–26:29:12.0</td>
<td>11.65</td>
<td>2.146</td>
<td>20</td>
<td>1</td>
<td>2062.10</td>
<td>Harm.</td>
<td>mh</td>
</tr>
<tr>
<td>W1763</td>
<td>16:23:29.81</td>
<td>–26:23:25.6</td>
<td>12.51</td>
<td>0.7836</td>
<td>8</td>
<td>0.4</td>
<td>2060.63</td>
<td>Harm.</td>
<td>mh</td>
</tr>
<tr>
<td>W1848</td>
<td>16:23:30.51</td>
<td>–26:23:57.9</td>
<td>17.59</td>
<td>0.4486</td>
<td>4</td>
<td>3</td>
<td>2060.33</td>
<td>Harm.</td>
<td>?</td>
</tr>
</tbody>
</table>

Note. — Table is continued in Table 3.4.

$^a$ The identifier by which this object is known in this work (see Table 3.1).

$^b$ J2000.0; data taken from Gaia DR2 (Lindegren et al., 2018a). All entries in this table are DR2 sources, so none of the information presented is from Gaia DR1.

$^c$ Gaia $G$ magnitude taken from Gaia DR2 (Riello et al., 2018). All entries in this table are DR2 sources, so none of the information presented is from Gaia DR1.

$^d$ The period of the variability in days.

$^e$ The uncertainty of the period of the variability in $10^{-4}$ days (see Section 3.3.8 for details on how this is measured.)

$^f$ The amplitude of the variability in millimagnitudes (see Section 3.3.8 for details on how this is measured). A negative amplitude means that the light curve shows a box-like signal that is a brightening, rather than the more common eclipse-based dimmings for such signals.

$^g$ The epoch of the minimum of the variability, expressed in KBJD (BJD – 2454833.0). See Section 3.3.8 for details on how this is measured.

$^h$ Method used for determining amplitude and epoch. “Harm.” means a harmonic fit was used, and “Trap.” means a trapezoid fit was used.

$^i$ Classification based on the GCVS Variability Types, fourth edition (Samus et al., 2017), where possible. Additional designations used: “mmRR,” millimagnitude RR Lyrae; “mh,” multiharmonic variability; “shortperiod,” sinusoidal variability of <$0.1$-day period; “xrb,” a likely X-ray binary, but not classified as “X” since we do not know of variability in the X-ray emission.

$^j$ Six other stars observed with same variability; this chosen as variable since it was the most robust detection (see text for details).

$^k$ These two stars (W689 and W1154) are 27 pixels apart but have consistent periods and, based on our analysis, may phase with each other.

$^l$ Slightly blended with V8. This detected variability is not a (sub)harmonic of that variability, so we are confident this belongs to the star itself.

$^m$ Slightly blended with V10. This detected variability is not a (sub)harmonic of that variability, so we are confident this belongs to the star itself.
intrinsic variability. Likewise, W1628 and W1643 are blended with V61 and V9 and may require a more careful analysis.

Interpreting the previously identified mmRRs in the context of these additional HB variables is informative. Given that the periodogram structures seem to form a continuum between the strongly mono-periodic W2015 and the rich, very multi-periodic periodogram of W521, it is possible that what we have called mmRRs are a transition between the asteroseismic variability of HB stars outside of the instability strip and the RR Lyrae pulsators inside. We note that W2015/mmRR 1 and W3125 are blueward of the instability strip, W4081 is inside the strip, and the remaining objects are redward. There still remain many questions. Why does W2015 (mmRR 1) have such a single dominant period whereas the other HBs do not have any periods with such great prominence? What causes the range of periods seen? What causes W4081’s striking even-odd amplitude modulation, and why is it found in the instability strip but not pulsating like the RR Lyrae variables? Certainly the K2 photometric precision and the observations of concentrations of HB stars in GCs allow for an unprecedented look at the asteroseismic variations of HB stars outside the instability strip in addition to the RR Lyrae variables themselves. We also echo our previous caveat that other HB stars with rich periodogram structures may have been missed by our period search method, and these may not be the only HB stars with detectable oscillations.
### Table 3.4. Newly Discovered Cluster Variables, Continued

<table>
<thead>
<tr>
<th>ID*</th>
<th>R.A.*</th>
<th>Decl.*</th>
<th>G*</th>
<th>Period*</th>
<th>Per. Unc.*</th>
<th>Amp.*</th>
<th>Epoch*</th>
<th>Method*</th>
<th>Type*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(hh:mm:ss)</td>
<td>(dd:mm:ss)</td>
<td>(mag)</td>
<td>(day)</td>
<td>(10^-4 day)</td>
<td>(mmag)</td>
<td>(KBJD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

...continuing from Table 3.3

#### Variables

<table>
<thead>
<tr>
<th>ID*</th>
<th>R.A.*</th>
<th>Decl.*</th>
<th>G*</th>
<th>Period*</th>
<th>Per. Unc.*</th>
<th>Amp.*</th>
<th>Epoch*</th>
<th>Method*</th>
<th>Type*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(hh:mm:ss)</td>
<td>(dd:mm:ss)</td>
<td>(mag)</td>
<td>(day)</td>
<td>(10^-4 day)</td>
<td>(mmag)</td>
<td>(KBJD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Suspected Variables

<table>
<thead>
<tr>
<th>ID*</th>
<th>R.A.*</th>
<th>Decl.*</th>
<th>G*</th>
<th>Period*</th>
<th>Per. Unc.*</th>
<th>Amp.*</th>
<th>Epoch*</th>
<th>Method*</th>
<th>Type*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(hh:mm:ss)</td>
<td>(dd:mm:ss)</td>
<td>(mag)</td>
<td>(day)</td>
<td>(10^-4 day)</td>
<td>(mmag)</td>
<td>(KBJD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.4 (cont’d)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>W1953</td>
<td>16:23:31:68</td>
<td>−26:28:06.8</td>
<td>18.00</td>
<td>2.34</td>
<td>400</td>
<td>1°</td>
<td>2061.42</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W2127</td>
<td>16:23:31.45</td>
<td>−26:29:20.7</td>
<td>17.98</td>
<td>...</td>
<td>...</td>
<td>~40</td>
<td>~2096</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>W2233</td>
<td>16:23:34.48</td>
<td>−26:28:29.6</td>
<td>18.91</td>
<td>0.46817</td>
<td>0.3</td>
<td>5</td>
<td>2060.29</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W2272</td>
<td>16:23:34.85</td>
<td>−26:29:04.6</td>
<td>18.74</td>
<td>2.223</td>
<td>70</td>
<td>26</td>
<td>2062.05</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W2499</td>
<td>16:23:37.11</td>
<td>−26:28:45.6</td>
<td>16.72</td>
<td>3.832</td>
<td>40</td>
<td>10</td>
<td>2062.90</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W2515</td>
<td>16:23:37.28</td>
<td>−26:28:08.5</td>
<td>19.11</td>
<td>1.208</td>
<td>10</td>
<td>8</td>
<td>2061.36</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W2556</td>
<td>16:23:37.7</td>
<td>−26:27:20.4</td>
<td>16.03</td>
<td>33.96</td>
<td>900</td>
<td>1</td>
<td>2079.36</td>
<td>Trap.</td>
<td></td>
</tr>
<tr>
<td>W2577</td>
<td>16:23:37.94</td>
<td>−26:28:41.7</td>
<td>13.01</td>
<td>3.20</td>
<td>100</td>
<td>0.4</td>
<td>2062.23</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W2616</td>
<td>16:23:38.3</td>
<td>−26:29:03.4</td>
<td>18.45</td>
<td>3.81</td>
<td>200</td>
<td>40</td>
<td>2060.81</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W2641</td>
<td>16:23:38.58</td>
<td>−26:29:11.7</td>
<td>14.43</td>
<td>0.8678</td>
<td>7</td>
<td>2</td>
<td>2061.11</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W2753</td>
<td>16:23:39.78</td>
<td>−26:29:42.1</td>
<td>18.44</td>
<td>0.5474</td>
<td>5</td>
<td>9</td>
<td>2060.42</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W2790</td>
<td>16:23:40.27</td>
<td>−26:37:37.8</td>
<td>17.42</td>
<td>6.36</td>
<td>200</td>
<td>~2</td>
<td>2061.21</td>
<td>Trap.</td>
<td></td>
</tr>
<tr>
<td>W2800</td>
<td>16:23:40.4</td>
<td>−26:29:20.6</td>
<td>17.87</td>
<td>1.84</td>
<td>200</td>
<td>5</td>
<td>2061.85</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W2819</td>
<td>16:23:40.61</td>
<td>−26:29:02.1</td>
<td>18.31</td>
<td>1.66</td>
<td>200</td>
<td>120</td>
<td>2060.77</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W2876</td>
<td>16:23:41.22</td>
<td>−26:28:53.0</td>
<td>16.64</td>
<td>4.203</td>
<td>90</td>
<td>5</td>
<td>2062.41</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W2893</td>
<td>16:23:41.41</td>
<td>−26:23:18.0</td>
<td>17.83</td>
<td>1.28</td>
<td>200</td>
<td>1</td>
<td>2060.63</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W2966</td>
<td>16:23:42.25</td>
<td>−26:27:42.2</td>
<td>15.76</td>
<td>0.372</td>
<td>20</td>
<td>0.2</td>
<td>2060.63</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W3105</td>
<td>16:23:43.88</td>
<td>−26:27:36.8</td>
<td>16.83</td>
<td>0.6060</td>
<td>3</td>
<td>4</td>
<td>2060.33</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W3125</td>
<td>16:23:44.21</td>
<td>−26:28:44.4</td>
<td>13.00</td>
<td>2.66</td>
<td>500</td>
<td>1</td>
<td>2061.01</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W3282</td>
<td>16:23:46.11</td>
<td>−26:25:34.7</td>
<td>17.37</td>
<td>2.50</td>
<td>200</td>
<td>1</td>
<td>2061.92</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W3371</td>
<td>16:23:47.31</td>
<td>−26:22:30.4</td>
<td>13.91</td>
<td>0.966</td>
<td>20</td>
<td>0.2</td>
<td>2060.42</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W3521</td>
<td>16:23:49.59</td>
<td>−26:29:30.8</td>
<td>18.84</td>
<td>2.692</td>
<td>100</td>
<td>22</td>
<td>2060.45</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W3887</td>
<td>16:23:55.53</td>
<td>−26:28:44.0</td>
<td>17.26</td>
<td>2.443</td>
<td>30</td>
<td>1</td>
<td>2062.30</td>
<td>Trap.</td>
<td></td>
</tr>
<tr>
<td>W3901</td>
<td>16:23:55.82</td>
<td>−26:29:20.8</td>
<td>17.11</td>
<td>0.522</td>
<td>30</td>
<td>0.9</td>
<td>2060.79</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W4014</td>
<td>16:23:58.03</td>
<td>−26:23:30.4</td>
<td>18.47</td>
<td>1.88</td>
<td>400</td>
<td>12</td>
<td>2060.28</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W4250</td>
<td>16:24:04.81</td>
<td>−26:24:17.4</td>
<td>19.01</td>
<td>1.784</td>
<td>20</td>
<td>33</td>
<td>2060.90</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W4268</td>
<td>16:24:05.39</td>
<td>−26:29:16.2</td>
<td>17.31</td>
<td>1.983</td>
<td>10</td>
<td>2</td>
<td>2061.75</td>
<td>Harm.</td>
<td></td>
</tr>
<tr>
<td>W4301</td>
<td>16:24:06.43</td>
<td>−26:28:53.3</td>
<td>18.70</td>
<td>4.283</td>
<td>20</td>
<td>5</td>
<td>2062.69</td>
<td>Trap.</td>
<td></td>
</tr>
</tbody>
</table>

Note. — Continued from Table 3.3. Classifications are not attempted for the suspected variables. Explanations regarding why these are reported as suspected instead of discovered variables can be found in Appendix 3.B.

aSee table notes for the equivalent columns in Table 3.3.

bThe trapezoid model appeared to fail to fit the full amplitude of the signal. Actual amplitude may be ~2–3 times larger.

cEpoch and possibly amplitude may be inaccurate, owing to PDM being employed to fold these transits and a harmonic fit being used to determine epoch and amplitude.

dSingle event.
Tables 3.3 and 3.4 show information for the variable cluster members, both proper and suspected variables. The suspected variables are more thoroughly discussed and presented in Appendix 3.B. Figures 3.11, 3.12, and 3.13 show the phase-folded light curves for the variables. We discuss here and in Section 3.4.5 some of the more notable cluster-member variables.

W4490 has a particularly interesting light curve: a 1.959 day period triangular-shaped increase in brightness, with an amplitude of $\sim 20$ mmag$^4$. Figure 3.13 plots the phase-folded raw light curve instead of the processed, final light curve. We found that the processing cut its amplitude approximately in half. The raw light curve has systematic noise, most likely due to this object’s period being very close to the resaturation period (1.962 days) and nearly an integer multiple of the drift correction and observing cadence. Verbunt (2001) reports a ROSAT X-ray source detection 2″8 away from this object (object X8 in NGC6121/M4), with a reported position statistical error on the X-ray source of 2″6 and an additional projection error of $\sim 5″$ also at play. This spatially coincident X-ray source with the reported variability period have informed our classification of this object as an X-ray binary. This portion of M4 unfortunately has not been included in fields of view of previous Chandra observations, which have been primarily focused on the cluster’s core (e.g. Bassa et al., 2004). Its unusual photometry was noted in Section 3.4.1 and Figure 3.6. As measured by Gaia DR2, this object is much more red than we would expect for a star of its luminosity in the cluster.

Of the other cluster-member variables in Figures 3.11–3.13, most are low-amplitude sinusoids, possibly including some ellipsoidal or rotational variables. Many are giant stars showing mmRR or multiharmonic asteroseismic variability. For those objects,

\[ \text{The value quoted here and seen in Figure 3.13 is different from that reported in Table 3.4, since the former is taken from the raw and the latter from the final light curves.} \]
Figure 3.11: Phase-folded light curves for cluster members that, other than W2665 and W3033, are newly identified as variable stars in this work. W2665 and W3033 were previously identified by Miglio et al. (2016). The panels are ordered by the target identifier. Here we show the first 15 cluster variables. Additional variables are shown in Figures 3.12 and 3.13. Gray points show the individual magnitude measurements, and the black points are binned-median values. The y-axis shows $K_p$, in units of millimagnitude, relative to the median $K_p$ magnitude. In each panel, the identifier of the star is shown in the upper-left corner, and (from top to bottom) the folding period and subtracted median magnitude are shown in the upper-right corner. For the listed period, “d” stands for “day.”
Figure 3.12: Same as Figure 3.11, but for additional cluster-member variables.
Figure 3.13: Same as Figures 3.11 and 3.12, but for additional cluster-member variables. The data from W4490 are taken from its raw light curve.
the periods shown in the figures are typically just the dominant sinusoidal component. In Figure 3.7, it can be seen in the top-left panel that these stars appear to extrapolate the period–luminosity relationship of the RR Lyrae, with variables of longer periods than the RR Lyrae variables continuing the relation of the RRABs (the cluster of diamonds with period greater than 0.4 days), the handful of objects with periods less than the RRCs (the two diamonds with periods ~0.3 days) seeming to form a parallel trend, and objects falling into the period range of the RR Lyrae variables themselves having similar G magnitudes as them. Since G is correlated with the evolutionary state of these stars, and thus with stellar density, it is not surprising that the oscillation periods, which are determined in part by stellar densities, are correlated with G even for oscillators with smaller amplitudes than the RR Lyrae variables. The scatter seen in the relation is probably due to the picking up of different modes for different stars as the dominant cause of the photometric variability. We also note an apparent correlation between amplitude and period in the lower-right panel of Figure 3.7 for the multiharmonic and mmRR stars.

There were a number of variable signals that were indistinguishably blended between two or more stars and that were not able to be disentangled either from our data or from referencing some other previous work of which we knew. Table 3.5 lists these objects, both cluster members and nonmembers, and Figure 3.14 shows the associated light curves. W283/W293, and W1318/W1335/W1246, both EWs, are discussed in Section 3.4.5.

### 3.4.5 Cluster Eclipsing and Contact Binaries

W1601, shown in Figure 3.11, is a detached eclipsing binary with a 4.6337 day period. The phase difference between the primary and secondary eclipses reveals the system to be slightly eccentric. The system appears to be grazing, with a primary eclipse depth of $0.038 \pm 0.005$ mag, a fractional duration of $0.046 \pm 0.003$, and an eclipse ingress
Figure 3.14: Same as Figure 3.11, but for stars with signals that are indistinguishably blended in our data. In each case, only one star from each set of blended stars is chosen to represent the light curve. See Table 3.5 for more information.
Table 3.5. Newly Discovered Variables that Are Indeterminable Blends

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(hh:mm:ss)</td>
<td>(dd:mm:ss)</td>
<td>(mag)</td>
<td>(day)</td>
<td>(10^{-4} day)</td>
<td>(mmag)</td>
<td>(KBJD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W283</td>
<td>16:23:06.14</td>
<td>-26:27:45.8</td>
<td>18.28</td>
<td>0.20450</td>
<td>0.3</td>
<td>6</td>
<td>2060.40</td>
<td>EW</td>
<td>1.00</td>
</tr>
<tr>
<td>W293</td>
<td>16:23:06.46</td>
<td>-26:27:50.9</td>
<td>19.08</td>
<td>0.20450</td>
<td>0.1</td>
<td>23</td>
<td>2060.40</td>
<td>EW</td>
<td>1.00</td>
</tr>
<tr>
<td>W1129</td>
<td>16:23:22.23</td>
<td>-26:28:05.4</td>
<td>17.71</td>
<td>0.967</td>
<td>20</td>
<td>7</td>
<td>2060.60</td>
<td>?</td>
<td>1.00</td>
</tr>
<tr>
<td>W1136</td>
<td>16:23:22.29</td>
<td>-26:28:03.7</td>
<td>15.80</td>
<td>0.967</td>
<td>10</td>
<td>0.8</td>
<td>2060.55</td>
<td>?</td>
<td>1.00</td>
</tr>
<tr>
<td>W1146</td>
<td>16:23:22.38</td>
<td>-26:27:59.3</td>
<td>17.77</td>
<td>0.969</td>
<td>20</td>
<td>3</td>
<td>2060.46</td>
<td>?</td>
<td>1.00</td>
</tr>
<tr>
<td>W2262</td>
<td>16:23:34.78</td>
<td>-26:29:15.1</td>
<td>17.38</td>
<td>0.5003d</td>
<td>10</td>
<td>4</td>
<td>2060.63</td>
<td>?</td>
<td>1.00</td>
</tr>
<tr>
<td>W2282</td>
<td>16:23:34.95</td>
<td>-26:29:14.2</td>
<td>18.32</td>
<td>0.50016</td>
<td>0.5</td>
<td>8</td>
<td>2060.61</td>
<td>?</td>
<td>1.00</td>
</tr>
<tr>
<td>W2289</td>
<td>16:23:34.99</td>
<td>-26:29:10.4</td>
<td>18.86</td>
<td>. . . . . .</td>
<td>. . . . . .</td>
<td>. . . . . .</td>
<td>. . . . . .</td>
<td>. . . . . .</td>
<td>. . . . . .</td>
</tr>
<tr>
<td>W2300</td>
<td>16:23:35.08</td>
<td>-26:29:12.7</td>
<td>16.97</td>
<td>0.4989</td>
<td>1</td>
<td>2</td>
<td>2060.65</td>
<td>?</td>
<td>1.00</td>
</tr>
<tr>
<td>W2826</td>
<td>16:23:40.66</td>
<td>-26:29:37.8</td>
<td>18.87</td>
<td>0.998</td>
<td>40</td>
<td>2</td>
<td>2060.53</td>
<td>?</td>
<td>1.00</td>
</tr>
<tr>
<td>W2830</td>
<td>16:23:40.69</td>
<td>-26:29:30.4</td>
<td>17.63</td>
<td>2.00</td>
<td>100</td>
<td>4</td>
<td>2060.54</td>
<td>?</td>
<td>1.00</td>
</tr>
<tr>
<td>W3431</td>
<td>16:23:48.31</td>
<td>-26:28:13.4</td>
<td>17.41</td>
<td>0.5887</td>
<td>4</td>
<td>3</td>
<td>2060.47</td>
<td>EW?</td>
<td>1.00</td>
</tr>
<tr>
<td>W3436</td>
<td>16:23:48.42</td>
<td>-26:28:13.7</td>
<td>17.51</td>
<td>0.5889</td>
<td>4</td>
<td>3</td>
<td>2060.47</td>
<td>EW?</td>
<td>1.00</td>
</tr>
<tr>
<td>W3456</td>
<td>16:23:48.64</td>
<td>-26:28:14.8</td>
<td>17.87</td>
<td>0.5887</td>
<td>4</td>
<td>4</td>
<td>2060.46</td>
<td>EW?</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Note. — All amplitudes and epochs calculated using a harmonic fit, compared to Tables 3.3, 3.4, and 3.6 where some were determined with a trapezoid fit

a See table notes for Table 3.3 for details on these columns.

b The epoch of the minimum of the variability, expressed in KBJD (BJD−2454833.0). See Appendix 3.B for details on how this is measured. Significant differences in epochs between blended objects are due to differences in the fitted harmonics for each case; these objects do phase up.

c Cluster membership probability; see Appendix A.

d The best period found for this object in our period search was 1.500 days, which may be a modulation of the ~0.5 day period.

The best period found for this object was 0.38587 days. We were unable to get a good fit on the ~0.5 day period, but this object does have visible variability when folded on this period and has an image location very close to the other stars in this blended group. It may be that there is more than one variable in this group.

f No proper motion data available; probable photometric cluster member.

Both W3883 and W3894 have similar periods and epochs as W4084, but are ~66 pixels away.

b Probable photometric cluster member.
fractional duration of $0.016 \pm 0.004$, making the eclipse very triangular. There is also a sinusoidal variability on top of the eclipses, suggesting ellipsoidal variability, which is not terribly surprising considering the short period of the binary. This informs our classification of this as an Algol-type eclipsing binary (EA). Based on the Clement et al. catalog, this is the sixth EA known in M4, with the note that the two EAs of Safonova et al. (2016) are not cluster members based on the proper motions reported there.

W4361, shown in Figure 3.13, is possibly another eclipsing binary. In this case, the system appears to be semi-detached or maybe even a contact binary. The eclipses are very triangular. The depth of the primary eclipse based on our trapezoid fit is $0.72 \pm 0.05$ mmag, with a fractional eclipse duration of $0.21 \pm 0.01$ and fractional ingress duration of $0.07 \pm 0.01$. This is a red giant star, with a radius that should be much larger than the $\sim 15 \, R_\odot$ implied by the orbital period and the $\sim 0.8 \, M_\odot$ maximum expected masses for each of the stars given their membership in the cluster. Perhaps W4361 is simply blended with a background eclipsing binary or even another binary in the cluster.

W293, blended with W283, is a clear example of an EW, having a period of $0.20450$ days and a primary eclipse depth of $\sim 30$ mmag and a secondary eclipse depth of approximately half that. Both stars are cluster members. Similarly, W1318, blended with W1335 and W1346, is also a clear EW. The orbital period is $0.277389$ days and the primary eclipse depth is $\sim 20$ mmag and the secondary eclipse depth $\sim 10$ mmag. W1318 and W1346 are proper motion members of the cluster, but W1335 does not have reported proper motions in Gaia DR2. However, based on its CMD location ($G \approx 18.4$, $G_{BP} - G_{RP} \approx 1.44$), it is a probable cluster member, and so we report a high degree of certainty that this EW also belongs to the cluster. There are also two other suspected EWs, both of which are hopelessly blended in our data. The first of these is one of the three blended stars W3431, W3436, or W3456 (all three are
cluster members) and the second is one of the blended stars W2761, W2779, W2793, or W2813 (all but W2813 are cluster members).

### 3.4.6 Variables not in M4

Included with the rich variety of cluster-member variables are many variables that were not cluster members. Table 3.6 shows information for these variable stars, and Figures 3.15 and 3.16 show the phase-folded light curves. The suspected variables will be more thoroughly discussed in Appendix 3.B.

At ~1.8 kpc in distance, and also being relatively close to the Galactic center ($l\approx351^\circ, b\approx16^\circ$), the non-cluster-member stars in the direction of M4 are a mixture of both foreground and background objects. We will touch on only two of the field variables here.

W1189 is also HD 147491 and V972 Sco of the GCVS. Yao & Tong (1989) reported this star as being a DSCUT variable with a ~0.02 day period; however, we do not see any ~0.02 day variability, and the 1.5097 day period we find is too long for a DSCUT. We think it is more likely that this is a GDOR. This is also the brightest star in the M4 superstamp, with $Gaia$ DR2 $G = 9.46$.

W3756 is also V1331 Sco of the GCVS. Yao et al. (2006a) identified a ~15 mmag, 1.03 day period variability in this star based on $V$-band observations taken in 1990 and 1991 and classified it as a GDOR. We see a ~1 mmag amplitude and a 0.634 day period. There is also power in our GLS, PDM, and BLS periodograms for this object at a period ~0.97 days (compare with the original 1.03 day period in the discovery), which is the dominant periodogram peak when the main period and its harmonics are removed. GDOR variability can change in amplitude and dominant frequency over time. This combined with the differences between the $K_p$ and $V$ bandpasses make it unsurprising for us to see a different amplitude and dominant period relative to the Yao et al. (2006a) observations, made over 23 years prior the K2 observations.
Figure 3.15: Same as Figure 3.11, but for variables that are not cluster members. 12 variables are shown here, and Figure 3.16 shows the remaining 10.

W2203 is a detached eclipsing binary with a 21.72 day period and what appears to be reflections or other brightening events just before and after both the primary and secondary eclipses. The primary eclipse depth is $0.013 \pm 0.001$ mag, with a fractional eclipse duration of $0.028 \pm 0.003$ and a fractional ingress duration of $0.005 \pm 0.002$.

Finally, we remind the reader of the blended variables in Table 3.5 and Figure 3.14 that are not cluster members: W2013 (blended with W2006) and W2813 (blended with W2761, W2779, and W2793).
Figure 3.16: Same as Figure 3.15, but for additional variables that are not cluster members.
Table 3.6. Newly Discovered Variables That Are Not Cluster Members

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(hh:mm:ss)</td>
<td>(dd:mm:ss)</td>
<td>(mag)</td>
<td>(day)</td>
<td>(10^{-4} day)</td>
<td>(mmag)</td>
<td>(KBJD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W92</td>
<td>16:22:58.62</td>
<td>-26:28:59.7</td>
<td>19.09</td>
<td>1.958b</td>
<td>60</td>
<td>24</td>
<td>2061.98</td>
<td>Harm.</td>
<td>?</td>
</tr>
<tr>
<td>W137</td>
<td>16:23:00.83</td>
<td>-26:25:22.2</td>
<td>15.10</td>
<td>1.023</td>
<td>20</td>
<td>0.4</td>
<td>2061.26</td>
<td>Harm.</td>
<td>?</td>
</tr>
<tr>
<td>W576</td>
<td>16:23:13.64</td>
<td>-26:26:27.8</td>
<td>18.89</td>
<td>0.977</td>
<td>10</td>
<td>5</td>
<td>2061.13</td>
<td>Harm.</td>
<td>?</td>
</tr>
<tr>
<td>W676</td>
<td>16:23:15.54</td>
<td>-26:27:46.3</td>
<td>13.03</td>
<td>2.098</td>
<td>10</td>
<td>0.4</td>
<td>2061.24</td>
<td>Harm.</td>
<td>?</td>
</tr>
<tr>
<td>W1543</td>
<td>16:23:27.47</td>
<td>-26:23:11.9</td>
<td>14.44</td>
<td>0.7840</td>
<td>7</td>
<td>0.3</td>
<td>2060.94</td>
<td>Harm.</td>
<td>?</td>
</tr>
<tr>
<td>W1750</td>
<td>16:23:29.66</td>
<td>-26:21:13.9</td>
<td>17.63</td>
<td>0.3225</td>
<td>2</td>
<td>2</td>
<td>2060.55</td>
<td>Harm.</td>
<td>?</td>
</tr>
<tr>
<td>W2028</td>
<td>16:23:32.46</td>
<td>-26:26:45.0</td>
<td>11.18</td>
<td>0.6021</td>
<td>5</td>
<td>0.4</td>
<td>2060.44</td>
<td>Harm.</td>
<td>?</td>
</tr>
<tr>
<td>W2290d</td>
<td>16:23:34.14</td>
<td>-26:25:50.2</td>
<td>15.65</td>
<td>2.0990</td>
<td>10</td>
<td>5</td>
<td>2060.29</td>
<td>Harm.</td>
<td>?</td>
</tr>
<tr>
<td>W2996</td>
<td>16:23:42.58</td>
<td>-26:23:43.9</td>
<td>14.42</td>
<td>0.6873</td>
<td>4</td>
<td>0.5</td>
<td>2060.76</td>
<td>Harm.</td>
<td>?</td>
</tr>
<tr>
<td>W3079</td>
<td>16:23:43.58</td>
<td>-26:26:18.4</td>
<td>13.12</td>
<td>0.33134</td>
<td>1</td>
<td>0.2</td>
<td>2060.55</td>
<td>Harm.</td>
<td>?</td>
</tr>
<tr>
<td>W3756</td>
<td>16:23:53.21</td>
<td>-26:22:44.5</td>
<td>12.47</td>
<td>0.63400</td>
<td>1</td>
<td>0.8</td>
<td>2060.47</td>
<td>Harm.</td>
<td>GDOR?</td>
</tr>
<tr>
<td>W3862</td>
<td>16:23:54.98</td>
<td>-26:10:10.8</td>
<td>12.83</td>
<td>1.270</td>
<td>30</td>
<td>0.7</td>
<td>2061.27</td>
<td>Harm.</td>
<td>?</td>
</tr>
<tr>
<td>W3984</td>
<td>16:23:57.56</td>
<td>-26:24:24.5</td>
<td>14.02</td>
<td>1.791</td>
<td>10</td>
<td>1</td>
<td>2061.13</td>
<td>Harm.</td>
<td>?</td>
</tr>
<tr>
<td>W4072</td>
<td>16:23:59.18</td>
<td>-26:23:30.8</td>
<td>13.92</td>
<td>0.22783</td>
<td>0.5</td>
<td>0.3</td>
<td>2060.47</td>
<td>Harm.</td>
<td>mh</td>
</tr>
<tr>
<td>W4084</td>
<td>16:23:59.36</td>
<td>-26:29:10.5</td>
<td>14.26</td>
<td>2.547</td>
<td>30</td>
<td>0.8</td>
<td>2062.71</td>
<td>Harm.</td>
<td>?</td>
</tr>
<tr>
<td>W4161</td>
<td>16:24:01.82</td>
<td>-26:32:24.4</td>
<td>13.37</td>
<td>2.576</td>
<td>40</td>
<td>1</td>
<td>2061.39</td>
<td>Harm.</td>
<td>mh</td>
</tr>
<tr>
<td>W4434</td>
<td>16:24:11.6</td>
<td>-26:26:26.1</td>
<td>15.15</td>
<td>3.010</td>
<td>80</td>
<td>0.7</td>
<td>2063.07</td>
<td>Harm.</td>
<td>?</td>
</tr>
</tbody>
</table>

Suspected Variables

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(hh:mm:ss)</td>
<td>(dd:mm:ss)</td>
<td>(mag)</td>
<td>(day)</td>
<td>(10^{-4} day)</td>
<td>(mmag)</td>
<td>(KBJD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W55</td>
<td>16:22:57.19</td>
<td>-26:29:09.1</td>
<td>18.22</td>
<td>1.85</td>
<td>100</td>
<td>7</td>
<td>2061.59</td>
<td>Harm.</td>
<td>...</td>
</tr>
<tr>
<td>W126</td>
<td>16:23:00.36</td>
<td>-26:27:32.7</td>
<td>18.92</td>
<td>0.041004</td>
<td>0.03</td>
<td>8</td>
<td>2060.29</td>
<td>Harm.</td>
<td>...</td>
</tr>
<tr>
<td>W951</td>
<td>16:23:19.93</td>
<td>-26:27:54.2</td>
<td>17.13</td>
<td>5.8</td>
<td>1000</td>
<td>1</td>
<td>2064.83</td>
<td>Harm.</td>
<td>...</td>
</tr>
<tr>
<td>W1779</td>
<td>16:23:29.93</td>
<td>-26:53:53.5</td>
<td>15.88</td>
<td>4.76</td>
<td>700</td>
<td>4</td>
<td>2063.40</td>
<td>Harm.</td>
<td>...b</td>
</tr>
<tr>
<td>W2571</td>
<td>16:23:37.87</td>
<td>-26:21:57.6</td>
<td>18.77</td>
<td>2.02</td>
<td>300</td>
<td>6</td>
<td>2061.34</td>
<td>Harm.</td>
<td>...</td>
</tr>
<tr>
<td>W2588a</td>
<td>16:23:38.06</td>
<td>-26:28:21.7</td>
<td>18.09</td>
<td>2.36</td>
<td>100</td>
<td>5</td>
<td>2060.74</td>
<td>Harm.</td>
<td>...</td>
</tr>
<tr>
<td>W3311i</td>
<td>16:23:40.5</td>
<td>-26:29:08.7</td>
<td>15.18</td>
<td>15.07</td>
<td>300</td>
<td>0.8</td>
<td>2073.72</td>
<td>Trap.</td>
<td>...</td>
</tr>
<tr>
<td>W3717</td>
<td>16:23:52.62</td>
<td>-26:23:14.4</td>
<td>19.22</td>
<td>0.11776</td>
<td>0.3</td>
<td>3</td>
<td>2060.35</td>
<td>Harm.</td>
<td>...</td>
</tr>
<tr>
<td>W3989</td>
<td>16:23:57.62</td>
<td>-26:29:13.2</td>
<td>18.79</td>
<td>0.112</td>
<td>...1</td>
<td>16.69</td>
<td>2060.32</td>
<td>Trap.</td>
<td>...</td>
</tr>
<tr>
<td>W4337</td>
<td>16:24:07.75</td>
<td>-26:27:37.3</td>
<td>15.45</td>
<td>0.4821</td>
<td>3</td>
<td>0.3</td>
<td>2060.72</td>
<td>Harm.</td>
<td>...</td>
</tr>
</tbody>
</table>

Note. — Classifications are not attempted for the suspected variables. Explanations regarding why these are reported as suspected instead of discovered variables can be found in Appendix 3.B.

aSee table notes for Table 3.3 for details on these columns.
bPeriod very close to a systematic period, but this object was kept as a variable owing to the strength of the signal.
cClassified as DSCUT by Yao & Tong (1989), but we do not observe the same variability that they report, and we think a GDOR classification is more likely to be correct.
dBlended with V19; this period appears in the data only after removing V19’s blended signal.
eLacked proper motion data to calculate membership probability.
fClassification from Yao et al. (2006a).
gBlended with W3825, which our code also marked as a variable; however, using a small aperture to evaluate differences in local flux amplitudes revealed this star to be the source of the variability.
hCluster membership probability is 0.067.
iEclipse was not identified by our main period-finding pipeline but was noticed in our by-eye vetting.
jThe trapezoid model struggled to fit well, and the calculated uncertainty on the period was unrealistically small, and so we decided to not report it.
3.5 Discussion

To our knowledge, only two other published works (other than Wallace et al. 2019b, which is based on Chapter 2) have presented results based on the K2 superstamp images of M4. Miglio et al. (2016) performed asteroseismology of K giants in M4, and Kuehn et al. (2017) looked at the RR Lyrae variables. We have already compared our results with those of Kuehn et al. (2017) in Section 3.4.2, and we compare our results with those of Miglio et al. (2016) here. Miglio et al. (2016) found evidence of solar-like oscillations in eight stars from their chosen set of 28 (chosen based on $B - I > 1.7$ and $V < 14$), or 29% of the stars. Making comparable cuts based on Gaia DR2 magnitudes and colors, $G_{BP} - G_{RP} > 1.25$ and $G < 14.0$, as well as including only those stars that have a >99% membership probability (see Table 3.1 and Appendix A), we end up with 55 stars in our chosen sample. Out of those stars, we find asteroseismic variability in 24 of them (W491, W508, W521, W799, W869, W1091, W1165, W1349, W1582, W1608, W1735, W1763, W2162, W2386, W2631, W2665, W2678, W2772, W2887, W3033, W3073, W3480, W3742, and W3996), or 44% of the stars, plus four suspected variables (W1068, W1717, W2577, and W3371). Note that five of these variables—W521, W799, W1608, W2386, and W2887—are included in the presentation of the HB variables in Section 3.4.3 and Figure 3.10. Restricting further to focus only on the largest giants, selecting those stars with $G < 12.7$ with the same color and membership cut as before, we end up with 18 stars in our sample, of which eight are identified as asteroseismic (multiharmonic) variables (W869, W1165, W1735, W1763, W2631, W2665, W2772, and W3996), or 44%. It would appear we were able to identify more asteroseismically active stars, both in number and in percentage, than Miglio et al. (2016). Of the eight stars they identified, their S1, S6, and S7 are in our edge region, so we do not have light curves for them. For the others, we match their S2 to our W2022, S3 to W2665, S4 to W760, S5 to W3033, and S8 to W3929. Our procedure detected variability
for only S3/W2665 and S5/W3033, though looking at the periodogram results for the other three, we would have definitely caught them had their periodograms been presented during a manual variability vetting. These objects did not make it to the by-eye portion of our variability search because they did not have sufficiently large periodogram SNRs, probably because of the very rich structure of the periodograms and the small differences in amplitude between the top periodogram peak and nearby peaks.

Other than these two papers, and our previous work in Wallace et al. (2019b) and Chapter 2, no other published work has used the M4 superstamp data. Given that it has been publicly available for over four years and has such rich potential, of which we believe this work has only scratched the surface, this is surprising. More generally, the cluster superstamps of K2 have received rather sparse attention, at least in terms of general variable searches (there have been a good number of searches targeted at specific stars). To our knowledge, the exhaustive list of general variability searches among K2 cluster superstamps is: work by LaCourse et al. (2015), Libralato et al. (2016a), and Soares-Furtado et al. (2017) for M35 and NGC 2185 in K2 Campaign 0; the work of Nardiello et al. (2016) for M67 in K2 Campaign 5; and the work of Libralato et al. (2016b) for Praesepe (M44) in K2 Campaign 5. Similar, though limited, work has been done for the K2 Campaign 9 microlensing superstamp (e.g. Zhu et al., 2017).

The incredible results from these cluster superstamp searches speak for themselves: Libralato et al. (2016a) presented a list of 2133 variables (out of 60,000 stars searched) for M35 and NGC 2158, and the work of Soares-Furtado et al. (2017) found 1151 variable stars from the same data (M. Soares-Furtado 2019, private communication). Libralato et al. (2016b) found 1680 variable stars—of which 1071 were new discoveries—in M44, and Nardiello et al. (2016) found 451 variable stars—of which 299 were new discoveries—in M67, not to mention the 94 variables in this work
(including the two mmRRs of Chapter 2), of which 76 are new, and 67 suspected variables, all of which are new. These new discoveries are valuable not just for better understanding the variable phenomena and/or the associated stars themselves, but with many belonging to either open or GCs, they can also help us learn more about these unique and astrophysically important environments. Focusing specifically on GCs like M4, eclipsing binaries—sometimes referred to as the “royal road” to stellar astrophysics (Russell, 1948)—can shed important light on the precise masses and radii of stars belonging to a (more or less) monolithic, metal-poor environment. Asteroseismic measurements can provide similar constraints on stellar properties for the evolved stars. Additionally, the as-yet elusive detection of a transiting exoplanet in a GC (despite previous efforts made by Gilliland et al. 2000; Weldrake et al. 2005, 2008 and Nascimbeni et al. 2012) could provide valuable clues on the dynamical and environmental histories of GCs. Our transiting exoplanet search for M4 is presented in Chapter 5.

Even more, M4 is not the only GC that has been observed by K2. M80 was observed concurrently with M4 during Campaign 2; M9, M19, NGC 6293, NGC 6355, and Terzan 5 were all observed during Campaign 11; and NGC 5897 was observed during Campaign 15. Given the increased distance of all of these clusters relative to M4, the data will be of lower quality and more crowded, but these are still potentially rich datasets nonetheless, for the giant stars if not for anything else. This untapped potential of the K2 cluster superstamps was recognized by Barentsen et al. (2018). Despite the crowding and the distance, the continuous nature and high precision of the observations make them very valuable datasets.

And finally, K2 will not be the end of such crowded, low-resolution, continuously observed data. The full-frame images from the Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2015) are providing similar data that, by the primary mission’s end, will cover nearly the whole sky. At an approximately five times larger pixel scale
than *Kepler*, observations of objects with a similar degree of crowding as M4 will
probably be hopelessly blended, but the outskirts of such objects as well as the cores
of sufficiently less compact objects (such as open clusters) will provide rich datasets,
with important discoveries for the making, if we can learn how to deal with such
crowding at scale.

To this end, we wish to reiterate some of the weaknesses of our present approach.
We do this not just to provide caveats to our present analysis but also to provide a
springboard for the community to improve upon our and others’ approaches as we
look to make the best use of TESS’s crowded data.

• Our roll-decorrelation procedure, based on that of Vanderburg & Johnson (2014),
does not work with large-amplitude variables, so our analysis of all of the
Clement et al. (2001) variables (see Table 3.2) could be improved by, e.g., a
simultaneous fit of the variability signal with the roll pattern.

• Also, our roll-decorrelation procedure fits out a B-spline with breakpoints set
nominally every 1.5 days, which we do not add back into the light curve. This
is likely to remove any long-term variability that may exist, and indeed in two
of the cases we examined (V13 and SC3), our final light curves did not exhibit
long-period variability that was apparent in the raw light curves.

• Since our primary variability selection criterion was based on periodogram SNR,
those objects with significant variability at a variety of periods may have low
periodogram SNR for otherwise robust variability due to extra noise included
in the calculation. As was discussed earlier in this section, we know this is a
problem for at least three of the asteroseismic oscillators that our code did not
mark as robustly variable (W760, W2022, and W3929) and some HB stars (see
Section 3.4.3), and we expect there are others.
• The blend identification and removal procedure in our code can be improved. One such improvement would be a more nuanced selection of fit for removal of signals—whether intrinsic or blended—for searching for additional variability. We employed an 11-harmonic Fourier series fit for removal of these signals. The reason we chose such a high-harmonic fit was to fit RRAB signals well, but in many cases having so many harmonics led to overfitting of the signal and introduced spurious signals of the same period but different shape into our light curves. Another improvement would be to include a more precise determination of variability period during the period search instead of after, since we found some cases where the detected period was off slightly from the true variability period, leaving significant signal of similar period in the residual due to the not-quite-correct period being used.

• Fainter stars that were very closely blended with considerably brighter, large-amplitude variable stars, as a result of the image subtraction photometric calculation, often had light curves with exceptionally high scatter. The stars also had many light-curve points that were unable to be calculated (e.g., image subtraction determined that at a variable minimum, the fainter star would have to have a negative flux to match the observed flux deficit and the calculation would thus fail). This itself is expected as a part of the image subtraction. However, because of this high scatter and systematically missing data, our seven-harmonic Fourier fit to determine the variability amplitude would often give an egregiously large value for the amplitude that would far exceed the amplitude for the variable itself. This meant that many of our highest-amplitude variables, for which detection should be most robust, were being marked as blends. We fixed this by requiring \( \Delta F/F_0 < 3 \) for an amplitude measure to be considered realistic and ignoring the amplitude otherwise. Better ways of avoiding this situation could certainly be implemented, such as determining from the light curves and
position information prior to the variability search which objects are likely to have these hopelessly blended, extreme light curves that would produce poor results in an amplitude determination.

- The harmonic fit method of amplitude determination did not always work well for the eclipsing binaries with narrow eclipses. For BLS searches, a more robust determination of signal amplitudes for comparison with neighbors and blend determination would be eclipse depth, determined either from the BLS fit itself or from another model fit, e.g. a trapezoid model. We reran our BLS search with eclipse depth as the amplitude determination but did not find any additional variables. This modification to \texttt{simple_deblend} is not yet implemented in the main branch, which is why we mention it here.

- While our selection of which aperture to use for an object of a given magnitude was based on a superstamp-wide evaluation of light-curve scatter versus magnitude, it may be that in the more crowded regions, smaller-than-globally expected apertures produce less scatter. A more robust determination of this could be useful.

- We do not treat saturated stars in any special way.

- Our variability search produced 1310 objects (out of 4554 searched) with purported robust variability. Our by-eye selection and manual blend determination reduced this to 161. Relying so heavily on a manual and qualitative final vetting step is less than ideal and likely to lead to incorrect determinations in some of the marginal cases. Reducing the amount of manual work involved in variable identification and classification is, of course, a long-standing problem in variable astronomy, and much headway is being made. Specifically for these data, it is likely that additional quantitative quality cuts could be determined to further pare down the number of objects that need to be searched by eye.
• We only examined objects with a Gaia DR1 $G < 19$. While in the crowded regions all fainter objects were essentially included since the apertures for the included sources overlapped and covered the whole image, many stars of potential interest in the less crowded regions of the images were not included. Since we discovered variables all the way down to the $G = 19$ cut we made (see, e.g., the blended pair W283 and W293 in Table 3.5), there may very well be other variables, both cluster members and nonmembers, to be discovered in this fainter population.

As mentioned in Section 3.2, this chapter is intended primarily as a work of breadth rather than depth. The light-curve processing and results are presented, but analysis of the individual variable objects is limited to only a very few of them, and the analysis is very limited at that. There is much that could be done with these data, and since our light curves are publicly available at Wallace et al. (2019a), we invite any and all interested in these objects to perform their own analyses in greater depth. Some potential jumping off points include: detailed analysis of the RR Lyrae variables and further comparison with Kuehn et al. (2017); detailed analysis of the asteroseismically active giants and comparison with Miglio et al. (2016); further analysis of the asteroseismically active HB stars and their connections with what we have called mmRRs, and what connection (if any) these may have with the RR Lyrae variables; cross-matching our identified non-cluster-member variables with available photometric catalogs to see if their variability could be classified; searching for long-period variables via a different light-curve processing pipeline; observational follow up on our blended objects (Table 3.5) to determine which are the actual sources of variability; radial velocity follow up of the eclipsing binaries; follow up, perhaps with an X-ray telescope, of our likely X-ray binary; and spectroscopic follow up and characterization of all new variables presented in this work. These light curves
represent the longest continuously observed GC with reduced data and, as such, have a myriad of potential uses.

This work and the others mentioned here that worked on the K2 open clusters demonstrate the efficiency of superstamp-style observations of crowded regions. For the 40,000 pixels of the K2 superstamp, we derived light curves for 4554 objects, or \( \sim 8.8 \) pixels per object. This is not including the objects in our edge region for which one could still extract light curves. To be comparably efficient, the stamp size for observing isolated targets would have to be \( \sim 3 \) pixels by \( \sim 3 \) pixels. This demonstrates how, for missions with limited data downlink bandwidth, observations of crowded regions can be an efficient way to maximize stars observed per pixel of data, with the tradeoff of blending.

### 3.6 Conclusions

We extracted light curves for 4554 objects in the GC M4 from the K2 superstamp data of the cluster. With \( \sim 78 \) days of continuous observations represented in the final light curves, these are, by far, the longest continuous light curves ever reduced for a GC, and monitored at the high precision that Kepler/K2 provides. We employ image subtraction to extract our raw light curves, then clean up the data using a roll-decorrelation procedure based on that of Vanderburg & Johnson (2014) and removing common trends in the data using TFA. Our final photometric precision is 0.2 mmag for \( G \approx 12 \), 1 mmag for \( G \approx 15 \), and 10 mmag for \( G \approx 18 \) objects, with M4’s main-sequence turnoff being around \( G \approx 16–17 \). We make these light curves publicly available (Wallace et al., 2019a).

We also searched for periodic variability in our light curves using the GLS, PDM, and BLS algorithms. We find 66 variables and 57 suspected variables that are cluster members, 24 variables and 10 suspected variables that are not cluster members, and
four where cluster membership is ambiguous. Of these, 52 cluster members (when including the two mmRRs of Chapter 2) and 20 cluster nonmembers, as well as all four of the variables with ambiguous membership and all of the 67 suspected variables, are new discoveries. Our number of newly discovered cluster-member variables is three times greater than the total number of cluster-member variables discovered in this area of the sky (K2 superstamp minus the edge region) in all previous surveys. Of note among cluster members are seven asteroseismically variable HB stars, a slightly eccentric $\sim 4.6$ day eclipsing binary cluster member, a $\sim 0.20$ day EW binary, a likely X-ray binary with quiescent periodic optical variability, and a $\sim 0.27$ day EW binary that is highly likely to be a cluster member. Among non-cluster members, we discover a slightly eccentric $\sim 22$ day eclipsing binary with apparent reflection effects just before and after transits.

This is just the starting point for the analysis of many of these objects. Miglio et al. (2016) performed an asteroseismic analysis for two of the asteroseismically active giants we identified, but there remain over 20 from this work to be analyzed, and more to be identified. The asteroseismic variability of the HB stars in particular are of interest in understanding the mmRRs first presented in Wallace et al. (2019b) (and also presented here in Chapter 2), and none of the seven variable non-RR-Lyrae HB stars (see Figure 3.10 and Section 3.4.3) have received an asteroseismic analysis. Additional analysis is needed to understand the large number of unclassified variables we present in this work, both in and out of the cluster. The results of this work are the longest continuously observed light curves ever derived for general GC stars, and we anticipate much to come from the data.

The research in this chapter utilized the following software: astrobase (Bhatti et al., 2017), astropy (Astropy Collaboration et al., 2018), FITSH (Pál, 2012), k2mosaic (Barentsen, 2016), matplotlib (Hunter, 2007), numpy (Oliphant, 2006), scikit-learn
(Pedregosa et al., 2011), simple deblend (Wallace & Hoffman, 2019), scipy (Jones et al., 2001), and VARTOOLS (Hartman & Bakos, 2016).

Appendix 3.A Notes on Identified Blends

This appendix provides a detailed look into blends that were manually assessed and removed by us after the automatic processing described in Section 3.3. This discussion is intended primarily as a record of the blends we manually assessed and/or a reference for those who wish to more completely understand the systematics in our search.

Despite the reasonably robust performance of our automated blend detection method, there still remained many blends in the final set of detected periods. Some reasons for the residual blends include: blending with or photometric footprinting by a variable object that was further away than our chosen search radius of 12 pixels or objects with particularly small separations ending up with similar flux amplitudes in their variability due to the amount of overlap in their apertures. In the latter case, there were some objects for which we were able to disentangle which was the real variable, while Table 3.5 records those objects which we were not able to disentangle.

Though the accounting here is fairly exhaustive, we did not record all instances of stars that were clear blends with the RR Lyrae variables based on proximity, period, and light-curve properties. Despite choosing the 12 pixel blend search radius based on results in the neighborhood of two RR Lyrae variables in our images, there were still some stars outside this radius for other RR Lyrae variables that were blended with those variables.

Many stars had similar variability and the same period and phase as V19. These were all $\sim$12–18 pixels away from the star and predominantly clustered together. We do not know for sure what caused this relatively distant blending. We checked all of the stars with periods and phases that matched V19 to make sure none were
obviously their own variable before excluding them from further consideration. The stars thus excluded were W1820, W1836, W1838, W1995, W2007, W2205, W2264, W2316, W2381, W2413, W2420, W2439, W2467, W2540, W2583, W2600, W2626, W2695, W2701, W2748, W2774, W2776, and W2777. There were also three stars that were 38–41 pixels away in rough relative proximity to each other that were 180° out of phase with V19 and had the same period. These were also excluded after a visual check of their light curves: W1948, W1960, and W2201.

The following stars were all blended with each other and all have the same period as V27. They are also all ~33-36 pixels away from V27. The signals look like inverted RRAB signals, so it may be some systematic from our data reduction. All of these were removed from consideration: W3232, W3234, W3246, W3248, W3262, W3285, W3296, and W4540.

W3623 has the same period and nearly same phase as V9 with a similar shape, despite being over 80 pixels away. We removed W3623 from consideration because of this.

W285 has the same period as V35 from Clement et al. (2001) and also looks like an RRAB, which V35 is. Thus, we consider W285 as a blend with V35 even though we do not have a light curve for V35.

W2398 is blended with ~0.47 day period V19 and thus its ~0.12 day variability detected by GLS is discounted by us, and we marked it as not a variable. Closer examination may be able to determine whether this is a correct call or not.

There were several stars in close proximity to each other with variabilities of approximately the same period as V29 but that did not phase up with V29, and were also ≥100 pixels away from V29. However, V28 in the catalog of Clement et al. (2001) has nearly the same period as V29 and, while not included in the K2 superstamp of M4, is only ~11–15 pixels away from most of these stars (one was 27 pixels away). Based on this, we decided to mark the following stars as blends with V28 given the
proximity, after a visual check of their light curves: W3678, W3735, W3796, W3811, W3848, W3854, W3880, and W3914. Additionally, W2709 phased up with V29 and was marked as a blend despite being ~125 pixels away.

W1097 is hopelessly blended with the bright variable W1165. Looking at the respective light curve, W1097’s light curve was excessively noisy (likely due to blending with the much brighter star) and the variability was not nearly as apparent as for W1165. We thus removed W1097 from consideration.

Many stars shared a similar ~1.95 day period and phased up with each other. This period is approximately the same period (1.962 days) as the resaturation events, producing a blank image at this period. These stars were all assumed to share a common systematic based on the resaturation events and removed from further consideration. These were W144, W221, W335, W338, W391, W528, W678, W2098, W2286, W2694, W3178, W3785, and W3955. Additionally, other stars were found with this similar period that did not quite phase up with the others (though some were 180° out of phase) but were still assumed to have a similar systematic unless visual inspection of their light curve revealed otherwise. These objects were W83, W610, W2040, W2309, W3306, W3779, W4062, W4083, W4096, W4177, W4293, W4318, and W4534. Upon visual inspection of the light curves, W92 and W4268 were kept as a variable (W92) or suspected variable (W4268) owing to the strength of their signal despite having periods around this systematic. W4490 was also kept as a variable owing to its high-amplitude variability.

W321, W470, W548, W566, W569, W645, and W692 all had the same variability period, phase, and shape, and were all in about the same area of the image. The apertures were not all quite overlapping. Of these, W566 had the most robust detection of the variability (detected by both GLS and PDM instead of just PDM, and also had the highest periodogram SNR) and so we decided to call that the variable
but wanted to record here the other stars that were blended with it. All are \(\sim 6-13\) pixels away from W566.

W1938, W2805, and W4143 all have \(\sim 3.4\) day transits. W1938 and W2805 even phase up based on a sine curve fit to the variability. However, these stars are all very separated. W1938 and W4143 are included as suspected variables in this work, in Table 3.4.

### Appendix 3.B Suspected Variables

This appendix presents results for our suspected variables. The suspected variables can be found in the corresponding sections of Tables 3.4 and 3.6. The phase-folded light curves are shown in Figures 3.17, 3.18, 3.19, 3.20, and 3.21. There are a few objects of particular note in this collection.

W1834 in Figure 3.18 is a cluster member with a \(\sim 5\) mmag box-shaped brightening in the light curve, occurring at a 9.29 day period. We consider the possibility that this is a gravitational self-lens from a neutron star/black hole in a binary with a main-sequence star. Figure 1 from Masuda & Hotokezaka (2018) shows that the amplitude and period are consistent with self-lensing from a \(\sim 10\ M_\odot\) black hole; however, their Equation (7) reveals that for a circular orbit, such a system would have a signal duration of \(\sim 1\) hr, much shorter than the \(\sim 20\) hr observed. If this is a self-lensing black hole system, it would have to be very eccentric. We would expect ellipsoidal variability in such a case during a pericenter passage, but we do not see anything larger than our \(\sim 0.1\) mag floor in the raw light curves.

Similarly, W2127 in Figure 3.21 is a cluster member that has a single observed \(\sim 50\) mmag brightening event over a \(\sim 5\) day period. Extrapolating from their Figure 1 and again using the Equation (7) from Masuda & Hotokezaka (2018) as before, a \(\sim 10\ M_\odot\) black hole on a \(\sim 250\) day circular orbit would broadly match the observed light
Figure 3.17: Same as Figure 3.11, but for suspected variables and for a mixture of cluster members and nonmembers. The first 15 suspected variables are shown in this figure, with the rest of the suspected variables shown in Figures 3.18, 3.19, 3.20, and 3.21. Cluster membership is indicated below the object identifier in the upper right corner of each panel: “M” means cluster member (specifically, that the cluster membership probably is >99%), while “NM” means not a cluster member (specifically, that the cluster membership probably is <1%).
Figure 3.18: Same as Figure 3.17, but for additional suspected variables. Instead of indicating “M” or “NM” for W1799’s cluster membership, we record the membership probability since it was not <1% or >99%. 
Figure 3.19: Same as Figure 3.17, but for additional suspected variables. “N/A” for W2588’s cluster membership status means cluster membership information not available since there are not Gaia DR2 proper motions reported for this object.
Figure 3.20: Same as Figure 3.17, but for additional suspected variables.
Figure 3.21: Same as Figure 3.17, but for additional suspected variables. W2127's light curve is unphased since only a single event was found.
curve. Of course, these situations would require an orbital inclination near 90°, which for the wide orbit of W2127 presents something of a fine-tuning problem, as does the large eccentricity needed for W1834. We merely present these as possible scenarios and do not conclude anything on the nature of the variability on these objects.

We list here the reasons we have for marking each of the suspected variables as suspected rather than definite variables.

- W55: Noisy periodogram; low-amplitude phase-folded light curve.
- W58: Many light-curve points from second half of campaign are missing due to blending with bright star.
- W126: Very short period, ~0.3% away from twice the cadence period.
- W267: Phase-folded light curve has low amplitude.
- W371: Noisy periodogram.
- W435: Noisy periodogram and phase-folded signal has low amplitude.
- W461: Very nearby to W491 and might be blended, W461’s period is a bit more than 14 times the period of W491.
- W829: Small transit depth compared to light-curve scatter.
- W901: Noisy periodogram.
- W920: Noisy periodogram and phase-folded signal has low amplitude.
- W951: By-eye judgment call that it is unclear whether this could be a real transit or not.
- W1056: By-eye evaluation makes it unclear whether this could be a real transit or not.
- W1068: Noisy periodogram.
- W1208: Noisy periodogram.
- W1222: Binned-median points show some bright points in transit in addition to the dimmer points filling out the transit.
- W1263: Noisy periodogram.
- W1539: Noisy periodogram and phase-folded signal has low amplitude.
• W1717: Noisy periodogram and phase-folded signal has low amplitude.
• W1725: Phase-folded light curve has low amplitude, ambiguous by eye.
• W1779: Near a saturated star; similar period to W1864, which is also near the same saturated star.
• W1809: Small transit depth compared to light-curve scatter.
• W1834: Scatter in anti-transit portion of phase appears to be smaller than the rest of the light curve.
• W1864: Near a saturated star; similar period to W1779, which is also near the same saturated star.
• W1938: Noisy periodogram.
• W1947: In a very crowded area of the image; rich, possibly noisy, periodogram.
• W1953: Possible transit, but depth is not large and not very wide.
• W2109: Periodogram peak similar in amplitude to other periodogram peaks, but phase-folded signal looks like it could be real.
• W2126: Noisy periodogram; phase-folded signal has low amplitude.
• W2127: Signal occurs close to the time the spacecraft’s roll changed directions, producing systematics in other light curves around this time, but this is a stronger signal than those other systematics.
• W2233: Noisy periodogram; looks like an RRAB signal and has close to the same period as V9, but they do not quite phase up.
• W2272: Blending with bright object, producing differing noise characteristics in second half of data relative to first half, may be producing some kind of unique systematic.
• W2324: Noisy periodogram; period matches W1189 and is 180° out of phase, but it is over 55 pixels away.
• W2499: Noisy periodogram.
• W2515: Noisy periodogram.
• W2543: Transit not very deep compared to noise.
• W2556: Only two transits observed.
• W2571: Noisy periodogram; strange shape to periodogram peak.
• W2577: A bright star blended with another bright star for which we do not have light curves since they are not Gaia DR1 sources; thus unsure whether this is the source of variability (though very likely it is).

• W2588: Noisy periodogram.

• W2616: Noise characteristics changed halfway through campaign.

• W2641: Noisy periodogram.

• W2747: Low-amplitude transit signal.

• W2753: Based on period, it might be a transformed blend of V29.

• W2790: Phase-folded light curve has low amplitude.

• W2800: Phase-folded light curve has low amplitude.

• W2819: Maybe a transit present, but differing noise characteristics in second half of data relative to first half may be producing some kind of unique systematic.

• W2876: Noise characteristics change slightly halfway through campaign; noisy periodogram.

• W2893: Noisy periodogram.

• W2966: Phase-folded light curve has low amplitude.

• W3105: Six pixels away from and similar variability to V27, but does not phase up. However, we have seen our light curve processing transform blended RRAB signals into sinusoidal signals with slightly different periods.

• W3125: Noisy periodogram.

• W3282: Phase-folded light curve of particularly small amplitude.

• W3311: Noisy periodogram.

• W3313: Low-amplitude transit signal.

• W3371: Phase-folded light curve has low amplitude.

• W3521: Noisy periodogram.

• W3552: Noisy periodogram.

• W3717: Noisy periodogram; phase-folded light curve has low amplitude.

• W3887: Small transit depth compared to light-curve scatter.
• W3901: Low-amplitude signal. Period matches V29, but does not phase up, and is over 120 pixels away.

• W3989: Noisy periodogram.

• W4014: Noisy periodogram; phase-folded signal has low amplitude.

• W4143: Small transit depth compared to light-curve scatter.

• W4250: Near to a bright star that was in the edge region. We do not have the light curve for the bright star to see if this signal is a blend.

• W4268: Period falls within the 1.95 day systematic range, but we still decided to keep as a suspected variable based on signal strength.

• W4301: Noisy periodogram.

• W4337: Noisy periodogram; phase-folded light curve has low amplitude.
Chapter 4

Collisional Fragmentation is not a Barrier to Close-in Planet Formation

4.1 Introductory Note—Why Is This Chapter Included?

This chapter (also published as Wallace et al. 2017) is sufficiently distinct from the other chapters that some explanation regarding its inclusion is merited. It presents theoretical work investigating rocky planet formation at the most close-in regions around a host star, determining whether or not collisional fragmentation could prevent planet formation in these regimes. It is included because it is part of the motivation for the planet search in Chapter 5. Due to the stellar densities in globular clusters, one might expect that formation of planets further out from host stars is less likely to occur because of increased close encounters with neighbor stars relative to the field population, among other reasons (see Section 5.6 for a more complete discussion). Even if they are able to form, their longevity is likely to be compromised, again...
because of close encounters with neighbor stars. Thus a search for planets in globular clusters might expect to only find planets in short-period orbits, but the physical possibility of forming such planets in situ is a still a theoretical question. If they did not form in situ, then they would have to form further out, where globular cluster stars might have difficultly forming planets. Thus the primary result of this chapter—that collisional fragmentation is not a barrier to planet formation—increases our confidence that we might find close-in planets around stars in a globular cluster. Since this is precisely the data available for our planet search (i.e., data that allows discovery of short-period planets), it also increases our confidence that we might find planets there. Unfortunately, the ultimate sensitivity of our planet search in Chapter 5 was not good enough to detect the small rocky planets investigated in this chapter.

4.2 Abstract

Collisional fragmentation is shown to not be a barrier to rocky planet formation at small distances from the host star. Simple analytic arguments demonstrate that rocky planet formation via collisions of homogeneous gravity-dominated bodies is possible down to distances of order the Roche radius ($r_{\text{Roche}}$). Extensive N-body simulations with initial bodies $\gtrsim 1700$ km that include plausible models for fragmentation and merging of gravity-dominated bodies confirm this conclusion and demonstrate that rocky planet formation is possible down to $\sim 1.1 r_{\text{Roche}}$. At smaller distances, tidal effects cause collisions to be too fragmenting to allow mass build-up to a final, dynamically stable planetary system. We argue that even differentiated bodies can accumulate to form planets at distances that are not much larger than $r_{\text{Roche}}$. 
4.3 Introduction

Exoplanetary systems possess a large and surprising diversity of architectures. In particular, many systems contain one or more planets with periods $\lesssim 10$ days. A partial list of examples includes Kepler–42, a $\sim 0.13 \, M_\odot$ star with three known planets, all rocky, with periods of $\sim 0.5$, $\sim 1.2$, and $\sim 1.9$ days (Muirhead et al. 2012); Kepler–32, a $\sim 0.58 \, M_\odot$ star with five known planets, having periods between $\sim 0.7$ and $\sim 22$ days (Muirhead et al. 2012); and TRAPPIST–1, a $\sim 0.08 \, M_\odot$ star with seven known planets, all with radii $\sim 1 \, R_\oplus$ and periods between $\sim 1.5$ and $\sim 19$ days (Gillon et al. 2017).

A central question is whether (i) such close-in planetary systems can form in situ, or (ii) they formed further out from their host star and migrated inwards to their current configuration. There are difficulties with both mechanisms. (i) Assuming solar metallicity, the gas surface density of a disk containing sufficient metals for in situ formation of close-in planets such as those highlighted above would be gravitationally unstable (Raymond & Cossou, 2014; Schlichting, 2014). Such disks should therefore form giant rather than small rocky planets. Moreover, even if the disk were stable, the high temperatures ($\sim 2000$ K) expected in the inner disk would prevent the condensation of dust within 0.1 AU (D’Alessio et al. 1998). A possible solution is that inward migration of solid material through the protoplanetary disk could allow for sufficient planet-building material to accrue at small semimajor axes without an unstable buildup of gas (e.g., Youdin & Shu, 2002; Youdin & Chiang, 2004; Chiang & Youdin, 2010; Chatterjee & Tan, 2014, 2015). (ii) Although the physics of planet migration is robust, the behavior of migrating planets depends sensitively on the properties and physics of protoplanetary disks (e.g., Baruteau et al., 2014). Simple models of migration incorrectly predict that most short-period planets should be in orbital resonances, and cannot explain the large numbers of short-period planets discovered by the *Kepler* spacecraft (e.g., Benz et al., 2014).
In this chapter, we examine another possible barrier to in situ formation of close-in planets: collisional fragmentation. Even if planetesimals or planetary embryos can form at or migrate to these short-period orbits, would the relative velocities between such bodies allow for merging and growth, or would they be sufficiently large that collisions would be primarily fragmenting? In the latter case, planetesimals would not be able to grow into planets. We address this question both analytically and with N-body simulations.

The destruction of bodies due to collisions has been studied in depth in the context of asteroids, the Kuiper belt, and debris disks (e.g., Dohnanyi 1969; Wyatt & Dent 2002; Dominik & Decin 2003; Kobayashi & Tanaka 2010; Kenyon & Bromley 2017b; Pan & Sari 2005). However, the bodies focused on in these studies are $\lesssim 100$ km in size (and many are $\lesssim 1$ km), much smaller than the bodies that will be examined in this chapter. The physics of fragmentation is different in small and large bodies, with small bodies ($\lesssim 0.1$–1 km for rocky bodies) being held together primarily by their internal strength (and thus affected by internal flaws and cracks) and large bodies being held together primarily by their gravitational self-attraction (although many small bodies are actually rubble piles with little internal strength; we note Pan & Sari 2005 assumed this in their collision models). The results of a fragmenting collision also depend on body size: larger bodies have significant escape velocities and thus some of the fragments from an impact can be reaccumulated. Collision velocities between large bodies, especially late in the planet-formation process when they are on well-separated and nearly stable orbits, can also differ from those of small bodies, which have random velocities that are themselves often set by stirring from the large bodies. In particular, the models of debris-disk production via collisional cascades are generally based on a bimodal mass distribution with the largest bodies having undergone runaway growth and stirring up the smaller bodies to sufficiently high random velocities to have destructive collisions. In this picture, the larger remnants
from runaway growth are what eventually form the final planets and are the focus of the present study. For these reasons, results from collisional cascade models of small bodies cannot be assumed to carry over to the large bodies that eventually form planets.

In Section 4.4 we provide an analytic motivation for our study, then in Section 4.5 we describe the methods of our N-body simulations. We present the results from the simulations in Section 4.6, provide a discussion in Section 4.7, and conclude in Section 4.8.

## 4.4 Analytic Motivation

### 4.4.1 Preliminaries

Here we present an order-of-magnitude argument describing the necessary conditions for collisions between gravity-dominated bodies to be erosive. We define “erosive” to mean that the largest remnant after a collision is less massive than the larger of the two colliding bodies. Throughout this chapter, we always assume that the colliding bodies have sufficient mass that their gravitational binding energy is greater than their molecular binding energy (the gravity-dominated regime). We also assume that collisions are between two spherical bodies of the same density, having masses $m_1$ and $m_2$ and radii $R_1$ and $R_2$. We specify $m_1 \geq m_2$ and, following common convention, call the larger body involved in the collision the “target” and the smaller body the “projectile”. In this section, we assume all orbits are coplanar (no mutual inclination).

The amount of fragmentation that occurs in a collision depends on the ratio between the specific impact energy in the center-of-mass frame $Q$ and some threshold energy $Q^*$, defined to be the specific impact energy necessary to disperse half the
mass involved in the collision. The impact energy is

\[ Q = \frac{\mu (\Delta v)^2}{2(m_1 + m_2)}, \quad (4.1) \]

where \( \mu \) is the reduced mass of the two colliding bodies and \( \Delta v \) their collision velocity (the relative velocity just before impact). From, e.g., Stewart & Leinhardt (2012)

\[ Q^* = \left( \frac{\mu}{\mu_\beta} \right)^{3/2} \left[ \frac{(1 + \alpha)^2}{4 \alpha} \right] \frac{4}{5} C^* \pi \rho_1 G R_{C1}^2, \quad (4.2) \]

where \( \mu_\beta \) is the reduced mass of the target body and the fraction \( \beta \) of the projectile body that intersects the target body during a collision, which for off-center collisions may be less than unity (see Stewart & Leinhardt 2012 for details), \( \alpha = m_2/m_1 \), \( C^* \) is a dimensionless factor that accounts for the dissipation of energy in the target body, \( \rho_1 = 1 \text{ g cm}^{-3} \), \( G \) is the gravitational constant, and \( R_{C1} \) is the radius of a body with density \( \rho_1 \) and a mass equal to the sum of the projectile and target masses,

\[ \frac{4\pi \rho_1}{3} R_{C1}^3 = m_1 + m_2. \quad (4.3) \]

Leinhardt & Stewart (2012) (hereafter LS12) found the best-fit value of \( C^* \) for planet-sized bodies is 1.9±0.3; Chambers (2013) (hereafter C13) used \( C^* = 1.8 \) and we use that value in this chapter. Note that the expressions of Equation (4.2) in Stewart & Leinhardt (2012) (primarily their equation 10) are more general because they contain an extra parameter, the velocity exponent \( \mu \). Following C13 and consistent with the best-fit values found by Stewart & Leinhardt (2012), we use \( \mu = 1/3 \) to get Equation (4.2).

In this section, we assume that collisions are between bodies on initially well-separated, near-circular orbits whose eccentricities have been gradually excited by mutual perturbations. This situation arises, for example, in the late stages of planet
formation after oligarchic growth is complete and the population of residual small bodies has decayed. (This is in contrast to the early stages of planet formation, where there is a much higher density of colliding bodies and the random velocities are set by the escape velocity of the largest bodies.) If the two bodies initially have semimajor axes $a \pm \Delta a/2$, with $\Delta a \ll a$, then their relative speed $(\Delta v)_{\text{rel}}$—when they are close to colliding but far enough away that gravitational focusing is unimportant—will be

$$(\Delta v)_{\text{rel}} = f_1 \left( \frac{GM_*}{a} \right)^{1/2} \frac{\Delta a}{a} = f_1 v_c \frac{\Delta a}{a}, \quad (4.4)$$

where $M_*$ is the mass of the central star, $v_c = \sqrt{GM_*/a}$ is the Keplerian circular velocity at $a$, and $f_1$ is a factor of order unity. Two such bodies are unable to collide unless the eccentricity $e \gtrsim \Delta a/a$. Including gravitational focusing, the collision velocity $\Delta v$ between these two bodies will be

$$\Delta v = [(\Delta v)_{\text{rel}}^2 + v_e^2]^{1/2}, \quad (4.5)$$

where $v_e$ is the mutual escape velocity, which we define as

$$v_e = \sqrt{2G \frac{m_1 + m_2}{R_1 + R_2}}. \quad (4.6)$$
4.4.2 Equal-mass Case

Assuming the bodies have the same mass \( m \) and density \( \rho \) (and thus the same radius \( R \)) and have a head-on collision so \( \mu = \mu_\beta \), the specific impact energy scaled by \( Q^* \) is

\[
\frac{Q}{Q^*} = \frac{\mu}{2(m_1 + m_2)} \frac{(\Delta v)^2}{Q^*} - \frac{1}{8} \left[ \frac{5f_1^2 M_s (\Delta a)^2}{2\pi \rho_1 C^* a^3 R_{C1}^2} + \frac{v_e^2}{Q^*} \right]
\]

(4.7)

\[
\frac{Q}{Q^*} = \frac{1}{8} \left[ \frac{5f_1^2 M_s (\Delta a)^2}{2^{4/3} 3^{2/3} \pi^{1/3} \rho_1^{1/3} C^* a^{2/3} m^{2/3}} + \frac{v_e^2}{Q^*} \right].
\]

(4.8)

Eliminating \( \rho_1 \) using Equation (4.3), we get

\[
\frac{v_e^2}{Q^*} = \frac{5(m_1 + m_2)}{2\pi \rho_1 C^* (R_1 + R_2) R_{C1}^2}.
\]

(4.11)

For arbitrary masses,

\[
R_{C1} = \left[ \frac{3(m_1 + m_2)}{4\pi \rho_1} \right]^{1/3},
\]

(4.13)

\[
R_1 + R_2 = \left( \frac{3}{4\pi \rho} \right)^{1/3} (m_1^{1/3} + m_2^{1/3}).
\]

(4.14)
Hence, for equal-mass bodies,

\[
\frac{Q}{Q^*} \approx \frac{5f_1^2 M_* (\Delta a)^2}{2^{11/3} 3^{1/3} \pi^{1/3} C^* a^3 m^{2/3}} + \frac{5}{3C^* 2^{8/3}} \left( \frac{\rho}{\rho_1} \right)^{1/3}.
\]  

(4.15)

Using the mutual Hill radius, \(r_{\text{Hill}}\),

\[
r_{\text{Hill}} = a [ (m_1 + m_2) / (3M_*) ]^{1/3},
\]  

(4.16)

Equation (4.15) can be rewritten in terms of \(r_{\text{Hill}}\),

\[
\frac{Q}{Q^*} \approx \frac{5f_1^2}{2^{11/3} 3^{1/3} \pi^{1/3} C^*} \left( \frac{M_*}{\rho_1 a^3} \right)^{1/3} \frac{(\Delta a)^2}{r_{\text{Hill}}^2} + \frac{5}{3C^* 2^{8/3}} \left( \frac{\rho}{\rho_1} \right)^{1/3}.
\]  

(4.17)

The criterion that head-on collisions of equal-mass bodies are non-erosive is \(Q < Q^*\). For later use we shall write this as \(Q < f_2 Q^*\) with \(f_2 = 1\) for equal-mass bodies. Thus we obtain a necessary condition for accumulation,

\[
\frac{5f_1^2}{2^{11/3} 3^{1/3} \pi^{1/3} C^*} \frac{M_*}{\rho_1 a^3} \left( \frac{\Delta a}{r_{\text{Hill}}^2} \right)^{1/3} \approx f_2 - \frac{5}{3C^* 2^{8/3}} \left( \frac{\rho}{\rho_1} \right)^{1/3}.
\]  

(4.18)

Equation (4.18) sets a maximum value for the initial separation \(\Delta a\) at which two colliding bodies avoid an erosive collision. For future use, we note that the relation between \(f_2\) and the minimum collision velocity for erosive head-on collisions between equal-mass bodies can be written

\[
\frac{(\Delta v)^2}{v_e^2} = \frac{3 \cdot 2^{8/3}}{5} f_2 C^* \left( \frac{\rho_1}{\rho} \right)^{1/3}.
\]  

(4.19)

110
If the typical separation in the post-collision system (which we call $\Delta a'$) is too small, then the system will be dynamically unstable and eventually undergo another collision. An approximate stability criterion is $\Delta a' \gtrsim f_3 r_{\text{Hill}}'$, where $f_3 \simeq 10$ for stability over $10^{10}$ orbits (Pu & Wu 2015) and the primes are used to distinguish post-collision properties from pre-collision properties. We note the chaotic nature of collisional growth, but still choose to employ this criterion as it is a necessary one for stability in the long-lived post-collision stage, not the short-lived chaotic collision stage. This gives us a lower bound on dynamically stable $\Delta a'$, while Equation (4.18) gives us an upper bound on $\Delta a$ for collisions to be non-erosive. If we assume a system of bodies with equal masses $m$, separated from their nearest companions in semimajor axis by $\Delta a$ just prior to their last collision before stability, and if we also assume that adjacent bodies collide and merge pairwise (this ignores the possibility of higher-eccentricity collisions between non-adjacent bodies), then the post-collision bodies will be separated from each other by $\Delta a' \simeq 2\Delta a$ and will have $r_{\text{Hill}}' \simeq 2^{1/3} r_{\text{Hill}}$. Thus, the post-collision stability condition $\Delta a' \gtrsim f_3 r_{\text{Hill}}'$ can be rewritten in terms of pre-collision quantities, $2^{2/3} \Delta a \gtrsim f_3 r_{\text{Hill}}$. Plugging this into Equation (4.18) to eliminate $\Delta a$, we obtain a range for $a$ in which collisions will not be erosive up to the final collision before long-term dynamical stability,

$$a \gtrsim \frac{5 f_1^2 f_3^2}{32 \pi^{1/3} 3^{4/3} C^*} \left( \frac{M_*}{\rho_1} \right)^{1/3} \times \left[ f_2 - \frac{5}{3 C^* 2^{8/3}} \left( \frac{\rho}{\rho_1} \right)^{1/3} \right]^{-1}.$$  \hspace{1cm} (4.20)

Remarkably, this criterion is independent of the planetesimal mass $m$. 

111
The Roche radius $r_{\text{Roche}}$ for a homogeneous body of density $\rho$ is

$$r_{\text{Roche}} = 1.523 \left( \frac{M_*}{\rho} \right)^{1/3}$$

$$= 0.0089 \text{ AU} \left( \frac{M_*}{M_\oplus} \frac{3 \text{ g cm}^{-3}}{\rho} \right)^{1/3}. \quad \text{(4.21)}$$

If a gravity-dominated body on a circular orbit has $a < r_{\text{Roche}}$, it is disrupted by tidal forces. Writing Equation (4.20) in terms of $r_{\text{Roche}}$ and using $C^* = 1.8$, we get

$$a \gtrsim \frac{5 f_1^2 f_3^2}{32 \pi^{1/3} 3^{1/3} C^*} \left( \frac{\rho}{\rho_1} \right)^{1/3} \frac{r_{\text{Roche}}}{1.523}$$

$$\times \left[ f_2 - \frac{5}{3 C^* 2^{8/3}} \left( \frac{\rho}{\rho_1} \right)^{1/3} \right]^{-1}$$

$$\gtrsim 0.013 \frac{f_1^2 f_3^2 \eta^{1/3}}{f_2 - 0.21 \eta^{1/3} r_{\text{Roche}}}, \quad \text{(4.22)}$$

where $\eta = \rho/3 \text{ g cm}^{-3}$ and $\rho/\rho_1 = 3\eta$. This is the semimajor axis range in which the last collisions before long-term dynamical stability will not be erosive. Apart from the dimensionless factors $f_1, f_2, f_3$, the value of $a$ at which a system transitions from non-erosive to erosive collisions is less than $r_{\text{Roche}}$.\(^1\) Since gravity-dominated bodies cannot survive within $r_{\text{Roche}}$, to decide whether collisional fragmentation is an additional barrier to growth we must determine whether the transition point is inside or outside $r_{\text{Roche}}$.

We now determine reasonable values for $f_1$, $f_2$, and $f_3$. As noted previously, Pu & Wu (2015) found that $f_3 \simeq 10$. A linear expansion in $a$ of orbital velocity $v$ around $v_c$ gives $f_1 \simeq 0.5$. As mentioned earlier, for equal-mass bodies, $f_2 = 1$. If we scale $f_1$,

\(^1\)One caveat is that we do not consider the effects of the strong tides present near $r_{\text{Roche}}$, which may reduce $Q^*$. The effect of tides is discussed further in Section 4.5.1.
$f_2$, and $f_3$ by these values, then from Equation (4.22) we find

$$a \gtrsim a_{\text{frag}} = 0.41 \ r_{\text{Roche}}$$

$$= 0.0036 \ \text{AU} \ \left( \frac{M_*}{M_{\odot}} \right)^{1/3}$$

$$\times \left( \frac{f_1}{0.5} \right)^2 \left( \frac{f_3}{10} \right)^2 \frac{1 - 0.21}{f_2 - 0.21 \eta^{1/3}}$$

(4.23)

$$\times \left( \frac{f_1}{0.5} \right)^2 \left( \frac{f_3}{10} \right)^2 \frac{1 - 0.21}{f_2 - 0.21 \eta^{1/3}},$$

(4.24)

where $a_{\text{frag}}$ is the semimajor axis at which collisions between equal-mass bodies transition from non-erosive to erosive. For comparison, the solar radius is 0.0047 AU. For these values of $f_1$, $f_2$, and $f_3$, the semimajor axis range in which collisions are primarily erosive will always be less than $r_{\text{Roche}}$, i.e., the bodies will be tidally disrupted before they are close enough to their host star to reach high enough collision velocities to erode one another. If we have chosen correct values for $f_1$, $f_2$, and $f_3$, then Equation (4.23) says that gravity-dominated bodies, once formed, are able to engage in non-erosive collisions for semimajor axes all the way down to $r_{\text{Roche}}$, independent of stellar mass and the mass and density of the bodies. However, modifications to $f_1$, $f_2$, and $f_3$ could lead to erosive collisions outside $r_{\text{Roche}}$. We give two brief examples:

- Kenyon & Bromley (2017b) model collisional cascades in strength-dominated bodies and find that $(\Delta v)^2/Q^* \geq 5$ led to erosion of the largest bodies. Equation (4.1) implies that $Q < (\Delta v)^2/8$ for the largest body in a continuous mass distribution, so this would mean that $f_2 \lesssim 0.6$. A value $f_2 = 0.6$, with the fiducial parameters in Equation (4.23), leads to $a_{\text{frag}} = 0.83 \ r_{\text{Roche}}$.

- Our results only hold for homogeneous bodies. Using hydrodynamical simulations of collisions between differentiated bodies, Asphaug (2010) found that equal-mass head-on collisions were non-erosive if $\Delta v < k\nu_e$ and $k \simeq 2.9$. For our fiducial parameters $C^* = 1.8$ and $\rho/\rho_1 = 3$, Equation (4.19) implies that...
$k^2 = 4.75 f_2$, so this value of $k$ implies $f_2 = 1.8$. However, Asphaug also found that the crusts and mantles of differentiated bodies are more easily stripped in collisions (particularly off-center collisions) than the outer parts of homogeneous bodies. His simulations show that collisions with an angle $\theta$ between the relative velocity vector and relative center-of-mass vector $\sim 30^\circ$ have $k \approx 1.5$ ($f_2 = 0.5$ for our fiducial parameters) and collisions with $\theta \gtrsim 45^\circ$ have even lower values, $k \approx 1.2 - 1.4$ ($f_2 = 0.3 - 0.4$). If $f_2 = 0.3 - 0.4$ and $f_1$, $f_3$, and $\eta$ respectively remain 0.5, 10, and 1, then from Equation (4.23), $a_{\text{frag}}/r_{\text{Roche}} = 3.6 - 1.7$. However, the fractional mass loss in these off-center collisions remains quite low over a large range of $k$. For $\theta = 30^\circ$ the fractional mass loss is $\lesssim 5\%$ for $k \lesssim 1.7$ ($f_2 \simeq 0.61$), which occurs for $a > 0.81 r_{\text{Roche}}$. Additionally, the mass loss for a given $k$ decreases with increasing $\theta$. Equal-mass collisions with $\theta = 45^\circ$ have $\lesssim 5\%$ fractional mass loss for $k \lesssim 2.2$ ($f_2 \simeq 1.1$), and $\theta = 60^\circ$ gives $\sim 2\%$ fractional mass loss even for $k = 3$ ($f_2 = 1.9$). The shallow slope of mass loss versus collision velocity for off-center collisions is balanced by the mass growth from low-velocity off-center and low- and high-velocity head-on collisions. These arguments suggest that our conclusion for homogeneous equal-mass bodies, that collisional fragmentation is not a barrier to planet formation all the way down to $r_{\text{Roche}}$, likely holds for differentiated bodies as well.

### 4.4.3 Unequal-mass Case

The preceding arguments can be generalized to bodies with unequal masses, though not quite as cleanly. Assuming that the collision is sufficiently close to head-on that $\mu = \mu_\beta$ (and with $m_1 \geq m_2$ and $\alpha = m_2/m_1$ as before), we can write a more general
form of Equation (4.10),

\[
\frac{Q}{Q^*} = \frac{\mu}{2(m_1 + m_2)} \left( \frac{(\Delta v)^2_{\text{rel}} + v_e^2}{Q^*} \right)_{(\Delta r)^2_{\text{rel}}} + v_e^2 \]  
\[
= \frac{\alpha}{2(1 + \alpha)^2} \left[ \frac{5\alpha f_1^2 M_*(\Delta a)^2}{(1 + \alpha)^2 \pi \rho_1 C^* a^3 R_1^2} + \frac{v_e^2}{Q^*} \right] \]

\[
= \frac{\alpha}{2(1 + \alpha)^2} \times \left[ \frac{5 \cdot 2^{4/3} \alpha f_1^2 M_*(\Delta a)^2}{3^{2/3}(1 + \alpha)^{8/3} \pi^{1/3} \rho_1^{1/3} C^* a^3 m_1^{2/3}} + \frac{v_e^2}{Q^*} \right]. \]

(4.26)

Still assuming that both bodies have the same density, we may follow steps similar to those in Equations (4.11)–(4.15) to obtain

\[
\frac{Q}{Q^*} = \frac{\alpha}{2(1 + \alpha)^2} \left[ \frac{5 \cdot 2^{4/3} \alpha f_1^2 M_*(\Delta a)^2}{3^{2/3}(1 + \alpha)^{8/3} \pi^{1/3} \rho_1^{1/3} C^* a^3 m_1^{2/3}} + \frac{v_e^2}{Q^*} \right]. \]

(4.27)

We substitute \( r_{\text{Hill}} \) in Equation (4.16), now expressed in terms of \( m_1 \) and \( \alpha \), into Equation (4.28) and obtain more general versions of Equation (4.17),

\[
\frac{Q}{Q^*} = \frac{\alpha}{(1 + \alpha)^4} \times \left[ \frac{5 \cdot 2^{1/3} f_1^2}{3^{4/3} \pi^{1/3} C^*} \left( \frac{M_*}{\rho_1 a^3} \right)^{1/3} \frac{(\Delta a)^2}{r_{\text{Hill}}^2} \right]
\]

\[
+ \frac{20 \cdot (1 + \alpha)^{1/3}}{3 C^*} \left( \frac{\rho}{\rho_1} \right)^{1/3}, \]

(4.29)

and Equation (4.18),

\[
\frac{(1 + \alpha)^4}{\alpha^2} f_2 \frac{2}{3 C^*} \left( \frac{\rho}{\rho_1} \right)^{1/3} \gtrsim \frac{5 \cdot 2^{1/3} f_1^2}{3^{4/3} \pi^{1/3} C^*} \left( \frac{M_*}{\rho_1 a^3} \right)^{1/3} \frac{(\Delta a)^2}{r_{\text{Hill}}^2}. \]

(4.30)
We use similar arguments as before regarding the pre- and post-collision values for $\Delta a$ and $r_{\text{Hill}}$, except that we relax the assumption that the masses of the bodies are all the same. As before, $\Delta a' \approx 2\Delta a$. However, the post-collision mutual Hill radius $r_{\text{Hill}}'$ depends not only on the mass of the newly merged body but also the masses of the nearest neighbors, which, in principle, could take any value. Using $c$ to express the mass ratio between the newly merged body (of mass $m_1 + m_2$) and the mass $m_3$ of a nearby body of interest, $m_3 = c(m_1 + m_2)$, we can write $r_{\text{Hill}}' \approx (1 + c)^{1/3}r_{\text{Hill}}$. Thus, the condition $\Delta a' \gtrsim f_3r_{\text{Hill}}'$ becomes $2\Delta a \gtrsim (1 + c)^{1/3}f_3r_{\text{Hill}}$. From this, we obtain an analog to Equation (4.20) in the case of arbitrary masses,

$$a \gtrsim \frac{5f_1^2f_3^2(1 + c)^{2/3}}{2^{5/3}3^{4/3}\pi^{1/3}C^*} \left(\frac{M_*}{\rho_1}\right)^{1/3} \times \left[\frac{(1 + \alpha)^4}{\alpha^2}f_2 - \frac{20 (1 + \alpha)^{1/3}}{3C^*} \left(\frac{\rho}{\rho_1}\right)^{1/3} \left(1 + \alpha^{1/3}\right)\right]^{-1}. \quad (4.31)$$

Expressing the result in units of $r_{\text{Roche}}$, we find

$$a \gtrsim \frac{5f_1^2f_3^2(1 + c)^{2/3}\eta^{1/3}r_{\text{Roche}}}{2^{5/3}3\pi^{1/3}C^*} \times \left[\frac{(1 + \alpha)^4}{\alpha^2}f_2 - \frac{20\eta^{1/3} (1 + \alpha)^{1/3}}{3^{2/3}C^*} \left(1 + \alpha^{1/3}\right)\right]^{-1}. \quad (4.32)$$

We now find the value of $\alpha$ that maximizes the right side of Equation (4.32). The parameter $f_2$ is a function of $\alpha$ (see the first line of Equation (8) in C13) and is expressed as such in our maximization. The value of $\alpha$ that maximizes Equation (4.32), as well as the maximum value, depend on the values of the various parameters in Equation (4.32). The factor $f_1$ does not depend on the masses of the bodies, and thus will keep its same value. We note that most of the work on orbital stability has focused on equal-mass bodies, so it is not known whether $f_3$ depends on relative
masses; we assume $f_3$ does not depend on mass. Thus, for the maximization, we assume as before $f_1 = 0.5$, $f_3 = 10$, $\eta = 1$, and $C^* = 1.8$, and we also assume $c = 1$. Since we have assumed $m_1 \geq m_2$, $\alpha$ must be in the interval $(0, 1]$. The local maximum in this domain using the assumed values of the other parameters is found at $\alpha = 0.49$ and gives

$$a \gtrsim a_{\text{frag, } \alpha} = 0.51 \ r_{\text{Roche}},$$

(4.33)

where $a_{\text{frag, } \alpha}$ is the same as $a_{\text{frag}}$ but generalized to arbitrary $\alpha$. Thus our conclusions from the homogeneous equal-mass case should remain approximately valid for arbitrary mass ratios. For differentiated bodies, the arguments presented before based on the results of Asphaug (2010) still hold, since for head-on collisions, $f_2 \simeq 1.3$ for $\alpha = 0.5$ and $f_2 \simeq 1.5$ for $\alpha = 0.1$ (which are the only two unequal mass ratios examined in Asphaug 2010), and for off-center collisions, all $f_2$ values are greater for $\alpha = 0.5, 0.1$ collisions than the $\alpha = 1$ collisions with corresponding $\theta$.

Our statement that collisions lead to growth if $Q < f_2 Q^*$ ignores any dependence of $f_2$ on stellar tides. Close to $r_{\text{Roche}}$, tides from the star will make it easier for fragments to escape the gravitational influence of whatever coalesced body may be left after a collision, and thus $f_2$ will take a smaller value than at larger semimajor axes. This is accounted for in our numerical work by modifying the escape speed, as described in Section 4.5.1. An additional complication not addressed by our analytic work is that collisions will take on a variety of impact parameters and have different possible collision outcomes than the pure fragmentation outcome considered in this section. Because of these additional considerations, and because of the sensitivity of our analytic result to even small uncertainties in $f_1$, $f_2$, and $f_3$, it is essential to perform N-body simulations of systems of rocky bodies to test the analytic results of

---

\[^2\text{The dependence on the relative mass is likely to be weak, since Hill’s equations describing the interactions of two small bodies orbiting a much larger third body depend only on the sum of the masses of the two small bodies (e.g., Hénon & Petit 1986).} \]
this section. The remainder of this chapter presents the methods and results of such calculations.

4.5 Numerical Methods

Our approach is very similar to that of C13, who studied the effect of collisional fragmentation on the formation of rocky planets in a solar system context. Our simulations were carried out using the hybrid-symplectic integrator in the \textit{Mercury} N-body integrator package (Chambers 1999), including the modifications introduced in C13 to describe collisional fragmentation. We first describe the collision algorithm and then discuss the initial conditions of our simulations.

4.5.1 Collision Algorithm

The prescription used for collision outcomes is that of LS12 as implemented in C13, with modifications described below to account for stellar tides. Here we briefly describe the possible outcomes of a collision and direct the reader to C13 for a full description of the algorithm. When a collision is identified, the relative velocity (including gravitational focusing) at the time of impact is calculated. If the impact velocity is less than a modified mutual escape velocity $v'_e$, then the two bodies merge. This modified mutual escape velocity takes into account the relatively small Hill spheres of the bodies at the small semimajor axes present in our simulations. Instead of fragments needing to escape to infinity to become unbound from the bodies, they need only escape to the edge of the Hill sphere. This is accounted for by defining $v'_e$ as follows,

$$v'_e^2 = \max \left\{ v_e^2 \left( 1 - \frac{R_t + R_p}{r_{\text{Hill}}} \right), 0 \right\},$$

(4.34)

where $R_t$ and $R_p$ are the radii of the target and the projectile, $v_e$ is the escape speed defined in Equation (4.6), and $r_{\text{Hill}}$ is the mutual Hill radius defined Equation (4.16).
If the impact velocity is greater than $v'_e$, then fragmentation occurs, and the largest remnant mass is calculated as outlined in Equation (8) of C13.

The mass from the impacting bodies not included in the largest remnant is then divided into one or more equal-mass fragments, with masses as close as possible to but not below a minimum fragment mass (MFM) that is specified for each simulation. If the remaining mass is less than the specified MFM, then the collision is treated as a merger and the total mass is placed in a single body.

The details of how non-head-on collisions are handled can be found in C13 and LS12. High impact-parameter collisions have other possible outcomes in addition to the merger–fragmentation scenario presented above for head-on collisions. These are:

- hit-and-run collisions, where the target remains intact and the projectile is disrupted, in some cases adding mass to the target;
- graze-and-merge collisions, where an initial off-center collision saps sufficient kinetic energy from the system for an eventual merger.

We use the boundary identified by Genda et al. (2012) between the hit-and-run and graze-and-merge regimes, replacing the escape velocity $v_{\text{esc}}$ in their Equation (16) with our modified escape velocity $v'_e$ to account for stellar tides.

As was mentioned before, the amount of mass in a collision that is dispersed by fragmentation, if any, depends on the ratio between $Q$ (given by Equation 4.1) and $Q^*$ (given by Equation 4.2). Stewart & Leinhardt (2009) found that, in the catastrophic disruption regime, the largest remnant mass (and thus the mass dispersed by fragmentation) is linear in this ratio. However, this simple relation appears not to hold in regions close to the Roche radius (e.g., Karjalainen, 2007; Hyodo & Ohtsuki, 2014; see also Kenyon & Bromley 2017a for an application to collisional cascades). Instead, the mass of the largest remnant depends on the shapes of the colliding bodies and the relative magnitudes of the azimuthal, radial, and vertical components of
the impact velocity (for the radial velocity component in particular, the dependence is not monotonic on specific impact energy). Since this study is focused more on answering the question of whether Keplerian orbit-driven collision velocities are ever large enough to prevent planet formation and not on precisely how tidal effects affect collisional planet formation, we approximate the tidal effects by replacing $Q^*$ with $Q_{\text{tidal}}^*$, defined as

$$Q_{\text{tidal}}^*(r) = \begin{cases} 
Q^* \left[ 1 - \left( \frac{r_{\text{Roche}}}{r} \right)^3 \right] & r > r_{\text{Roche}} \\
0 & r \leq r_{\text{Roche}} 
\end{cases}$$

where $r$ is the distance between the central star and the center of mass of the colliding bodies. For $r > r_{\text{Roche}}$, we found that Equation (4.35) predicts collision energies required to nearly entirely disrupt bodies that are broadly consistent with the simulations of Hyodo & Ohtsuki (2014). Equation (4.35) also converges to the LS12 results far away from the Roche radius, a result also found by Hyodo & Ohtsuki (2014). In the case $Q_{\text{tidal}}^* = 0$, any collision will cause the bodies to fragment into as many fragments as possible, as limited by the MFM. Thus, collisions that occur inside $r_{\text{Roche}}$ will lead to complete disruption of the colliding bodies. In our algorithm, disruption inside the Roche radius only occurs during a collision, i.e., if a body migrates in to $r < r_{\text{Roche}}$, it will not be disrupted unless it suffers a collision while inside $r_{\text{Roche}}$.

We also note that we are using the classical Roche radius, which was derived for a strengthless, co-rotating body on a circular orbit and that more general forms of the disruption radius have been developed (e.g., Holsapple & Michel, 2006, 2008). However, the strengthless model approximates collisional fragmentation well, since after a collision the fragments are no longer bound to each other by tensile forces, only gravity. For this reason, and for simplicity, we use the classical Roche radius in Equation (4.35) and throughout the chapter. We believe that the qualitative picture
our results paint is correct while quantitative details that depend on the Roche radius may well be inaccurate.

The choice of MFM for our collision algorithm represents a compromise between computational cost and realism: if the MFM is too large, a large fraction of fragmenting collisions will be erroneously classified as mergers; if the MFM is too small, then fragmentation will create so many fragments that the numerical calculations will slow to a crawl. To check that our results are not strongly dependent on the choice of MFM, each set of initial conditions was run with three different values of the MFM, as detailed in Section 4.5.2.

If multiple collisions involving the same body occur in the same time step, only one of the collisions (chosen at random) was considered to have happened. For a typical run, this situation happened at most a few times over the course of the run. Any bodies that pass inside the solar radius (\(\sim 0.0047\) AU) were assumed to merge with the central star. Over the course of a typical run, this happened only a few times. Virtually all such events involved bodies with small masses (very near the MFM), and the resulting total mass loss was at most a few percent of the total mass of the initial bodies.

### 4.5.2 Initial Conditions

We carried out several sets of N-body integrations. We first describe the initial conditions of our fiducial set of simulations then describe the variations we made to this fiducial set. All of our systems orbit a 1 M\(_\odot\) star and we assume a fixed density of 3 g cm\(^{-3}\) for all bodies, independent of mass. A time step of \(6 \times 10^{-3}\) days was used. For simplicity, we assume that the star has a radius of 1 R\(_\odot\), but note that the young stars in systems such as those we simulate may have considerably larger radii (e.g., calculations by Baraffe et al. 2015 show that a 1 M\(_\odot\) star with an age of 3 Myr has a radius of 1.7 R\(_\odot\)). Our 15 fiducial simulations start with 150 equal-mass
bodies of individual mass 0.02 $M_{\oplus}$ ($\sim$1.6 times the mass of the Moon), corresponding to a radius of $\sim$2100 km with the 3 g cm$^{-3}$ density used in this chapter). This mass is roughly what is expected for bodies that form during the runaway growth phase of planet formation (e.g., Wetherill & Stewart 1989) and thus our initial conditions can be thought of as the beginning of the oligarchic growth or giant-impact stage of planet formation.

We ran our simulations assuming no gas was present, i.e., we assumed that the protoplanetary gas disk at these radii had entirely disappeared by the beginning of our simulations. Thus our initial conditions can be thought of as corresponding to planetesimals/planetary embryos that were prevented from evolving into orbit-crossing trajectories by eccentricity damping while the gas was present, or planetesimals/planetary embryos that underwent disk-driven migration and were frozen in semimajor axis as the disk evaporated away. We caution that this assumption may not be realistic in some or perhaps all disks. The existence of planets like GJ 1214 b (Charbonneau et al. 2009)—a $\sim$7 $M_{\oplus}$ planet at a period of 1.58 days that is believed to have a substantial gas envelope—implies that in at least some cases the formation of planets with $\sim$1 day orbital periods has occurred before the gas has completely disappeared.

The initial semimajor axes of the bodies were distributed between 0.005 and 0.04 AU using a similar disk surface-density profile $\sigma(a)$ as Chambers (2001) and C13: $\sigma \propto a^{-3/2}$ from 0.02 AU to 0.04 AU and, inside 0.02 AU, $\sigma$ linear in $a$, starting from $\sigma = 0$ at 0.005 AU and increasing to a value matching the $a^{-3/2}$ profile at 0.02 AU. A cumulative distribution function was calculated from this disk profile and its inverse was sampled in 150 equally spaced locations to determine the initial semimajor axes of the bodies. This semimajor axis range was chosen to straddle the Roche radius (0.0089 AU for a 1 $M_{\odot}$ star with planet density $\rho$ = 3 g cm$^{-3}$), with the lower bound chosen to be close to the solar radius (0.0047 AU). We note that one
would not expect large rocky bodies to be found inside $r_{\text{Roche}}$, so it is unphysical to include such bodies in our initial conditions but doing so is useful to verify our analytic results. Initial eccentricities and inclinations were drawn from a Rayleigh distribution, $f(x) = x/\sigma^2 \exp[-x^2/(2\sigma^2)]$, with $\sigma = 0.01$ for eccentricity and $\sigma = 0.5$ for inclination. The initial arguments of periapsis, ascending nodes, and mean anomalies were chosen at random from uniform distributions in the interval $[0, 2\pi)$.

As described in Section 4.5.1, the MFM sets a floor to the size of fragments in our simulations. To investigate how the choice of MFM affects our results, we employed three different MFMs, which are 5, 7.5, and 15 times smaller than the initial masses of the bodies used in our fiducial simulations, or, respectively, $4.0 \times 10^{-3}$, $\sim 2.67 \times 10^{-3}$, and $\sim 1.33 \times 10^{-3}$ $M_\oplus$. At the density used in this chapter ($3 \text{ g cm}^{-3}$), these bodies have radii of $\sim 1200$, $\sim 1100$, and $\sim 900$ km. We ran five simulations for each combination of initial conditions and MFM. A smaller MFM means that more fragments, on average, are produced per collision. Since the time required to calculate gravitational forces goes as the square of the number of bodies, increasing the number of fragments by decreasing the MFM significantly slows down the calculation. This slowdown is exacerbated by the accumulation of fragments inside $r_{\text{Roche}}$, which are unable to merge and thus persist for the duration of the simulation. Therefore the choice for the smallest MFM was limited by computation time. The simulations were run for $3 \times 10^5$ years with the exception of the runs with the smallest MFM, $\sim 1.33 \times 10^{-3}$ $M_\oplus$, which lasted for only $1 \times 10^5$ years, due to limits on computation time.

We adopted a naming convention to organize our numerical results. A name defines a “set of simulations”, by which we mean simulations that have identical initial conditions and MFM but use different seed values for the random number generator. The initial orbital parameters that are chosen at random are eccentricity, inclination, argument of periapsis, ascending node, and mean anomaly. Thus, simulations in a set share identical initial body masses, numbers of bodies, semimajor axes, and MFM.
Our three fiducial sets of simulations are described above. The fiducial set with \( M_{\text{FM}} = 4.0 \times 10^{-3} M_{\odot} \) is named \textit{fiducial\textunderscore\text{MFM\textunderscore large}}, the fiducial set with \( M_{\text{FM}} \simeq 2.67 \times 10^{-3} M_{\odot} \) is named \textit{fiducial\textunderscore\text{MFM\textunderscore mid}}, and the fiducial set with \( M_{\text{FM}} \simeq 1.33 \times 10^{-3} M_{\odot} \) is named \textit{fiducial\textunderscore\text{MFM\textunderscore small}}. Each set in this chapter consists of five individual simulations.

To understand how changes to the initial number of bodies affect the outcomes, we ran three sets of simulations with initial conditions and MFM that matched the three fiducial sets except that the initial number of bodies was 75 (instead of 150), as well as three sets of simulations matching the fiducial sets except with 300 initial bodies. The sets with 75 initial bodies have names prepended \textit{num\textunderscore down} and the sets with 300 initial bodies have names prepended \textit{num\textunderscore up}.

To understand how changes to the total mass present in the initial bodies affected the outcomes, we ran three sets of simulations with initial conditions and MFM that matched the three fiducial sets except that the initial body masses were 0.01 \( M_{\odot} \) (instead of 0.02 \( M_{\odot} \)), as well as three sets of simulations matching the fiducial sets except with initial body masses of 0.04 \( M_{\odot} \). The sets with 0.01 \( M_{\odot} \) initial bodies have names prepended \textit{mass\textunderscore down} and the sets with 0.04 \( M_{\odot} \) initial bodies have names prepended \textit{mass\textunderscore up}.

We also ran a few sets of control simulations. Three sets had identical initial conditions and MFM as the fiducial sets but the effects of stellar tides were ignored, i.e., the collision algorithm of C13 was followed exactly and Equations (4.34) and (4.35) were ignored (i.e., \( v_{\epsilon} \) and \( Q^* \) were used instead of \( v'_{\epsilon} \) and \( Q^*_{\text{tidal}} \)). These sets have names prepended \textit{not\textunderscore tidal}. A fourth set of control simulations had perfect mergers for all collisions, i.e., all collisions were perfectly inelastic and no fragmentation occurred. The name of this set is prepended \textit{merge}.

Table 4.1 summarizes the prefixes used in our naming scheme for the sets of simulations.
Table 4.1. Naming Convention for Simulations

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>fiducial</td>
<td>The fiducial set of runs: $150 \times 0.02 , M_\oplus$ bodies</td>
</tr>
<tr>
<td>numdown</td>
<td>Initial number of bodies half that of the fiducial runs: $75 \times 0.02 , M_\oplus$ bodies</td>
</tr>
<tr>
<td>numup</td>
<td>Initial number of bodies twice that of the fiducial runs: $300 \times 0.02 , M_\oplus$ bodies</td>
</tr>
<tr>
<td>massdown</td>
<td>Initial body mass half that of the fiducial runs: $150 \times 0.01 , M_\oplus$ bodies</td>
</tr>
<tr>
<td>massup</td>
<td>Initial body mass twice that of the fiducial runs: $150 \times 0.04 , M_\oplus$ bodies</td>
</tr>
<tr>
<td>notidal</td>
<td>Same as the fiducial set of runs, but with tidal effects ignored</td>
</tr>
<tr>
<td>merge</td>
<td>Same as the fiducial set of runs, but all collisions assumed to be mergers</td>
</tr>
</tbody>
</table>

We found that \( \sim 10\% \) of the runs failed to finish. An investigation showed that these runs were effectively frozen immediately after collisions that produced \( \gtrsim 90 \) fragments. We believe that the large number of bodies produced in close proximity to each other after these particularly fragmenting collisions have caused these runs to get bogged down in the Bulirsch-Stoer portion of the hybrid-symplectic integration of Mercury. We neglect these failed-to-complete runs in our analysis, but believe that doing so should not bias our conclusions since the results do not appear to depend on the details of the collision history during the run. We also ran supplementary simulations, of which a sufficient number completed to make up for the runs that failed to complete. For supplementary numup and massup runs, it was necessary to set a maximum number of fragments that could be produced in a single collision to allow any of the runs to successfully complete. The maximum number of fragments used for these runs was 50.
4.6 Results

We present the results from our fiducial runs in Section 4.6.1 and then describe the results from our variations on the fiducial runs in Section 4.6.2.

4.6.1 Fiducial runs

Figure 4.1 shows several snapshots of eccentricity vs. semimajor axis from one of the fiducial MFM large simulations. Within 100 years, several of the bodies have already built up to large fractions of their final masses. This may seem fast, but by 100 years bodies with $a = 0.01$ AU have gone through $10^5$ orbits. Over the remainder of the simulation, the planets continue building up to their final masses by accumulating the remaining smaller bodies and merging with each other. The fragments inside (and just outside) the Roche radius never consolidate, as expected from the collision algorithm, and end up with larger (but still small) eccentricities at the end of the simulation than the planets outside $r_{\text{Roche}}$. Figure 4.2 is the same as Figure 4.1 but for one of the fiducial MFM small simulations, which has an MFM that is three times smaller than the run shown in Figure 4.1. Other than a larger buildup of fragments inside $r_{\text{Roche}}$ and an increased number of fragments outside $r_{\text{Roche}}$ in the intermediate stages, the simulation proceeds qualitatively the same as that in Figure 4.1. In both runs, the fragments are unable to accumulate for $a \lesssim 0.01$ AU, which is $\sim 10\%$ larger than $r_{\text{Roche}}$. This behavior occurs in all fiducial MFM large, fiducial MFM mid, and fiducial MFM small runs. The lack of planet formation between $r_{\text{Roche}}$ and $\sim 1.1 r_{\text{Roche}}$ is presumably due to the large tidal forces acting on the bodies in that regime, which enhance the amount of fragmentation that occurs and makes it more difficult for bodies to merge, especially for collisions between bodies of similar size ($\alpha \approx 1$). This behavior has been studied in detail in the context of ring particles by, e.g., Canup & Esposito (1995) and Yasui et al. (2014). Our control runs (merge and
Figure 4.1: Eccentricity versus semimajor axis of all the bodies for one of the fiducial_MFMlarge simulations, taken at six different snapshots. The symbol radius is proportional to the radius of each body. The minimum fragment mass (MFM) is $4.0 \times 10^{-3} M_{\oplus}$, the largest value used in this chapter. At the end of the run, after $3 \times 10^5$ years, the system contains 16 bodies on low-eccentricity orbits outside $r_{\text{Roche}} = 0.0089$ AU and 18 fragments inside $r_{\text{Roche}}$.

notidal), which ignore tidal effects, form planets at all semimajor axes that were populated in the initial state, even well inside $r_{\text{Roche}}$.

The eccentricities of the planets in Figure 4.1 get damped over time, and the final mean eccentricity of the 16 bodies outside $r_{\text{Roche}}$ is only 0.0028. To determine the source of the eccentricity damping, we re-ran some of the fiducial runs with a single modification: at $1.5 \times 10^4$ years, all bodies inside a radius $r_{\text{clear}}$ were removed. We experimented with $r_{\text{clear}} = r_{\text{Roche}}$ and $r_{\text{clear}} = 1.1 r_{\text{Roche}}$. Eccentricity damping still occurred in the simulations of the $r_{\text{clear}} = r_{\text{Roche}}$ sets but the eccentricity damping in
Figure 4.2: Same as Figure 4.1, but for one of the fiducial_MFMsmall simulations, with MFM set to $\sim1.33\times10^{-3} M_\oplus$, a factor of three smaller than in Figure 4.1. Note that the final snapshot is at 100 kyr instead of 300 kyr as in Figure 4.1; this simulation was not run as long due to the larger computational demands of runs with smaller MFM. At the end of the run, the system contains 24 bodies on low-eccentricity orbits outside $r_{\text{Roche}}$ (several of which are fragments inside $1.1 r_{\text{Roche}}$) and 45 fragments inside $r_{\text{Roche}}$. 

128
Figure 4.3: The number of bodies as a function of time for the five simulations in each of the three fiducial sets. The top panel shows the total number of bodies, while the bottom panel shows only those bodies that are outside $r_{\text{Roche}}$. The simulations from each set are encoded by color: “large MFM” corresponds to $\text{fiducial\_MFM\_large}$ (MFM = $1.2 \times 10^{-8} \, M_\odot$) and is shown in red, “mid MFM” corresponds to $\text{fiducial\_MFM\_mid}$ (MFM = $0.8 \times 10^{-8} \, M_\odot$) and is shown in cyan, and “small MFM” corresponds to $\text{fiducial\_MFM\_small}$ (MFM = $0.4 \times 10^{-8} \, M_\odot$) and is shown in black. All simulations converge on a final set of planets outside $r_{\text{Roche}}$ with similar properties ($N \simeq 15, \langle e \rangle \simeq 0.003$).

the $r_{\text{clear}} = 1.1 \, r_{\text{Roche}}$ sets ceased at the time of body removal. Thus, the eccentricity damping seems to be due mainly to the bodies between $r_{\text{Roche}}$ and $1.1 \, r_{\text{Roche}}$.

Figure 4.3 shows the number of bodies as a function of time for all the simulations in the $\text{fiducial\_MFM\_large}$, $\text{fiducial\_MFM\_mid}$, and $\text{fiducial\_MFM\_small}$ sets. The top panel shows the total number of bodies while the bottom panel shows only those bodies outside $r_{\text{Roche}}$. In the latter case, the number of bodies converges to a constant value of $\sim 15$ bodies outside $r_{\text{Roche}}$ by $\sim 100$ kyr, independent of MFM. In what follows, we regard the bodies present outside $r_{\text{Roche}}$ at the end of the integrations as the “final planets” (even though those inside $\sim 1.1 \, r_{\text{Roche}}$ have not built up to typical planetary masses), while the bodies present inside $r_{\text{Roche}}$ at the end of the integrations are referred to as “fragments”, because these bodies have fragmented down to nearly the
MFM. The number of fragments has also converged to a constant value in each run, as can be seen in the top panel of Figure 4.3.

In both panels of Figure 4.3, the number of bodies shows a persistent trend downward with time in each simulation after the first 30 years or so. The downward trend is interrupted occasionally by collisions that produce a large number of fragments, especially for lower-MFM runs. However, these fragments get quickly reaccumulated and there is never any runaway fragmentation. In all of our simulations, a system of 14–24 planets is always formed outside \( r_{\text{Roche}} \) (with 11–17 bodies outside \( 1.1 \ r_{\text{Roche}} \)).

The left panel of Figure 4.4 shows a sample of the collision outcomes in one of the fiducial_MFMlarge simulations and the right panel shows a sample of the subset of collision outcomes involving the bodies that become the final planets. Figure 4.5 shows the same for one of the fiducial_MFMsmall runs. In these figures, and in the discussion that follows, only collisions that occurred outside \( r_{\text{Roche}} \) are considered. The collisions are classified into one of four categories: “merger,” “grow,” “erode,” and “hit & run.” Collisions that were treated in the code as mergers or as graze-and-merge collisions are here collectively classified as “merger,” hit-and-run collisions are classified as “hit & run,” while among non-hit-and-run fragmenting collisions those where the largest body loses mass are classified as “erode” and those where the largest body comes out of the collision with at least as much mass as before are classified as “grow.” In the right panel of Figure 4.4, 6% of collisions were classified as “merger” and 1% were classified as “grow,” compared to 5% and 0.9% for all collisions. These percentages are characteristic of all of the fiducial_MFMlarge runs. Remarkably, the bodies that became the final planets had a larger percentage of “erode” collisions than did all the bodies: 52% for the run in Figure 4.4 for the final planets as compared with 46% among all the collisions. For all five fiducial_MFMlarge runs, the percentage of “erode” collisions was higher among the bodies that became the final planets than among all the bodies. Similar patterns were found among the fiducial_MFMsmall...
Figure 4.4: A random sample of collision outcomes for the same fiducial MFMLarge simulation shown in Figure 4.1. The plots show impact velocity versus impact parameter for collisions occurring outside $r_{\text{Roche}}$. The impact parameter $b$ is defined as the distance between the centers of the two bodies at the time of the collision, projected perpendicular to the relative velocity vector. Only a random sample of 2000 collision outcomes is shown in each panel to prevent overcrowding. The left panel shows a random sampling of all collisions (total number: 88,659) and the right panel shows a sampling of the collisions involving the bodies that become the final planets (total number: 4,625). In the legend, “merger” refers to collisions that were treated in the code as mergers or as graze-and-merge collisions, “hit & run” refers to collisions classified as hit-and-run (high impact parameter, no mass loss for target body), while among non-hit-and-run fragmenting collisions those with the largest body losing mass in the collision are labeled as “erode” and those with the largest body coming out of the collision with at least as much mass as before the collision are labeled as “grow.” The division at $b/R_{\text{target}} = 1$ between “grow” collisions and “hit & run” collisions is due to the definition of hit-and-run collisions, as detailed in LS12 and C13. There are far more erosive and hit-and-run collisions than collisions that lead to mass growth; however, a closer look reveals that most these are within $\sim 1.1 \, r_{\text{Roche}}$. See text for details.
Figure 4.5: Same as Figure 4.4, but for a simulation from the fiducial_MFMsmall set (the same simulation as shown in Figure 4.2). The total number of collisions is 344,847 and the total number of collisions involving the bodies that become the final planets is 10,003.

runs, one of which is shown in Figure 4.5: all five of the runs had a greater fraction of “merger” and “grow” collisions among collisions involving just the bodies that became the final planets than all the bodies, and four of the five runs had a greater fraction of “erode” collisions among collisions involving just the bodies that became the final planets than all the bodies. For the run shown in Figure 4.5, 7% of collisions involving bodies that became the final planets were “merger,” 0.7% were “grow,” and 55% were “erode,” as compared to 4% “merger,” 0.3% “grow,” and 54% “erode” collisions among all the bodies. For all ten of the runs in the fiducial_MFMlarge and fiducial_MFMsmall sets, the fraction of “hit & run” collisions is lower among collisions involving just the bodies that became the final planets than among all the bodies.

It may seem strange that the collisions involving just the bodies that became the final planets had a larger fraction of “erode” collisions than for all the bodies, since those bodies were able to gain sufficient mass to become planets by the end of the simulations. However, most of these erosive collisions occurred close to $r_{\text{Roche}}$, while further out from $r_{\text{Roche}}$ a larger fraction of collisions lead to mass growth. Close to $r_{\text{Roche}}$, 

132
tidal effects cause a larger fraction of the collisions to be fragmenting. Specifically, for the bodies that became the final planets of the fiducial_MFMlarge run shown in Figure 4.4, 99.5% of “erode” collisions occurred inside 0.01 AU (\(\sim 1.13 \, r_{\text{Roche}}\)) as well as 86% of hit & run collisions, while only 8.8% of “merger” and 3.4% of “grow” collisions occurred inside 0.01 AU. Thus, the vast majority of erosive collisions occur close to \(r_{\text{Roche}}\), where tides are the strongest, and the majority of collisions leading to mass growth occur outside 0.01 AU, where the final planetary system forms. Comparable numbers are found for all runs in the fiducial_MFMlarge, fiducial_MFMmid, and fiducial_MFMsmall sets. In short, the bodies outside \(\sim 1.1 \, r_{\text{Roche}}\) that become the final planets mostly engage in collisions that lead to mass growth, while most of the fragmenting collisions occur inside \(\sim 1.1 \, r_{\text{Roche}}\), where tidally enhanced fragmentation prevents a steady accumulation of mass.

Figure 4.6 shows the masses of the bodies that become the final planets as a function of time. In this chapter and for this figure, the largest remnant in a collision adopts the name of the more massive of the two impacting bodies. The final planets outside \(1.1 \, r_{\text{Roche}}\) have masses between \(\sim 0.05 \) and \(\sim 0.4 \, M_{\oplus}\). Each vertical “jump” in mass represents a collision. The bodies that undergo significant mass growth (i.e., the final planets outside of \(\sim 1.1 \, r_{\text{Roche}}\)) have nearly uninterrupted marches towards higher masses. We must keep in mind, though, that the bodies in Figure 4.6, by virtue of their growth and their surviving to the end, exhibit a selection bias in which kinds of collisions they experienced. The planets assemble very quickly, with most reaching their final masses within \(\sim 10^{3}\) years.

Figure 4.7 shows the separation between adjacent bodies at the end of the simulations in units of \(r_{\text{Hill}}\). Here, separation is defined as the difference of the apastron of the inner body and the periastron of the outer body. The condition for long-term stability (over \(10^{10}\) orbits, or 10 million years at 0.01 AU) of a final planetary configuration is that the separations exceed \(\sim 10 \, r_{\text{Hill}}\) (Pu & Wu 2015). For bodies outside
Figure 4.6: Mass as a function of time for bodies that become the final planets. The top panel shows one of the simulations from the fiducial_MFMlarge set (the same simulation as in Figures 4.1 and 4.4) while the bottom panel shows one of the simulations from the fiducial_MFMsmall set (the same simulation as in Figures 4.2 and 4.5). The final planets outside 1.1 $r_{\text{Roche}}$ exhibit a steady march to their final masses, with relatively few erosive collisions. Note that the simulations of the bottom panel were run for less time than those of the top panel, as described in Section 4.5.2.
Figure 4.7: The separation between adjacent bodies (defined as the difference between apastron of the inner body and periastron of the outer body) in units of $r_{\text{Hill}}$ as a function of semimajor axis at the end of the simulation. Here, the semimajor axis is calculated as the mean of the apastron of the inner body and the periastron of the outer body. Each color corresponds to bodies from a single simulation. The vertical dashed line shows the location of $r_{\text{Roche}}$. The left panel shows results from the fiducial_MFMlarge simulations while the right panel shows results from the fiducial_MFMsmall simulations. The minimum separation required for stability over $10^{10}$ orbits (10 million years at 0.01 AU) is $\sim 10$ $r_{\text{Hill}}$ (Pu & Wu 2015), marked by a horizontal dashed line. The separations of bodies inside $r_{\text{Roche}}$ are smaller, partly because the eccentricities are higher and partly because we do not allow these bodies to merge to form more stable systems.

$r_{\text{Roche}}$ (denoted with the vertical dashed line), this condition is met, with only a small fraction of the bodies having separation $< 10$ $r_{\text{Hill}}$. This result confirms our use of $f_3 = 10$ in Section 4.4. Inside $r_{\text{Roche}}$, the fragments show a large spread in separations and the separations are, on average, smaller than those of the planets outside $r_{\text{Roche}}$. In particular, separations as small as $r_{\text{Hill}} \approx 0$ are seen. The small values of these separations are partly due to the relatively high eccentricities (see Figures 4.1 and 4.2) and partly due to the inability of the bodies inside $r_{\text{Roche}}$ to merge and evolve into a dynamically stable system.

Figure 4.8 shows the final planetary configurations for our fiducial runs. The semi-major axes of the planets with the most mass in each system are clustered around $\sim 0.2$–$0.3$ AU. As was noted earlier, tidally enhanced fragmentation prevents planet
Figure 4.8: The final planetary configurations for all the fiducial sets of simulations. The top panel shows final configurations from the fiducial_MFMlarge set, the middle panel shows final configurations from the fiducial_MFMmid set, and the bottom panel shows final configurations for simulations from the fiducial_MFMsmall set. Each row of circles corresponds to the final bodies of a single simulation. The radius of each circle is scaled by the final radius of the corresponding body, assuming the same density for all bodies. The blue circle in the bottom right of each panel shows the radius of a 1 M⊕ body at this density. The vertical dashed line shows the location of r_{Roche}. All of our fiducial simulations form a stable planetary system at all orbital radii greater than ∼1.1 r_{Roche}. The gray horizontal line to the left of ∼0.01 AU is a series of overlapping circles representing the closely spaced fragments found at these semimajor axes.
formation out to $\sim 0.01$ AU, which is $\sim 1.1 \, r_{\text{Roche}}$. All of the fiducial runs form a final, stable planetary system. The planet-formation efficiency (the fraction of mass initially present that is incorporated into the final planets) for the mass initially outside $\sim 1.1 \, r_{\text{Roche}}$ is $\sim 100\%$, while inside it is much lower because the tidally enhanced fragmentation prevents planets from growing, decreasing with radius to be approximately zero near $r_{\text{Roche}}$. The high planet-formation efficiency outside $\sim 1.1 \, r_{\text{Roche}}$ is partly an artifact of the MFM, which prevents mass loss through collisional cascades.

4.6.2 Other scenarios

Our other sets of runs with tidal effects ($\text{numdown}$, $\text{numup}$, $\text{massdown}$, $\text{massup}$; see Table 4.1) also formed planets all the way down to $\sim 1.1 \, r_{\text{Roche}}$. The distribution and masses of the planets in the final systems of these sets differed from our fiducial sets in ways that related to the differing initial conditions: runs with a larger total initial mass ($\text{numup}$ and $\text{massup}$) produced a smaller number ($\sim 10$) of higher-mass planets in their final systems than did the fiducial sets, while the runs with a smaller total initial mass ($\text{numdown}$ and $\text{massdown}$) produced about the same number of planets in their final systems as our fiducial sets, but with lower masses. As mentioned previously, our control runs ($\text{notidal}$ and $\text{merge}$) also form planets and are able to do so at semimajor axes as small as $\sim 0.6 \, r_{\text{Roche}}$ (planet formation at smaller semimajor axes is suppressed only because our initial conditions do not have mass interior to 0.005 AU). The details of how the various runs got to their final configurations also differed from the fiducial sets; for instance, more fragmentation occurred in the $\text{numup}$ runs than in the fiducial set—likely due to the increased mass available to be turned into fragments—and the numbers of bodies in the $\text{numup}_\text{MFMsmall}$ runs was as high as $\sim 700$ before accumulation was able to take over and reduce the total number of bodies. The planet formation efficiency for these runs for the mass initially outside $\sim 1.1 \, r_{\text{Roche}}$ is also $\sim 100\%$, as it was for the fiducial runs. Figure 4.9 shows the
**Figure 4.9:** Same as Figure 4.8 but for final planetary systems from our initial setups other than fiducial.
final configurations of a random selection of the numdown, numup, massdown, massup, notidal, and merge runs.

4.7 Discussion

The analytic work of Section 4.4 concluded that collisional fragmentation was not a barrier to planet formation outside of the Roche radius. Our numerical integrations, described in Sections 4.5 and 4.6, enhanced our analytic work by including the effects of stellar tides and off-center collisions. Even with the richer physics, the results of our numerical integrations agree with the analytic arguments, except for a small region between $r_{\text{Roche}}$ and $\sim 1.1 r_{\text{Roche}}$, where stellar tides are strong enough that even relatively weak collisions lead to fragmentation.

Other works have examined in situ formation of rocky planets at small semimajor axes (Hansen & Murray, 2012; Ogihara et al., 2015; Dawson et al., 2016; Moriarty & Ballard, 2016; Matsumoto & Kokubo, 2017). However, these investigations (i) did not consider collisional fragmentation; (ii) examined only initial semimajor axes larger than 0.04 AU, whereas we focused on smaller semimajor axes. Our fiducial simulations start with $3 M_\oplus$ of solid material, about three times larger than the mass expected in the minimum-mass solar nebula between 0.01 AU and 0.04 AU (Weidenschilling, 1977; Hayashi, 1981), which is comparable to the disk profiles used by the referenced authors (of course, the applicability of the minimum-mass solar nebula to these small radii is highly uncertain). Their simulations typically ran for $\sim 1$–10 Myr. Our simulations ran for only 0.3 Myr because the dynamical time at these small radii is much shorter; the shorter integrations are justified because the numbers and properties of the planets seemed to be stable well before the end of the integration (see Figure 4.6).
Not considered in this chapter is the potential loss of mass due to Poynting–Robertson (PR) drag on dust and small pebbles produced in collisional cascades. Bodies of 1 cm radius with density of 1 g cm$^{-3}$ around solar-type stars have PR drag lifetimes of $\sim$10 kyr at 0.04 AU and $\sim$1 kyr at 0.01 AU, with smaller bodies having shorter lifetimes proportional to their radius (e.g., Wyatt & Whipple 1950). If a sufficient number of other small bodies are present, lifetimes due to collisional cascades can be even less than PR lifetimes (e.g., Wyatt 2008). Depending on the amount of small bodies produced in fragmenting collisions, these processes may lead to significant mass loss, especially near $r_{\text{Roche}}$ due to the enhanced fragmentation that occurs there.

Also not considered in this chapter is the effect of large planets on wider orbits on the formation and evolution of inner planetary systems. A sufficiently close and massive gas giant could stir up the bodies and produce larger eccentricities and higher collision velocities while the planets are forming. The potential effects of gas giants depend on the relative times at which they form; the giant planets would need to be in place before inner planet formation has ceased in order to have an effect. Stirring from giant planets would also increase the separation necessary between bodies for long-term stability.

This chapter assumes no tidal damping of eccentricity or semimajor axis. This assumption is justified based on calculations of the tidal damping timescales (e.g., Jackson et al. 2008). For these calculations, we assumed a Sun-like star, a 0.5 M$_\odot$ planet with $a = 0.01$ AU, and $Q_{\text{star}} = 10^{5.5}$ and $Q_{\text{planet}} = 100$ for the tidal dissipation parameters of the star and planet. For these values, the eccentricity damping timescale is $\sim 10^5$ yr and the semimajor axis damping timescale is $\sim 10^{10}$ yr, compared to the $\sim$100–1000 year timescale for planets to form in our simulations and the $3\times10^5$ year length of our longest simulations. So, although eccentricity damping is negligible on the timescales of planet formation, it could affect the longer-term evolution and
cleanup of the system. Since eccentricity damping leads to lower-velocity collisions on average, it should not modify the main conclusion of this chapter, that planet formation is not suppressed by fragmentation.

This chapter also assumes there is no gas left in the disk at the later stages of planet formation examined here. With the short $\sim 100$ year formation timescale found in our simulations starting from lunar-sized bodies, though, this assumption may be incorrect. This, of course, depends on the gas dissipation timescale at small radii, which is highly uncertain.

Future work could improve on the collision algorithm of C13 by implementing a size distribution for the fragments resulting from a collision, as was done by LS12. This would lead to a richer variety of mass ratios in collisions than were present in this chapter. Additionally, our initial conditions assumed that all the bodies started out with the same mass. A more detailed follow-up could look at a more realistic size distribution for the initial bodies, as well as taking into account the continued formation of large bodies from accumulating planetesimals (similar to McNeil et al. 2005), which would also lead to a richer distribution in mass ratios. Since, as was shown in Section 4.4.3, collisional growth is robust across all mass ratios, we do not expect that a more realistic size distribution of collision fragments or initial bodies will affect our main conclusions.

As was pointed out in Section 4.4, differentiated bodies behave differently in collisions than homogeneous bodies, and bodies of the masses encountered in this chapter ($\gtrsim$lunar mass) are likely to have significant differentiation. Head-on collisions of differentiated bodies have higher values of $f_2$ relative to homogeneous bodies and thus require higher-velocity collisions to lead to fragmentation, while off-center collisions have lower values of $f_2$ (Asphaug 2010), so these two effects will partially cancel each other. Even if the off-center collisions with a lower $f_2$ dominate, the analytic calculation and arguments of Section 4.4.3 give $a_{\text{frag}} \approx 1.1 \ r_{\text{Roche}}$ as likely for equal-mass
bodies; thus, differentiation of bodies is not expected to modify our conclusions significantly. A collision prescription that combines the general applicability and scaling laws of LS12 with results of collisions between differentiated bodies similar to Asphaug (2010) would allow for simulating systems of differentiated bodies in similar style as this chapter.

4.8 Conclusion

We have carried out both analytic and numerical investigations of collisional rocky planet formation at small semimajor axes (∼0.01 AU) to determine whether collisional fragmentation is a barrier to planet formation. Our analytic argument (which ignores tidal effects) predicts that collisions leading primarily to mass growth are possible all the way down to the Roche radius and, thus, collisional fragmentation is not a barrier to planet formation. Our numerical integrations (which include tidal effects), starting with ∼lunar-sized bodies, are able to form planets all the way down to \( a \simeq 1.1 r_{\text{Roche}} \). Control integrations that ignore the effects of tides are able to form planets at even smaller semimajor axes, consistent with our analytic result. Our numerical results thus confirm that collisional fragmentation is not a barrier to rocky planet formation, except perhaps in a narrow range of distances within 10% of \( r_{\text{Roche}} \). The resulting planetary systems are expected to be stable over long time scales (∼10^{10} orbits).

The research in this chapter utilized the following software: matplotlib (Hunter, 2007), Mercury (Chambers, 1999), and numpy (Oliphant, 2006).
Chapter 5

A Search for Transiting Planets in the Globular Cluster M4 with K2: Candidates and Occurrence Limits

5.1 Abstract

We perform a search for transiting planets in the NASA K2 observations of the globular cluster (GC) M4. This search is sensitive to larger orbital periods ($P \lesssim 35$ days, compared to the previous best of $P \lesssim 16$ days) and, at the shortest periods, smaller planet radii ($R_p \gtrsim 0.3 R_J$, compared to the previous best of $R_p \gtrsim 0.8 R_J$) than any previous search for GC planets. Seven planet candidates are presented. An analysis of the systematic noise in our data shows that most, if not all, of these candidates are likely false alarms. We calculate planet occurrence rates assuming our highest signal-to-pink-noise candidate is a planet and occurrence rate upper limits assuming no detections. We calculate 3σ occurrence rate upper limits of 6.1% for 0.71–2 R$_J$ planets with 1–36 day periods and 16% for 0.36–0.71 R$_J$ planets with 1–10 day periods. The occurrence rates from *Kepler, TESS*, and RV studies of field stars
are consistent with both a non-detection of a planet and detection of a single hot Jupiter in our data. Comparing to previous studies of GCs, we are unable to place a more stringent constraint than Gilliland et al. (2000) for the radius–period range they were sensitive to, but do place tighter constraints than both Weldrake et al. (2008) and Nascimbeni et al. (2012) for the large-radius regimes to which they were sensitive.

5.2 Introduction

The globular cluster (GC) M4 (NGC 6121) was observed by the K2 mission (Howell et al. 2014) during its Campaign 2. These data underwent a blanket search for variable objects in our previous work (Chapter 3), but did not receive a focused search for planetary transits. Any constraints that could be put on planet occurrence rates in a GC would be of scientific interest. GCs provide more-or-less homogeneous populations of metal-poor stars—M4 in particular has a metallicity of \([\text{Fe/H}] \approx -1.2\) (Harris 1996, 2010 edition). As such, they would provide valuable test beds for theories about planet formation and its dependence on stellar metallicities (e.g., Ida & Lin 2004; Johansen et al. 2009; Ercolano & Clarke 2010; Johnson et al. 2010; Johnson & Li 2012; also see Chapter 4), assuming such formation mechanisms also take into account the denser stellar environment. These high stellar densities could also provide a fruitful testbed of dynamical planet formation and evolution theories. The relatively large number of stellar encounters in GCs, due to both their old ages and high stellar densities, are thought to kick planets out of planetary systems, especially those on wide orbits (Sigurdsson, 1992; Davies & Sigurdsson, 2001; Bonnell et al., 2001; Fregeau et al., 2006; Spurzem et al., 2009). However, stellar encounters are also expected to increase the probability of formation of hot Jupiters (HJs) via high-eccentricity migration in some cases (Hamers & Tremaine 2017). For reference for this work, a typical
definition of an HJ is a planet with a radius $\gtrsim 0.8 \, R_J$ and an orbital period $\lesssim 10$ days. An enhanced HJ occurrence rate in GCs may suggest a preference for dynamical formation pathways in our own neighborhood for close-in planets. And finally, HJs are expected to undergo tidal orbital decay on Gyr timescales (e.g., Penev et al. 2018), and the old ages of GCs may be helpful in testing this theoretical expectation.

M4 holds the distinction of possessing the only planet known in a GC, PSR 1620-26 b, a planetary-mass object orbiting a pulsar–white dwarf binary (Backer et al. 1993; Thorsett et al. 1993; Michel 1994; Rasio 1994; Arzoumanian et al. 1996; Thorsett et al. 1999; Ford et al. 2000; Richer et al. 2003), with its mass measured by Sigurdsson et al. (2003) to be $2.5 \pm 1 \, M_J$. Work since its discovery has shown that this planet may have formed later in the cluster’s life, rather than from a protostellar disk. For example, Beer et al. (2004) propose a model where a stellar encounter during the common envelope phase that led to the formation of the pulsar caused a dynamical instability in the dense equatorial wind formed as part of the common envelope phase. As such, PSR 1620-26 b may not be able to provide constraints on planet formation processes in GCs that are contemporaneous with star formation.

Given the relatively large distance to M4 of 1.8 kpc (Hendricks et al. 2012; Kaluzny et al. 2013b; Braga et al. 2015; Neeley et al. 2015), a wide-scale radial velocity (RV) survey to search for planets is impractical, but a photometric survey to search for transits is feasible. Despite its distance, M4 is nevertheless the closest GC, and it has a relatively sparse core, so it is perhaps the best target for discovering additional GC planets.

Previous searches for transiting planets have been made in GCs. The largest to date are the HST campaign of Gilliland et al. (2000) and the ground-based campaign of Weldrake et al. (2005), both searching for planetary transits in 47 Tuc. Gilliland et al. (2000) state that the reason for choosing this cluster (its distance is about twice that of M4) is because its spatial and main-sequence brightness distributions matched
well the capabilities of HST. They obtained high-precision photometry for ∼34,000 stars over an 8.3 day observing campaign. With the then-current understanding of HJ occurrence rates, they had expected to find ∼17 planets; however, they found none. More recently, Masuda & Winn (2017) revised the expected number of planets that Gilliland et al. would have found to $2.2^{+1.6}_{-1.1}$, based on an updated understanding of planet occurrence from the Kepler mission. This revised number additionally does not account for the lower metallicity of 47 Tuc ([Fe/H] $\approx$ −0.7; McWilliam & Bernstein 2008) relative to the Kepler stars, which are primarily field stars. This is expected to revise the number even lower due to the metallicity dependence of the occurrence rate of HJs (Santos et al. 2001; Fischer & Valenti 2005; Petigura et al. 2018). Weldrake et al. (2005) used the Australian National University 40 in telescope at Siding Spring observatory over 33 nights and obtained a much wider field of view than Gilliland et al.’s HST observations, observing out to 60% of 47 Tuc’s tidal radius. They obtained light curves for ∼110,000 stars (though only ∼20,000 of these had sufficiently low scatter to be sensitive to HJs) and could detect giant planets with periods up to 16 days. Their calculated expected planet yield was ∼7 planets (based on 1 R$_J$ planets with periods less than 16 days and an intrinsic formation rate of 0.8%), but they found none. Masuda & Winn (2017) revised the expected number of planets from this survey down by about a factor of four. Both the Gilliland et al. (2000) and Weldrake et al. (2005) surveys were not sensitive to much other than HJs. Other searches for transiting exoplanets in GCs include the search of Weldrake et al. (2008) in ω Cen, which for the most part was sensitive only to planets with radii >1.5 R$_J$ and had no detections, and Nascimbeni et al. (2012) in NGC 6397, which with a null detection and ∼5000 light curves, were not able to derive constraints on planet occurrence that fell below the occurrence rates measured by Kepler.

M4 was in the field of view of the Kepler telescope during Campaign 2 (running from 2014 August 23 to 2014 November 10) of the K2 mission and, as mentioned,
continuous observations of a portion of this cluster were included in the data downloaded from the telescope. Though the original proposals to obtain these data were focused on observing RR Lyrae variables in the cluster, the excellent photometry and long-term coverage of M4 allows for detecting other variable objects, potentially including transiting HJs. The Kepler telescope and detector were not designed with GC observations in mind: the $\sim4''/\text{pixel}$ image resolution leads to significant blending in the images, and the periodic telescope drift experienced during the K2 mission produces systematic noise in the photometry.

Despite these problems, the longer observation span of K2 Campaign 2 relative to the previous GC transiting planet searches potentially opens a new regime of planetary orbital period to explore and to place constraints on GC planet occurrence. The longer observations also increase the number of observed transits for orbits of a given period relative to the previous surveys, helping to boost sensitivity to smaller radius planets in the period ranges that have already been probed by other GC studies. Given the scientific motivation for finding planets in GCs and the new parameter space opened for exploration by these data, and since the reduced data have scientific utility in addition to permitting a transit search (see the variable catalog in Chapter 3), there is more than sufficient merit to motivate the effort for the search. We summarize our photometric reductions and explain our transit search methodology in Section 5.3, present the results of our transit search in Section 5.4, and provide planet occurrence rates and limits in Section 5.5. We then discuss the results in Section 5.6 and conclude in Section 5.7.
5.3 Method

5.3.1 Photometric Processing

Our photometric processing pipeline is fully described in Chapter 3—Section 3.3 in particular—and is similar to that of Soares-Furtado et al. (2017). We provide a brief summary here. M4 was observed by K2 for \( \sim 79 \) days in 2014, during the mission’s Campaign 2. Given the high degree of blending in the images, we decided to use image subtraction (Alard & Lupton, 1998) to extract light curves for the objects. The Gaia first data release (DR1) source catalog (Gaia Collaboration et al. 2016a,b) was used as an astrometric and photometric reference catalog. This was used instead of the Gaia second data release (DR2) catalog owing to our beginning this study prior to Gaia DR2. We included all sources with \( G < 19 \). In using Gaia DR1 as a photometric reference catalog, we had to convert from \( G \) to our Kepler instrumental magnitudes. We found that a simple additive conversion was all that was needed, likely due to the similar bandpasses of the two instruments. The photometry was extracted for apertures of 1.5, 1.75, 2.0, 2.25, 2.5, 2.75, and 3.0 pixels in radius. At the end of our processing, we found which aperture radius minimized the photometric scatter as a function of \( G \) and used the data from the corresponding aperture for all objects of a given magnitude throughout our analysis. For reference, the typical full-width at half maximum value for the point spread function of the images was \( \sim 1.5 \) pixels.

After extracting this raw photometry, systematic variability from the spacecraft roll was cleaned up based our implementation of the algorithm developed by Vanderburg & Johnson (2014) and Vanderburg et al. (2016). The light curves were then further cleaned of common systematic trends using the trend filtering algorithm (TFA; Kovács et al. 2005) as implemented in VARTOOLS (Hartman & Bakos 2016). In this work, all objects are referred to by the identifiers assigned them in Chapter 3; see
Table 3.1. All of the raw and processed light curves are available at Wallace et al. (2019a)\textsuperscript{1}.

5.3.2 Transit Search

The 4554 light curves produced were searched for planet transits using the \textsc{varTools} implementation of the box-fitting least squares (BLS; Kovács et al., 2002) algorithm. The light curves were sigma clipped prior to the search (5\(\sigma\), three iterations). We ran some injected transits through our pipeline to ensure that they would be recoverable even with the photometric post-processing and sigma clipping. We searched periods between one day and the maximum observation length of the given light curves (most having the maximum length of \(\sim 78\) days, which is slightly shorter than the full span of observations owing to our need to trim the first \(\sim 1\) day of data). A one day orbit around the most massive and evolved cluster-member stars under consideration (0.81 \(M_{\odot}\)) has an orbital semimajor axis of 3.9 \(R_{\odot}\), compared to a stellar radius of 4.9 \(R_{\odot}\) for these same stars. Thus for the handful of the most evolved stars under consideration, one day orbits are not possible, but they are possible for the vast majority of the stars we consider.

The range of values for the fractional transit duration \(q\) used in the search varied between 0.1 \(\times q_{\text{exp}}\) and 2 \(\times q_{\text{exp}}\), with \(q_{\text{exp}}\) being the expected transit duration at a given period based on the density of the given star (see the next two paragraphs), assuming a circular orbit, and that the impact parameter \(b\) is zero. The minimum \(q\) searched was adjusted as necessary so that it was never less than \((\Delta t)_{\text{min}}/P\), with \((\Delta t)_{\text{min}}\) being the minimum time between observations (the Kepler cadence, 29.4 minutes, adjusted slightly based on the actual BJD values of the observations) and \(P\) being the period being searched. The number of phase bins used was set to \(2/q_{\text{min}}\).

\textsuperscript{1}Published at Princeton University’s DataSpace and licensed under a Creative Commons Attribution 4.0 International License, accessible at http://arks.princeton.edu/ark:/88435/dsp01h415pd368
with \( q_{\text{min}} \) here being \( 0.1 \times q_{\text{exp}} \), up to a maximum value of 2100. This value is based on the expected transit duration of the shortest-period planets around the smallest-radius stars in our sample. With having less than 4000 measurements per light curve, a greater number of phase bins would not have been useful anyway.

Stellar densities were calculated based on an isochrone fit for the cluster and the \( V \) magnitudes taken from Mochejska et al. (2002). For objects not found in the Mochejska et al. (2002) catalog, their \( V \) magnitudes were converted from \( G \) based on a second-degree polynomial fit between \( G \) and the Mochejska et al. (2002) \( V \) magnitudes of the matched objects. The isochrone was taken from the calculations of Yi et al. (2001). For the isochrone, we assumed a metallicity of \([\text{Fe/H}]=-1.2 \) (Harris 1996, 2010 edition) and an estimated age of 11.3 Gyr based on our fit to the data. The \( V \) magnitudes were used instead of \( \text{Gaia} \ G \) because this calculation was performed earlier in \( \text{Gaia}'s \) mission and our isochrone database had not yet incorporated results that used the \( \text{Gaia} \) bandpass. The \( V \) magnitudes were converted to absolute magnitudes assuming a cluster distance of 1.8 kpc (Hendricks et al., 2012; Kaluzny et al., 2013b; Braga et al., 2015; Neeley et al., 2015) and an \( A_V \) extinction of 1.24 mag based on our fit to the data, a value \( \sim 0.15 \) less than the mean value found by Hendricks et al. (2012). We also note that there is differential reddening across the cluster, with Hendricks et al. (2012) finding the difference between the lowest and the highest \( E(B-V) \) values to be \( \sim 0.2 \) mag. Our isochrone density determination produced incorrect densities for objects that were not cluster members, but since they were not included in our final analysis this did not matter.

The use of stellar density in the BLS calculation tailors the transit duration range that is searched to those expected for a planet around a star of the given density. This boosts sensitivity to physically likely transit scenarios. This is in contrast to our previous search in Chapter 3, which was not restricted to only planetary transit-minded transit durations, and is part of the reason we were better able to find planet...
candidates in this search (the previous search found none). Given the fairly wide
distribution in $q$ that we search ($0.1 \times q_{\text{exp}}$ to $2 \times q_{\text{exp}}$), the search is insensitive to
modest errors in density.

The phase-folded light curves and periodograms for the periods of the top five pe-
riodogram peaks were then examined by eye for significance. We used the checkplot
module of astrobase (Bhatti et al. 2017) for this by-eye examination. As part of
this, obvious non-planet-transit physical signals were excluded (e.g., RR Lyrae and
eclipsing binaries, as well as objects blended with them). We also found that a large
and temporary systematic variation in brightness occurred about halfway through the
observation in the light curves of many of the objects, at the point in the campaign
when the roll direction of the spacecraft changed. This light curve systematic in some
cases phased up with other outliers to produce large-period BLS signals; cases where
this happened were determined by eye and removed. In ambiguous cases, we erred
on the side of completeness and included the objects in our subsequent consideration,
since we did not want to impose an unquantified limit on signal-to-pink-noise (S/PN)
that was stricter than the hard limit that we used—only those objects with a S/PN
greater than eight were examined by eye. Approximately half of the total number of
objects passed this limit. We wish to note that, at this stage, all of the objects were
considered without respect to their cluster membership status. Although our initial
S/PN threshold was set to eight, we found later that a higher threshold should be
used, as detailed subsequently in Section 5.3.3.

5.3.3 Signal-to-pink Noise Threshold Determination

Given the residual systematic noise left in our data—largely leftover from the roll-
correlated variability that was mostly but not entirely removed by our processing
pipeline—we found S/PN to be a useful metric in evaluating signal significance. We
used the S/PN as calculated by VARTOOLS, which is based on the definition of Pont
et al. (2006). The signal value used in this calculation is the BLS transit depth and the pink noise is a quadrature sum of the light curve white noise divided by the number of points in transit and the light curve red noise (calculated from the RMS of the binned light curve with bin size equal to the transit duration) divided by the number of transits. After an initial search through the BLS search results, we saw transit-like signals that had lower S/PN than would be expected based on the observed white noise, suggesting a significant amount of correlated noise that could mimic transits. To better understand how well the correlated noise could mimic transits in our data, we reran our BLS search with the same parameters as before, but this time looking instead for “anti-transits”, periodic box-shaped brightenings in the data instead of dimmings. Since there are no common astrophysical phenomena that can produce such brightening signals at the $\sim 10$ mmag level that Jupiter-sized planets produce for dimmings, these presumably are all due to noise. Gravitational self-lenses from binary systems consisting of a neutron star/black hole and a main sequence star can produce such signals, but the occurrence rates for such objects are expected to be low (e.g., Masuda & Hotokezaka 2018 expect the TESS survey to have a detection rate of such objects of $\sim 10^{-4}$).

The by-eye vetting was performed again, with both the transit and anti-transit results presented. The light curves presented for the anti-transit search results were inverted so that the signals would appear as transits. No special indication was given during the manual vetting as to whether a given signal was a transit or anti-transit, permitting a blind vetting of the signals. We implemented several cuts based on the BLS statistics. Only those signals with $S/PN>8$, $q/q_{\text{exp}} \geq 0.25$ (or 0.5 if $8 < S/PN < 9$), number of transits $n_t \geq 3$, and number of points in transit $n_{\text{pit}} \geq 15$ were examined. We recorded all signals that we thought were possible transits, identifying 27 transits and 17 anti-transits as planet candidates (though here we call the anti-transits “planet candidates,” we note that since these are not transits they
cannot be actual planet candidates). We then examined the distributions in S/PN for both the transits and anti-transits. These distributions are shown in Figure 5.1. A Kolmogorov-Smirnov (KS) test of the distributions for our 27 transit and 17 anti-transit planet candidates has a $p$-value of 0.51 and a KS statistic of 0.24, indicating that we cannot reject the hypothesis that the planet-candidate transits and anti-transits are drawn from different distributions of S/PN.

Based on these results, it is possible our planet candidate signals are due to correlated noise and are thus false alarms. Despite this, we present our strongest candidates in Section 5.4. Based on our results, we decided that a S/PN cutoff value of 12 would be used in our transit–injection–recovery pipeline to quantify our sensitivity to planetary transits. We do have one candidate with S/PN > 12, W2282, but with an S/PN value of 12.3, it is still only of marginal significance.

### 5.3.4 Occurrence Rate Calculation

We focused our occurrence rate calculation only on stars that are likely cluster members by including only those objects with membership probabilities greater than 99% as calculated by Wallace (2018a) (see also Appendix A) using Gaia DR2 proper motions. There were 3784 such objects. W2282, the S/PN > 12 star, is among the cluster members. For our occurrence rate calculation, we also decided to focus only on main sequence and subgiant stars. We imposed a cutoff of $G > 14$ to focus on these objects, leaving us with 3704 objects for the calculation. Figure 5.2 shows a color–magnitude diagram (CMD) of the cluster members in our analysis with this cutoff indicated.

As a first step to calculating occurrence rates from our results, we quantified our transit detection efficiency. To do this, we injected transits into our light curves to test how well we could recover them. The transits were injected using VARTOOLS, based on the transit model of Mandel & Agol (2002). The injected periods and planet radii were taken from a 5x5 grid, with periods drawn uniformly from uniform bins in period
Figure 5.1: Histograms of signal-to-pink noise (S/PN) for all signals that exceed our thresholds and for our selected planet candidates. The distributions for the transits are shown in blue and the distributions for the anti-transits (labeled as “Anti-tr.” in the figure) are shown in orange. The top panel shows all the values that cross our S/PN, transit duration, number of transits, and number of points in transit thresholds (see text) and the bottom panel shows just those candidate signals that were selected in our by-eye vetting. The vertical line in the bottom panel shows our chosen S/PN cutoff value of 12.
Figure 5.2: Color–magnitude diagram for objects in our analysis with a cluster membership probability >99%. The photometric data are taken from Gaia DR2 (Brown et al. 2018; Riello et al. 2018). The horizontal dashed line shows our magnitude cut for objects considered in our occurrence rate calculation, with objects below the line being included. The solid black line shows the isochrone fit used for transit injection and recovery, as described in the text.
between one and 36 days and planet radii drawn log-uniformly from log-uniform bins between 0.3 and 2.0 R_J.

The stellar radii and masses were determined from the PARSEC stellar evolution models (Marigo et al., 2017), obtained through the CMD v3.2 web interface\(^2\), this time making direct use of the Gaia DR2 G magnitudes and using the bolometric corrections for the Gaia band-passes from Maíz Apellániz & Weiler (2018). Through trial-and-error we determined the best fit to the G vs. G_BP − G_RP CMD to be provided by a PARSEC isochrone with an age of 12.5 Gyr, a metallicity of [Fe/H]=−1.2, a distance of 1.8 kpc, extinction in the G band A_G = 1.4, and reddening E(G_BP − G_RP) = 0.57. This isochrone is shown in Figure 5.2. The stellar parameters were determined from the isochrone using just G magnitudes. There were seven objects with 14.0 < G < 14.28 for which the isochrone interpolation as we had implemented it failed: W364, W642, W1643, W1898, W1912, W2757, and W3684 (using the identifiers from Chapter 3). We exclude these objects from the occurrence rate calculation.

Random eccentricities, phases, longitudes of periapsis, and inclinations were chosen for each injected transit. Phases and longitudes of periapsis were chosen uniformly between zero and 2π, inclinations were chosen uniform in \(\cos i\) subject to the constraint that transits actually occur, and eccentricities were drawn from a Beta distribution, with parameters as determined by the empirical fit of Kipping (2013) to his short-period planets. The planet mass was fixed at 0.8 M_J for all transit injections independent of injected radius, as the simulated transit signal is effectively independent of the planetary mass. Limb darkening was incorporated with a quadratic model, using the parameters determined by Claret (2018) for Kepler using the PHOENIX-COND model (Husser et al., 2013).

The transit-injected light curves were then run through the same photometric processing as the light curves searched for planetary transits: decorrelation of systematic

\(^2\)http://stev.oapd.inaf.it/cgi-bin/cmd
brightness variations against the telescope roll and TFA. Due to time constraints, we were unable to run a full BLS search for each transit-injected light curve. Instead, we used the \texttt{-BLSFixPer} option of \textsc{VARTOOLS} to perform a BLS search at only the injected period in order to get the BLS statistics. Our S/PN cut of 12 was applied, as well as the additional cuts used in our planet search (see Section 5.3.3), namely: $q/q_{\text{exp}} > 0.25$, $n_t \geq 3$, and $n_{\text{pit}} \geq 15$. Using \texttt{-BLSFixPer} is effectively a conservative approach, as targets that have S/PN$<12$ at the injected frequency, but S/PN$>12$ at other frequencies (such as a harmonic of the transit period), will be excluded in our search whereas they may have been recovered in a full BLS search.

We ran some initial reconnaissance runs of our transit–injection–recovery pipeline with a coarser period–radius grid consisting of three period bins and four fixed planet radii and 12 samples from each bin. The periods were not sampled uniformly from each period bin, but rather from a range of the smallest periods in each bin. The purpose of these runs was to determine, star-by-star, parameter ranges in which we might expect to have a near-0% transit recovery rate. This information could then be used to accelerate the subsequent calculations. Applying just the S/PN cut and not the other cuts in $q/q_{\text{exp}}$, $n_t$, and $n_{\text{pit}}$, we found 442 objects for which none of the 144 injected transits were recovered. Many of these objects were significantly blended with brighter and/or variable objects and their light curves had very large scatter. These objects were removed from subsequent consideration, which, with the seven objects that were not fit by the isochrone, left us with 3255 objects that were included in our final transit–injection–recovery analysis. Additionally, for a given object, those period bin/planet radius pairs that had no recovered transits were recorded. Injected transits with periods equal to or longer and planet radii equal to or smaller than the values represented by these period bin/planet radius pairs were automatically recorded as non-recoveries. This was about 50% of our injected transits. If this approximation lead to us missing some injected transits that may have been recovered, then our final
occurrence rate upper limits will be higher than we would have otherwise calculated, making this a conservative approximation for the upper limits.

We then injected 56 transits into each of the raw light curves for each radius–period bin in our $5 \times 5$ grid and ran them through our photometric processing pipeline. Recovered transits were then determined based on the BLS statistics and associated cuts as discussed. Then for each radius–period bin, we calculated the number, $N$, of expected planets that we would have detected if every star hosted one planet in that bin, using (from, e.g., Ford et al. 2008)

$$N = \sum_{i}^{n_*} \frac{1}{n_i} \sum_{j}^{n_i} \delta_{ij} \frac{(R_{*,i} + R_{p,ij})(1 - e_{ij} \cos \varpi_{ij})}{a_{ij} \times (1 - e_{ij}^2)}, \quad (5.1)$$

where $i$ is an index over the stars examined, $n_*$ is the number of stars examined, $j$ is an index over the individual transit injections, $n_i$ is the number of transit injections performed for the star, $\delta_{ij}$ is one if the particular injected transit is recovered and zero if not, $R_{*,i}$ is the stellar radius of the star $i$ (based on the isochrone interpolation; this is the same stellar radius used for the transit injection), $R_{p,ij}$ is the radius of the planet for the given injected transit (taken as the actual value used for the transit injection rather than a calculated radius recovered from the transit signal), and $e_{ij}$, $\varpi_{ij}$, and $a_{ij}$ are respectively the eccentricity, longitude of periapsis, and semimajor axis of the orbit of the injected transit. The quantity $(R_{*,i} + R_{p,ij})(1 - e_{ij} \cos \varpi_{ij})/[a_{ij} \times (1 - e_{ij}^2)]$ accounts for the probability of transit given the random inclinations of orbits.

Once $N$ is calculated for a given radius–period bin, the $3\sigma$, 99.73% confidence interval upper limit for the occurrence rate assuming no detections is calculated using the binomial distribution, with $N$ rounded to the nearest integer. For the bin in which our $S/\mathrm{PN} > 12$ planet candidate falls, the $3\sigma$ confidence interval for the occurrence rate is also calculated. When calculating occurrence rates and limits for comparisons with other works, Equation (5.1) is again used to calculate the expected number of
planets but with injected transits chosen from a selected radius–period range instead of just the fixed bins we drew from for the transit injections.

In performing the calculation as we have, there is an implicit assumption that 100% of our injected transits would appear in the five highest BLS peaks in the full BLS search (since that is the number of peaks used in our planet search). There is also an assumption that 100% of injected transits that exceed our cutoff values would be identified in our by-eye analysis. We ran a full BLS calculation on a subset of our injected transits and found that 97.5% of injected transits that exceed our cutoff values appear in the five highest BLS peaks. We also performed a by-eye vetting of approximately 500 injected transits that exceed our cutoff thresholds with \( 12 < \frac{S}{PN} < 12.1 \) and found 98.8% passed our vetting. Presumably an even larger fraction of those with higher \( \frac{S}{PN} \) values would pass the by-eye vetting. Based on these results, we decided to maintain our assumptions of 100% recovery for both of these steps.

5.4 Transit Search Results

Figure 5.3 shows the phase-folded light curves of our seven most promising planet candidates. We choose not to present the other 20 candidates that initially passed our by-eye vetting as we now consider these to almost certainly be false alarms. Table 5.1 presents information on the stars hosting these planet candidates and Table 5.2 presents information for each of the transit signals and calculated planet properties. All of the planet candidates presented are proper motion members of the cluster (Wallace 2018a). Except for W2282, all of these objects fall below our \( \frac{S}{PN} \) threshold of 12, and W74 is brighter than our \( G \) threshold of 14. Owing to the potential scientific impact of discovering a transiting exoplanet in a GC, we choose to present the most promising candidates we found irrespective of these cuts. That being said, the
Figure 5.3: Phase-folded light curves for the transit signals of the best planet candidates. Each panel is for a different candidate. In the upper-left corner of each panel is shown, from top to bottom, the object’s identifier and the S/PN. In the upper-right corner is shown, from top to bottom, the period in days and the median magnitude subtracted off for the light curve. In each panel, the gray points are the individual measurements (subject to a $5\sigma$ sigma clipping with three iterations) and the blue points are binned-weighted-mean values. The red line shows the BLS fit to each phase-folded light curve.
Table 5.1. Information on Stars Hosting Planet Candidates

<table>
<thead>
<tr>
<th>ID</th>
<th>Gaia DR2 ID</th>
<th>R.A.</th>
<th>Decl.</th>
<th>G(^d)</th>
<th>Radius(^e)</th>
<th>Mass(^f)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(hh:mm:ss)</td>
<td>(dd:mm:ss)</td>
<td>(mag)</td>
<td>(R(_\odot))</td>
<td>(M(_\odot))</td>
<td></td>
</tr>
<tr>
<td>W2282</td>
<td>6045466502667197056</td>
<td>16:23:34.95</td>
<td>-26:29:14.2</td>
<td>18.32</td>
<td>0.68</td>
<td>0.66</td>
</tr>
<tr>
<td>W2863</td>
<td>6045466640106160128</td>
<td>16:23:41.10</td>
<td>-26:28:04.2</td>
<td>15.07</td>
<td>3.4</td>
<td>0.80</td>
</tr>
<tr>
<td>W74</td>
<td>6045477635223138432</td>
<td>16:22:57.99</td>
<td>-26:28:46.8</td>
<td>12.30</td>
<td>13</td>
<td>0.86</td>
</tr>
<tr>
<td>W1955</td>
<td>6045503091478311808</td>
<td>16:23:31.71</td>
<td>-26:22:33.7</td>
<td>18.23</td>
<td>0.69</td>
<td>0.67</td>
</tr>
<tr>
<td>W3055</td>
<td>6045501996278191104</td>
<td>16:23:43.33</td>
<td>-26:25:06.1</td>
<td>18.57</td>
<td>0.64</td>
<td>0.64</td>
</tr>
<tr>
<td>W1184</td>
<td>6045478597295755520</td>
<td>16:23:22.89</td>
<td>-26:27:04.2</td>
<td>15.02</td>
<td>3.5</td>
<td>0.80</td>
</tr>
<tr>
<td>W3128</td>
<td>6045466429642662272</td>
<td>16:23:44.26</td>
<td>-26:29:12.0</td>
<td>18.13</td>
<td>0.71</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Note. — All of these stars are proper motion cluster members (see Appendix A).

\(^a\) The identifier by which the object is known in this work, the same as in Chapter 3.

\(^b\) Gaia DR2 source ID.

\(^c\) J2000.0; data taken from Gaia DR2 (Lindegren et al., 2018b).

\(^d\) Gaia G magnitude from Gaia DR2 (Riello et al., 2018).

\(^e\) The radius of the star in units of solar radii, determined from an isochrone fit.

\(^f\) The mass of the star in units of solar mass, determined from an isochrone fit.
Table 5.2. Information on Planet Candidates

<table>
<thead>
<tr>
<th>ID</th>
<th>Period&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Epoch&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Depth&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Radius&lt;sup&gt;e&lt;/sup&gt;</th>
<th>q&lt;sup&gt;f&lt;/sup&gt;</th>
<th>q/q&lt;sub&gt;exp&lt;/sub&gt;&lt;sup&gt;g&lt;/sup&gt;</th>
<th>S/PN&lt;sup&gt;h&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>W2282</td>
<td>1.29</td>
<td>2061.96</td>
<td>22</td>
<td>1.0</td>
<td>0.047</td>
<td>0.82</td>
<td>12.3</td>
</tr>
<tr>
<td>W2863</td>
<td>29.1</td>
<td>2078.59</td>
<td>1.4</td>
<td>1.3</td>
<td>0.011</td>
<td>0.36</td>
<td>9.9</td>
</tr>
<tr>
<td>W74</td>
<td>18.8</td>
<td>2071.69</td>
<td>0.21</td>
<td>1.9</td>
<td>0.049</td>
<td>0.33</td>
<td>9.9</td>
</tr>
<tr>
<td>W1955</td>
<td>3.66</td>
<td>2062.66</td>
<td>7.5</td>
<td>0.62</td>
<td>0.028</td>
<td>1.0</td>
<td>9.0</td>
</tr>
<tr>
<td>W3055</td>
<td>2.23</td>
<td>2062.00</td>
<td>8.9</td>
<td>0.62</td>
<td>0.037</td>
<td>1.0</td>
<td>9.0</td>
</tr>
<tr>
<td>W1184</td>
<td>8.07</td>
<td>2066.85</td>
<td>0.51</td>
<td>0.82</td>
<td>0.087</td>
<td>1.2</td>
<td>8.6</td>
</tr>
<tr>
<td>W3128</td>
<td>10.1</td>
<td>2063.45</td>
<td>21</td>
<td>1.1</td>
<td>0.017</td>
<td>1.1</td>
<td>8.2</td>
</tr>
</tbody>
</table>

Note. — All of these stars are proper motion cluster members (see Appendix A).

<sup>a</sup>The identifier by which the object is known in this work, the same as in Chapter 3.

<sup>b</sup>The period of the transit signal in days.

<sup>c</sup>The epoch of the transit signal, taken at the center of the transit.

<sup>d</sup>The depth of the transit signal in millimagnitudes.

<sup>e</sup>The calculated radius of the planet in Jupiter radii based on the transit depth and isochrone-based stellar radius.

<sup>f</sup>The fractional transit duration.

<sup>g</sup>The ratio of the fractional transit duration with the expected $b = 0$ transit duration.

<sup>h</sup>The signal-to-pink noise of the signal.
relatively low S/PN values these signals have indicates that most, if not all, of these candidates are likely false alarms. The phase-folded light curves of W74, W1184, W2863, and W3128 appear to be the most robust of the seven, while the other three appear less robust. Given the high probability that these signals are false alarms, follow up is needed before they are confirmed. The next step in following these up would be to confirm the transits and then look for background objects to ensure these objects are not blended eclipsing binaries. For ∼1-R_J objects, additional RV follow up would be needed to measure the masses to identify them as planets, brown dwarfs, or late M-dwarfs. For Neptune-sized objects and smaller, the follow-up photometric data may be sufficient to classify the objects as planets without RV data.

Owing to the crowded nature of the cluster and of the K2 observations in particular, blending is a virtually unavoidable aspect of the data. We confirmed that W1184, W1955, W2282, W3055, and W3128 had the largest signal amplitudes of all nearby objects at the respective transit periods, for those objects for which we had light curves. The results for W74 and W2863 were more ambiguous, likely owing to their brighter magnitudes impacting larger areas of the images than the typical fainter stars. However, these two stars are each the brightest stars in their areas of the images.

W74 merits some additional comments. Its CMD position puts it on the red giant branch and it is asteroseismically active. A Generalized Lomb–Scargle (GLS, Lomb 1976; Scargle 1982) search reveals significant sinusoidal variability at a variety of periods (though not the ∼19 days found by BLS), with the strongest variability at 0.77 and 1.69 days. Pre-whitening the light curve by running LS three times and removing a two-harmonic and one-subharmonic fit to the peak period each time prior to running BLS recovers a similar period as before—18.843 days—and a comparable though lesser S/PN value of 9.6. It also bears mentioning that the ∼18.8 day period of this object, based on our Yi et al. (2001) isochrone fit for the mass and radius of
the star, has a semimajor axis of $\sim 28 \, R_\odot$, compared to the calculated stellar radius of $\sim 13 \, R_\odot$. This is a physically plausible scenario, but again this is a blended object and we were not able to conclusively determine if the transit belonged to this object. A blended eclipsing binary scenario is also possible. The implied planet radius based on the calculated stellar radius is $1.9 \, R_J$.

A few of our “planet candidate” anti-transits are shown in Figure 5.4 as examples of the kinds of false alarms the correlated noise in our light curves can produce. While the S/PN values are comparable to those of our prospective planet candidates, we think that a qualitative, by-eye evaluation of the signals show W74, W1184, W2863, and W3128 in particular to be more physical and transit-like than even the highest S/PN anti-transits. Also, those four objects have much longer periods than any of the anti-transits we identified, suggesting that the transit-mimicking correlated noise may exist only at shorter periods and that these longer-period signals may be more likely to be real.

5.5 Occurrence Rate Results

5.5.1 Transit Recovery Results

Figure 5.5 shows the recovery efficiency of our injected transits across our radius and period bins. We define recovery efficiency as the fraction of transits that were successfully recovered, and in Figure 5.5, this is the efficiency across all of the injected transits and all of the stars. As would be expected, the recovery efficiency trends towards higher values for larger planets and smaller orbital periods. Of note, though, is that for period bins greater than 8 days, the recovery efficiency is higher for our second-largest planet radius bin ($\sim 0.9$ to $\sim 1.4 \, R_J$) than for our largest planet radius bin ($\sim 1.4$–2 $R_J$). A possible explanation for this is that the deeper transits produced by the larger radius planets were more likely to be distorted and diminished by our
Figure 5.4: Same as Figure 5.3 but for a few representative anti-transits. These are presented as examples of the false alarms that can be produced by the systematic noise that exists in our data.
Figure 5.5: Recovery efficiency of our transit–injection–recovery pipeline. Each bin shows the fraction of injected transits that were successfully recovered across all the stars. The number in the lower-right corner of each bin shows the efficiency value, which is also represented by the color of the bin and the associated color bar. In cases where the recovery efficiency was less than 1%, an upper limit of 1% is shown.
Figure 5.6: Recovery efficiency of our transit–injection–recovery pipeline, broken down by stellar magnitude and bins of injected planet radius and orbital period. Magnitude is represented along the horizontal axis, planet radius by the three panels, and orbital period by the color. The lines show the median recovery efficiency as a function of magnitude across all stars for a given radius–period bin. In the rightmost panel, the 22–29 day (red) and 29–36 day (blue) lines fall behind the 29–36 day (yellow) line and thus do not appear. The planet radius range for each panel is indicated in the bottom of the panel. The legend in the rightmost panel shows the color representation of the orbital period bins (“d” in the legend stands for “day”) and applies to all three panels. G converted to stellar radius via an isochrone fit is shown on the top of each panel.

The shorter-duration transits at smaller periods may be less likely to be impacted by the processing pipeline, which would explain the higher recovery efficiency seen for the larger planets for periods shorter than eight days. Also, we found that some of the deepest transits had the bottom portions of the transits trimmed by the sigma clipping. Such transits were still detectable by BLS but had a lower S/PN due to the diminished apparent transit depth.

Figure 5.6 shows the recovery efficiency broken down by G magnitude, orbital period, and planet radius. We see (as expected) that shorter-period planets have a higher recovery efficiency than longer-period planets. We also see lower recovery effi-
ciencies for the brightest stars relative to the peak efficiencies reached (usually around \( G \approx 16–17 \)). This is due to the large radii of the brightest stars (see top axis of Figure 5.6) diminishing the transit depth and thus the signal size and recoverability. In the leftmost panel of Figure 5.6, corresponding to the largest-radius injected planets, we also note that detection efficiencies tend to be higher at both \( G = 16 \) and \( G = 19 \) than at \( G = 18 \), particularly in the 1–15 day period range. The non-monotonic variation in detection efficiency with magnitude is due to the different magnitude dependencies of two competing effects. Brighter stars in the cluster have higher precision light curves, which tend to increase the signal to noise of the transits. However, fainter stars in the cluster have smaller stellar radii, which leads to deeper transits for a given planetary radius. In the middle and rightmost panels of Figure 5.6, this increase in recovery efficiency at the faintest magnitudes is not seen. In the rightmost panel, we see a large drop in recovery efficiency overall for 0.64–0.94 R_J planets relative to the other two panels—the larger radii planets.

### 5.5.2 Planet Occurrence Rates and Limits

We now present our calculated occurrence rate limits and compare with other published occurrence rates. Figure 5.7 shows the calculated occurrence rate upper limits across our radius–period bins, and in the case of the bin containing our single S/PN > 12 planet candidate (W2282), the range for the occurrence rate if the planet candidate is real. For our shortest period bins, we are able to get down to limits of 1.6–3.5% for bins with planet radius larger than 0.64 R_J. To put these limits in context, Table 5.3 compares our occurrence rate limits with those of works using Kepler, TESS, or RV surveys for field stars. These previous works are: Kepler-based occurrence rates from Howard et al. (2012), Fressin et al. (2013), Masuda & Winn (2017), and Petigura et al. (2018); a TESS-based occurrence rate from Zhou et al. (2019); and RV occurrence rates from Mayor et al. (2011) and Wright et al. (2012). Table 5.4 com-
Figure 5.7: Calculated upper limits on occurrence rates for our radius–period bins. The lower-right hand of each bin shows the 3σ upper limit in the fraction of stars having at least one planet in that bin as calculated using a binomial distribution based on our determined detection efficiencies and transit probabilities. Those bins marked “N/A” either had too low of detection efficiencies for us to calculate any occurrence rate or had a rate that was indistinguishable from 100%. The color of each bin is a representation of the occurrence rate upper limits, based on the color bar at the bottom of the figure. The white point represents the one planet candidate we found that passes our S/PN threshold, W2282, and the range in the upper-right corner of the associated bin is the 3σ range on the occurrence rate assuming the planet candidate is real.
Table 5.3. Comparison with *Kepler*, *TESS*, and RV Occurrence Rates for Field Stars

<table>
<thead>
<tr>
<th>Per. Range</th>
<th>Rad. Range</th>
<th>Reference</th>
<th>Published Rate</th>
<th>Our Upper Limit</th>
<th>Our Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>(day)</td>
<td>(R$_J$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8–10</td>
<td>0.71–2.85</td>
<td>Howard et al. (2012)</td>
<td>0.4 ± 0.1%</td>
<td>&lt; 2.2%</td>
<td>0.38–3.3%</td>
</tr>
<tr>
<td>0.8–50</td>
<td>0.71–2.85</td>
<td>Howard et al. (2012)</td>
<td>1.3 ± 0.2%</td>
<td>&lt; 6.1%</td>
<td>1.1–9.1%</td>
</tr>
<tr>
<td>0.8–10</td>
<td>0.36–0.71</td>
<td>Howard et al. (2012)</td>
<td>0.5 ± 0.1%</td>
<td>&lt; 16%</td>
<td>...</td>
</tr>
<tr>
<td>0.8–17</td>
<td>0.54–1.96</td>
<td>Fressin et al. (2013)</td>
<td>0.43 ± 0.05%</td>
<td>&lt; 2.6%</td>
<td>0.44–3.9%</td>
</tr>
<tr>
<td>0.8–29</td>
<td>0.54–1.96</td>
<td>Fressin et al. (2013)</td>
<td>0.70 ± 0.08%</td>
<td>&lt; 3.8%</td>
<td>0.66–5.7%</td>
</tr>
<tr>
<td>0–10</td>
<td>0.8–2</td>
<td>Masuda &amp; Winn (2017)</td>
<td>0.93 ± 0.10%</td>
<td>&lt; 6.0%</td>
<td>1.0–8.9%</td>
</tr>
<tr>
<td>0–10</td>
<td>0.8–2</td>
<td>Masuda &amp; Winn (2017)</td>
<td>0.43$^{+0.07}_{-0.08}$%</td>
<td>&lt; 2.1%</td>
<td>0.36–3.2%</td>
</tr>
<tr>
<td>1–10</td>
<td>0.71–2.14</td>
<td>Petigura et al. (2018)</td>
<td>0.57$^{+0.14}_{-0.13}$%</td>
<td>&lt; 2.2%</td>
<td>0.38–3.3%</td>
</tr>
<tr>
<td></td>
<td>0.9–10</td>
<td>Zhou et al. (2019)</td>
<td>0.41 ± 0.10%</td>
<td>&lt; 2.1%</td>
<td>0.36–3.2%</td>
</tr>
</tbody>
</table>

**KEPLER Studies**

<table>
<thead>
<tr>
<th>(day)</th>
<th>(R$_J$)</th>
<th>Reference</th>
<th>(R$_J$)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–10</td>
<td>0.8–2</td>
<td>Masuda &amp; Winn (2017)</td>
<td>0.43$^{+0.07}_{-0.08}$%</td>
<td>&lt; 2.1%</td>
</tr>
<tr>
<td>1–10</td>
<td>0.71–2.14</td>
<td>Petigura et al. (2018)</td>
<td>0.57$^{+0.14}_{-0.13}$%</td>
<td>&lt; 2.2%</td>
</tr>
</tbody>
</table>

**TESS Study**

<table>
<thead>
<tr>
<th>(day)</th>
<th>(R$_J$)</th>
<th>Reference</th>
<th>(R$_J$)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9–10</td>
<td>0.8–2.5</td>
<td>Zhou et al. (2019)</td>
<td>0.41 ± 0.10%</td>
<td>&lt; 2.1%</td>
</tr>
</tbody>
</table>

**RV Studies**

<table>
<thead>
<tr>
<th>(day)</th>
<th>(R$_J$)</th>
<th>Reference</th>
<th>(R$_J$)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–11</td>
<td>0.72–2</td>
<td>Mayor et al. (2011)</td>
<td>0.89 ± 0.36%</td>
<td>&lt; 2.4%</td>
</tr>
<tr>
<td>0–10</td>
<td>0.55–2</td>
<td>Wright et al. (2012)</td>
<td>1.20 ± 0.38%</td>
<td>&lt; 2.5%</td>
</tr>
</tbody>
</table>

---

*a* The period range used in the comparison work for the occurrence rate calculation. Note that the smallest period used in this work is 1 day and so our calculation truncates smaller period ranges at 1 day.

*b* The planet radius range used in the comparison work for the occurrence rate calculation. Several references used R$_J$ as their unit of radius and these values have been converted to R$_J$ and rounded. Note that the largest radius examined in this work is 2 R$_J$ and so our calculation truncates larger radius ranges at 2 R$_J$.

*c* The comparison work’s planet occurrence rate as published.

*d* Our calculated occurrence rate upper limit for the same period and radius range.

*e* Our calculated occurrence rate assuming W2282 is a planet. This value is not included if W2282 does not fall in the given period and radius ranges.

*f* Our calculation truncates at 36 days.

*g* For this value, Masuda & Winn (2017) restricted their analysis to the *Kepler* stars that were in the same range of masses as the stars search in 47 Tuc for planets by Gilliland et al. (2000).

*h* These were limits in mass rather than radius. We converted the lower mass limit to a radius using the empirical relation derived by Chen & Kipping (2017) and imposed our default upper limit of 2 R$_J$. 

---

170
Table 5.4. Comparison with Globular Cluster Occurrence Rates

<table>
<thead>
<tr>
<th>Per. Range (day)</th>
<th>Rad. Range (R\textsubscript{J})</th>
<th>Reference</th>
<th>Published Rate(^c)</th>
<th>Our Upper Limit(^d)</th>
<th>Our Rate(^e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–8</td>
<td>0.64–2</td>
<td>Gilliland et al. (2000)(^f)</td>
<td>(\lesssim 0.7%)</td>
<td>(&lt; 2.1%)</td>
<td>0.35–3.1%</td>
</tr>
<tr>
<td>1–16</td>
<td>1–2</td>
<td>Weldrake et al. (2005)</td>
<td>(\ldots)</td>
<td>(&lt; 2.7%)</td>
<td>0.47–4.1%</td>
</tr>
<tr>
<td>1–5(^g)</td>
<td>1.5–2(^g)</td>
<td>Weldrake et al. (2008)</td>
<td>(&lt; 1.7%)</td>
<td>(&lt; 0.81%)(^h)</td>
<td>(\ldots)</td>
</tr>
<tr>
<td>1–14(^i)</td>
<td>0.94–1.37(^i)</td>
<td>Nascimbeni et al. (2012)</td>
<td>(&lt; 9.1%)</td>
<td>(&lt; 0.93%)(^h)</td>
<td>0.31–1.7%</td>
</tr>
</tbody>
</table>

\(^a\)The period range used in the comparison work for the occurrence rate calculation.

\(^b\)The planet radius range used in the comparison work for the occurrence rate calculation. The radius ranges used were not always clear in the comparison works, so we made our best guess, taking into account our radius–period grid boundaries. Note that the largest radius examined in this work is 2 R\textsubscript{J} and so our calculation truncates larger radius ranges at 2 R\textsubscript{J}.

\(^c\)The comparison work’s planet occurrence rate upper limit as published.

\(^d\)Our calculated occurrence rate upper limit for the same period and radius range.

\(^e\)Our calculated occurrence rate assuming W2282 is a planet. This value is not included if W2282 does not fall in the given period and radius ranges.

\(^f\)Provided courtesy K. Masuda (private communication) based on the work in Masuda & Winn (2017). The calculated 3\(\sigma\) occurrence rate upper limit is based on the same stellar mass range described in Table 5.3, note g.

\(^g\)The published rate is for their 1–5 day calculation; we performed our calculation over 1–8 days to guarantee at least one of our radius–period bins be included. Similarly, the quoted rate is only for \(>1.5\ R\textsubscript{J}\) objects, but we had to use 1.37–2 R\textsubscript{J} objects to cover a whole bin.

\(^h\)95% confidence instead of our typical 3\(\sigma\), to match the confidence level used by both Weldrake et al. (2008) and Nascimbeni et al. (2012).

\(^i\)The upper end of the period range was arrived at dividing their total observation duration (28 days) in half; planet radius range chosen to span their single injected planet radius, 1 R\textsubscript{J}.
pares our occurrence rate limits with those of previous GC planet searches: Gilliland et al. (2000) and Weldrake et al. (2005) for 47 Tuc, Weldrake et al. (2008) for ω Cen, and Nascimbeni et al. (2012) for NGC 6397.

As seen in Table 5.3, in no case were we able to set an upper limit that shows an occurrence rate smaller than what we would expect from the field population. And even if W2282 or a comparable planet candidate in the HJ regime is shown to actually be a planet, most of the Kepler- or TESS-based occurrence rates would be consistent with the lower end of our calculated occurrence rate ranges, though the upper ends of our ranges are inconsistent in all those cases. Thus based on these previous studies, and ignoring the metallicity dependence of HJ occurrence as seen in the field, we would expect to have a non-vanishing probability of finding a planet. Comparing with the RV studies, the RV rates fall within our W2282-based occurrence rate ranges, also suggesting from these results that there is some meaningful, if small, probability of finding a planet.

Comparing with the previous searches in GCs, and focusing first on the 47 Tuc surveys as those were the most constraining, Masuda & Winn (2017) showed that Gilliland et al. (2000) should have found $2.2^{+1.6}_{-1.1}$ planets in their survey, and that Weldrake et al. (2005) should also have found $\sim 2$ planets in their survey. Thus their results (possibly) show a lower occurrence rate in 47 Tuc than found by Kepler for the field population for the period ranges searched. Our results do not reach such a constraining level for HJs. However, our sensitivity reaches further in planet radius and period than either of those two previous surveys. Our $\sim 78$ day baseline and the nearly continuous nature of the observations would virtually guarantee us three visible transits for orbital periods up to $\sim 26$ days and for some cases out to $\sim 39$ days. This is compared to the 8.3 day baseline of Gilliland et al. (2000) and the 33 day baseline of Weldrake et al. (2005). Additionally, Gilliland et al. (2000) were insensitive to nearly all planets with a radius below 0.8 R$_J$ and had at best 40% recovery of 1 R$_J$ objects.
for the optimal stellar magnitude. Our work was still reasonably sensitive down to $\sim 0.6 R_J$ for small periods and somewhat sensitive down to $\sim 0.4 R_J$.

In Table 5.4, we do have a more constraining upper limit than the work in $\omega$ Cen by Weldrake et al. (2008) for their very limited period and radius range. We also improve on the limit determined by Nascimbeni et al. (2012). They were able to put a (95% confidence) upper limit of 9.1% on the occurrence of $\sim 1 R_J$ objects with periods between 1 and $\sim 14$ days, while we are able to put an upper limit of 0.93% for the same period range and a similar planet radius range at the same confidence level. Weldrake et al. (2005) did not provide any quantification of their sensitivity to planet radius, but their calculations assumed a relatively large radius of 1.3 $R_J$, and we assume they were sensitive to planets of that radius and larger.

The primary contribution of this work is the new parameter range it explores for planet occurrence rates in GCs, in both planet radius and orbital period. The occurrence rate limits in these new parameter ranges ($0.3 \lesssim R_p \lesssim 0.8 R_J$ at short periods and $P \lesssim 36$ days for large-radius planets) are shown in Table 5.3 in comparison with the field occurrence rates. In particular, we set occurrence rate upper limits of 16% for $0.36$–$0.71 R_J$ planets with 1–10 day periods and 6.1% for $0.71$–$2 R_J$ planets with 1–36 day periods. While these numbers may not seem impressive when compared to the equivalent occurrence rates determined from Kepler (0.5 ± 0.1% for $0.36$–$0.71 R_J$ planets with 0.8–10 day periods and 1.3 ± 0.2% for $0.71$–$2 R_J$ planets with 1–50 day periods; Howard et al. 2012), these are the first limits set for planets in a GC in these period and radius regimes. These limits demonstrate that the occurrence of planets in M4 just outside the HJ regime (in terms of period or radius) is not ubiquitous, and, for the $0.71$–$2 R_J$, 1–36 day range, is at most a factor of about five higher than has been found for the field population.
5.6 Discussion

This work represents the first look at a planet occurrence rate for the GC M4, and the fifth photometry-based examination of a planet occurrence rate for a GC. Its sensitivity reaches further in orbital period and planet radius than any of the previous works (Gilliland et al. 2000, Weldrake et al. 2005, Weldrake et al. 2008, Nascimbeni et al. 2012) by approximately a factor of two in each relative to the previous best. The occurrence limits we were able to calculate for HJs are not more constraining than the best constraints already determined for GCs by the previous works, and our sensitivities in the extra ranges in orbital period and planet radius that this work reaches do not permit constraints tighter than those previously determined for field stars in those regimes. However, we do note from Table 5.3 that the occurrence rate of HJs in M4 is no larger than \( \sim 5 \) times what have been determined by Kepler and TESS at 3\( \sigma \) confidence, and no larger than \( \sim 2-3 \) times what have been determined by the RV surveys. For the previous GC studies, in Table 5.4 we are able to obtain a tighter constraint than Weldrake et al. (2008) for the largest radius planets and a better limit than Nascimbeni et al. (2012) for \( \sim 1 R_J \) planets. Kepler was not designed or optimized for looking at GCs—in particular, the \( \sim 4''/\text{pixel} \) image resolution led to significant blending in the images—and the superstamp observations of M4 were originally intended for observing RR Lyrae variables. Obtaining even the level of constraints we did from a telescope and observations not originally intended for a GC planet search is yet another demonstration of an unanticipated scientific result from Kepler.

Though our constraints cannot rule out planet occurrence rates for M4 matching those of the field stars, given the current uncertainties on planet formation—particularly the formation of close-in giant planets—and uncertainties on GC formation, obtaining any constraints on planet occurrence in GCs for new regimes of planet
radius and period is useful. It may be that the occurrence of certain kinds of close-in planets in GCs is more common due to some unique aspect of GCs.

There are some reasons we might expect the occurrence of close-in planets to be higher in a GC than in the field. For example, Hamers & Tremaine (2017) demonstrated that the increased number of close stellar encounters experienced by GC stars over their lifetimes could enhance the HJ occurrence rate for certain stellar densities (peak formation occurred at a density of $\sim 4 \times 10^4$ pc$^{-3}$) if HJs are formed through high-eccentricity migration. The formation of GCs themselves is still something of a mystery (see Gratton et al. 2012 for a review), and perhaps there is something unique about the formation of stars in GCs that would increase the formation of close-in planets.

The results of Gilliland et al. (2000) and Weldrake et al. (2005), even with the reinterpretation of Masuda & Winn (2017), indicate that at most the occurrence rate of HJs in GCs is not likely to be greater than that seen by Kepler or RV studies, but are unable to say much about the occurrence of planets with $R_p \lesssim 0.9$ $R_J$ or with periods $\gtrsim 16$ days. Our constraints show that short-period planets with $0.4$ $R_J \lesssim R_p \lesssim 0.9$ $R_J$ and Jupiter-sized planets with periods between 16 to $\sim 30$ days are not ubiquitous in GCs. Our results, with those of the previous GC planet occurrence works, provide constraints on just how enhanced a planet occurrence rate might be should there indeed be enhanced close-in planet formation in GCs.

On the other hand, much work has been done to show why specifically HJ occurrence in GCs might be suppressed, motivated in part by the original report of Gilliland et al. (2000) that HJs are much less common in GCs than in the field (though, as previously mentioned, Masuda & Winn 2017 showed that their non-detection is not as robust as originally thought). The occurrence of HJs is known to correlate with host star metallicity (e.g., Fischer & Valenti 2005) and this has been used to argue that the low metallicities of GCs would inhibit HJ formation (for example, Santos
et al. 2001 showed the known metallicity dependence as being able to explain the 47 Tuc planet non-detection of Gilliland et al.), but the reason behind a metallicity–occurrence connection is not well understood and it may be that the underlying cause of this connection does not apply in the unique environments of GCs. Additionally, the dense stellar environment of GCs and the associated levels of radiation from particularly the nearby massive stars may inhibit giant-planet formation (Armitage, 2000; Adams et al., 2004; Thompson, 2013). Also, in addition to enhancing the rate of close-in planets, dynamical interactions with passing stars can also remove planets from planetary systems (Sigurdsson, 1992; Davies & Sigurdsson, 2001; Bonnell et al., 2001; Fregeau et al., 2006; Spurzem et al., 2009), particularly planets on wide orbits. Interactions between stars and protoplanetary disks lead to decreases in disk sizes as well (Breslau et al., 2014). Until better constraints or actual occurrence rates are determined for GCs, for a larger range of planet radii and orbital periods than are presently accessible from existing data, it will be difficult to determine the precise impact a GC environment has on planet formation and occurrence.

As an analysis of how an improvement on our light curves and/or noise characterization and removal could improve our occurrence rate limits, we show a forecast in Figure 5.8 of the limits that would be set if an S/PN threshold of eight could be imposed instead of 12 and assuming no planets were found. W2282 and the associated occurrence rate is still included for comparison with Figure 5.7. Our upper limits in the HJ regime would not improve by very much, but we would be able to place more stringent constraints for $0.64 \, R_J \lesssim R_p \lesssim 0.94 \, R_J$ across all the periods examined, and for $0.3 \, R_J \lesssim R_p \lesssim 0.64 \, R_J$ for the shortest periods examined here. Even if our limits in the HJ regime do not improve much, a better understanding of the noise would allow for an improved vetting of the current planet candidates.

As limited as our constraints are, they may be the best to come for a while. The only near-term continuous photometric survey is TESS, but with its $\sim 20''$/pixel
Figure 5.8: Same as Figure 5.7, but instead showing a forecast of the limits that could be set if a S/PN threshold of eight was able to be used instead of 12 and no additional planet candidates were found.
image resolution it will leave most of the stars in GCs hopelessly blended. Moreover, the \( \sim 1 \) month observation span most of its survey field will be covered by is only about a third the span of what is available in this work with K2. An \( HST \) campaign similar to that of Gilliland et al. (2000) for M4—being about half the distance as 47 Tuc—should permit a factor of two increase in the signal-to-noise ratio for stars of comparable masses and evolutionary state as in 47 Tuc; a campaign along these lines might be considered. The main limiting factor in setting the HJ occurrence limit from the K2 data is the relatively small number of cluster stars observed, \( \sim 4000 \) compared to \( \sim 34,000 \) GC stars in Gilliland et al. 2000 and \( \sim 20,000 \) GC stars in Weldrake et al. 2005). The K2 superstamp covered a relatively small fraction of the stars in the cluster, so a survey that covers more of the cluster could be useful.

Despite the low S/PN of our planet candidates, given the scientific impact of discovering and characterizing a transiting planet in a GC, we argue that it is worth the effort to confirm whether these are real planet signals. Unfortunately, the data are already five years old, and the uncertainties on the present transit epochs are large. While we do not formally calculate the period uncertainties, our estimates show that current uncertainties of transit times could be anywhere from a few hours to about a day. In this, though, the crowded nature of the cluster and the manageably sized field of view of the K2 superstamp \( (\sim 10' \text{ by } \sim 20') \) are advantages. Many available wide field imagers can cover a large fraction of or even the entire superstamp, allowing for simultaneous observation of more than one planet candidate, and the data would also have the advantage of observing other interesting variables guaranteed to be present (see Chapter 3 for a catalog).
5.7 Conclusion

We searched for transiting planets in the GC M4 using data from the K2 mission. These data represent the longest continuous observation of a GC, permitting a search for the longest-period planets ever searched in a GC. The data are also of sufficient quality to be sensitive to planets of smaller radii than any previous transit search in a GC. From 3784 light curves extracted from the data, with a maximum observation duration of \(\sim 78\) days, and using a BLS transit search followed by a by-eye vetting, we identified 27 planet candidates in the data. An analysis of the systematic noise in the light curves revealed that a S/PN cutoff value of 12 should be used to remove probable false positives, with only one of the planet candidates exceeding this cutoff value, yet there still remains uncertainty as to whether this might be a false alarm. Despite this, information on this and six other of our most promising candidates are presented. The light curves are publicly available at Wallace et al. (2019a).

We calculated \(3\sigma\) occurrence rate upper limits based on a non-detection of planets and occurrence rate ranges assuming our S/PN \(> 12\) planet candidate as real, for a variety of period and planet radius ranges. Comparing these limits and rates to the literature, for previous GC works, we find a factor of two lower occurrence rate limit than was calculated by Weldrake et al. (2008) for \(\omega\) Cen for \(R_p > 1.5\) \(R_J\) objects. We also improve on the Nascimbeni et al. (2012) limit for \(\sim 1\) \(R_J\) planets with \(\lesssim 14\) day orbits, obtaining a \(2\sigma\) limit of \(< 0.93\%\). Our limit for a similar period and radius range as the landmark study of Gilliland et al. (2000) was sensitive to, 1–8 days and 0.64–2 \(R_J\), was \(< 2.1\%\), compared to the \(< 0.7\%\) limit determined by Masuda & Winn (2017) using the Gilliland et al. (2000) data. Comparing with occurrence rates calculated from field star transit surveys, our HJ occurrence rate limits are factors of about four to six larger than the \textit{Kepler} and \textit{TESS} rates. Similarly, for RV studies, our HJ occurrence limits are about a factor of two higher than the rate of Wright et al. (2012) and about a factor of three higher than the rate of Mayor et al. (2011).
Our rate upper limits for longer period orbits ($\gtrsim 15$ days) of $\sim 1 \, \text{R}_J$ objects and for smaller planets ($\sim 0.4 \, \text{R}_J$ and larger) are much larger than the rates known for those regimes from Kepler and are not very constraining, but are the first such limits ever set for a GC.

Future work that could be done to build off these results includes photometric follow up of the planet candidates to confirm the transits and improving the systematic noise characterization and abatement in the light curves to permit greater sensitivity to lower S/PN transits. Lowering the S/PN threshold would allow us to put significantly better constraints on $P \lesssim 8$ day planets with $0.3 \, \text{R}_J \lesssim R_p \lesssim 0.6 \, \text{R}_J$ planets and for $0.6 \, \text{R}_J \lesssim R_p \lesssim 0.9 \, \text{R}_J$ planets across all periods examined.

The research in this chapter utilized the following software: *astrobasis* (Bhatti et al., 2017), *astropy* (Astropy Collaboration et al., 2018), *FITSH* (Pál, 2012), *k2mosaic* (Barentsen, 2016), *matplotlib* (Hunter, 2007), *numpy* (Oliphant, 2006), *scikit-learn* (Pedregosa et al., 2011), *scipy* (Jones et al., 2001), and *VARTOOLS* (Hartman & Bakos, 2016).
Appendix A

Cluster Membership from Gaia

Proper Motions

The relatively nearby GC M4 (NGC 6121) has a proper motion that is well separated from the interloping field stars, permitting proper motion to be a useful determiner of cluster membership. The first M4 proper motion membership catalog was that of Cudworth & Rees (1990) (updated in Peterson et al., 1995), a probabilistic catalog from ~90 years of photographic observations for 530 stars down to $V \approx 16$ and, for the faintest stars, within ~$6'5$ of the cluster center. Zloczewski et al. (2012) (hereafter Z12) derived an M4 proper motion catalog with discrete classifications for 13,036 stars in a ~$10' \times 10'$ square from ~10 years of ground-based observations, with ~25% of detected $V=20$ objects having successfully measured proper motions. Watkins & van der Marel (2017) used the relatively shallow Tycho-Gaia Astrometric Solution catalog and its ~30-year baseline to identify five of M4’s brightest cluster members.

This appendix presents a new probabilistic proper motion membership catalog for M4 calculated using a Gaussian mixture model (GMM) fit to Gaia DR2 proper motions (Brown et al., 2018; Lindegren et al., 2018a), using only DR2’s 22-month baseline. The sources within 30' of the cluster’s center (RA: 16h 23m 35.22s, decl-
nation: -26° 31’ 32” 7; Goldsbury et al., 2010) and brighter than \( G = 19 \) were used. For reference, M4’s main sequence turnoff is at \( G \approx 16 \). Out of 30,540 objects, 14,642 were identified as having membership probabilities >99%. This is the widest proper motion catalog of M4 ever assembled, though the data exists to extend it to arbitrary size. I chose 30’ as a cutoff since that covers most of the cluster without overwhelming the fit with field stars, as well as that being more than sufficient to cover the entire K2 M4 superstamp that is the focus of this dissertation. Though \textit{Gaia} DR2 proper motion measurements reach roughly at least as deep as Z12 (and fainter than the \( G = 19 \) magnitude cutoff) in the outskirts of the cluster, near the cluster center the depth of \textit{Gaia} DR2 suffers due to crowding (quantified in Arenou et al., 2018). I did not evaluate the completeness of this catalog relative to that of Z12.

The \texttt{scikit-learn} (Pedregosa et al., 2011) GMM implementation was used, fitting two components to the proper motion distribution. The only preprocessing step was to exclude objects with no reported proper motions, which were 5.0% of the \( G < 19 \) sources. Though many \( G \approx 20 \) objects had measured proper motions (but 0% of \( G = 21 \) objects), the GMM fit was carried out only for objects with \( G < 19 \). The larger proper motion error bars and the increasing number of field stars at fainter magnitudes affected the overall fit and made a magnitude cut prudent. A more detailed catalog could bin objects in magnitude and perform a separate fit for each bin. The code, fitted models, and catalog are available on GitHub\footnote{https://github.com/joshuawallace/M4_pm_membership} or Zenodo (Wallace, 2018a). Figure A.1 shows the proper motions and membership probabilities for the \( G < 19 \) objects. The individual proper motion measurement errors are small: \( \sim 1 \) mas/yr for \( G = 19 \) objects, and \( \sim 0.1 \) mas/yr for \( G = 15 \) objects.

GMMs incorporate prior probabilities based on the relative numbers of members of each component. If sources within 60’ of the cluster center were used instead of 30’, a larger fraction of the sources would be field stars and the final fit would be
Figure A.1: The Gaia DR2 proper motion data and calculated membership probabilities for objects with $G<19$. Proper motions are shown in the large panel, with membership probability encoded by color as shown in the color bar in the upper right. Proper motions of M4 members form the cluster in the lower left. The top and right panels show distributions of the points in the large panel in, respectively, $\mu_\alpha$ and $\mu_\delta$. The units for the proper motions are milliarcseconds per year, and the distribution histograms show the number of objects per proper motion bin.
affected. I found the fits to be very similar whether I included sources within 6′, 30′, or 60′ of the cluster center, though including sources within 60′ actually decreased the number of sources with membership probabilities >99% (13,132 instead of the 14,642 for 30′). Since a greater fraction of stars in the wider sample were field stars, the GMM fit took this into account and decreased the membership probabilities overall. A more detailed catalog could bin objects by sky position in annuli centered on the cluster center and perform a separate fit for each bin. This catalog does not take into account radial velocity measurements (which are available from Gaia DR2 for sources with $G \lesssim 12$), parallaxes (which are relatively well measured for the brightest stars but rather poorly measured for the fainter stars), or positions of the sources in color–magnitude space.

The research in this appendix utilized the following software: astropy (Astropy Collaboration et al., 2018), matplotlib (Hunter, 2007), numpy (Oliphant, 2006), and scikit-learn (Pedregosa et al., 2011).
Bibliography


Asphaug, E. 2010, Chemie der Erde / Geochemistry, 70, 199


Cannon, A. J. 1907, Annals of Harvard College Observatory, 55, 1

Canup, R. M., & Esposito, L. W. 1995, Icarus, 113, 331

Carroll, B. W., & Ostlie, D. A. 2007, An Introduction to Modern Astrophysics (Addison-Wesley)


—. 2001, Icarus, 152, 205
Davis, H. 1917, PASP, 29, 260
Hartman, J. D., & Bakos, G. Á. 2016, Astronomy and Computing, 17, 1
Hayashi, C. 1981, Progress of Theoretical Physics Supplement, 70, 35
Hénon, M., & Petit, J.-M. 1986, Celestial Mechanics, 38, 67
Herschel, W. 1789, Philosophical Transactions of the Royal Society of London Series I, 79, 212
Holsapple, K. A., & Michel, P. 2006, Icarus, 183, 331
—. 2008, Icarus, 193, 283
Hunter, J. D. 2007, Computing In Science & Engineering, 9, 90
Kaluzny, J., Thompson, I. B., Rozyczka, M., & Krzeminski, W. 2013a, AcA, 63, 181
Karjalainen, R. 2007, Icarus, 189, 523
Kobayashi, H., & Tanaka, H. 2010, Icarus, 206, 735
Leavitt, H. S., & Pickering, E. C. 1904, Harvard College Observatory Circular, 90, 1
Messier, C. 1771, Mémoires de Mathématiques & de Physiques Tirés des registres de l’Académie Royale des Sciences pour l’année, 435

190


Oliphant, T. 2006, Guide to NumPy (Trelgol Publishing)


Pan, M., & Sari, R. 2005, Icarus, 173, 342


Russell, H. N. 1948, HarMo, 7, 181


Samus, N. N., Kazarovets, E. V., Durlevich, O. V., Kireeva, N. N., & Pastukhova, E. N. 2017, Astronomy Reports, 61, 80

191
Shapley, H. 1918, PASP, 30, 42
Sigurdsson, S., Richer, H. B., Hansen, B. M., Stairs, I. H., & Thorsett, S. E. 2003, Science, 301, 193
Smolec, R., & Moskalik, P. 2008, AcA, 58, 193
Wallace, J., Hartman, J., Bakos, G., & Bhatti, W. 2018, Light curves for M4 milli-
magnitude RR Lyrae (mmRR), doi:10.5281/zenodo.2220532


Wallace, J. J. 2018b, Research Notes of the AAS, 2, 213


Watson, C., Henden, A. A., & Price, A. 2017, VizieR Online Data Catalog, B/vsx


—. 2007, Acta Astronomica Sinica, 48, 18

Yao, B. A., & Tong, J. H. 1989, Information Bulletin on Variable Stars, 3334, 1


—. 1981b, ChA&A, 5, 476


193
Zloczewski, K., Kaluzny, J., Rozyczka, M., Krzeminski, W., & Mazur, B. 2012, AcA, 62, 357