12-2022

# Investigating the O'Connell Effect in Kepler Eclipsing Binaries

Matthew Frederick Knote

# Investigating the O'Connell Effect in *Kepler* Eclipsing Binaries

by

Matthew Frederick Knote

Master of Science in Physics
Department of Physics and Astronomy
Ball State University
2015

Bachelor of Science in Physics
Department of Physics and Astronomy
Ball State University
2012

A dissertation
submitted to the College of Engineering and Science
at Florida Institute of Technology
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in
Space Sciences

Melbourne, Florida
December, 2022

---

We the undersigned committee
hereby approve the attached dissertation

Investigating the O'Connell Effect in *Kepler* Eclipsing Binaries by Matthew Frederick Knote

| | |
|---|---|
| Eric S. Perlman, Ph.D.<br>Professor<br>Aerospace, Physics, and Space Sciences<br>Major Advisor | Saida M. Caballero-Nieves, Ph.D.<br>Graduate Faculty<br>Aerospace, Physics, and Space Sciences |
| Jean Carlos Perez, Ph.D.<br>Associate Professor<br>Aerospace, Physics, and Space Sciences | Isaac Silver, Ph.D.<br>Graduate Faculty<br>College of Aeronautics |
| Vayujeet Gokhale, Ph.D.<br>Associate Professor<br>Physics<br>Truman State University | David C. Fleming, Ph.D.<br>Associate Professor and Department Head<br>Aerospace, Physics, and Space Sciences |

ABSTRACT

Title:

Investigating the O'Connell Effect in *Kepler* Eclipsing Binaries

Author:

Matthew Frederick Knote

Major Advisor:

Eric S. Perlman, Ph.D.

Advisor:

Saida M. Caballero-Nieves, Ph.D.

The O'Connell effect – the presence of unequal maxima in eclipsing binaries – remains an unsolved riddle in the study of close binary systems. The *Kepler* space telescope produced high precision photometry of nearly 3,000 eclipsing binary systems, providing a unique opportunity to study the O'Connell effect in a large sample and in greater detail than in previous studies. I have characterized the observational properties – including temperature, luminosity, and eclipse depth – of a set of 212 systems (7.3% of *Kepler* eclipsing binaries) that display a maxima flux difference of at least 1%, representing the largest sample of O'Connell effect systems yet studied. I explored how these characteristics correlate with each other to help understand the O'Connell effect's underlying causes. In studying these systems, I found that ~30% of my sample belonged to four system classes with peculiar light curve features aside from the O'Connell effect: systems with temporal variation, systems with asymmetric minima, systems with a concave-up region, and a white dwarf. I studied the characteristics and correlations of the first three of these system classes to better understand how they differed from other O'Connell effect systems. Finally, I observed ten systems in my sample as a follow-up to *Kepler*'s

observations. I found that the O'Connell effect size's correlations with period and temperature are inconsistent with Kouzuma's (2019) starspot study. Up to 20% of systems display the parabolic eclipse timing variation signal expected for binaries undergoing mass transfer. Most systems displaying the O'Connell effect have the brighter maximum following the primary eclipse, suggesting a fundamental link between which of the maxima is brighter and the O'Connell effect's physical causes. The systems displaying an asymmetric minimum are split into two fundamentally different subsets: temporally stable systems and temporally varying systems. Chromospheric activity largely explains the features observed in the peculiar system classes, aside from the temporally stable asymmetric minima systems. Most importantly, I found that the O'Connell effect occurs exclusively in systems where the components are close enough to significantly affect each other, suggesting that the interaction between the components may ultimately be responsible for causing the O'Connell effect.

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| AP | Asymmetry Parameter |
| DFT | Discrete Fourier Transform |
| ETV | Eclipse Timing Variation |
| FFT | Fast Fourier Transform |
| KEBC | *Kepler* Eclipsing Binary Catalog |
| KPNO | Kitt Peak National Observatory |
| LCA | Light Curve Asymmetry |
| LCP | Left Concavity Parameter |
| ODR | Orthogonal Distance Regression |
| OER | O'Connell Effect Ratio |
| OES | O'Connell Effect Size |
| ORM | Observatorio del Roque de los Muchachos |
| PHOEBE | PHysics Of Eclipsing BinariEs |
| RCP | Right Concavity Parameter |
| SARA | Southeastern Association for Research in Astronomy |
| SVP | Scaled Variance Parameter |
| UVP | Unscaled Variance Parameter |

# Acknowledgements

I would like to think my two advisors, Dr. Saida Caballero-Nieves and Dr. Eric Perlman. Over the last several years, they have shaped me into the scientist I am today. I greatly appreciate their enduring patience during the long and often troubled course of my graduate career. I would also like to thank my first advisor here at FIT, Dr. Véronique Petit, for helping set me on the course my studies took. I would like to thank my other committee members, Dr. Jean Carlos Perez, Dr. Issac Silver, and Dr. Vayujeet Gokhale, for their assistance through my dissertation preparation.

I would like to thank many friends I have made here at FIT, who have helped to keep me sane and social in an insane and asocial world. To Craig Kolobow, Keefe Kamp, Ashley Hughes, Ramana Sankar, Athela Frandsen, Lauren Persons, Eric Burns, Katie Davis, Sydney Tobolski, and many others over the years, thank you so much for being there! I also want to thank FIT's Society for Science Fiction and Fantasy (FITSSFF), the resident club of all things nerdy that I had the privilege to be vice-president of for three years. FITSSFF's monthly Board Game Day was something bright to look forward to after entirely too much grading. I want to thank my brother in all but blood, Caleb Mosier, and my best friend, Monique Gabb, for their enormous (albeit long-distance) support throughout my time here.

I would like to thank the support staff in the Department of Aerospace, Physics, and Space Sciences, particularly our previous administrative assistants, Leslie Smith and Kathy Alger, and our current one, Paula Couch. They all helped extensively with the copious paperwork required during my seven years. I would also like to thank our laboratory director, James Gering, who I would turn to whenever my students became unruly.

Finally, I would like to thank my parents Jan and Steve Knote for supporting my curiosity throughout my life, as well as for their unending support throughout my years at FIT. I would also like to thank my wonderful fiancée, Sheri Sagall, for her continued support.

# Dedication

I would like to dedicate this dissertation to my fiancée, Sheri, my parents, Jan and Steve, and my late dog, Jean-Luc.

Sheri and I met during the latter part of my studies here, a time that is not known to be conducive to a stable relationship. Further complicating matters was the COVID-19 pandemic that began only three months into our relationship. Still Sheri has been unyielding in her support of my ambitions, and the pandemic only pushed us closer together. As I close the chapter on my formal education, I look forward to what lies ahead in the chapter we have already begun to write together. Thank you, Sheri, my love, for being beside me every step of the way.

As an adopted only child, I have often felt like my parent's Chosen One. My parents have always supported my boundless curiosity. They have always pushed me to follow my dreams, and my dreams have led me to where I am today. They have loved me and supported me in more ways than I will ever know over the last thirty-four years, and I will be forever grateful to them for that. Mom, Dad, I love you. Thank you for choosing me to be your son.

Finally, I lost my beloved Bichon Frisé Jean-Luc just over five years ago. Named for the Star Trek captain he happened to share a birthday with, Jean-Luc taught this lonely autistic child how to love during one of the darkest chapters of his life.

Even in his old age, when he could no longer see and I had moved away two years prior, he still recognized me every time I saw him. Although he only lived 14 years, he shaped my life as much as any human. There will always be an empty space in my heart for him. Thank you, Jean-Luc.

# Chapter 1

# Introduction

About a third of all stars exist as part of a binary or multiple star system (Lada 2006). Binary systems represent the only method to measure the masses of stars directly. They therefore serve a critical role in determining the relationship between mass and other fundamental physical properties of stars. Accurate measurements of stellar masses are critical to constrain theoretical stellar atmosphere and evolution models.

Stars in close binary systems (which de Boer & Seggewiss 2008 defines as a system wherein at least one of the stars will fill its Roche lobe at some point in its life) are expected to evolve differently from isolated stars or stars in more widely separated binaries. Nearly a quarter of solar-type stars are part of a close binary, with that fraction increasing with decreasing metallicity (Moe et al. 2019; Bate 2019). Mass transfer between two stars can occur when one of the stars fills its Roche lobe, which will cause the two stars to evolve dramatically differently than they otherwise would. In the most extreme cases, these systems are thought to be the progenitors of type Ia supernovae and the source of the recently discovered

gravitational wave events. Therefore, understanding the interaction between stars in close binary systems has applications in several branches of astronomy, including stellar evolution, gravitational wave astronomy, and cosmology.

## 1.1 Eclipsing Binaries

Eclipsing binaries are binary systems wherein the stars periodically occlude each other from our vantage point as they orbit, causing variation in the observed brightness over time. Eclipsing binaries are vital because spectroscopic and photometric data can be used to determine a wide range of parameters, including stellar mass and temperature. Therefore, much of our knowledge of stellar evolution and its relation to mass comes from studying eclipsing binaries (Kallrath & Milone 2009). Astronomers have studied individual eclipsing binary systems in great detail over the last 200 years since the 1782 discovery of the first eclipsing binary, Algol ($\beta$ Persei; Goodricke 1783).

There are several, often contradictory schemes used to classify eclipsing binaries. Payne-Gaposchkin & Gaposchkin (1938), Kopal (1955), and Svechnikov et al. (1980, for which Avvakumova et al. 2013 give an English translation) define three of these schemes. Dryomova (2014) gives an excellent overview of these three and more while expressing a strong preference toward the scheme Svechnikov et al. (1980) introduced. Additionally, there is ambiguity regarding the term Algol-type binary. For instance, Papageorgiou et al. (2018) uses Algol-type to refer to light curves with well-defined eclipses and minimal out-of-eclipse variations, while Negu & Tessema (2018) uses Algol-type to refer to systems in a similar evolutionary state to Algol. Budding et al. (2004) and Mennickent & Djurašević (2021) explic-

Table 1.1: Light curve classification of eclipsing binaries.

| Light Curve Class | Light Curve Characteristics | Typical Morphology Class |
|---|---|---|
| Algol | Minimal variation outside of eclipses; sharp, well-defined eclipses | Detached, semi-detached |
| $\beta$ Lyrae | Continuously changing light curve; poorly defined eclipses of unequal depth | Semi-detached |
| W Ursae Majoris | Continuously changing light curve; poorly defined eclipses of equal depth | Overcontact |

itly mention the Algol-type ambiguity. To avoid these ambiguities, I will define the two classification schemes I adopt in this dissertation: the light curve classification and the morphological classification.

I adopted a light curve classification scheme mirroring the one introduced by Payne-Gaposchkin & Gaposchkin (1938). Table 1.1 gives a brief description of each class along with the morphological classes they typically (but not exclusively) correspond to, while Figure 1.1 shows an example of each light curve class from my sample. The scheme consists of three classes: Algol-type (EA in some literature), $\beta$ Lyrae-type (EB), and W Ursae Majoris-type (EW). Algol-type systems have sharp, well-defined eclipses and minimal out-of-eclipse variations. Meanwhile, $\beta$ Lyrae- and W Ursae Majoris-type systems have continuously variable light curves, but the former have minima of significantly unequal depth while the latter have minima of equal or nearly equal depth. My scheme is based solely on the appearance of the light curve and does not consider spectral type. For instance, I would classify the O-type binary V382 Cygni as a W Ursae Majoris-type system, rather than a $\beta$ Lyrae-type system like Landolt (1975).

To understand the morphological classification scheme I adopted, I must first

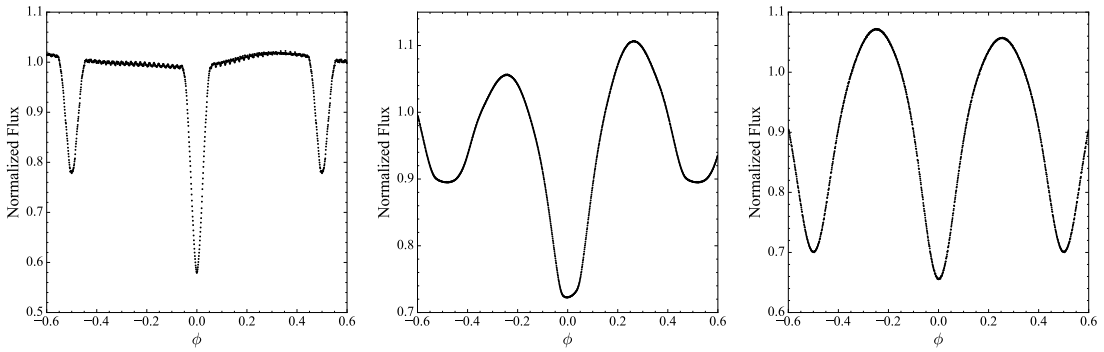


Figure 1.1: Example light curves of an Algol-type system (KIC 8608490, left), a $\beta$ Lyrae-type system (KIC 9164694, center), and a W Ursae Majoris-type system (KIC 7871200, right).

discuss the concept of the Roche lobe. Consider a system of two point masses $M_1$ (the primary component) and $M_2$ (the secondary component) separated by a distance $D$ and orbiting each other. The point masses rotate synchronously, i.e., with a period equal to their orbital period. The origin of the rotating coordinate system is at the location of $M_1$, the x-axis connects the two masses, and the z-axis is perpendicular to the orbital plane. Kopal (1955) gives the total (gravitational and centrifugal) potential $W$ acting on a point $P(x, y, z)$ as:

$$W = G\frac{M_1}{r} + G\frac{M_2}{r'} + \frac{\omega^2}{2}\left[\left(x - \frac{DM_2}{M_1 + M_2}\right)^2 + y^2\right] \tag{1.1}$$

where $G$ is the universal gravitational constant, $r = \sqrt{x^2 + y^2 + z^2}$ and $r' = \sqrt{(D - x)^2 + y^2 + z^2}$ are the distances between $P(x, y, z)$ and the primary and secondary components, respectively, and:

$$\omega^2 = G\frac{M_1 + M_2}{D^3} \tag{1.2}$$

is the Keplerian angular velocity. Kopal (1955) first defined *Roche equipotentials*

as surfaces of constant potential defined by:

$$(1+q)C = \frac{2}{r} + 2q\left(\frac{1}{r'} - x\right) + (1+q)(x^2+y^2) + \frac{q^2}{1+q} \tag{1.3}$$

where $q = \frac{M_2}{M_1}$ is the mass ratio and the dimensionless Roche constant $C$ is:

$$C = \frac{2DW}{G(M_1+M_2)} \tag{1.4}$$

Plavec (1958) introduced a modified potential defined by (Plavec 1958; Kruszewski 1963):

$$\Omega = \frac{(1+q)}{2}C - \frac{q^2}{2(1+q)} \tag{1.5}$$

resulting in the *modified Kopal potential*:

$$\Omega = \frac{1}{r} + q\left(\frac{1}{r'} - x\right) + \frac{(1+q)}{2}(x^2+y^2) \tag{1.6}$$

Stellar surfaces conform to the Roche equipotentials given by Equations 1.3 and 1.6.

The *Lagrange points* (Euler 1767; Lagrange 1772) are the five unique points wherein $\frac{\partial\Omega}{\partial x} = \frac{\partial\Omega}{\partial y} = \frac{\partial\Omega}{\partial z} = 0$, i.e., the five critical points of $\Omega$ (Plavec 1958). These five points are labeled $L_1$ to $L_5$. $L_1$ is the pertinent Lagrange point for this discussion, with the distance $r_{12}$ between $L_1$ and $M_2$ given by solving the quintic equation[1]:

$$\frac{M_1}{(D-r_{12})^2} = \frac{M_2}{r_{12}^2} + \frac{M_1+M_2}{D^3}\left(\frac{DM_1}{M_1+M_2} - r_{12}\right) \tag{1.7}$$

---

[1] `https://datagenetics.com/blog/august32016/index.html`

$L_1$ lies between the two stars. The Roche equipotential passing through $L_1$ is known as the *Roche lobe*. The Roche lobe is the *critical lobe* of a synchronously-rotating star, where the critical lobe defines the maximum extent of a star (Wilson 1979). Therefore, the Roche lobe defines the maximum extent of a synchronously-rotating star before it begins transferring matter to its companion. The Roche equipotential passing through $L_2$ defines the maximum extent of a Roche lobe overflowing binary before it loses matter through $L_2$. Figure 1.2's top panel labels the Roche lobes and the outer critical surface along with $L_1$ and $L_2$.

I adopted a morphological classification mirroring the one Kopal (1955) introduced. His scheme consists of three classes: detached, semi-detached, and contact. Stars in detached systems are smaller than their Roche lobes, while the less massive secondary star exactly fills its Roche lobe in semi-detached systems. Finally, both stars in contact systems exactly fill their Roche lobes and touch at $L_1$. Wilson (2001) presents a convincing argument that contact systems are better described as *overcontact* systems, wherein both stars exceed their Roche lobes and create a common envelope. Wilson (2001) also mentions a fourth morphological class: double contact, wherein both stars exactly fill their critical lobes but are not in physical contact due to one star rotating super-synchronously (which shrinks the critical lobe of that star; Wilson 1979). I will not discuss double contact systems further in this dissertation because my method of estimating a system's morphology (the morphology parameter described in Section 2.2.3) does not distinguish double contact systems from other morphology classes. Additionally, unlike Kopal (1955), I allow either star to fill its Roche lobe in semi-detached systems. I also make no assumptions about either star's mass or evolutionary state. In summation, my classification scheme consists of three classes: detached, semi-detached, and over-

contact. Figure 1.2 shows examples of all four morphological classes created in BinaryMaker3 (BM3; Bradstreet & Steelman 2002) using data from CALEB[2].

Both the light curve and morphological classification schemes have advantages and disadvantages. The former can be applied at a glance of the light curve and provides clues as to the system's physical state. For instance, an overcontact system displaying an Algol-type light curve does not make physical sense because overcontact systems necessarily have continuously changing light curves. However, the light curve classification is subjective (although I present a quantifiable scheme in Section 2.2.2.2). Additionally, only a vague link with the system's physical state exists, and most combinations of light curve and morphological classes are possible (indeed, an overcontact Algol is the only non-physical combination). The morphological classification, by contrast, is intimately tied to the shape and evolutionary state of the stars. Unfortunately, the only definitive method of determining the morphological class (aside from direct imaging of the stellar surfaces) is by light curve modeling (see Section 1.3), a computationally intensive process subject to degenerate solutions.

Finally, there is another ambiguity regarding the definition of the primary and secondary components. Prša (2006) defines the primary component as the star eclipsed at the deeper minimum, or equivalently, as the hotter star. Kallrath & Milone (2009) notes that other definitions exist, however, including designating the more massive star as the primary component. I will follow the definition of Prša (2006) for this dissertation.

---

[2]Catalog and AtLas of Eclipsing Binaries; `http://caleb.eastern.edu/`

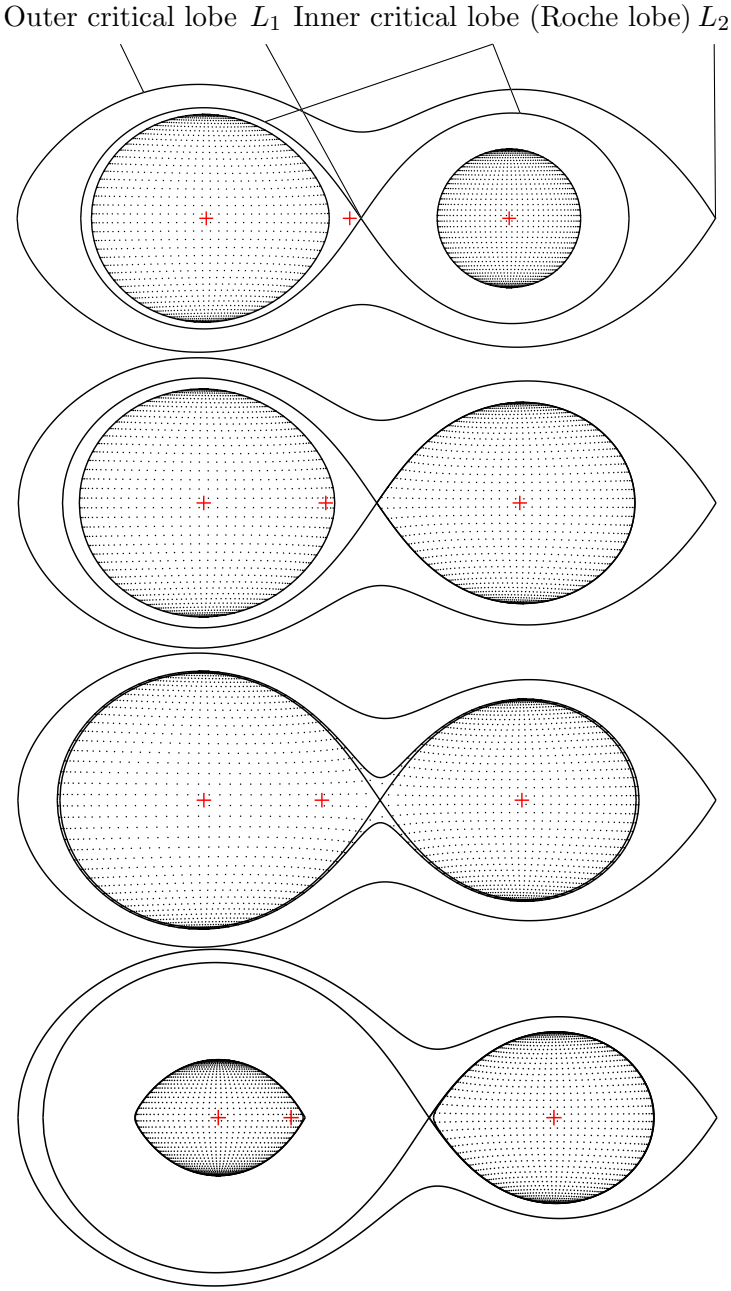


Figure 1.2: Example stellar configurations of a detached system (V499 Scorpii, top), a semi-detached system (AX Draconis, second from top), an overcontact system (S Antliae, second from bottom), and a double contact system (RZ Scuti, bottom). The top panel labels the inner (Roche) and outer critical lobes along with $L_1$ and $L_2$. The outer red markers indicate each star's center, while the central red marker indicates the binary's center of mass. The primary star is on the left-hand side and is moving toward the observer.

## 1.2 The O'Connell Effect

Eclipsing binary light curves are typically shown in phase space, as in Figure 1.1. The phase $\phi$ measures where the components are in their orbit, with $\phi = 0$ (or $\phi(0)$) defined as the deepest eclipse's center and $\phi(0) = \phi(1)$. For circular orbits, $\phi(0.25)$ and $\phi(0.75)$ are the quadrature phases where the components appear side-by-side, as in Figure 1.2 (where $\phi = 0.25$). These phases are (approximately) where the flux is maximum for systems with out-of-eclipse variations. If each star's leading and trailing hemispheres are equally bright, then the observed flux $F(\phi)$ at the two maximal phases $\phi_{\max 1}$ and $\phi_{\max 2}$ (where max 1 follows the primary eclipse and max 2 precedes it) should be identical. However, several systems do not behave this way. The phenomenon wherein $F(\phi_{\max 1}) \neq F(\phi_{\max 2})$ is called *the O'Connell effect* (O'Connell 1951; Milone 1968). Figure 1.3 shows an example of the O'Connell effect from my sample. Roberts (1906) was the first to remark on the O'Connell effect, where he attributed it to stars in an eccentric orbit becoming tidally distorted near periapsis. However, this "periastron effect" does not explain the O'Connell effect in systems with circular orbits. The O'Connell effect remains mysterious over a century later, with Liu & Yang (2003) stating that it is "one of the celebrated difficult problems in the field of close binary systems," and Papageorgiou et al. (2014) stating that "the O'Connell effect is still one of the most perplexing challenges in binary studies."

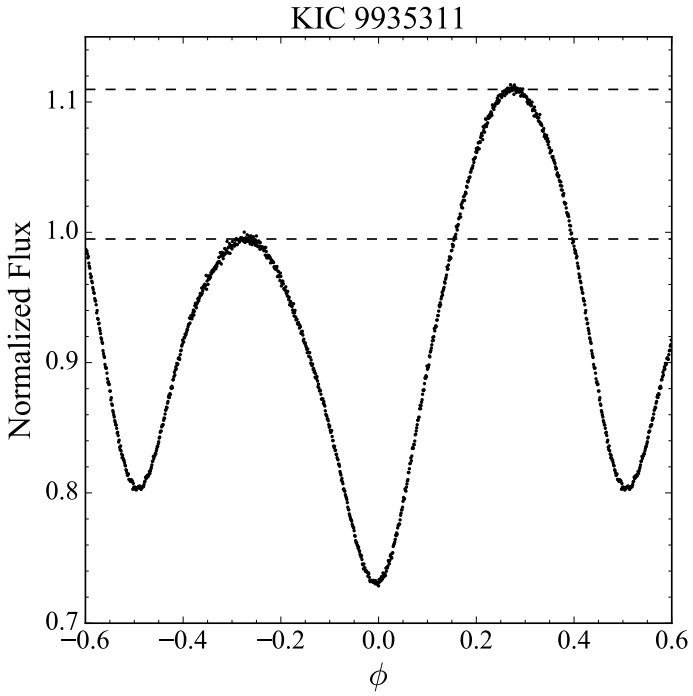


Figure 1.3: Averaged light curve of KIC 9935311 showing the O'Connell effect. The maximum following the primary eclipse (indicated with the upper dashed line) is significantly brighter than the maximum preceding it (indicated by the lower dashed line).

### 1.2.1 O'Connell (1951): Characterizing the O'Connell Effect

O'Connell (1951), hereafter O51, was the first major study of O'Connell effect systems. He presented a list of 87 eclipsing binaries systems along with their known properties, including the O'Connell effect size (OES, defined as $F(\phi_{\max 1}) - F(\phi_{\max 2})$ and known as $\Delta m$ in O51 and most other literature), spectral type, surface gravity, size, and absolute magnitude. Of these 87 systems, 75 display the O'Connell effect. He analyzed only those systems with photographic or photoelectric measurements due to their centimagnitude accuracy. He excluded from his

analysis 2 of the 75 systems (RS Canum Venaticorum and U Cephei) because "the asymmetry of the curves appears to be of a different type ... than that considered here." He also excluded W Ursae Majoris-type systems because it was difficult to differentiate the primary and secondary eclipses from the available data. Finally, he excluded systems with light curves (and therefore an OES) that varied with time, although he does not specify why.

O51 used Pearson's coefficient (r; Pearson 1895) to determine correlations between the OES and the following characteristics:

- Primary Ellipticity, $\left(\frac{b}{a}\right)_1$
- Secondary Ellipticity, $\left(\frac{b}{a}\right)_2$
- Relative Component Distance, $d$
- Absolute Component Distance, $\log D$
- Component Radius Ratio, $\log K$
- Larger Component Radius, $\log R_l$
- Smaller Component Radius, $\log R_s$
- Heaver Component Mass, $\log M_h$
- Lighter Component Mass, $\log M_l$
- Component Density Ratio, $\left|\log \frac{\rho_1}{\rho_2}\right|$
- Denser Component Density, $\log \rho_d$
- Less Dense Component Density, $\log \rho_l$

Ellipticity is the ratio of the equatorial semi-major and semi-minor axes ($a$ and $b$, respectively) of the stars, which are ellipsoidal due to the models available in 1951. Relative component distance measures the distance between the surfaces of each component relative to the distance between their centers, which is defined as one unit. By contrast, the absolute component distance measures the distance between each component's center in units of $R_\odot$. Finally, he (apparently) defines the component radius ratio as the larger star's radius divided by the smaller star's radius. O51 separately used 23 systems with a known OES in multiple colors to test the OES's color dependency.

O51 found strong correlations between the OES and the following parameters:

the difference in OES at different wavelengths, the ellipticity of both components, the ratio of stellar radii, and the ratio of stellar densities. The OES's color dependency was the strongest correlation ($r = 0.88$) and indicated that the OES becomes larger at shorter wavelengths as the OES increases. The OES's correlations with the component radius ratio ($r = 0.79$) and component density ratio ($r = 0.70$) were the next strongest correlations. The correlations with the component ellipticities ($r = -0.56$ for the primary, $r = -0.62$ for the secondary) are also strong. Finally, O51 noted that 55 of the 58 binaries with OES $\neq 0$ have OES $> 0$. Of the three exceptions, two are W Ursae Majoris-type systems (and thus beyond the scope of his paper), while Z Vulpeculae has an OES ($-0.01$) smaller than what he considers significant.

Finally, O51 looked at three hypotheses for the O'Connell effect. Mergentaler (1950) introduced the first hypothesis, which proposed that a layer of gas surrounding the stars – thicker on one side – causes the flux difference between the leading and trailing hemispheres. Mergentaler (1950) then concludes that the variation observed is due to selective absorption from $H^-$ ions. O51 states that this absorption would have to occur on the cooler star to explain the predominance of a positive OES. The second hypothesis, introduced by Struve (1941), proposed a circumstellar gaseous ring surrounding a component of $\beta$ Lyrae to explain emission features in its spectrum. This ring takes the form of a "twin-stream" of matter, with one stream flowing from the hotter to the cooler star and the other flowing the other way. O51 notes that the stream visible after the primary eclipse would be the one flowing from the hotter to the cooler star, which agrees with the observation that the system is bluer at that time. His study therefore agrees well with the "twin-stream" theory of Struve (1941). Finally, Kuiper's (1941) hypothesis proposed

that matter is transferred across a Lagrange point (either $L_1$ or $L_2$) in systems where at least one star touches a critical lobe. This matter stream obscures the binary at one quadrature but not the other, accounting for the asymmetric light curve. O51 claimed his work further supports this hypothesis.

### 1.2.2 Davidge & Milone (1984): Continuing O'Connell's Work

Davidge & Milone (1984), hereafter D84, was a continuation of O51's work. They presented a list of 103 O'Connell effect binaries and their known properties, among which were the OES in unspecified[3] UBV passbands. They analyzed only systems with multi-color photometry available and modern light curve solutions, meaning those analyzed by the Wilson–Devinney program (Wilson & Devinney 1971), the Wood model (Wood 1972), or the Russell–Merrill model (Russell & Merrill 1952). Like O51, D84 excluded overcontact systems, with the reasoning being that the parameters for those systems were not well-established at that time. They also excluded systems with a variable OES, and like O51, they do not provide a reason for doing so in their paper. In private correspondence, they stated that they removed these systems to remain consistent with O51. They also suggested that O'Connell's logic was that including these systems would blur correlations (T. Davidge, E. Milone, private communication). Finally, D84 restricted their analysis to BV data due to the small number of systems with non-zero $\mathrm{OES_U}$. These

[3]Inspection of D84's citations (particularly Mancuso et al. 1977 and de Bernardi & Scaltriti 1979) implies that the UBV filters were not the standard Johnson UBV filters (Johnson & Morgan 1953) in all cases.

restrictions reduced the sample they analyzed to 18 systems.

D84 studied all of the characteristics O51 did. Like O51, D84 used Pearson's coefficient to determine correlations between the OES and other characteristics. However, D84 also used the non-parametric Spearman's coefficient ($\rho$; Spearman 1904), which, unlike the parametric Pearson's coefficient, does not assume a linear correlation between characteristics, only a monotonic one. Additionally, it is not sensitive to outliers like Pearson's coefficient. D84 applied a linear regression to significant correlations to find the equation of best fit. They also tested the OES's color dependency using 54 systems, but unlike O51, they included systems with varying OES for this portion of their analysis.

By far the strongest correlation D84 found was $\mathrm{OES_V}$ with $\mathrm{OES_B} - \mathrm{OES_V}$, with $r = -0.82$ and $\rho = -0.87$. Here, $\mathrm{OES_B}$ and $\mathrm{OES_V}$ refer to the OES in the B and V filters, respectively. The values of the coefficients indicate strong anticorrelation. Other significant correlations with $\mathrm{OES_V}$ they found using their parametric analysis were those with (using my notation above) $\log D$ ($r = -0.66$) and $\log R_l$ ($r = -0.71$). With their non-parametric analysis, they found correlations between $\mathrm{OES_V}$ and $\left(\frac{b}{a}\right)_1$ ($\rho = -0.65$), $d$ ($\rho = -0.64$), and $\log K$ ($\rho = -0.67$). The only significant correlation they found with $\mathrm{OES_B}$ was with $\mathrm{OES_B} - \mathrm{OES_V}$ ($r = -0.59$). They find the lack of correlations with $\mathrm{OES_B}$ "puzzling."

D84 also searched for correlations for characteristics O51 did not. These included size- and temperature-differentiated correlations, which they do not explicitly define, but I take their terminology to mean comparing the hotter/larger star to the cooler/smaller star in each system based on their notation. They also checked correlations with the orbital period. They found no significant correlations with $\mathrm{OES_B}$ but several with $\mathrm{OES_V}$, just as with the characteristics O51 studied.

The strongest size-differentiated correlations were the equatorial semi-minor radius ($b$ in my notation above) of the smaller star ($\rho = 0.69$) and the ratio of $b$ between the hotter and cooler stars ($\rho = 0.67$), the latter of which they label as "somewhat redundant" due to the prior correlation with $\log K$. The strongest temperature-differentiated correlation was with the semi-minor radius of the hotter star ($\rho = 0.72$). Finally, they found a very strong correlation between $\mathrm{OES_V}$ and period ($r = -0.65$ and $\rho = -0.79$). Since their database is larger and the data more precise than O51's, they suggest that their results are more general. Additionally, they find a more even distribution of positive (~60% of their sample) and negative (~40%) OESs than O51.

While the correlations between the OES and color were the strongest ones in O51 and D84, they had opposite sign ($r = 0.88$ in the former, but $r = -0.82$ and $\rho = -0.87$ in the latter). Therefore, whereas O51 found that the OES becomes larger at shorter wavelengths as the OES increases, D84 found that the OES becomes larger at *longer* wavelengths as the OES increases. One potential explanation for this they give is that their sample has only six systems in common with O51's. Furthermore, their color analysis included systems with varying OES, unlike O51's. They state that Struve's (1941) "twin-stream" model is a likely mechanism for systems where the brighter maximum is bluer, but Mergentaler's (1950) $\mathrm{H^-}$ model is more appropriate for systems where the brighter maximum is redder, as in their study. They concluded that there are likely multiple causes of the O'Connell effect.

### 1.2.3 Wilsey & Beaky (2009): Discussing the O'Connell Effect's Causes

Wilsey & Beaky (2009), hereafter W09, provided an overview of the O'Connell effect's proposed explanations. They stated that Roberts's (1906) "periastron effect," Mergentaler's (1950) $H^-$ absorption theory, and Struve's (1941) "twin-stream" theory are no longer considered likely explanations of the O'Connell effect. The reason a single explanation of the O'Connell effect eludes us, they speculated, is due to its wide variety of manifestations. For instance, the O'Connell effect is temporally quite variable in some systems but non-variable in others. Additionally, it appears in systems of all morphological types despite fundamental differences in the structure and evolutionary state of stars in each type. Like D84, they concluded that there are multiple causes of the O'Connell effect.

W09 focused on four proposed explanations of the O'Connell effect: starspots, gas stream impacts, circumstellar material impacts, and asymmetric circumfluence in overcontact systems due to the Coriolis force.

- **Starspots**: This idea states that hot or cool spots caused by chromospheric activity on one component create the observed flux difference. W09 stated that starspots are the most common explanation of the O'Connell effect. They also noted that the WD program (the most commonly used program for modeling eclipsing binaries) incorporates starspots. The WD spot models can be used to explain both chromospheric starspots like those found on the Sun and starspots arising from the impact of a matter stream. However, W09 stated that the spots need to be unrealistically large to explain the observed O'Connell effect in many systems, and Maceroni & van't Veer (1993)

stated that the photometric light curves used to build these models are often insufficient to justify the existence of starspots. Furthermore, W09 stated that starspots would not be expected to remain stable in the long term, and Drake et al. (2014) found "no evidence for changes in the maxima that are expected as star spot numbers or sizes vary."

- **Mass Transfer**: This idea states that the impact of a matter stream onto the stellar surface or a circumstellar accretion disk creates a hotspot responsible for the observed flux difference. The matter stream originates from the Roche lobe-filling component of a semi-detached system. The stream is comprised of matter lost through the $L_1$ point. The kinetic energy of the matter heats the material surrounding its impact point, creating the hotspot that increases the observed flux when visible. W09 noted that "[t]he location of this point and the circumstances required for mass transfer are defined by the geometry of a close binary system," which implies that the hotspot's location is stable over long periods. V361 Lyrae (Andronov & Richter 1987; Hilditch et al. 1997) and GR Tauri (Gu et al. 2004) are two systems W09 noted were explained by this idea, and the model Djurašević et al. (2008) applied to RY Scuti also used this idea to explain its O'Connell effect. W09 do not consider the possibility of mass transfer due to winds.

- **Circumstellar Material**: This idea (Liu & Yang 2003) states that free-floating circumbinary material impacts the stars as they orbit, heating the leading hemispheres and causing the observed flux difference. This idea has the advantage that it can be used to theoretically predict the OES, although this requires knowing the difficult-to-measure density of the surrounding ma-

terial. Liu & Yang (2003) used this capability to compare their predictions to observed data. According to W09, Liu & Yang (2003) found "reasonable but not exceptional" agreement with their comparison. However, W09 stated that their model's assumptions are unrealistic. Liu & Yang (2003) do not posit a source for the material, but mass transfer could provide such a source.

- **Asymmetric Circumfluence**: This idea (Zhou & Leung 1990) states that hot matter is preferentially sent to one side of an overcontact binary and causes the observed asymmetry. The hot material accumulates on one side because the Coriolis force deflects the flow of material. Unlike the Liu & Yang (2003) model, their model does not predict the OES, and W09 stated that Zhou & Leung (1990) provided no method of verifying their idea.

Finally, W09 presented BVRI[4] observations of the eclipsing binaries V573 Lyrae and UV Monocerotis, both of which show the O'Connell effect. They state that UV Monocerotis bears a striking resemblance to V361 Lyrae, which they noted earlier had its O'Connell effect explained by the mass transfer idea. In addition to measuring the OES of each system, they applied the two statistical measures McCartney (1999) introduced, O'Connell Effect Ratio (OER) and Light Curve Asymmetry (LCA). They analyzed the light curves by fitting a Fourier series to them and determining the OES, OER, and LCA using the Fourier series. The use of a Fourier series to determine OER and LCA represented a novel improvement over the original method used by McCartney (1999). W09 noted that the $b_1$ Fourier coefficient is related to the OES (specifically, OES $\approx -2b_1$), the $a_1$ coefficient is

[4]While not stated in W09, the observations were conducted using Bessell filters (V. Gokhale, private communication).

related to the difference in minima depth, and $a_2$ and $a_4$ help determine the light curve class (Rucinski 1997). Section 2.2.2 describes OER, LCA, and the use of the Fourier series in further detail.

## 1.3 Eclipsing Binary Modeling

Eclipsing binary modeling refers to the process of determining the system parameters (such as orbital inclination, eccentricity, component temperatures, and component masses) from observational data (Kallrath & Milone 2009). Complete system characterization requires three data sets (except for totally eclipsing semi-detached and overcontact systems that require only one radial velocity curve; see Terrell & Wilson 2005 and Wilson 2006): a complete photometric light curve and complete radial velocity curves of each component. The light curve supplies the orbital inclination, temperature ratio, and each component's Roche equipotential. Meanwhile, the radial velocity curves fix the system's absolute parameters, including mass, radius, and semi-major axis.

There are two concepts associated with modeling: the direct problem and the inverse problem. The direct problem is taking a set of observables and generating an observable curve, which could be a light curve or a radial velocity curve. Solving the direct problem requires an understanding of three key components:

1. The system's geometry, including factors such as the orbital inclination and eccentricity and the sizes and shapes of the component stars.

2. The stellar radiative properties, including temperature, gravity brightening, limb darkening, and effects from the stellar atmosphere.

3. The observed flux, taking into account eclipses as well as possible circumstellar material.

It is necessary to have a model that can solve the direct problem before solving the inverse problem.

The inverse problem is taking an observable curve and generating a set of observables. It is a non-linear least-squares problem (Kallrath & Milone 2009, pg. 169). Solving the inverse problem is done by performing an iterative least-squares fit between a set of observable and an observable curve to minimize their difference. The global minimum of the parameter space is assumed to represent the system's parameters. However, the existence of local minima complicates the analysis, as resolving whether the found minimum is global or local requires multiple models. If several models produce similar observables, it increases the confidence that they describe the system's true properties. In addition to this uniqueness problem, there are strong correlations between various parameters (such as the mass ratio and inclination), wherein reducing one parameter can be compensated for by increasing the other. These correlations make it more difficult to find the parameter space's global minimum by increasing the problem's complexity.

The earliest eclipsing binary models were purely geometrical regarding star shape. The most notable of these was the Russell–Merrill model (RM; Russell & Merrill 1952). RM used a truncated Fourier series to "rectify" the observed light curve to remove the effects of ellipticity and irradiation by companions. It then modeled the stars as spheres and produced a light curve using the Stefan–Boltzmann law and eclipse occlusion. Such models failed spectacularly when applied to systems with strongly distorted components, especially overcontact systems (Kallrath & Milone 2009). Nonetheless, RM's simplicity allowed for manual

computation of system parameters, which was a distinct advantage in the pre-computer era.

As the computer age began, geometric models like RM were replaced by models using Roche equipotentials like Equation 1.6 to define stellar surfaces. While more complex than geometrical models, Roche geometry is more accurate and naturally handles distorted components. The Wilson–Devinney program (WD; Wilson & Devinney 1971) was the most prominent Roche geometry model and remains the most used eclipsing binary model (Kallrath & Milone 2009). WD includes many observables, including orbital (such as inclination, eccentricity, and argument of periapsis), component (such as temperature, albedo, and Kopal potential value), and spectral (i.e. radial velocity) observables, and it incorporates a least-squares algorithm to solve the inverse problem.

PHysics Of Eclipsing BinariEs (PHOEBE; Prša & Zwitter 2005; Prša et al. 2016) is a graphical user interface (GUI) eclipsing binary program. PHOEBE Legacy (Prša & Zwitter 2005) is based on WD, while PHOEBE 2's (Prša et al. 2016) light curve synthesis code is written from scratch but based on WD's principles. I have used PHOEBE Legacy v0.31a and v0.32 to produce the binary models in this dissertation due to its ease of use compared to WD. PHOEBE 2 promises several innovations lacking in WD, including the addition of an accretion disk, pulsations, and accounting for interstellar reddening and extinction. However, PHOEBE 2 is still under development and currently lacks support for features critical to this dissertation, most notably the ability to accurately model overcontact systems.

## 1.4 Genesis: NSVS 7322420

The genesis of this project stemmed from my desire to understand NSVS 7322420, one of two eclipsing binary systems I studied for my master's thesis (Knote M.S. Thesis; Knote et al. 2019; in prep.). The models I produced for my thesis and Knote et al. (2019) indicated that NSVS 7322420 is a semi-detached system with the primary component filling its Roche lobe. My early observations of the system revealed several peculiar features, including a pronounced O'Connell effect as well as a noticeable "kink" around the secondary eclipse. This "kink" was a sudden, nearly discontinuous change in flux. The primary minimum was also asymmetric, sloping upward as the system passed the primary eclipse. Finally, the light curve changed noticeably over time scales as short as half a year, particularly regarding the minima depth and maxima height.

PHOEBE could not accurately reproduce the observed light curve because of these features. The model Knote et al. (2019) presented only used data from January 2014 due to the aforementioned temporal variance, and Figure 1.4 shows that this model only marginally fits the data. Reproducing the "kink" feature required two spots on the secondary component (plus the one explaining the O'Connell effect), and it is difficult to find a physical rationale for their existence. Due to the additional free parameters starspots introduce (as discussed in Section 1.2.3) and poor fit to the observed data, the Knote et al. (2019) model is not a realistic description of the system. The middle column of Table 1.2 gives this model's final output parameters.

I obtained more photometric data of NSVS 7322420 in late 2015 and early 2016, while several collaborators (Thomas Boudreaux and Drs. Brad Barlow and Patricia Lampens) obtained concurrent spectroscopic data. I produced a second

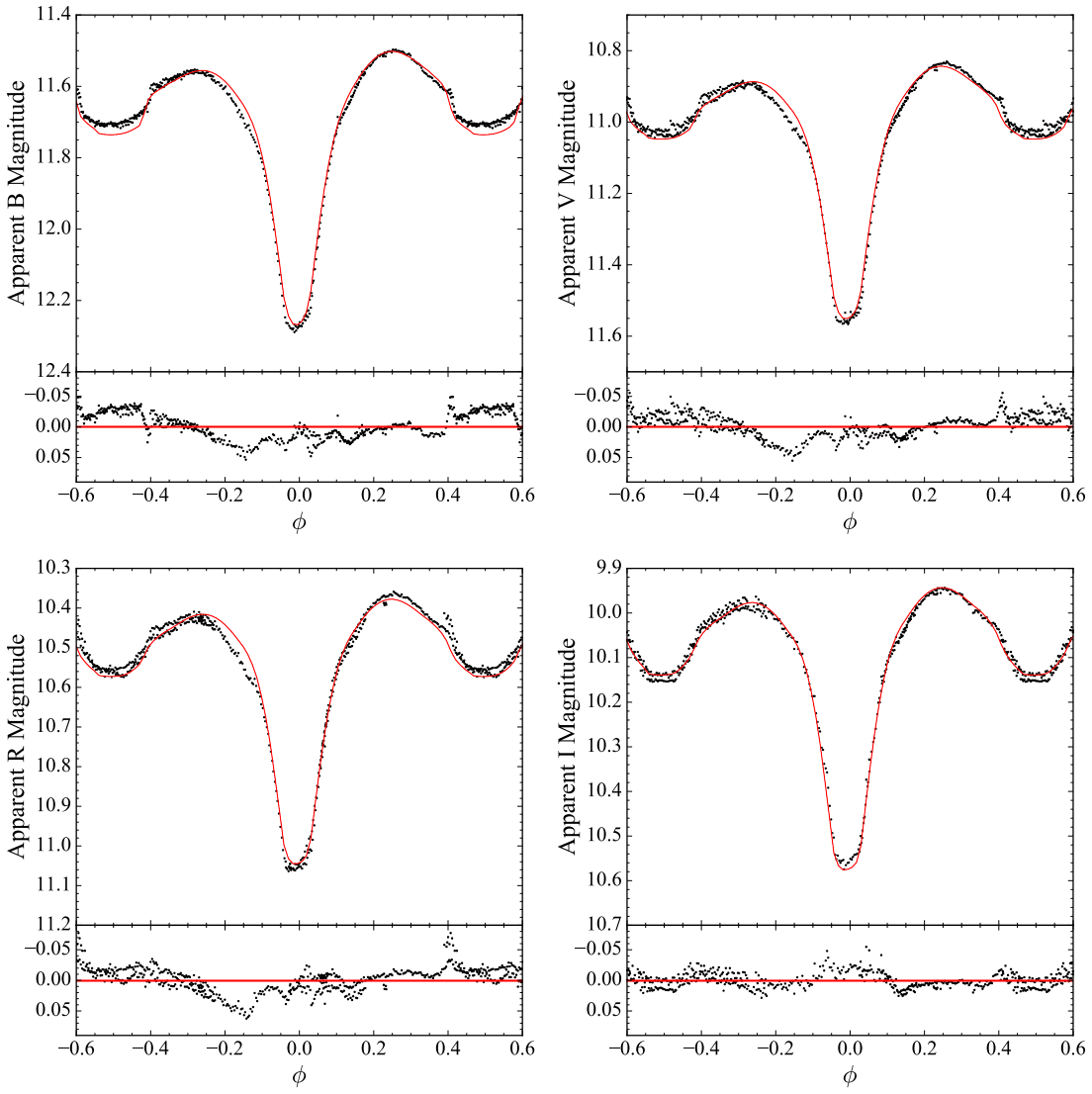


Figure 1.4: B (top left), V (top right), $R_C$ (bottom left), and $I_C$ (bottom right) filter light curves for NSVS 7322420 using January 2014 data. The red line shows PHOEBE's synthetic model produced using this data. The lower panel shows the observed data's divergence from the synthetic model as a function of phase.

model of the system using this data, which Figures 1.5 and 1.6 show. This model fits the observed data much better than the Knote et al. (2019) model, and is more physically plausible due to the reduced number of free parameters. I was able to model the "kink" with only one spot (plus the one explaining the O'Connell effect),

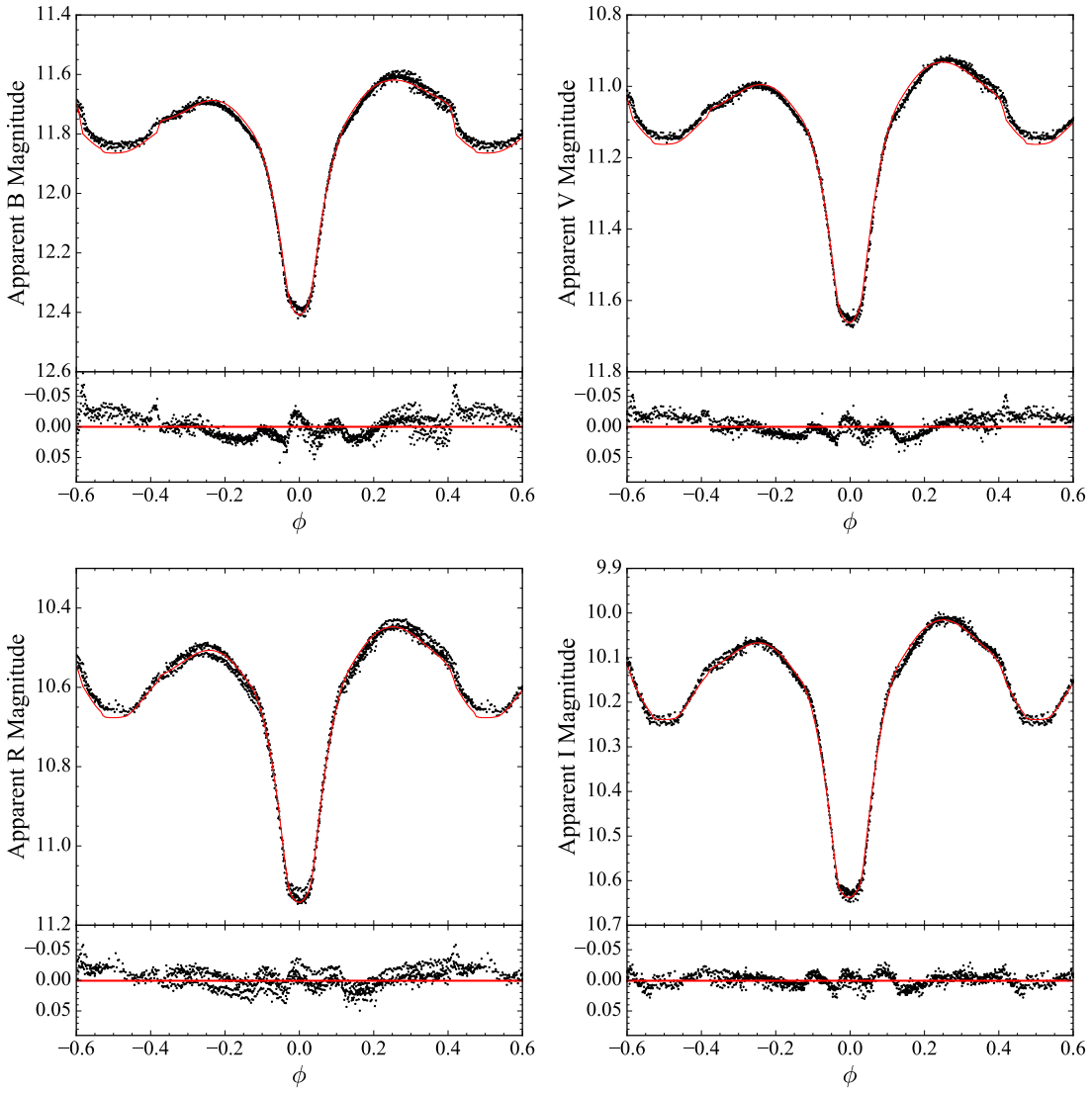


Figure 1.5: B (top left), V (top right), $R_C$ (bottom left), and $I_C$ (bottom right) filter light curves for NSVS 7322420 using January and February 2016 data. The red line shows PHOEBE's synthetic model produced using this data. The lower panel shows the observed data's divergence from the synthetic model as a function of phase.

and this first spot can be interpreted as the initial impact of the matter stream on the secondary star. The second spot at longitude 90° can be interpreted as a second impact zone if we postulate that the matter hitting the first spot has

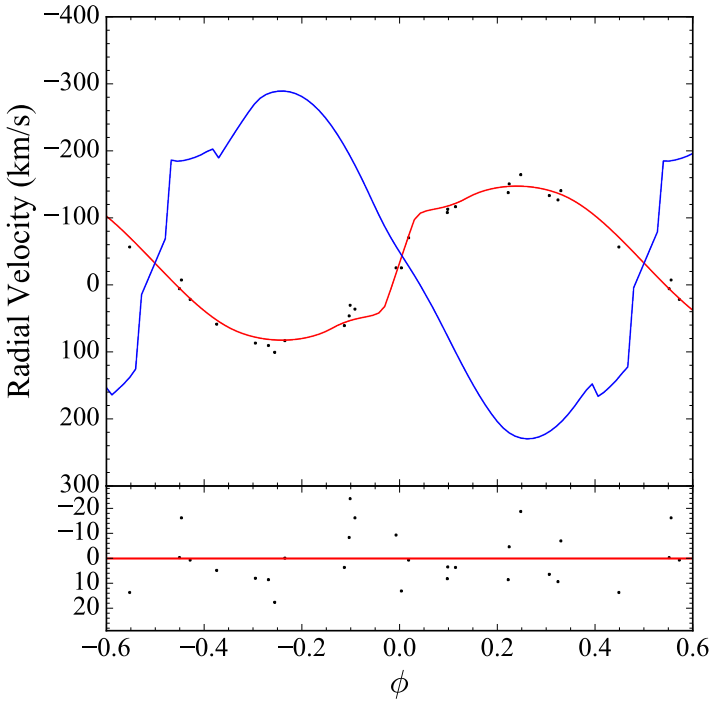


Figure 1.6: Radial velocity curve for NSVS 7322420. The lines show PHOEBE's synthetic model for the primary (red) and secondary (blue) components produced using this data. The lower panel shows the observed data's divergence from the primary component's synthetic model as a function of phase.

sufficient momentum to "bounce" off the star and impact it a second time. This "bounce" could explain why the second spot is much larger, cooler, and at a higher longitude than the first. The "kink" itself is a result of the first spot's small size but large luminosity, meaning that when it disappears or reappears, it causes a large and rapid change in flux. The rightmost column of Table 1.2 gives this model's final output parameters. Note that the spot used to explain the "kink" is described as the first spot in this paragraph, but is Spot 2 in Table 1.2.

NSVS 7322420's spectra indicate that the system is a single-lined spectroscopic binary (SB1), meaning that its spectra only show the spectral lines from the brighter star. I determined the component stars' masses, absolute sizes, and

Table 1.2: Results of NSVS 7322420's modeling with the 2014 data set (middle column) and the 2016 data set (right column).

| Parameter | 2014 Model | 2016 Model |
|---|---|---|
| Orbital Period (d) | $0.467467 \pm 0.000015$ | |
| Orbital Inclination (°) | 91.66 ± 0.04 | 89.44 ± 0.05 |
| Mass Ratio | 0.3388 ± 0.0025 | 0.4371 ± 0.0015 |
| $T_{eff}$ of Primary Component (K, fixed) | 5,700 | 5,650 |
| $T_{eff}$ of Secondary Component (K) | 3,361 ± 29 | 3,622 ± 13 |
| Mass of Primary Component ($M_\odot$) | — | 2.0512 |
| Mass of Secondary Component ($M_\odot$) | — | 0.8964 |
| Radius of Primary Component ($R_\odot$) | — | 1.6497 |
| Radius of Secondary Component ($R_\odot$) | — | 0.9673 |
| Surface Potential of Primary Component | — | — |
| Surface Potential of Secondary Component | 2.6085 ± 0.0088 | 2.9619 ± 0.0056 |
| Surface Gravity of Primary Component ($\frac{m}{s^2}$) | — | 206.41 |
| Surface Gravity of Secondary Component ($\frac{m}{s^2}$) | — | 262.39 |
| Absolute Magnitude of Primary Component | — | 3.7517 |
| Absolute Magnitude of Secondary Component | — | 6.8406 |
| Magnitude Difference ($M_2 - M_1$) | 3.4974 | 3.0889 |
| Spot 1 Temperature (K) | 6,227 ± 89 | 5,433 ± 20 |
| Spot 1 Latitude (°, fixed) | 0 | 0 |
| Spot 1 Longitude (°, fixed) | 90 | 90 |
| Spot 1 Radius (°, fixed) | 16 | 25 |
| Spot 2 Temperature (K) | 6,722 ± 58 | 9,779 ± 35 |
| Spot 2 Latitude (°, fixed) | 0 | 0 |
| Spot 2 Longitude (°, fixed) | 30 | 10 |
| Spot 2 Radius (°, fixed) | 5 | 5 |
| Spot 3 Temperature (K) | 6,722 ± 58 | — |
| Spot 3 Latitude (°, fixed) | 0 | — |
| Spot 3 Longitude (°, fixed) | 340 | — |
| Spot 3 Radius (°, fixed) | 5 | — |

luminosities from the radial velocity data obtained from these spectra. This determination was possible despite NSVS 7322420 being an SB1 because it is a totally eclipsing, semi-detached system (Terrell & Wilson 2005; Wilson 2006). However, the masses and absolute magnitudes in Table 1.2 are at odds with the other observed properties. While the primary component fills its Roche lobe and is therefore probably evolved, the secondary's temperature is at odds with its mass. Addition-

ally, according to Dr. Eric Mamajek's online table[5] (Pecaut & Mamajek 2013), the primary is smaller than a main sequence star of its mass should be, rather than larger like a subgiant or giant would be. Finally, the primary's absolute magnitude does not agree with NSVS 7322420's apparent magnitude and 178 pc distance (Bailer-Jones et al. 2021), instead implying either a greater distance or a brighter apparent magnitude. All of these issues can ultimately be traced back to a mass ratio that is too small, and a manually input mass ratio of ~0.6 produces values more in line with the observed properties.

NSVS 7322420 changed in the time between the observations in 2014 (Figure 1.4) and 2016 (Figure 1.5). Most notably, the system systematically dimmed by 0.1 magnitudes in the interim. A more minor change is that the asymmetric minimum slopes upward toward the brighter maximum in Figure 1.4 but toward the dimmer maximum in Figure 1.5. Additionally, the brightness entering and exiting the secondary eclipse is similar in Figure 1.4, but the ingress is much brighter than the egress in Figure 1.5. Data taken in April 2018 showed that the system had returned to its original brightness displayed in Figure 1.4, and the "kink" around the secondary eclipse had entirely disappeared. Combined with the changes preventing me from combining all the data within a given year, these changes indicate that the system varies over short and long timescales.

While my second model better fits the observations, it still fails to explain certain features of the light curve, most notably the asymmetric primary minimum. Additionally, PHOEBE insists on a mass ratio that is small enough to give masses and luminosities incongruent with other observed properties. Therefore, this model

[5] http://www.pas.rochester.edu/~emamajek/EEM_dwarf_UBVIJHK_colors_Teff.txt

does not yet give a fully accurate description of NSVS 7322420. My desire to understand this system prompted me to search for other systems showing similar irregularities, which was the genesis of my dissertation project.

## 1.5 My Project and Goals

The focus of my dissertation is to characterize the observational properties (such as period, color, OES, and eclipse depth) of a complete and statistically significant sample of O'Connell effect systems. This characterization is meant to lay the foundation needed to tackle the deeper question of what causes the O'Connell effect. My sample of O'Connell effect systems consists of 258 binaries observed by the *Kepler* space telescope (Borucki et al. 2010). I divided this sample into a core sample of 212 systems showing an OES of at least 1% in normalized flux and a marginal sample of 46 systems where the OES was smaller than this cutoff.

The goal of this project is to learn more about O'Connell effect systems and answer questions raised in earlier works. One such question arises from the observation that the O'Connell effect is found exclusively in systems where the stars are close enough to distort or otherwise significantly affect each other. Is this observation a selection bias, or is the O'Connell effect related to binary interaction? Another observation is that O51 found that the brighter maximum always followed the primary minimum, while D84 found several systems where the brighter maximum preceded the primary minimum. What is the true proportion of where the brighter maximum occurs? Furthermore, are there differences between systems showing the former versus the latter? Another question strikes at the heart of the O'Connell effect's physical cause, particularly between the starspot and mass

transfer explanations: do these systems correlate in the manner Kouzuma (2019) suggests they should, or do they show the eclipse timing variation signal indicative of mass transfer? Are the O'Connell Effect Size and Light Curve Asymmetry statistical measures McCartney (1999) introduced a better way to measure the O'Connell effect than OES? The final question my project hopes to answer is determining how prevalent the O'Connell effect is in eclipsing binaries.

I discuss the methodology I use to select and analyze my sample in Chapter 2. This discussion includes the specifics of the *Kepler* satellite, how I defined my sample, the codes and methods I used to analyze the sample systems, and how I planned and conducted observations on select sample systems. In Chapter 3, I review the literature on my sample's systems to describe what is already known about my sample and the systems therein. Chapter 4 presents the results of my analysis as well as a discussion of those results. These results include the observational characteristics of the sample and analysis of the distributions and correlations of these characteristics. Chapter 5 discusses four subsets of my sample with peculiar features: systems showing strong temporal variation in their light curve, systems with an asymmetric minimum, the lone white dwarf in my sample, and systems displaying a concave-up region in their light curve. The first two subsets are defined by features also present in NSVS 7322420, so understanding these systems allows us to better understand that enigmatic system. I present my observational data on ten systems from my sample in Chapter 6. Finally, Chapter 7 summarizes this dissertation's major results and conclusions and details the future work that arises from my project.

# Chapter 2

# Methodology

I used several methods to define my sample and characterize the systems within it. Many of these methods involved analyzing the phased data from *Kepler* using Python codes (Van Rossum & Drake Jr 1995) I created for that purpose. I also conducted observations on several of the systems in my sample and used several methods to analyze the data I obtained. This chapter focuses on the methodology I used during this project.

## 2.1 Target Selection

My project's main goal is to characterize eclipsing binaries displaying a significant O'Connell effect. I selected my sample from the *Kepler* Eclipsing Binary Catalog (KEBC; Kirk et al. 2016), a set of 2,907 eclipsing binaries observed by the *Kepler* space telescope. I defined my sample using a single criterion: an O'Connell effect with an average |OES| greater than 1% in normalized flux (I detail the flux normalization in Section 2.1.4). There are two reasons for this criterion:

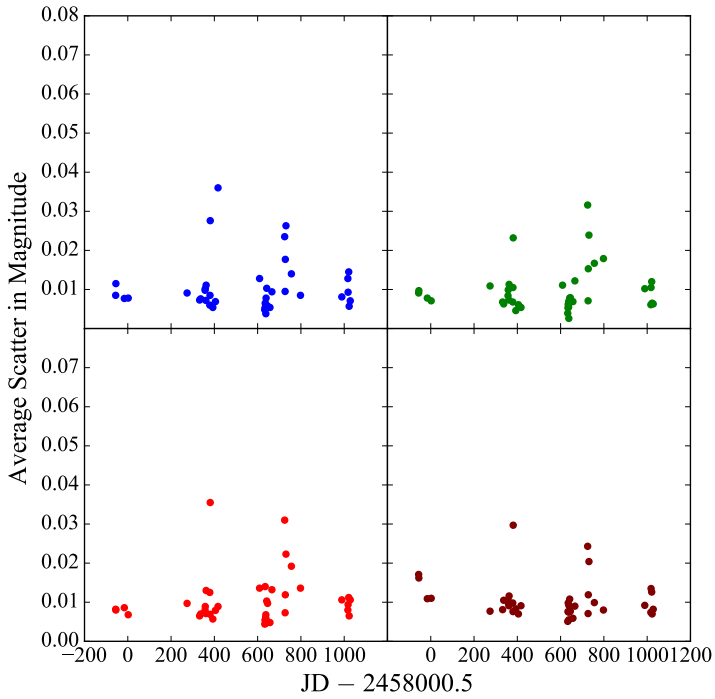


Figure 2.1: Plot showing the average scatter of my observational data for each night in the B (upper left), V (upper right), R (lower left), and I (lower right) filters. The typical scatter is about 1%.

- Sunspots cause up to a 0.3% difference in the luminosity of the Sun (Willson et al. 1981), which means that an |OES| over 1% requires either significantly more spotting than present on the Sun or a different explanation altogether. My project is therefore aimed at better understanding systems displaying an O'Connell effect that solar-type spots cannot easily explain.

- Follow-up observations are unlikely to have the same precision as the *Kepler* data, so an |OES| larger than 1% is easier to observe than a smaller |OES| would be. For reference, Figure 2.1 shows that my observational data's scatter is around 1% on most nights for targets ranging in magnitude from V = 11 to V = 15.

Table 2.1: Median, mean, and standard deviation of the KEBC's and core sample's OES and |OES| distributions.

| Sample | Median OES | Mean OES | $\sigma_{\mathrm{OES}}$ | Median \|OES\| | Mean \|OES\| | $\sigma_{\mathrm{\|OES\|}}$ |
|---|---|---|---|---|---|---|
| KEBC | 0.001% | 0.075% | 0.832% | 0.030% | 0.260% | 0.794% |
| Core | 1.314% | 0.713% | 2.874% | 1.622% | 2.099% | 2.088% |

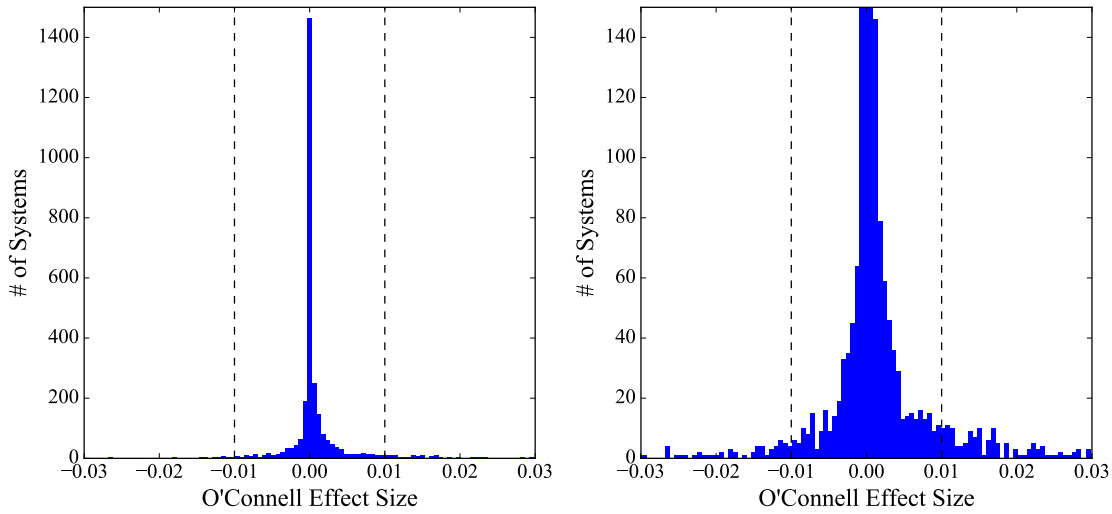


Figure 2.2: Histograms showing the OES distribution for the KEBC. The central peak at an OES of 0.00 rises to a value of 1,463 systems and is truncated in the right panel for clarity. The 26 systems with |OES| larger than 0.03 are also not displayed for clarity. The dashed lines indicate the |OES| = 0.01 cutoff defining the core sample.

I varied the criterion between 0.5% and 1.5% and found that the results of my analysis were largely independent of the exact cutoff value in that range.

I determined the median, mean, and standard deviation ($\sigma$) of the KEBC's and core sample's (defined in Section 2.1.5) OES and |OES|. Table 2.1 lists these values, while Figure 2.2 shows two histograms of the OES distribution for the KEBC with the |OES| = 0.01 cutoff indicated by dashed lines. Figure 2.2's left panel shows that the KEBC's OES distribution is significantly non-Gaussian. As

a result, $\sigma_{\rm OES}$ does not indicate the distance from the mean value wherein ~68% of the population lies, as it would if the distribution was Gaussian. Instead, using Chebyshev's inequality (Chebyshev 1867), it can be inferred that at most 25% of the population lies more than $2\sigma_{\rm OES}$ away from the mean OES. I determined that the core sample does not include 45 systems with an |OES| larger than the $1\sigma$ value but below the 1% cutoff (equivalent to $1.2\sigma_{\rm OES}$).

### 2.1.1 *Kepler* Space Telescope

The *Kepler* space telescope (Borucki et al. 2010) is a 0.95-meter telescope in an Earth-trailing orbit (Van Cleve & Caldwell 2016) launched on March 7, 2009, with the primary goal of detecting planetary transits. Once operational, it observed a field centered on the J2000 equatorial coordinates $19^{\rm h}22^{\rm m}40^{\rm s}$ $+44°30'00''$ in the constellation of Cygnus. With a 16.1° field of view, it also observed portions of Lyra and Draco. The *Kepler* field was previously observed in anticipation of the *Kepler* mission, resulting in the *Kepler* Input Catalog (KIC; Brown et al. 2011) of 13,161,029 (Thompson et al. 2016) objects. Only about a third of these objects ever fell on *Kepler*'s CCDs, however, and *Kepler* only observed about 150,000 objects at any given time. The observing seasons were called quarters (Q) and lasted three months, between which the *Kepler* spacecraft would roll 90° to keep its solar panels pointed at the Sun. *Kepler* ceased observing the Cygnus field when a second reaction wheel failed on May 11, 2013, during Q17 (Van Cleve et al. 2016), rendering the spacecraft incapable of maintaining its pointing to mission specifications.

*Kepler* is equipped with a 94.6-megapixel CCD detector array operated at a temperature of −85 °C (Van Cleve & Caldwell 2016). The array contains 42

CCDs arranged into 21 pairs. *Kepler* could only transmit full-frame images once per month due to limited bandwidth. Therefore, only the pixels surrounding selected targets were stored and transmitted to Earth. *Kepler*'s observations were accurate to within 29 parts per million (ppm) for a star with an apparent magnitude of 12 and 80.7 ppm at an apparent magnitude of 14.5 (Gilliland et al. 2011), which is superior to what can be achieved from the ground. For comparison, Tregloan-Reed & Southworth (2013) cites a precision of 258 ppm (of a V = 11.45 star using the 4-meter Mayall telescope; Gilliland et al. 1993) as the most precise ground-based observations known. Therefore, *Kepler* is much more precise than ground-based observations even for stars fainter than V = 11.45, so features like kinks or subtle changes over time are more apparent in the *Kepler* data. *Kepler*'s observations consisted of 6.02-second exposures that had a readout time of 0.52 s. *Kepler* observed targets in two cadences: short cadence and long cadence. A short cadence (SC) coadded 9 exposures, giving an effective integration time of 58.85 s. A long cadence (LC) coadded 270 exposures, giving an effective integration time of 1,765.5 s. While every target observed was a long-cadence target, *Kepler* could only observe 512 short-cadence targets at any given time.

The observed targets were drawn from the KIC, which estimated each star's spectral type and luminosity class based on photometrically determined colors. Batalha et al. (2010) describes the *Kepler* target prioritization and statistics for the operational target list. They assigned the highest priority to systems with a *Kepler* magnitude $K_p < 14$, and where an Earth-sized planet in the habitable zone would produce at least three transits in the 3.5-year mission with a signal-to-noise ratio greater than $7.1\sigma$. They chose the first condition to facilitate follow-up high-precision spectroscopy to confirm planet detections, while the second condition

eliminated most O- and B-type stars. As a result of this prioritization, the *Kepler* target list includes more G- and F-type stars than K- or M-type stars, and *Kepler* only observed ~3,000 M-type stars and fewer than 200 O- and B-type stars. The *Kepler* target list is also biased against giants and subgiants because planetary transits are harder to detect due to the smaller ratio between the star and planet radii. Comparing Tables 2 and 3 of Batalha et al. (2010) shows that the *Kepler* target list included fewer than 10% of giants located within the *Kepler* field of view. Because of these selection biases, my study undersamples the giants and subgiants as well as the low- and high-mass stellar population. Consequently, it provides less information on the O'Connell effect in systems containing those stars.

Bryson et al. (2020) and Wolniewicz et al. (2021) further investigated the *Kepler* selection function using *Gaia* DR2 data (Gaia Collaboration et al. 2018). Bryson et al. (2020) focused on exoplanet occurrence rates, while Wolniewicz et al. (2021) focused on stellar populations. Wolniewicz et al. (2021) found that the *Kepler* sample is nearly unbiased for stars brighter than $K_p = 14$. For fainter stars ($K_p > 14$), they find a bias toward main sequence and subgiant stars with late-F to early-M spectral types and a bias against cool giants. They state that the bias toward subgiants was because the KIC misidentified them as main sequence stars. They also found a bias against binaries by analyzing the re-normalized unit weight error (RUWE), which is the normalized $\chi^2$ obtained from fitting the point-spread function of *Gaia* sources re-normalized to correct for color-dependent biases (Lindegren 2018). They considered a system to be binary when its RUWE was greater than 1.2 (Kraus et al. in preparation) and found that *Kepler*'s completeness was 8% lower for main sequence binaries than it was for solitary main sequence stars for targets fainter than $K_p = 14$. It is unclear how significantly this impacts

the KEBC, however, because Wolniewicz et al. (2021) are considering systems with stellar separations on the order of tens of AU and above, while even the most widely separated systems in the KEBC ($P \sim$ 1,000 d) only have separations of order a few AU. *Gaia* EDR3 data (Gaia Collaboration et al. 2021) shows that some of the KEBC (24.5% of the 2,861 systems with *Gaia* parallaxes) have an RUWE above the 1.2 limit Wolniewicz et al. (2021) cites. That nearly a quarter of the KEBC shows a RUWE above this limit means that the *Kepler* target list likely excluded some close eclipsing binaries.

### 2.1.2 *Kepler* Eclipsing Binary Catalog

*Kepler* detected many eclipsing binaries during its original mission. These systems have extremely well-characterized light curves due to the photometric precision required for detecting planetary transits. After the final *Kepler* data release, Kirk et al. (2016) compiled the KEBC, a catalog of 2,920 entries about 2,907 unique eclipsing binary systems observed by *Kepler*. The KEBC has multiple entries for a handful of systems, indicating either a second eclipsing binary or a third body that eclipses the inner binary. Figure 2.3 shows the light curves of a system in the latter category, KIC 2856960 (Armstrong et al. 2012). The KEBC contains an automatically extracted, phased light curve for each entry and several parameters determined by the KEBC team, including period, morphology parameter (discussed in Section 2.2.3), and the eclipse timing variation (ETV, discussed in Section 2.2.4). The KEBC is accessible online[1] and can be searched based on parameters like period, temperature, and morphology parameter. The *Kepler* data

[1] `http://keplerebs.villanova.edu/`

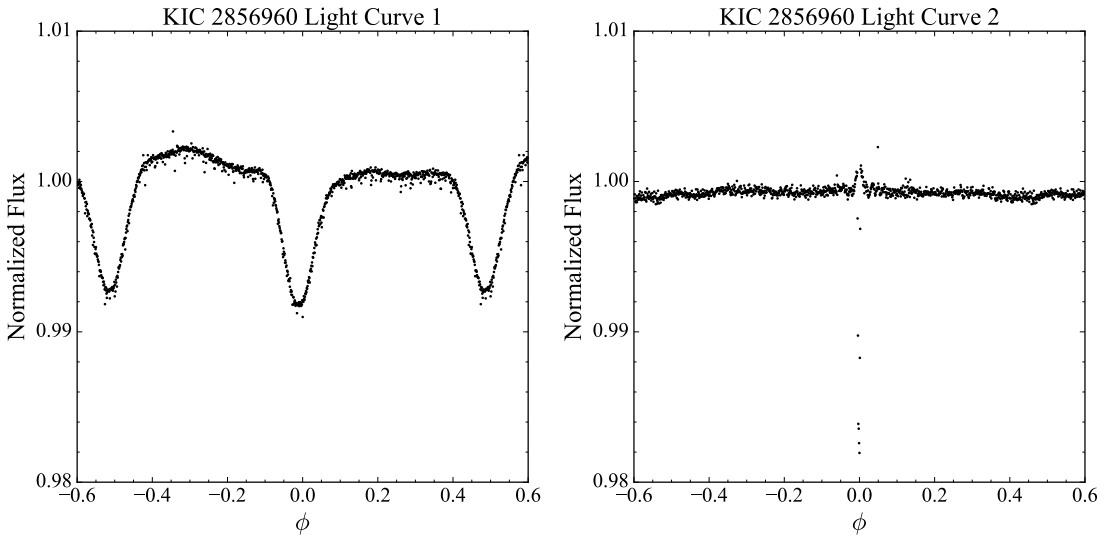


Figure 2.3: Averaged light curves of KIC 2856960, demonstrating that the system contains a third body eclipsing an inner eclipsing binary. The data was phased using a 0.258 d period (left) and a 204.256 d period (right).

for each system is available for download from the KEBC website.

The KEBC provides an excellent source for my study for three primary reasons: the large number of systems observed, the photometric precision of the data, and the time span of observation. The 2,907 eclipsing binaries in the KEBC represent a large, complete sample of eclipsing binaries. The KEBC contains eclipsing binaries with a wide variety of different characteristics and allows me to study a statistically significant sample. Kirk et al. (2016) states that the KEBC completeness is 89.1% for eclipsing binaries, with essentially 100% completeness for systems with periods of order one day. However, Bienias et al. (2021) recently found 547 short-period eclipsing binaries in the *Kepler* field not included in the KEBC. These new eclipsing binaries are fainter than the KEBC systems, with an average magnitude of 18.2. Nevertheless, the KEBC's high completeness for brighter ($K_p \lesssim 16$) means it is well-representative of the true eclipsing binary population (modulo the *Kepler*

selection function).

The excellent photometric precision of *Kepler*'s data, discussed in Section 2.1.1, allows me to study finer details in the light curve that statistical noise may have otherwise hidden, such as kinks or subtle changes over time. Additionally, as a result of *Kepler*'s precision, I am more confident that any unique features observed in a light curve are real structures rather than a result of statistical noise. Some targets have short-cadence data available, which allows me to further study short-timescale variations in those systems. However, this dissertation is focused on the analysis of the long-cadence data, and analysis of the short-cadence data is largely left for future work.

*Kepler* observed most systems nearly continuously for over three years, providing an opportunity to see how each system changes over that time. Such changes can be attributed to various factors, including starspot evolution or accretion disk instabilities. The long observation span maximizes the sensitivity to transient effects like flares, allowing me to characterize the prevalence of these events in my sample. Additionally, variations in the eclipse timing can indicate the presence of a third body in the system or an actual change in the period due to mass transfer. Finally, observing these targets for such a long, continuous time minimizes the risk of artifacts due to poor sampling or transient effects such as flares. The effects of such artifacts are unpredictable, with one such effect (an oscillatory signal introduced in systems with a period that is a near-integer multiple of the *Kepler* cadence) discussed at the end of Section 2.2.2.2.

I use the KEBC sample as a point of comparison throughout this dissertation. However, I use three separate subsets of the KEBC (each more restrictive than the last) in different contexts. The first subset – the comparative subset – removed

four systems, one due to data corruption and three because they are not in the KEBC. I used the comparative subset to create the histograms, color-magnitude diagram, and HR diagram in Chapter 4. The second subset – the trend subset – removed a further 228 systems due to missing certain parameters (specifically a *Gaia* parallax or a valid morphology parameter) or having multiple KEBC entries like KIC 2856960 (Figure 2.3). I use the comparative subset in the plots in Sections 4.2.1, 4.2.3, and 4.3.2.3. Finally, the third subset – the analysis subset – removed a still further 1,309 systems to select only systems similar in period and light curve shape to the core sample. I use the analysis subset in the statistical analysis explained in Section 2.2.5 and discussed in Sections 4.2.3 and 4.3.2.3. I describe these subsets further in Sections 4.1, 4.2.1, and 4.2.3, respectively.

### 2.1.3 Selection Methods

I tried several methods to select the targets in my sample before settling on a single, objective method. At first, I tried manually selecting the sample based on visual inspection of KEBC light curves, which resulted in an initial 22 systems, but this method is inherently subjective and non-reproducible. I next tried using the automated detector Johnston et al. (2019) describes, as they created this detector to define my sample. The automated detector selected nearly 70 new systems meeting the criterion, but it also selected almost 80 systems that did *not* meet the criterion. Additionally, the detector did not find several other systems that I knew from visual inspection should be in my sample, possibly due to a limited training sample. Therefore, I developed a code that directly calculated the OES, which I used to select my sample. In this section, I describe both the automated detector and the OES detection code.

#### 2.1.3.1 Automated Detector for Sample Selection

Kyle Johnston created the automated detector (Johnston et al. 2019) as part of his PhD dissertation. The detector is a similarity-based two-class detection algorithm, which means a system is either an O'Connell effect binary or it is not. It uses a waveform representation parameter space and a metric learning classifier. The waveforms chosen were the phased *Kepler* light curves obtained from the KEBC and scaled such that the normalized flux ranged from 0 to 1. Johnston selected this combination of components as it provides several advantages that are of interest to astronomers:

1. The detector uses metric learning, which defines a distance between two points in a parameter space. This means that any decision made, such as a target being identified or missed, can be linked back to the original training set, providing transparency in the decision-making process. In other words, the detector classifies a target "of interest" because it is most similar to a training point that was also "of interest."

2. The detector used representation learning, which applies the learning methodology to a transformed version of the data rather than the raw data. Lin et al. (2003) shows that symbolic representation like transformations can produce high accuracy rates while guarding against spurious noise in the waveforms. In other words, noise will not throw off the detector as easily using representation learning.

3. It is trivial (Gautam et al. 2019) to extend the similarity logic the detector uses (a $k$-nearest neighbor classifier) to an anomaly detection algorithm that discards outliers found by the detector. Introducing an anomaly detection is

vital to reduce the false alarm rate in an ever-expanding data set and grounds the detection system in "known space."

Given an initial training set of curves that are "of interest" versus curves that are "not of interest," the algorithm, when applied to targets of unknown interest class, will find new curves that will have a similar morphological shape with respect to the "of interest" training data. In other words, it will find new O'Connell effect binaries. The detector implements a $k$-nearest neighbor ($k$-NN; Altman 1992) classification method and an anomaly detection algorithm, allowing for direct control of the detector's constant false alarm rate (C-FAR) by imposing a maximum distance limit in parameter space. As the name would suggest, the C-FAR is just the constant rate at which false alarms occur. Direct control of the C-FAR is a desirable quality in a detection algorithm that can potentially be applied to ~1 million phased light curves. Johnston's detector limited the maximum parameter space distance for finding "of interest" systems to 75% of the maximum distance between systems in the training sample.

For this specific implementation, I manually selected an initial 22 targets of interest based on their phased light curves and 121 targets not of interest from the KEBC data set. The final run of the detector also included 8 systems found by earlier runs for a total of 30 targets of interest. The light curves of these 151 objects were then smoothed using Friedman's SUPERSMOOTHER (Friedman 1984) and then scaled to the flux range $[0, 1]$ using:

$$f(\phi)_N = \frac{f(\phi) - \min\left[f(\phi)\right]}{\max\left[f(\phi)\right] - \min\left[f(\phi)\right]} \tag{2.1}$$

where $f(\phi)$ is the smoothed, phased light curve. The detector then transformed

these light curves using a distribution field (DF; Helfer et al. 2015; Johnston et al. 2019), resulting in a symbolic, discretized representation of the phased light curve. A DF is defined as (Helfer et al. 2015; Sevilla-Lara & Learned-Miller 2012):

$$DF_{ij} = \frac{\sum_k^N [y_j < f(x_i \leq \phi_k \leq x_{i+1})_N < y_{j-1}]}{\sum_k^N [y_j < f(\phi_k)_N < y_{j-1}]} \tag{2.2}$$

where $N$ is the number of samples in the phased data, $\phi_k$ is the phase of the $k^{\text{th}}$ data point, $x_i$ and $y_j$ are the corresponding normalized amplitude and phase bins ($x_i = 0, 1/n_x, 2/n_x, ..., 1$, $y_i = 0, 1/n_y, 2/n_y, ..., 1$), respectively, $n_x$ is the number of time bins, $n_y$ is the number of amplitude bins, and [ ] is the Iverson bracket (Iverson 1962):

$$[P] = \begin{cases} 1 & P \text{ is true} \\ 0 & P \text{ is false} \end{cases} \tag{2.3}$$

The resulting DF is a matrix in which the rows sum to 1. These DFs represent the probability of a bin containing points that describe the light curve being analyzed and comprise the transformed data set used for the representation learning.

The bin numbers $n_x$ and $n_y$ and the number of nearest neighbors $k$ were optimized by cross-validation as part of the classification training process. The cross-validation procedure split the original 151 objects into a training group of 76 and a testing group of 75 and further split 70 objects in the training group into five subgroups of equal size. It then chose four subgroups to test $n_x$, $n_y$, and $k$ and comparing the results to the fifth, excluded subgroup. It performed five cycles for each combination of $n_x$, $n_y$, and $k$, with each cycle excluding a different subgroup. Out of the initial ranges $n_x, n_y \in [20, 40]$ and $k \in [1, 13]_{\text{odd}}$, Johnston determined $n_x = 25$, $n_y = 35$, and $k = 3$ were optimal.

Similarity is based on the distance between training and observation (the DFs $X$ and $Y$) defined as (Bellet et al. 2015; Helfer et al. 2015):

$$d(X,Y) = ||X - Y||_M^2 = \mathrm{tr}\{(X-Y)^T M (X-Y)\} \tag{2.4}$$

The closer in distance the two DFs are, the more similar they are, and the more the algorithm considers them members of the same class ("of interest" or "not of interest"). The matrix $M$ – the metric that is optimized – is based on the training data and the objective function defined in Johnston et al. (2019).

Based on the initial training data of 22 identified targets, this program identified an additional 144 potential targets. Of these 144, the detector's final run (Johnston & Haber 2019) found 129 – split into 8 unevenly-sized clusters – while an earlier run accounts for the remaining 15. The physical significance of the eight clusters is an open question, but systems in a given cluster tend to have similar light curves. Cluster six, for instance, contained all systems found by the detector where the maximum before the primary eclipse was brighter than the one following it. I only included 67 of the 144 potential targets in the sample after manually inspecting the light curves and finding that 77 did not meet my then-current criteria (which included a lower limit on the eclipse depth and an upper limit on period). The OES determination code discussed in Section 2.1.3.2 added 1 of the remaining 77 systems the detector found to the sample.

While the targets found were valid (~0 false detections based on the detector's scaled light curves), a cursory scan of the KEBC data set yielded numerous missed detections. Therefore, I did not utilize the detector to select my sample. Johnston and I attributed the missed detections to the small initial training data set mixed

with a highly diverse population of targets of interest, the fact that the desired characteristic (the OES) was removed by the light curve scaling, and a tuned low C-FAR. He designed the algorithm to be robust enough to handle and classify any type of binary, which Johnston et al. (2019) demonstrated by applying it to the LINEAR data set and discovering 24 eclipsing binaries. He also implemented a near-zero false alarm rate into the system by design because the detector is expected to be applied to high output (~1 million phased light curves) surveys where even a <1% FAR would result in thousands of mislabeled stars in the resulting sample.

#### 2.1.3.2 O'Connell Effect Size Determination

My final selection method directly determined the OES using an in-house Python code developed for this purpose. First, I downloaded the short- and long-cadence *Kepler* data for each eclipsing binary from the KEBC website. My code split this data into 1,000 bins in phase space for each system and cadence. The bins covered phase space over the range $\phi \in [-0.5, 0.5)$, with each equal-sized bin covering a phase of width 0.001, and the code assigned data points to the bin closest to the point's phase. It added an additional bin at phase 0.5 for symmetry – bringing the total number of bins to 1,001 – and included any points placed in the bins at either phase $-0.5$ or 0.5 in both bins. It then calculated each bin's weighted average flux $F$ using inverse-variance weighting (Hartung et al. 2008):

$$F = \frac{\sum_i f_i/\sigma_{f_i}^2}{\sum_i 1/\sigma_{f_i}^2} \tag{2.5}$$

where $f_i$ is the flux value of the $i^{\text{th}}$ data point and $\sigma_{f_i}$ is its error. The error $\sigma_F$ of this average is:

$$\sigma_F = \sqrt{\frac{1}{\sum_i 1/\sigma_{f_i}^2}} \tag{2.6}$$

The code produced a light curve from these averages, found the minimum value of this light curve, and introduced a phase offset to ensure this coincided with phase 0. It copied this light curve onto each side of the existing one to eliminate edge effects, extending the phase space range to $\phi \in [-1.5, 1.5]$. The resultant light curve represents an averaged light curve for the system, examples of which are shown in Figures 1.1 and 1.3. I removed all KEBC data from KIC 9164694 with BJD in the ranges [2455309.2870355, 2455336.8939506] and [2455432.2188883, 2455552.5485149] to correct an apparent data processing issue wherein the flux values during these intervals was systemically reduced by ~2%. I also removed a single outlier datum from KIC 8029708's data. Finally, 160 KEBC systems have the flag QAM (quarter amplitude mismatch), which indicates that these systems' amplitudes differ between *Kepler* quarters. Eighteen of these systems are in my sample. The KEBC corrected for this amplitude mismatch by rescaling the data to match between quarters, and I use this rescaled flux in my analysis. However, in at least one case (KIC 4474637), this rescaling did not correct the issue, and so this system exhibits significantly increased scatter in its data and, as a result, its averaged light curve.

The averaged light curve proved unsuitable for directly finding the OES due to data scatter. Therefore, the code produced a smoothed version of the light curve by convolving each point in the averaged light curve with the immediately adjacent points. The convolution reduces the effect of data scatter and oscillations on the light curve, giving a more accurate measure of the OES. Convolution in time

space is equivalent to multiplication in frequency space, making it computationally simpler to perform the convolution in frequency space (Press et al. 2007). My code accomplished this by applying a discrete Fourier transform (DFT, defined further in Section 2.2.2.2) to the system's averaged light curve using NumPy's (Oliphant 2006; Van Der Walt et al. 2011) `rfft` (real-valued fast Fourier transform, or real-valued FFT) function. It also defined a convolution kernel of window length $w = 3$. The kernel was an array the same length as the averaged light curve wherein all but $w$ elements were zero. The code arranged the $w$ elements – each equal to $1/w$ – such that $\lfloor w/2 \rfloor + 1$ elements lie at the beginning of the array and $\lfloor w/2 \rfloor$ lie at the end. It then applied `rfft` to the kernel and multiplied the transformed light curve with the transformed kernel 50 times (i.e. the kernel raised to the $50^{\text{th}}$ power, equivalent to iteratively applying the convolution 50 times in time space) on an element-by-element basis. Finally, it applied NumPy's `irfft` (inverse real-valued FFT) function to the resultant array, transforming the convolved curve back to time space. I determined the propagated error for each bin by performing the preceding procedure using a data array containing the errors given by Equation 2.6, although this changed the error values very little in practice, and the error is generally 2-3 orders of magnitude smaller than the OES regardless.

Finally, the code measured the maximum value of the convolved curve on both sides of the primary minimum in the phase range $\phi \in [-0.5, 0.5]$. The difference between the two maxima is the OES, which I define to be positive when the maximum after the primary eclipse is brighter than the one before it. A system with a positive OES is said to have a positive O'Connell effect. Conversely, a system where the maximum before the primary eclipse is brighter than the one following it has a negative OES and is said to have a negative O'Connell effect. The error

in OES is given by adding the propagated convolved error of each maximum in quadrature.

Because I determine the OES from the convolved curve - a derivative of the averaged light curve – this OES represents the average size of the O'Connell effect over the duration of *Kepler*'s observations. Averaging the light curves simplifies their analysis but reduces their utility in studying the temporally varying properties of the systems. I know that the non-averaged OES changes magnitude and even sign in several systems, as does the overall shape of the light curve. I will analyze the temporal variation in such systems separately by studying their eclipse timing variations (discussed in Section 2.2.4) and variance parameters (discussed in Section 2.2.6).

This method for finding O'Connell effect binaries is viable only because I applied it to a set of known eclipsing binaries. If I instead applied it to a general population of stars, it would find many non-eclipsing systems, such as pulsating variables and spotted stars. The KEBC therefore acts as a filter that allows this method to find O'Connell effect binaries. This method is also quite sensitive, and several systems near the |OES| cutoff of 0.01 would cross that threshold when I changed the method. For instance, when I began using Equation 2.6 instead of the standard definition of the mean, three systems with |OES| above the criterion threshold decreased to under it, while three systems under it increased to above it. This sensitivity to the precise method of determining the OES makes it more challenging to accurately reproduce the sample. It also explains why this method added systems to the sample that ultimately did not meet the criterion.

Figure 2.4 shows the averaged light curve of KIC 5700330, an Algol-type binary that is not in my sample. Also shown are three smoothed representations

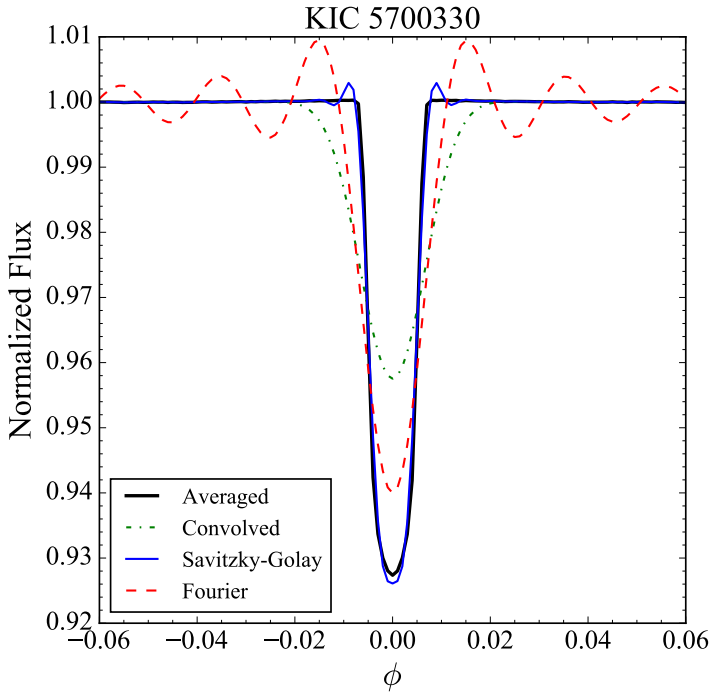


Figure 2.4: Averaged light curve (solid black) of KIC 5700330 (a non-sample Algol) showing three smoothed approximations of the light curve: the convolved curve (dot-dashed green), the Savitzky–Golay curve (solid blue), and the Fourier curve (dashed red).

of the light curve: the convolved curve, the Savitzky–Golay curve (discussed in Section 2.2.1), and the Fourier curve (discussed in Section 2.2.2.2). The convolved curve describes the out-of-eclipse variations well and shows little extraneous sinusoidal pattern, but it does not accurately describe the width or depth of the eclipse.

### 2.1.4 Incorporating KIC 7667885

*Kepler* observed KIC 7667885 in both the short and long cadences during Q14-17, but the KEBC does not include the system. Ramsay et al. (2014) discusses KIC 7667885 along with the core sample system KIC 9786165, and inspection of

KIC 7667885's short-cadence data from MAST[2] indicated that the system met my criterion and therefore should be included in my sample. Since KIC 7667885 is not in the KEBC, however, it did not have the detrending and phasing methods applied to it that other systems in the KEBC had. As such, I needed to perform these procedures myself.

I used the detrending process outlined in Section 4.2 of Slawson et al. (2011), which was the process used to produce the KEBC's photometric data. My code read in the raw *Kepler* data from MAST into a single NumPy array. It then extracted the relevant columns of data: the modified Barycentric Julian Date (BJD), the cadence number, the single-aperture photometry (SAP) flux and its error, and the pre-search data conditioning (PDC) SAP flux and its error. The modified BJD is BJD − 2454833.0. PDC detrends the raw SAP data, removing systematic errors such as pointing errors and focus changes (Jenkins et al. 2010). However, this procedure is optimized for detecting planetary transits and adversely affects the data for eclipsing binaries (Slawson et al. 2011). Therefore, my detrending process works solely with the SAP, and I include the PDCSAP flux only to remain consistent with the KEBC data format. The code removed any data points lacking SAP flux or PDCSAP flux and converted from modified BJD to BJD.

Next, the code divided the data into sections of continuous data. I took the cadence numbers corresponding to *Kepler*'s quarterly rolls from Van Cleve et al. (2016), and the code detected any further gaps by noting differences between the cadence numbers of successive data points corresponding to a time gap greater than a day. It detrended these continuous data sections independently by making

[2]Mikulski Archive for Space Telescopes; `http://archive.stsci.edu/kepler`

a least-squares fit to the data using Legendre polynomials up to an order $k$. I found that $k = 150$ (which I chose due to Slawson et al. 2011 showing it) produced a good fit for the KIC 7667885 data. The code then temporarily removes any data points that lie more than $3\sigma$ above the fit or $1\sigma$ below it. I chose this asymmetric $\sigma$ interval to keep consistent with Slawson et al. (2011), with the idea being that it will remove the eclipses that lie below the fit but keep more of the out of eclipse variations like flares that lie above it. The code then repeats the Legendre polynomial fitting and sigma clipping process, iterating until it removes no further points. The final Legendre polynomial fit represents the system trend, which the code removes by dividing the SAP flux and its error by the Legendre fit value at the time of that cadence. The code then divides the SAP flux, the PDCSAP flux, and the detrended flux by the median value of the respective flux in each time interval, then stitches the data together and exports it to a file. The output file's columns are: BJD, cadence number, SAP flux, SAP flux error, PDCSAP flux, PDCSAP flux error, detrended flux, and detrended flux error. I detrended KIC 7667885's short- and long-cadence data separately.

I also needed to find KIC 7667885's period, for which I used the software Peranso (Paunzen & Vanmunster 2016). Peranso is a GUI program designed to determine periodicity in variable astronomical sources. It allows the user to import time-series astronomical data and search for periodicity using methods including Deeming's implementation of DFT (Deeming 1975), Lomb-Scargle (Lomb 1976; Scargle 1982), phase dispersion minimization (Stellingwerf 1978), and analysis of variance (ANOVA; Schwarzenberg-Czerny 1989, 1996). I chose to use ANOVA as it is the method with which I am most familiar. ANOVA uses orthogonal polynomials

of order $2N$ defined as:

$$\begin{aligned}\Psi_{2N}(z) &= z^N F^{(N)}(t) \\ &= z^N \sum_{n=0}^{N} \left[a_n \cos\left(n\omega t\right) + b_n \sin\left(n\omega t\right)\right]\end{aligned} \tag{2.7}$$

where $z$ is a complex number and $F^{(N)}(t)$ is the $N$-term Fourier series. ANOVA uses these orthogonal polynomials to define the ANOVA statistic, which is the variance of the signal divided by the variance of the noise. This statistic is maximized when the test frequency $\omega$ equals the variability frequency.

Using ANOVA, I found a period $P$ of $0.314840 \pm 0.000004$ d for KIC 7667885. I also estimated a time of minimum $T_0$ of 2455000.087750 in BJD using my visually estimated time of minimum. Using this information, I found the orbital phase of the system for each data point using:

$$\phi = \frac{T - T_0}{P} \bmod 1 \tag{2.8}$$

where $T$ is the BJD of the observation. The phase overwrote the cadence number column in the data file, producing a file identical in structure to the KEBC data files. Figure 2.5 shows the resultant short- and long-cadence phased light curves for KIC 7667885.

I did not determine the morphology parameter for KIC 7667885 because I did not have access to the transformation Matijevič et al. (2012) used for the KEBC, and I could not reproduce the transformation without detailed knowledge of the KEBC subset they used to define it. I also did not perform an ETV analysis due to time constraints. Otherwise, I fully incorporated KIC 7667885 into the core

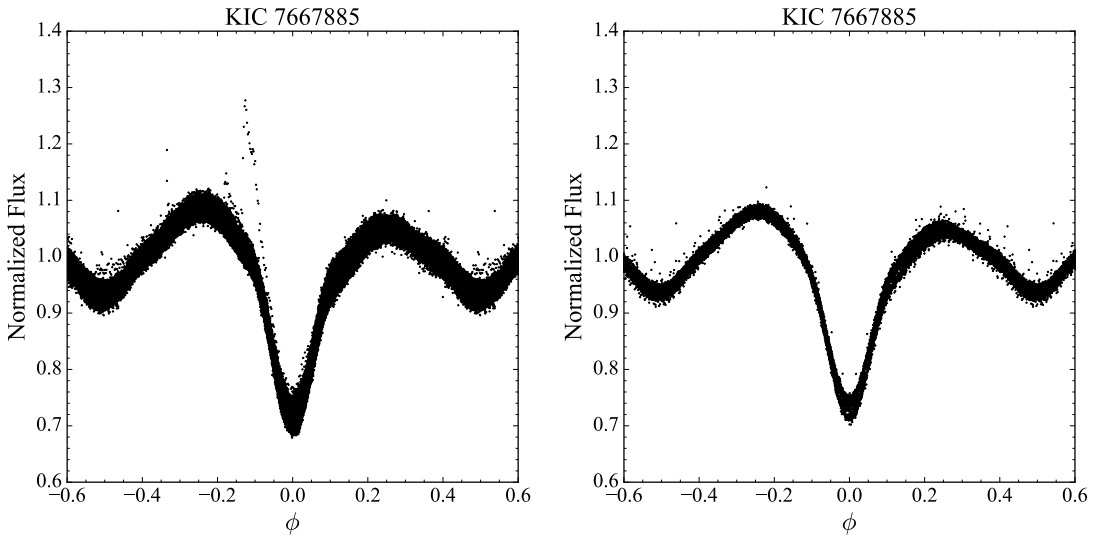


Figure 2.5: *Kepler* short-cadence (left) and long-cadence (right) light curves of KIC 7667885. The system displays a negative O'Connell effect and at least one significant flare event.

sample defined in the next section.

### 2.1.5 Target Sample

The final sample consists of 22 systems from the initial training set, 37 systems manually identified later, 67 systems identified by the automated detector, 129 systems identified by the OES measurement code, and 1 system (KIC 7667885) found in the literature, for a total of 258 systems. Of these 258 systems, 46 systems selected before I finalized the selection criterion and method have an average OES smaller than the cutoff. However, several of these systems show a larger O'Connell effect for short times. These systems constitute the marginal sample, which I describe separately in Section 4.3. The remaining 212 systems constitute the core sample. While I assembled my sample using different methods, the core sample consists only of the systems found by the OES determination code Section 2.1.3.2

described. The core sample systems represent 7.3% of the KEBC. Table A.1 lists all 258 systems in the complete sample.

The OES determination code found three additional systems that had an |OES| greater than the cutoff. Of these, KIC 7950964 is a duplicate entry of a system already in the sample (KIC 7950962), while KIC 9137819 appears to be an RR Lyrae variable and not an eclipsing binary at all. The KEBC team removed these systems from the KEBC after I consulted them (A. Prša, private communication). The final system, KIC 8456774, is a heartbeat star (Thompson et al. 2012), which are eccentric binaries that become tidally distorted at periapsis. KIC 8456774, like many heartbeat stars, does not appear to be eclipsing, and I do not consider it to have an O'Connell effect due to the fundamentally different origin of its asymmetry. Additionally, it is inappropriate to apply the phase offset I introduced to place the primary minimum at $\phi = 0$ to non-eclipsing systems like KIC 8456774, and the OES without this phase offset is negligible. I did not include these three systems in the sample, nor did I include systems like KIC 11560447 showing an OES above the cutoff in their short-cadence data but not in their long-cadence data.

One core sample system, KIC 7879399, lies only 4" from another *Kepler* system, KIC 7879404. As a result of their proximity, these two systems' data are blended, making it hard to determine which one is the eclipsing binary. The first two releases of the KEBC (Prša et al. 2011 and Slawson et al. 2011) identified KIC 7879404 as the binary, while the third release (Kirk et al. 2016 and Abdul-Masih et al. 2016) identified KIC 7879399 as the binary. Chuck Cynamon of the American Association of Variable Star Observers (AAVSO) observed these stars (C. Cynamon, private communication) for me in July 2021. I analyzed his data by performing photometry (described in Section 2.3.3) on each star separately, excluding as much of the other

star's light as possible using a judicious choice of annulus size and position. This analysis definitively showed that KIC 7879399 is the eclipsing binary, and so I used its *Kepler* and *Gaia* information in my analysis.

## 2.2 Characterization Methods

I needed to incorporate the means to analyze the sample once I defined it. I was interested in characteristics that were not readily available, including the eclipse depth, system luminosity, and two statistical measures introduced by McCartney (1999). I also wanted to determine how these characteristics related to each other and how the distributions differed between my sample and the KEBC.

### 2.2.1 Eclipse Depth Determination

The KEBC provides a value for the eclipse depth determined using the polyfit method described in Appendix A of Prša et al. (2008). The polyfit fits a chain of piecewise smooth $n^{\text{th}}$-order polynomials connected at a series of knots to fit the light curve. Using the given set of initial knots, the code performs a least-squares fit to the data on each interval defined by the knots while forcing the polynomials to connect at each knot and wrap phase space (i.e. $f(\phi = 0) = f(\phi = 1)$). It then calculates the sum of the squares of the residual before moving each knot slightly and repeating the procedure. At the end of the process, the code selects the polyfit with the lowest sum of residual squares. The KEBC uses four quadratic functions for the polyfit (Matijevič et al. 2012). The polyfit works for most systems in the KEBC but fails to provide an accurate measure of the eclipse depth for a small number of systems in my sample. Figure 2.6 demonstrates this inaccurate

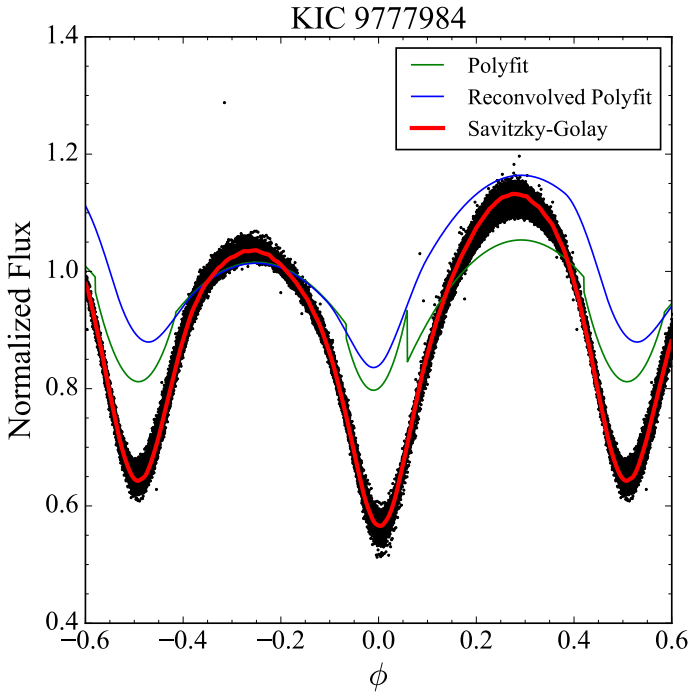


Figure 2.6: *Kepler* light curve of KIC 9777984 showing three fits to the system: the KEBC's original polyfit (green), the KEBC's reconvolved polyfit (blue) that corrects for *Kepler*'s non-zero integration time, and my fit using the Savitzky–Golay filter (red). The Savitzky–Golay curve better fits KIC 9777984's light curve and provides a more accurate measure of the eclipse depth for this system.

measurement with the polyfits of KIC 9777984, which Kirk et al. (2016) and the KEBC identified as KIC 9777987 but Abdul-Masih et al. (2016) reidentified as KIC 9777984. Furthermore, Kirk et al. (2016) explicitly states that the provided eclipse depths are approximate and only as accurate as the polyfits used to produce them.

Considering the polyfit issues for KIC 9777984 and other systems, I needed to define my own method for determining the eclipse depth. The convolved curve used to determine the OES consistently underestimates the eclipse depth, while the Fourier curve (described in Section 2.2.2.2) underestimates the eclipse depth in

Algol-type systems, so my code instead applied a Savitzky–Golay filter (Savitzky & Golay 1964) to the averaged light curve. The Savitzky–Golay filter smooths the data by fitting an $n^{\text{th}}$-order polynomial to $w$ adjacent points using the method of least squares. For my implementation, I used a window size of seven ($w = 7$) and quartic polynomials ($n = 4$). I chose this combination after testing $w \in [3, 15]_{\text{odd}}$ and $n \in [0, w - 3]_{\text{even}}$, with my choice of tested parameters guided by four facts about the Savitzky–Golay filter:

1. $w$ must be odd
2. $n$ must be less than $w$
3. The filter's output for $n$ and $n + 1$ is identical if $n$ is even
4. If $n = w - 1$, the filter's output is identical to its input

I found that $w = 7$ and $n = 4$ gave the best compromise between producing an accurate measure of the eclipse depth and not having such a large $n$ that it overfits the data.

My code applied the Savitzky–Golay filter in the same manner that it convolved the data in Section 2.1.3.2. The sole difference between the convolution and Savitzky–Golay implementations was in the kernel that I multiplied the data by in frequency space. With the Savitzky–Golay filter, my code used SciPy's (Virtanen et al. 2020) `savgol_coeffs` function, which gives the coefficients of the polynomial the filter applies to the data based on the values of $w$ and $n$. These coefficients are the non-zero kernel elements, which are arranged in the same manner as in the convolution kernel. The rest of the procedure was identical to my description in Section 2.1.3.2, including the error propagation. This process produced a

curve that approximated the eclipse depth better than the polyfits – as Figure 2.6 shows – despite retaining more of a sinusoidal pattern than the convolved curve – as Figure 2.4 shows. A limitation of this method is that Algol-type systems with narrow eclipses can have their eclipse depths underestimated by as much as 50%. Nevertheless, the Savitzky–Golay curve measures the eclipse depth more accurately than the Fourier curve, particularly for Algol-type systems like KIC 5700330 (Figure 2.4). The Savitzky–Golay curve is also more accurate than the convolved curve for all but a few edge cases involving totally eclipsing Algol-type systems.

Establishing the proper reference point from which to measure the depth is a non-trivial matter. The Villanova team that created the KEBC faces the same uncertainty and noted that modeling the light curve provided the only robust way of determining eclipse depth (A. Prša, private communication), which is an infeasible option in the time frame of this project considering the sample size. Neither maximum serves as a proper reference point because the process that causes the O'Connell effect affects at least one of them, and potentially both. Therefore, I decided to use a normalized flux value of one as the baseline flux value of each system. As explained in Section 2.1.4, Slawson et al. (2011) details the normalization scheme used for the KEBC data, and it is reasonable to assume that the data is generally median-normalized (A. Prša, private communication).

### 2.2.2 O'Connell Effect Ratio and Light Curve Asymmetry

#### 2.2.2.1 Definition

The O'Connell Effect Ratio (OER) and Light Curve Asymmetry (LCA) are two statistical measures defined by McCartney (1999), who used them with phased light curves broken into $\eta$ equally spaced phase bins. While originally introduced to characterize the O'Connell effect in W Ursae Majoris-type systems, they are valid statistics for all light curve classes. The original definitions found in McCartney (1999) are:

$$\mathrm{OER} = \frac{\sum_{i=1}^{\eta/2} I_i - I_0}{\sum_{i=(\eta/2)+1}^{\eta} I_i - I_0} \tag{2.9}$$

and:

$$\mathrm{LCA} = \sqrt{\sum_{i=1}^{\eta/2} \left( \frac{I_i - I_{\eta+1-i}}{I_i} \right)^2} \tag{2.10}$$

where $\eta$ is the number of phase bins, $I_i$ is the average intensity of the $i^{\mathrm{th}}$ bin, and $I_0$ is the minimum average intensity of the light curve. The LCA should contain a normalization factor to account for the number of bins, i.e.:

$$\mathrm{LCA} = \sqrt{\frac{1}{\eta} \sum_{i=1}^{\eta/2} \left( \frac{I_i - I_{\eta+1-i}}{I_i} \right)^2} \tag{2.11}$$

as in W09. W09 transformed these summations into integrals (taking $\eta \to \infty$), and the corresponding equations are:

$$\mathrm{OER} = \frac{\int_0^{1/2} [I(\phi) - I(0)] \, d\phi}{\int_{1/2}^{1} [I(\phi) - I(0)] \, d\phi} \tag{2.12}$$

and:

$$\mathrm{LCA} = \sqrt{\int_0^{1/2} \left( \frac{I(\phi) - I(1-\phi)}{I(\phi)} \right)^2 d\phi} \tag{2.13}$$

where $I(\phi)$ is an $N$-term Fourier series representation of the light curve:

$$I(\phi) = \frac{a_0}{2} + \sum_{n=1}^{N} \left[ a_n \cos\left( \frac{2\pi n \phi}{P} \right) + b_n \sin\left( \frac{2\pi n \phi}{P} \right) \right] \tag{2.14}$$

Since these calculations occur in phase space, $P = 1$ by definition, giving:

$$I(\phi) = \frac{a_0}{2} + \sum_{n=1}^{N} \left[ a_n \cos\left(2\pi n \phi\right) + b_n \sin\left(2\pi n \phi\right) \right] \tag{2.15}$$

I used the integral formulation for this project.

#### 2.2.2.2 Fourier Series Curve

I created a Fourier series representation of each system by applying a DFT to the system's averaged light curve using NumPy's `fft` (fast Fourier transform) function. A DFT like `fft` generates coefficients $X_n$ given by (Cochran et al. 1967):

$$X_n = \sum_{j=0}^{\eta-1} x_j e^{-\frac{2\pi i n j}{\eta}} \tag{2.16}$$

where $\eta$ is the number of data points (bins in this case) the DFT is applied to and $x_j$ are the data point values given by Equation 2.5. The inverse discrete Fourier transform is:

$$x_j = \frac{1}{\eta} \sum_{n=0}^{\eta-1} X_n e^{\frac{2\pi i n j}{\eta}} \tag{2.17}$$

The Fourier series coefficients $a_n$ and $b_n$ from Equation 2.15 can be found by reworking Equation 2.17 in the following manner. Since the data $x_j$ are real,

$x_j = \frac{1}{2}(x_j + x_j^*)$, and therefore:

$$x_j = \frac{1}{\eta}\sum_{n=0}^{\eta-1} X_n e^{\frac{2\pi i n j}{\eta}} \tag{2.18a}$$

$$= \frac{1}{2\eta}\sum_{n=0}^{\eta-1} \left(X_n e^{iw} + X_n^* e^{-iw}\right) \tag{2.18b}$$

$$= \frac{1}{4\eta}\sum_{n=0}^{\eta-1} \left(2X_n e^{iw} + 2X_n^* e^{-iw}\right) \tag{2.18c}$$

$$= \frac{1}{4\eta}\sum_{n=0}^{\eta-1} \left(X_n e^{iw} + X_n^* e^{iw} + X_n e^{-iw} + X_n^* e^{-iw} + X_n e^{iw} - X_n e^{-iw} + X_n^* e^{-iw} - X_n^* e^{iw}\right) \tag{2.18d}$$

$$= \frac{1}{\eta}\sum_{n=0}^{\eta-1} \left[\left(\frac{X_n + X_n^*}{2}\right)\left(\frac{e^{iw} + e^{-iw}}{2}\right) - \left(\frac{X_n - X_n^*}{2i}\right)\left(\frac{e^{iw} - e^{-iw}}{2i}\right)\right] \tag{2.18e}$$

$$= \frac{1}{\eta}\sum_{n=0}^{\eta-1} \left[\mathrm{Re}(X_n)\cos\left(\frac{2\pi n j}{\eta}\right) - \mathrm{Im}(X_n)\sin\left(\frac{2\pi n j}{\eta}\right)\right] \tag{2.18f}$$

where $w = \frac{2\pi n j}{\eta}$ and I have used the identities:

$$\mathrm{Re}(z) = \frac{1}{2}(z + z^*) \tag{2.19a}$$

$$\mathrm{Im}(z) = \frac{1}{2i}(z - z^*) \tag{2.19b}$$

$$\cos z = \frac{1}{2}\left(e^{iz} + e^{-iz}\right) \tag{2.19c}$$

$$\sin z = \frac{1}{2i}\left(e^{iz} - e^{-iz}\right) \tag{2.19d}$$

Defining $\phi = j/\eta$, $f(\phi) = x_j$, $a'_n = \mathrm{Re}(X_n)$, and $b'_n = -\mathrm{Im}(X_n)$ allows Equation 2.18f to be written as:

$$f(\phi) = \frac{1}{\eta}\sum_{n=0}^{\eta-1} \left[a'_n \cos(2\pi n\phi) + b'_n \sin(2\pi n\phi)\right] \tag{2.20}$$

There are obvious parallels between Equations 2.15 and 2.20, and it is natural to wonder if they are equivalent. To show that they can be made to be equivalent, I first note that they must have the same number of terms to be equal. Since a DFT cannot determine more coefficients than there are data points, I cannot use a Fourier series with more coefficients than $\eta - 1$. In other words, $N \leq \eta - 1$. Therefore, I set the upper bound on Equation 2.20 to $N$ and bring the $a'_0$ coefficient outside of the sum, giving:

$$f(\phi) = \frac{a'_0}{\eta} + \frac{1}{\eta}\sum_{n=1}^{N}\left[a'_n \cos\left(2\pi n\phi\right) + b'_n \sin\left(2\pi n\phi\right)\right] \tag{2.21}$$

Setting Equations 2.15 and 2.21 equal to each other gives:

$$\begin{aligned}\frac{a_0}{2} + \sum_{n=1}^{N}\left[a_n \cos\left(2\pi n\phi\right) + b_n \sin\left(2\pi n\phi\right)\right] \\ = \frac{a'_0}{\eta} + \frac{1}{\eta}\sum_{n=1}^{N}\left[a'_n \cos\left(2\pi n\phi\right) + b'_n \sin\left(2\pi n\phi\right)\right]\end{aligned} \tag{2.22}$$

I now postulate that $a_n \propto a'_n \ \forall\ n$ and $b_n \propto b'_n \ \forall\ n$, or $a_n = A_n a'_n$ and $b_n = B_n b'_n$ for some proportionality constants $A_n$ and $B_n$. Looking first at the $a_0$ term gives:

$$\frac{a_0}{2} = \frac{A_0}{\eta}a'_0 \Rightarrow a_0 = \frac{2A_0}{\eta}a'_0 = \frac{2A_0}{\eta}\mathrm{Re}(X_0) \tag{2.23}$$

From Equation 2.16:

$$X_0 = \sum_{j=1}^{\eta} x_j \tag{2.24}$$

where I have shifted the sum's limits without loss of generality. Therefore:

$$a_0 = \frac{2A_0}{\eta} \sum_{j=1}^{\eta} x_j \tag{2.25}$$

The $x_j$ values are simply $\eta$ samples of the function $f(\phi)$, so taking the limit as $\eta \to \infty$ converts Equation 2.25 into an integral:

$$a_0 = 2A_0 \int f(\phi)\ d\phi \tag{2.26}$$

Equation 2.26 is precisely the definition of the $a_0$ Fourier coefficient:

$$a_0 \equiv \frac{2}{P} \int f(\phi)\ d\phi \tag{2.27}$$

with period $P = 1$ and $A_0 = 1$. Therefore, Equation 2.23 implies:

$$a_0 = \frac{2\mathrm{Re}(X_0)}{\eta} \tag{2.28}$$

Looking next at the individual $a_n$ coefficients for $n > 0$ gives:

$$a_n \cos(2\pi n\phi) = \frac{A_n}{\eta} a'_n \cos(2\pi n\phi) \Rightarrow a_n = \frac{A_n}{\eta} a'_n = \frac{A_n}{\eta} \mathrm{Re}(X_n) \tag{2.29}$$

Substituting Equation 2.16 into Equation 2.29 gives:

$$a_n = \frac{A_n}{\eta} \mathrm{Re}\left( \sum_{j=1}^{\eta} x_j e^{-\frac{2\pi i n j}{\eta}} \right) \tag{2.30}$$

Recall:

$$e^{-\frac{2\pi i n j}{\eta}} = \cos\left(\frac{2\pi n j}{\eta}\right) - i \sin\left(\frac{2\pi n j}{\eta}\right) \tag{2.31}$$

Therefore, Equation 2.30 can be written as:

$$a_n = \frac{A_n}{\eta} \sum_{j=1}^{\eta} x_j \cos\left(\frac{2\pi n j}{\eta}\right) \tag{2.32}$$

Recalling the definition $\phi = j/\eta$, taking the limit as $\eta \to \infty$ turns Equation 2.32 into the integral:

$$a_n = A_n \int f(\phi) \cos\left(2\pi n \phi\right)\, d\phi \tag{2.33}$$

Again, Equation 2.33 is precisely the definition of the $a_n$ Fourier coefficient:

$$a_n \equiv \frac{2}{P} \int f(\phi) \cos\left(\frac{2\pi n \phi}{P}\right)\, d\phi \tag{2.34}$$

with $P = 1$ and $A_n = 2$. Therefore, Equation 2.29 implies:

$$a_n = \frac{2\mathrm{Re}(X_n)}{\eta} \tag{2.35}$$

which is identical in form to Equation 2.28. A similar process applied to the $b_n$ coefficients shows that $B_n = 2$ as well. Therefore, the general conversion equations are:

$$a_n = \frac{2\mathrm{Re}(X_n)}{\eta} \tag{2.36a}$$

$$b_n = -\frac{2\mathrm{Im}(X_n)}{\eta} \tag{2.36b}$$

Using Equations 2.36a and 2.36b, my code converted the complex-valued coefficients $X_n$ from `fft` into the Fourier series coefficients $a_n$ and $b_n$. It then evaluated the resultant function at the phases used in the averaged light curve to produce

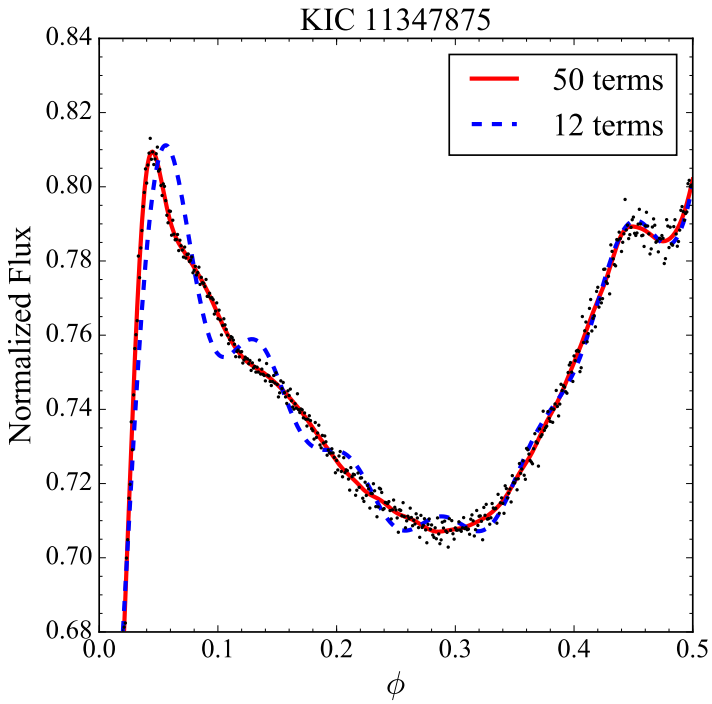


Figure 2.7: Averaged light curve of KIC 11347875 showing two Fourier series approximations to the system: the 50-term series I use (solid red) and a 12-term series (dashed blue) similar to one Akiba et al. (2019) would use. The 50-term series follows the averaged data more closely than the 12-term series while not capturing the statistical scatter.

the Fourier curve.

I chose to use a 50-term Fourier series based on a $\chi^2$ analysis of $N$ of the complete sample of 258 systems. My code calculated the $\chi^2$ statistic by comparing the Fourier curve with $N$ coefficients to the averaged light curve of the system, where $N \in [2, 500]$. This analysis indicated a median optimal $N$ of 59. I adopted an $N$ of 50 because I found that varying $N$ in the range $[40, 60]$ changes the results little. Figure 2.7 compares my 50-term Fourier series of the core sample system KIC 11347875 to a 12-term Fourier series like those used in Akiba et al. (2019) and Hahs et al. (2020). Figure 2.7 demonstrates that 50 terms are sufficient to capture

most features in the light curve without having so many terms that the series approximates small-scale statistical scatter. The 50-term Fourier series provides an excellent approximation of $\beta$ Lyrae- and W Ursae Majoris-type systems.

I also used the Fourier $a_n$ coefficients to quantify the light curve classification described in Section 1.1. My procedure is a modification of the one described in Section 3.2 of Akiba et al. (2019) and uses the following criteria to classify systems:

1. If $a_4 < a_2(0.125 - a_2)$, the system is classified as an Algol-type system.
2. If $a_4 \geq a_2(0.125 - a_2)$ and $|a_1/a_2| > 0.25$, the system is classified as a $\beta$ Lyrae-type system.
3. If $a_4 \geq a_2(0.125 - a_2)$ and $|a_1/a_2| \leq 0.25$, the system is classified as a W Ursae Majoris-type system.

Criterion 1 comes from Rucinski (1997) and is unchanged from Akiba et al. (2019), as are the first parts of criteria 2 and 3. However, the second parts of criteria 2 and 3 differ from Akiba et al. (2019), who instead used the criteria $|a_1| > 0.05$ to define $\beta$ Lyrae-type systems and $|a_1| < 0.05$ to define W Ursae Majoris-type systems. These criteria arise from W09's observation that $a_1$ measures the difference in eclipse depth. Unfortunately, $a_1$ does not contain any information about the eclipse depth *ratio*, which is the parameter related to the component temperature ratio (Kallrath & Milone 2009). I introduced $a_2$ (which W09 note is proportional to the light curve amplitude) to quantify the eclipse depth ratio.

The Fourier curve is not without issues. Its most prominent issue is that it struggles to fit Algol-type systems with sharp, narrow eclipses and minimal inter-eclipse variation, as Figure 2.4 shows. Another issue is that `fft`, like any DFT,

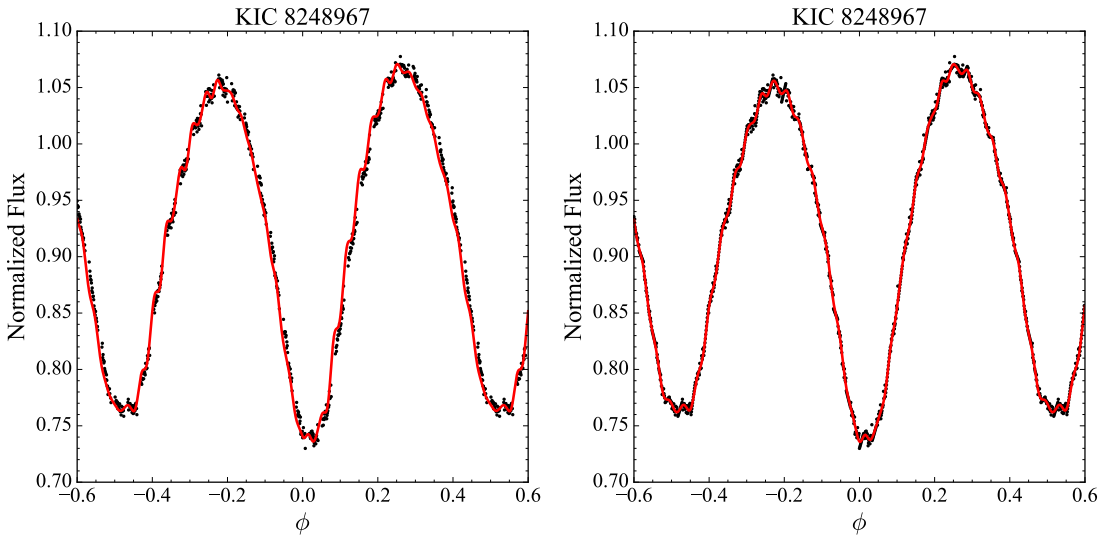


Figure 2.8: Averaged light curve of KIC 8248967 showing the system's uncorrected (left) and corrected (right) Fourier curves. The accuracy of the uncorrected Fourier curve is degraded by the missing data.

requires the data to be evenly spaced in time (Swan 1982)[3], meaning this procedure is only applicable to the averaged light curve, not the *Kepler* data itself. This issue presents a problem for any systems missing data at certain phases, which is most likely to occur because the orbital period is nearly resonant with the *Kepler* long-cadence integration time of 1,765.5 s. This issue is not merely academic, as the core sample system KIC 8248967 displays this phenomenon. The KEBC lists KIC 8248967 with a period of 0.5925091 d, or 51,193 s, 28.996 times the *Kepler* long-cadence integration time. Additionally, *Kepler* only observed KIC 8248967 for the last four months of its primary mission, which was insufficient time to get data at all phases due to the near-resonance of its period with the integration time. Figure 2.8's left panel shows the averaged light curve of KIC 8248967 along with its

[3]Swan (1982) does provide two methods to circumvent this limitation. I chose not to implement them because I could just apply `fft` to the averaged light curve.

Fourier curve. The regions following each eclipse have a "stairstep" pattern caused by the missing data. The oscillating feature most prominent near KIC 8248967's maxima and minima is another phenomenon associated with systems having an orbital period close to resonant with the *Kepler* cadence, which is seen even in systems like KIC 8696327 that have complete phase coverage. These oscillations are readily apparent in the averaged light curves of affected systems.

To mitigate the missing data issue, I filled in any missing data by linear interpolation. My code determined the phase $\phi_i$ of each bin containing no data by multiplying its bin number by the size of each bin. It then found the nearest bin at an earlier phase (point 1) and a later phase (point 2) containing data and determined their phases $\phi_1$ and $\phi_2$, their fluxes $F_1$ and $F_2$, and the error of their fluxes $\sigma_{F_1}$ and $\sigma_{F_2}$. It calculated the interpolated flux at $\phi_i$ by:

$$F_i = (\phi_i - \phi_1)\frac{F_2 - F_1}{\phi_2 - \phi_1} + F_1 \tag{2.37}$$

with error:

$$\sigma_{F_i} = \sqrt{\left(\frac{\phi_i - \phi_1}{\phi_2 - \phi_1}\right)^2 \left(\sigma_{F_1}^2 + \sigma_{F_2}^2\right) + \sigma_{F_1}^2} \tag{2.38}$$

Figure 2.8's right panel shows KIC 8248967's Fourier curve after interpolating missing data. While the oscillatory signal remains, the interpolation significantly reduces the "stairstep" pattern following each eclipse.

#### 2.2.2.3 Error Analysis

The error analysis for OER and LCA is complex, so I now give a detailed description of it. For readability, I use $\delta x$ in this section to indicate the error in $x$, whereas I use $\sigma_x$ in the rest of this dissertation. Throughout this section, I extensively use

the generalized error propagation equation (Bevington & Robinson 2003):

$$\delta f(x, y, ...) = \sqrt{\left(\frac{\partial f}{\partial x}\delta x\right)^2 + \left(\frac{\partial f}{\partial y}\delta y\right)^2 + ...} \tag{2.39}$$

I began by defining the errors of the Fourier coefficients $X_n$ given by Equation 2.16. The only variable with an error on the right-hand side of Equation 2.16 is $x_j$, so the variance of $X_n$ can be written as:

$$\mathrm{Var}\left(X_n\right) = \mathrm{Var}\left(\sum_{j=0}^{\eta-1} A_j x_j\right) \tag{2.40}$$

where $A_j = e^{-\frac{2\pi i n j}{\eta}}$ is a constant. Since $x_i$ and $x_j$ are uncorrelated for $i \neq j$, Equation 2.40 is equivalent to:

$$\mathrm{Var}\left(X_n\right) = \sum_{j=0}^{\eta-1} \mathrm{Var}\left(A_j x_j\right) \tag{2.41}$$

For complex constants such as $A_j$, the variance for $A_j x_j$ is not $A_j^2 \,\mathrm{Var}\left(x_j\right)$ as for real constants, but rather $A_j A_j^* \,\mathrm{Var}\left(x_j\right)$. Therefore:

$$\mathrm{Var}\left(X_n\right) = \sum_{j=0}^{\eta-1} A_j A_j^* \,\mathrm{Var}\left(x_j\right) \tag{2.42}$$

Since $A_j = e^{-\frac{2\pi i n j}{\eta}}$, $A_j A_j^* = 1$, while $\mathrm{Var}\left(x_j\right) = (\delta x_j)^2$. Therefore:

$$\mathrm{Var}\left(X_n\right) = \sum_{j=0}^{\eta-1} (\delta x_j)^2 \tag{2.43}$$

and:

$$\delta X_n = \sqrt{\sum_{j=0}^{\eta-1} (\delta x_j)^2} \tag{2.44}$$

Equation 2.44 indicates that the error in the $X_n$ coefficients is just the square root of the sum of squared errors of the data used to produce the Fourier transform, which are given by Equation 2.6. Note that Equation 2.44 has no dependence on $n$. I define the quantity $\delta X$ as:

$$\delta X = \frac{\sqrt{2}}{\eta} \sqrt{\sum_{j=0}^{\eta-1} (\delta x_j)^2} \tag{2.45}$$

for reasons that will become apparent in the next paragraph.

Next, I needed to find $\delta a_n$ and $\delta b_n$. Expanding Equation 2.36a using Equation 2.19a gives:

$$a_n = \frac{1}{\eta}(X_n + X_n^*) \tag{2.46}$$

Using Equation 2.39 gives Equation 2.46's error as:

$$\delta a_n = \frac{1}{\eta}\sqrt{(\delta X_n)^2 + (\delta X_n^*)^2} \tag{2.47}$$

While Equation 2.16 indicates that $X_n \neq X_n^*$, Equation 2.44's lack of an imaginary component implies that $\delta X_n = \delta X_n^*$. Therefore:

$$\delta a_n = \frac{1}{\eta}\sqrt{2(\delta X_n)^2} \tag{2.48}$$

or:

$$\delta a_n = \frac{\sqrt{2}}{\eta}\delta X_n \tag{2.49}$$

In other words, $\delta a_n = \delta X$ as defined by Equation 2.45. The same analysis shows that $\delta b_n = \delta X$ as well. Note that $\delta a_0$ is an exception to this relation due to $X_0$ being a real number instead of a complex number. Therefore, $X_0 = X_0^*$, and Equation 2.46 becomes:

$$a_0 = \frac{2}{\eta} X_0 \tag{2.50}$$

with an error given by:

$$\delta a_0 = \frac{2}{\eta} \delta X_0 \tag{2.51}$$

As a result, $\delta a_0 = \sqrt{2}\delta X$.

I next defined the error in the Fourier series $I(\phi)$. To do this, I used the method outlined in the appendices of Akiba et al. (2019), with help from one of the authors to correct mistakes in the published version (V. Gokhale, private communication). The corrected version of Equation B2 in Akiba et al. (2019) – modified to fit my nomenclature – is:

$$\delta I(\phi) = \sqrt{\left(\frac{\delta a_0}{2}\right)^2 + \sum_{n=1}^{N} \left\{ \left[\delta a_n \cos\left(2\pi n\phi\right)\right]^2 + \left[\delta b_n \sin\left(2\pi n\phi\right)\right]^2 \right\}} \tag{2.52}$$

Since $\delta X = \delta a_n = \delta b_n = \frac{1}{\sqrt{2}}\delta a_0$, the terms of Equation 2.52's sum become:

$$\begin{aligned} \left[\delta a_n \cos\left(2\pi n\phi\right)\right]^2 + \left[\delta b_n \sin\left(2\pi n\phi\right)\right]^2 &= (\delta X)^2 [\cos^2\left(2\pi n\phi\right) + \sin^2\left(2\pi n\phi\right)] \\ &= (\delta X)^2 \end{aligned} \tag{2.53}$$

and Equation 2.52 becomes:

$$\delta I(\phi) = \sqrt{\left(\frac{\delta X}{\sqrt{2}}\right)^2 + \sum_{n=1}^{N} (\delta X)^2} \tag{2.54}$$

$$= \delta X \sqrt{N + \frac{1}{2}} \tag{2.55}$$

This error is independent of $\phi$ and increases as more Fourier coefficients are added.

The error in OER is given by Equation D1 in Akiba et al. (2019):

$$\frac{\delta \mathrm{OER}}{\mathrm{OER}} = \sqrt{\left\{\frac{\delta \int_0^{1/2} [I(\phi) - I(0)] d\phi}{\int_0^{1/2} [I(\phi) - I(0)] d\phi}\right\}^2 + \left\{\frac{\delta \int_{1/2}^{1} [I(\phi) - I(0)] d\phi}{\int_{1/2}^{1} [I(\phi) - I(0)] d\phi}\right\}^2} \tag{2.56}$$

The path to $\delta$OER begins with Equation 2.12. The numerator and denominator of Equation 2.12 are of the form:

$$\int_b^a [I(\phi) - I(0)] \, d\phi \tag{2.57}$$

I define a function $H_1(\phi)$ such that $H_1(\phi) \equiv \int I(\phi) \, d\phi$:

$$H_1(\phi) = \frac{a_0 \phi}{2} + \sum_{n=1}^{N} \left[\frac{a_n}{2\pi n} \sin(2\pi n \phi) - \frac{b_n}{2\pi n} \cos(2\pi n \phi)\right] \tag{2.58}$$

where I have suppressed the constant of integration. I also define $H_0(\phi)$ such that $H_0(\phi) \equiv \int I(0) \, d\phi$, giving:

$$H_0(\phi) = \frac{a_0 \phi}{2} + \sum_{n=1}^{N} a_n \phi \tag{2.59}$$

again suppressing the constant of integration. Combining Equations 2.58 and 2.59

gives the function $H(\phi)$:

$$H(\phi) \equiv H_1(\phi) - H_0(\phi) = \sum_{n=1}^{N} \left[ \frac{a_n}{2\pi n} \sin(2\pi n\phi) - \frac{b_n}{2\pi n} \cos(2\pi n\phi) - a_n\phi \right] \quad (2.60)$$

$H(\phi)$ is the antiderivative of Expression 2.57 with the constant of integration suppressed. Applying Equation 2.39's partial derivatives to Equation 2.60 gives:

$$\frac{\partial H(\phi)}{\partial a_0} = 0 \quad (2.61a)$$

$$\frac{\partial H(\phi)}{\partial a_n} = \frac{1}{2\pi n} \sin(2\pi n\phi) - \phi \quad (2.61b)$$

$$\frac{\partial H(\phi)}{\partial b_n} = -\frac{1}{2\pi n} \cos(2\pi n\phi) \quad (2.61c)$$

for all $n > 0$. Therefore, applying Equation 2.39 to Equation 2.60 gives:

$$\delta H(\phi) = \delta X \sqrt{N\phi^2 + \sum_{n=1}^{N} \left[ \left( \frac{1}{2\pi n} \right)^2 - \frac{\phi}{\pi n} \sin(2\pi n\phi) \right]} \quad (2.62)$$

by analogy with Equation 2.55. However, unlike Equation 2.55, Equation 2.62 depends on $\phi$. I note that the limits of integration for OER are $\phi = 0, \frac{1}{2}$, and 1. For these values of $\phi$, $\sin(2\pi n\phi) = 0 \; \forall \; n$, so Equation 2.62 can be simplified to:

$$\delta H(\phi) = \delta X \sqrt{N\phi^2 + \sum_{n=1}^{N} \left( \frac{1}{2\pi n} \right)^2} \quad (2.63)$$

for the purposes of finding $\delta$OER. The error for Expression 2.57 is then:

$$\delta \int_b^a [I(\phi) - I(0)] \, d\phi = \delta H(\phi)|_b^a = \delta X \sqrt{Na^2 + Nb^2 + 2\sum_{n=1}^{N} \left( \frac{1}{2\pi n} \right)^2} \quad (2.64)$$

and the error in OER is:

$$\delta\text{OER} = \text{OER}\frac{\delta X}{2}\sqrt{\frac{N + 2\sum_{n=1}^{N}\left(\frac{1}{\pi n}\right)^2}{\left[H\left(\frac{1}{2}\right) - H(0)\right]^2} + \frac{5N + 2\sum_{n=1}^{N}\left(\frac{1}{\pi n}\right)^2}{\left[H(1) - H\left(\frac{1}{2}\right)\right]^2}} \tag{2.65}$$

by Equation 2.56.

The error analysis for LCA is more complicated. Like Akiba et al. (2019), I define several functions to simplify Equation 2.13, starting with $J(\phi)$:

$$J(\phi) = I(\phi) - I(1 - \phi) \tag{2.66}$$

then $K(\phi)$:

$$K(\phi) = \frac{J(\phi)}{I(\phi)} \tag{2.67}$$

and finally $L(\phi)$:

$$L(\phi) = [K(\phi)]^2 \tag{2.68}$$

Since $I(1-\phi)$ is just $I(\phi)$ mirrored about the line $\phi = \frac{1}{2}$, it is clear that $\delta I(1-\phi) = \delta I(\phi)$ and $\delta J(\phi) = \sqrt{2[\delta I(\phi)]^2}$, or:

$$\delta J(\phi) = \delta X\sqrt{2\left(N + \frac{1}{2}\right)} \tag{2.69}$$

using Equation 2.55. The error in Equation 2.67 is:

$$\delta K(\phi) = K(\phi)\sqrt{\left(\frac{\delta J(\phi)}{J(\phi)}\right)^2 + \left(\frac{\delta I(\phi)}{I(\phi)}\right)^2} \tag{2.70}$$

while the error in Equation 2.68 is:

$$\delta L(\phi) = 2K(\phi)[\delta K(\phi)] \tag{2.71}$$

Since LCA $= \sqrt{\int_0^{1/2} L(\phi)\ d\phi}$, $\delta$LCA $= \delta\sqrt{\int_0^{1/2} L(\phi)\ d\phi}$, and using Equation 2.39 gives:

$$\delta\sqrt{\int_0^{1/2} L(\phi)\ d\phi} = \frac{\delta \int_0^{1/2} L(\phi)\ d\phi}{2\sqrt{\int_0^{1/2} L(\phi)\ d\phi}} \tag{2.72}$$

Unfortunately, $\int_0^{1/2} L(\phi)\ d\phi$ is intractable, so I make the same approximation that Akiba et al. (2019) does:

$$\delta \int_0^{1/2} L(\phi)\ d\phi\ \approx \int_0^{1/2} \delta L(\phi)\ d\phi \tag{2.73}$$

Under this approximation and combining Equations 2.55 and 2.66-2.72, the error in LCA is:

$$\begin{aligned}\delta\text{LCA} \approx \frac{\delta X}{\text{LCA}} \int_0^{1/2} \Bigg[ &\left(\frac{I(\phi) - I(1-\phi)}{I(\phi)}\right)^2 \\ &\times \sqrt{\frac{2\left(N+\frac{1}{2}\right)}{\left[I(\phi) - I(1-\phi)\right]^2} + \frac{N+\frac{1}{2}}{I(\phi)^2}} \Bigg] d\phi\end{aligned} \tag{2.74}$$

### 2.2.3 Morphology Parameter

Matijevič et al. (2012) introduced the morphology parameter (which I represent with the symbol $\mu$ throughout this dissertation) to determine the morphology class of KEBC systems using the light curve. Bódi & Hajdu (2021) recently performed the same procedure on the OGLE database of eclipsing binaries (Soszyński et al. 2016). Matijevič et al. (2012) selected a set of 1,000 equally spaced samples

from the polyfit for each KEBC system and applied the method of local linear embedding (Roweis & Saul 2000) on the resultant data set. They represented each system by a point in 1,000-dimensional phase space. Their code characterized the local environment of each point by taking a linear combination of the $k$-NN. Their goal was to reduce the 1,000-dimensional space to a more easily visualized three-dimensional space by determining the weights that minimized the difference between the weighted cost function and the system's position in 1,000-dimensional space. They further reduced the dimensionality of their data set to a two-dimensional space by repeating this procedure. A spline fit on the resulting curve transformed it into a one-dimensional parameter called the morphology parameter. They used a subset of 1,572 KEBC systems to compute the transformation, but they mapped all 2,920 KEBC entries using this transformation. They assigned a value of $\mu = -1$ to systems that this procedure could not assign a valid $\mu$ to.

The morphology parameter correlated very well with their manually identified morphological classification for most KEBC systems. They determined that $\mu \leq 0.5$ implied detached systems, $0.5 < \mu \leq 0.7$ implied semi-detached systems, $0.7 < \mu \leq 0.8$ implied overcontact systems, and $\mu > 0.8$ implied ellipsoidal variables (non-eclipsing systems where the variation is due to the changing aspect of tidally distorted stars) as well as systems with uncertain classification. They note, however, that these are a "best guess" at the morphological classification and that system modeling may be required to accurately determine the correct morphological class.

Based on visual inspection of the systems in the sample, I estimated that Algol-type systems have $\mu < 0.6$, $\beta$ Lyrae-type systems have $0.6 \leq \mu < 0.75$, and W

Ursae Majoris-type systems have $\mu \geq 0.75$. As with Matijevič et al. (2012), these are not firm boundaries and there is overlap between classes. These boundaries agree with those in Matijevič et al. (2012) as Algol-type systems are generally detached or semi-detached, $\beta$ Lyrae-type systems can be of any morphological class, and W Ursae Majoris-type systems are predominantly overcontact, although detached and semi-detached eclipsing binaries, as well as ellipsoidal variables, may also resemble W Ursae Majoris-type systems.

The sample likely contains several ellipsoidal variables. It is difficult to differentiate between an ellipsoidal variable and an overcontact system due to their similar light curves. Preliminary modeling of four systems in my sample with PHOEBE v0.31a (Prša & Zwitter 2005) shows that the morphology parameter is not a completely effective way to identify ellipsoidal variables, either. As an example, KIC 8285349 has $\mu = 0.90$ but clearly shows an eclipse. Meanwhile, KICs 10815379 and 10979669 have $\mu = 0.79$ and 0.83, respectively, but are non-eclipsing according to my preliminary models. Identifying and removing ellipsoidal variables would therefore require large-scale modeling of the sample systems, which is beyond the scope of my project.

### 2.2.4 Eclipse Timing Variation

Eclipse timing variation (ETV, also known as observed − calculated or O − C) compares the observed time of an eclipse minimum to the predicted time of that minimum calculated from a known previous eclipse and the binary period. As Conroy et al. (2014) explains, differences between the observed and calculated times of minimum can reveal several processes occurring in the system. A third body in the system creates a light travel time effect (LTTE) as the three bodies

orbit their barycenter, causing an apparent change in the eclipsing binary's period. A third body, if close enough, can also cause a genuine change in the binary period through gravitational perturbations. Mass transfer between the binary components can increase or decrease the orbital period while mass loss increases it, which would be apparent by the increasing difference between the observed and calculated eclipse times. Eccentric systems undergo apsidal motion that changes the eclipse timings as the orbit's orientation changes. Coupling of the stars' gravitational quadrupoles exchanges angular momentum between the binary orbit and stellar rotation via the Applegate mechanism (Applegate 1992). Finally, anticorrelated primary and secondary eclipse ETVs can be an indicator of starspot activity, as I discuss in more detail in Section 3.2.

For this dissertation, I use the results of Conroy et al. (2014), which describes the ETVs of KEBC systems with $0.5 < \mu \leq 1.0$. They processed the *Kepler* data in the manner described in Section 2.1.4 and then phased using the ephemeris reported by Kirk et al. (2016). To measure each eclipse's ETV, Conroy et al. (2014) shifted the polyfits described in Section 2.2.1 in phase by $\pm 0.05$ and measured the $\chi^2$ at 20 evenly spaced points in that interval. They chose the point with the lowest $\chi^2$ as the center point of a bisection algorithm that determined the phase offset that produced the minimum $\chi^2$. The time that this phase offset represented served as the ETV of that eclipse.

After finding the ETV for all observed eclipses, Conroy et al. (2014) differentiated between a parabolic signal (as caused by mass transfer or the Applegate mechanism) and a long-period sinusoidal signal (as caused by a long-period third

body) using the Bayesian Information Criterion (BIC; Schwarz 1978):

$$\mathrm{BIC} = n \ln \left[ \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{x}_i)^2 \right] + k \ln n \tag{2.75}$$

where $n$ is the number of data points $x_i$, $x_i$ is the observed ETV for eclipse $i$, $\hat{x}_i$ is the ETV for eclipse $i$ predicted by one of three models (mass transfer, circular LTTE, and eccentric LTTE), and:

$$k = \begin{cases} 3 & \text{for the mass transfer model} \\ 4 & \text{for the circular LTTE model} \\ 6 & \text{for the eccentric LTTE model} \end{cases} \tag{2.76}$$

They calculated the BIC for all three models and chose the model with the lowest BIC as the most accurate fit to the observed data.

The KEBC categorizes each system using Conroy et al.'s (2014) results into one of four categories and flags them accordingly:

1. Systems with no apparent ETV pattern remain unflagged.

2. Systems showing an LTTE effect have the flag TM, indicating they are potential multiples.

3. Systems showing a parabolic ETV have the flag TT, indicating possible mass transfer.

4. Systems showing an ETV pattern that does not fall under either previous category have the flag TI, indicating an "interesting" ETV.

To further investigate the ETVs, I applied a digital Butterworth filter (But-

terworth 1930) of order one to the ETV data using SciPy's `butter` function to reduce noise from effects such as starspots. The sampling frequency is the orbital frequency, and I set the cutoff frequency to five cycles over the system's observation span, thus removing signals such as starspot modulation that recur more than five times during *Kepler*'s observations. Due to the need for continuous sampling, I removed any system with an observation gap longer than 20 d. I fit a parabola to each system's filtered data using SciPy's `curve_fit` function and calculated the $R^2$ value of the fit to test if the system's ETV could indicate mass transfer. I considered systems to have a roughly parabolic ETV (i.e. potentially indicating mass transfer) if $R^2 \geq 0.8$.

### 2.2.5 Statistical Analysis

I wanted to compare the characteristics of my sample with the characteristics of the entire KEBC. In doing this, I discover which characteristics change when looking at the subset of O'Connell effect systems – indicating a possible relationship between that characteristic and the presence of an O'Connell effect – and which characteristics do not change – indicating the characteristic is unrelated to the O'Connell effect. To determine this, I used the two-sample Kolmogorov–Smirnov test (K–S test; Kolmogorov 1933; Smirnov 1948), which compares the cumulative distribution function of two sample distributions $F_{1,m}(x)$ and $F_{2,n}(x)$ by defining the K–S statistic:

$$D_{m,n} = \sup |F_{1,m}(x) - F_{2,n}(x)| \tag{2.77}$$

where sup is the supremum function. The K–S test rejects the null hypothesis that the two samples are drawn from the same distribution at a level $\alpha$ if (Knuth 1997):

$$D_{m,n} > \sqrt{-\frac{m+n}{2mn} \ln \frac{\alpha}{2}} \tag{2.78}$$

where $m$ and $n$ are the sizes of the first and second samples, respectively. The K–S statistic describes the degree of difference between the populations the two samples were drawn from, while its $p$-value indicates how compatible the K–S test results are with the null hypothesis that the two samples were drawn from the same population.

I also wanted to compare the characteristics within a sample to each other, which allows me to look for correlations or anti-correlations that indicate a relationship between the physical mechanisms underlying these characteristics. Like D84, I used Spearman's rank correlation coefficient (Spearman 1904) which is also sensitive to non-linear correlations, unlike the Pearson correlation coefficient used in O51. The definition of Spearman's rank correlation coefficient $\rho$ is:

$$\rho = \frac{\mathrm{cov}(r_x, r_y)}{\sigma_{r_x} \sigma_{r_y}} \tag{2.79}$$

where $r_x$ and $r_y$ are the rank variables of the distributions of interest $x$ and $y$, respectively, and cov is the covariance. The ranks are assigned by ordering $x_i$ and $y_i$ from smallest to largest and assigning ranks based on that ordering. The form of Equation 2.79 is identical in form to the definition of the Pearson correlation coefficient except applied to the rank variable instead of the $x_i$ and $y_i$ data directly. The $\rho$ coefficient describes the correlation's strength, while its $p$-value indicates how compatible the Spearman test's results are with the null hypothesis that the

characteristics are uncorrelated.

I also used Kendall's rank correlation coefficient (Kendall 1938) as a second test of correlation. Given a set of $n$ ordered pairs $(x_n, y_n)$, check the pair of data points $(x_i, y_i)$ and $(x_j, y_j)$. If both $x_i > x_j$ and $y_i > y_j$ hold or both $x_i < x_j$ and $y_i < y_j$ hold, the pair is said to be concordant. Otherwise, the pair is said to be discordant. The definition of Kendall's rank correlation coefficient $\tau$ is:

$$\tau = \frac{n_c - n_d}{n_0} \tag{2.80}$$

where $n_c$ is the number of concordant pairs, $n_d$ is the number of discordant pairs, and $n_0 = \frac{n(n-1)}{2}$ is the binomial coefficient $\binom{n}{2}$. Equation 2.80 holds only when there are no ties (i.e. $x_i \neq x_j$ and $y_i \neq y_j \ \forall \ i, j$), however. I chose to use the so-called $\tau$-b formulation (Agresti 2010) to handle ties:

$$\tau_b = \frac{n_c - n_d}{\sqrt{(n_0 - n_1)(n_0 - n_2)}} \tag{2.81}$$

where $n_1 = \sum_i \frac{t_i(t_i-1)}{2}$, $n_2 = \sum_j \frac{u_j(u_j-1)}{2}$, $t_i$ is the number of ties among the $x$ variable, and $u_j$ is the number of ties among the $y$ variable. Like the $\rho$ coefficient, the $\tau$ coefficient describes the correlation's strength, while its $p$-value indicates how compatible the Kendall test's results are with the null hypothesis.

I created a code that used the SciPy functions `ks_2samp`, `spearmanr`, and `kendalltau` to implement the K–S test, Spearman's rank correlation coefficient, and Kendall's rank correlation coefficient, respectively. My code applied these three functions to data representing numerous characteristics about the KEBC

and my target sample. Some data came from online sources such as SIMBAD[4] and VizieR[5], while other data came from the analysis I have described throughout this chapter. The code sorted this data into NumPy arrays while keeping track of the data locations using Python dictionaries. Settings at the beginning of the code allow me to control parameters such as which data sets to include or whether to take the absolute value of the OES. The code exports the results of these statistical tests to a human-readable file.

In order to test the robustness of my correlation analysis results, I performed a "bootstrapping" procedure on the sample. My code randomly selected 20 subsets of 40 systems from the core sample (and 400 systems from the KEBC) and computed the Spearman's $\rho$ coefficients for each subset. I considered a correlation robust if at least 19 of the 20 subsets (or 95% of the subsets) recovered the correlation because the 95% value closely approximates the $2\sigma$ confidence interval.

My code also determined the functional form of the correlations using SciPy's implementation of orthogonal distance regression (ODR; Boggs et al. 1987), which differs from ordinary least squares in that it accounts for the error in the $x$ variable too. Ordinary least squares assumes that the data satisfies a mathematical model of form:

$$y_i = f(x_i, \hat{\beta}) + \sigma_{y_i} \tag{2.82}$$

where $\sigma_{y_i}$ are the errors in the dependent variable, $\hat{\beta}$ is a set of parameters to be estimated, and $f(x_i, \hat{\beta})$ is a smooth but not necessarily linear function. The least

---

[4] `https://simbad.u-strasbg.fr/simbad/`

[5] `https://vizier.u-strasbg.fr/viz-bin/VizieR`

squares problem seeks to minimize the expression:

$$\min \sum_{i=1}^{n} [y_i - f(x_i, \hat{\beta})]^2 \tag{2.83}$$

By contrast, ODR also allows errors $\sigma_{x_i}$ on the independent variable as well, which changes Equation 2.82 to:

$$y_i = f(x_i + \sigma_{x_i}, \hat{\beta}) + \sigma_{y_i} \tag{2.84}$$

and Expression 2.83 to:

$$\min \sum_{i=1}^{n} \{[y_i - f(x_i + \sigma_{x_i}, \hat{\beta})]^2 + \sigma_{x_i}^2\} \tag{2.85}$$

The code attempted to fit six different functions to characteristic pairs that had a Spearman's $|\rho| \geq 0.1$: linear ($y = ax+b$), quadratic ($y = ax^2+bx+c$), exponential ($y = ae^{bx} + c$), logarithmic ($y = a \log(x + b) + c$), power law ($y = ax^b + c$), and inverse ($y = \frac{a}{x+b} + c$). It attempted two fits for each pair of characteristics for all six equations, one for each combination of abscissa and ordinate.

### 2.2.6 Variance Parameter

Figure 2.9's left panel shows the strong temporal variation in KIC 7433513's *Kepler* data. I created the variance parameter to quantify this temporal variation. More specifically, I created two such parameters: the unscaled variance parameter (UVP) and the scaled variance parameter (SVP). The SVP is my quantification parameter. To calculate these parameters, I wrote a dedicated code.

The code begins by creating a full-data averaged light curve in the same fashion

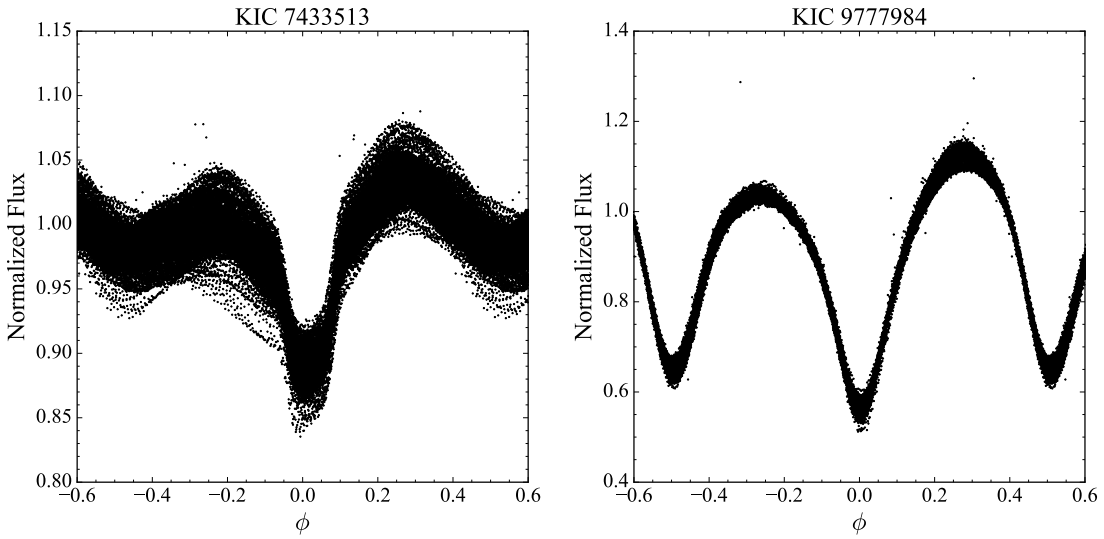


Figure 2.9: *Kepler* light curves of KIC 7433513 (left) and KIC 9777984 (right) showing the temporal variation in both systems.

as described in Section 2.1.3.2, but using $\eta = 51$ bins instead of 1,001 due to using fewer data points in subsequent steps. It then defines the number of data points $n$ that will constitute an arbitrary time interval $I_j$. The number of points $n$ equals four times either the number of data points in one orbital period or 50, whichever is larger. As such, $n$, $I_j$, and the total number of intervals $I$ vary between systems. The code imports the *Kepler* data again and separates sets of $n$ consecutive points into the time intervals $I_j$. It uses the $n$ points to create a 51-bin averaged light curve for each $I_j$, again in the same manner as Section 2.1.3.2. For a given $I_j$ and each bin $i$, the code calculates the absolute value of the flux difference between the full-data averaged light curve and the time interval-specific averaged light curve:

$$d_{I_{ij}} = |F_{0_i} - F_{I_{ij}}| \tag{2.86}$$

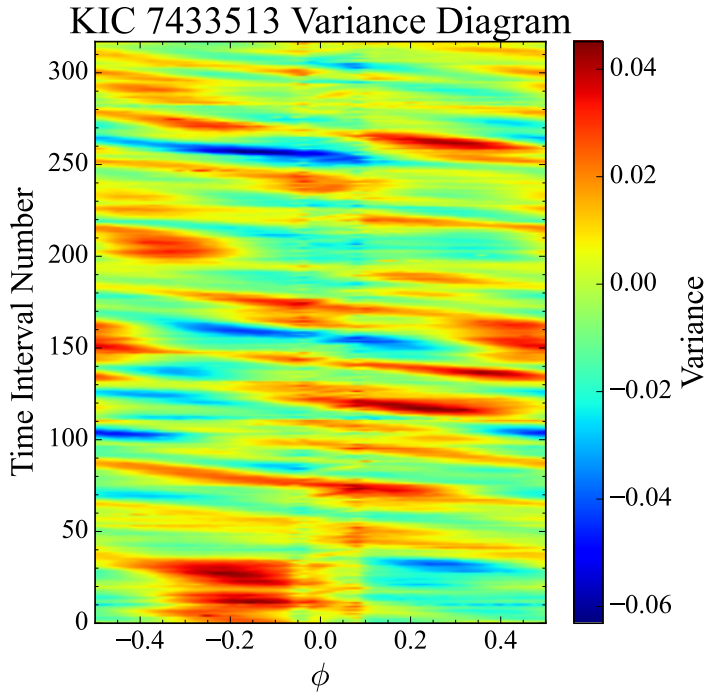


Figure 2.10: Variance diagram of KIC 7433513 showing the variance (color axis) as a function of phase and time interval number. The UVP is the average of the absolute value of the $\phi$-$I_j$ plane. KIC 7433513's time intervals are 4.087 d long.

It then calculates the average flux difference for that $I_j$:

$$\bar{d}_{I_j} = \frac{1}{\eta_{I_j}} \sum_{i=0}^{\eta_{I_j}} d_{I_{ij}} \tag{2.87}$$

where $\eta_{I_j} = 51 - N$ and $N$ is the number of bins containing no data in $I_j$. Finally, the code calculates the overall average flux difference for the system by finding the average of all time intervals' average flux differences:

$$\bar{d} = \frac{1}{I} \sum_{j=0}^{I} \bar{d}_{I_j} \tag{2.88}$$

This overall average flux difference $\bar{d}$ is the system's UVP. Figure 2.10, showing

KIC 7433513's variance diagram, illustrates the UVP. Figure 2.10 plots the variance (Equation 2.86 sans the absolute value), represented by the color axis, against the phase $\phi$ and the time interval number $I_j$. Equation 2.88 states that the UVP is the average of the absolute value of the $\phi$-$I_j$ plane shown in Figure 2.10.

The code also calculates the UVP error by propagating Equations 2.86-2.88 using Equation 2.39. Equation 2.86's error is:

$$\sigma_{d_{I_{ij}}} = \sqrt{\sigma^2_{F_{0_i}} + \sigma^2_{F_{I_{ij}}}} \tag{2.89}$$

while Equation 2.87's error is:

$$\sigma_{\bar{d}_{I_j}} = \frac{1}{\eta_{I_j}} \sqrt{\sum_{i=0}^{\eta_{I_j}} \sigma^2_{d_{I_{ij}}}} \tag{2.90}$$

Finally, the error in UVP is:

$$\sigma_{\mathrm{UVP}} = \frac{1}{I} \sqrt{\sum_{j=0}^{I} \sigma^2_{\bar{d}_{I_j}}} \tag{2.91}$$

In practice, the UVP error is insignificant enough for most systems to be zero at the number of significant figures I give.

Unfortunately, the UVP does not provide a good measure of the degree of temporal variation in the light curve. For instance, Figure 2.9's right panel shows the light curve of KIC 9777984, which indicates that it has a large amplitude but only modest temporal variation. However, KIC 9777984's UVP (0.0119) is slightly *larger* than KIC 7433513's (0.0116) despite the latter's much more pronounced temporal variation. This unexpected result stems from the amplitude difference

between the two systems. My solution to this problem was to scale the *Kepler* data such that the full-data averaged light curve spanned exactly the flux range $[0, 1]$. To do this, the code uses the full-data averaged light curve's maximum flux value $F_{max}$ and minimum flux value $F_{min}$ to calculate the scaling factor $S$:

$$S = \frac{1}{F_{max} - F_{min}} \tag{2.92}$$

It then multiplies the UVP and its error by this scaling factor, giving the SVP and its error. The SVP error is again insignificant for most systems.

This scaling removes the dependence on a system's amplitude, allowing me to directly compare two systems' variance parameters. KIC 9777984's SVP (0.022) is significantly smaller than KIC 7433513's (0.080), which properly reflects the degree of temporal variation observed in each. Furthermore, the SVP accurately represents the degree of variation seen in my sample. In other words, systems with a small SVP have the most temporally stable light curves, and as the SVP increases, the light curves become gradually more temporally variable. Figure 2.11 shows the *Kepler* light curves of the systems with the smallest (KIC 10857342, left panel; SVP = 0.002) and largest (KIC 3965242, right panel; SVP = 0.519) SVPs in my sample. I discuss these temporally varying systems and the SVP further in Section 5.1.

### 2.2.7 Asymmetry Parameter

Figure 2.12 shows the slightly asymmetric primary eclipse in the light curve of the non-sample system KIC 8265951. I created an asymmetry parameter (AP) to quantify the degree of asymmetry a given system shows. Unfortunately, my efforts

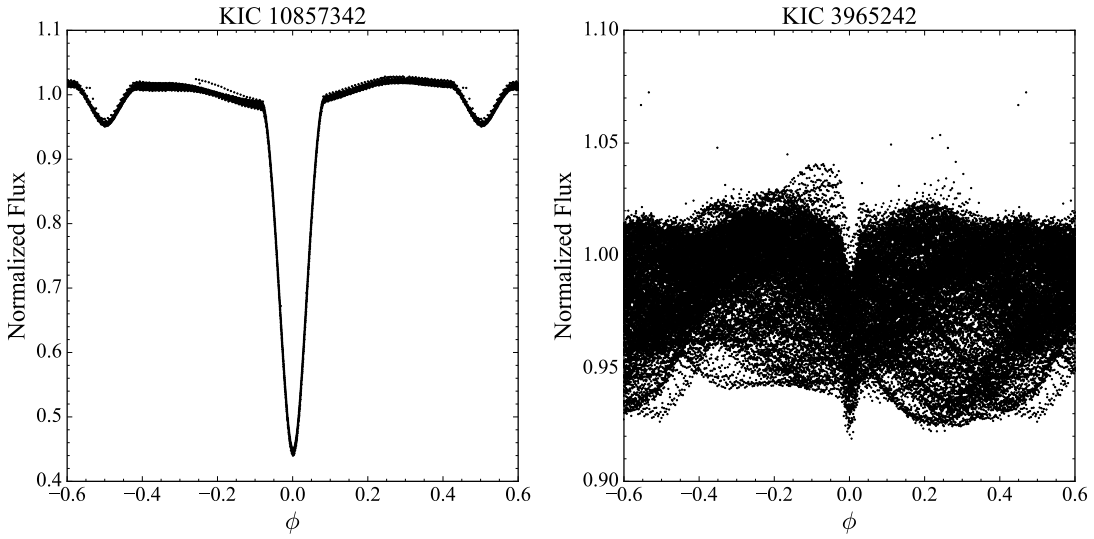


Figure 2.11: *Kepler* light curves of KIC 10857342 (left) and KIC 3965242 (right), which have the smallest (0.002) and largest (0.519) SVPs in my sample, respectively.

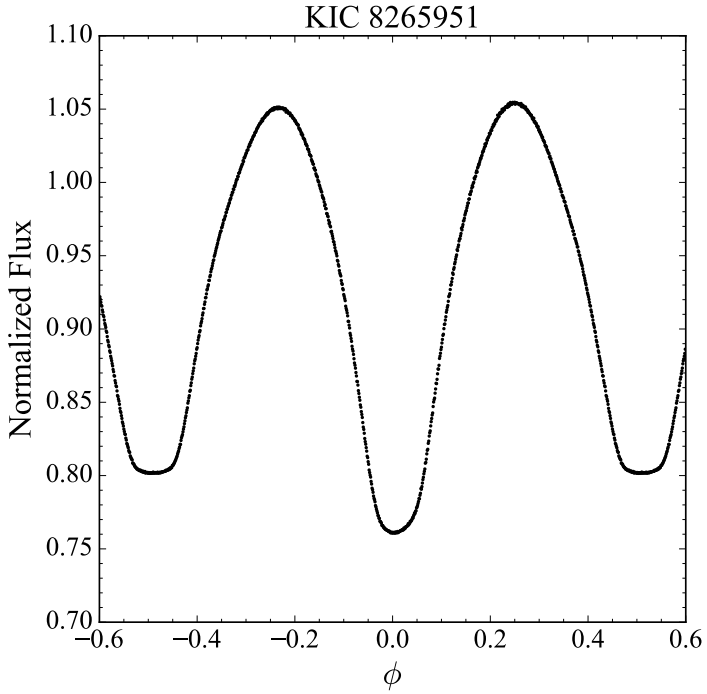


Figure 2.12: Averaged light curve of KIC 8265951 showing its asymmetric primary eclipse.

produced a less robust parameter than the SVP discussed in Section 2.2.6. I wrote a dedicated code to calculate the AP, which I will now describe. I will also detail the AP's issues and potential avenues for addressing these issues.

The code uses the same 1,001-bin averaged light curve Section 2.1.3.2 describes to calculate the Fourier coefficients as Section 2.2.2.2 describes. However, because the Fourier series is a function, I can determine its value at an arbitrary number of points, even if that number is greater than the number of points used to determine the Fourier coefficients. Because of this, I increased the number of bins from 1,000 to 10,000 (plus the additional bin for $\phi = 0.5$) for the Fourier curve, increasing its phase resolution by a factor of 10. The increased resolution allows me to determine the symmetry phase more precisely.

After creating this enhanced Fourier curve, the next step is to accurately determine the best zero phase from which to measure the eclipse's asymmetry. The proper phase to use is the primary conjunction phase, where the *Kepler* spacecraft is in syzygy with both binary components during the primary eclipse. This phase is the one about which an ideal eclipsing binary should be symmetrical, and using it for every system provides a consistent reference point that allows the AP to be compared between systems. I attempted six methods to determine this phase. Note that several of these methods implicitly assume a circular orbit, but remain flawed even with that restriction. The six methods are:

1. The code cross-correlates the given target system's light curve with the light curve of another eclipsing binary using NumPy's `correlate` function. My idea here is that aligning an O'Connell effect binary with a non-O'Connell effect binary should properly identify the conjunction phase. I chose the non-sample $\beta$ Lyrae-type system KIC 9592145 as the second eclipsing binary

because it has a symmetric light curve, meaning that no asymmetries exist to produce an artificial offset. The code sets the target system's zero phase to be where the cross-correlation is the largest.

2. The code calculates the first derivative of the Fourier series, which has its $m^{\text{th}}$ derivative given by the equation:

$$\frac{d^m}{d\phi^m} I(\phi) = \sum_{n=1}^{N} (2\pi n)^m \left[ a_n \cos\left(2\pi n\phi + \frac{m\pi}{2}\right) + b_n \sin\left(2\pi n\phi + \frac{m\pi}{2}\right) \right] \tag{2.93}$$

Using $m = 1$, Equation 2.93 gives the Fourier series' first derivative:

$$I'(\phi) = -\sum_{n=1}^{N} 2\pi n [a_n \sin(2\pi n\phi) - b_n \cos(2\pi n\phi)] \tag{2.94}$$

The code then finds the phase $\phi_{min}$ at $I'(\phi)$'s minimum in $\phi \in [-0.25, 0]$ and the phase $\phi_{max}$ at $I'(\phi)$'s maximum in $\phi \in [0, 0.25]$, representing where the ingress and egress of the eclipse are steepest. My idea here is that the conjunction phase should be precisely between these two phases, although whether this assumption holds for all systems is an open question. The code sets the target system's zero phase to the average of $\phi_{min}$ and $\phi_{max}$.

3. The code uses a 10-term Fourier series to create a lower-order Fourier curve of the target system. It then sets the target system's zero phase to the minimum of this curve. My idea here is that a 10-term Fourier series does not have enough terms to accurately approximate an asymmetric minimum. Therefore, the primary eclipse minimum of the low-order curve should occur where it would if the minimum was symmetric, which corresponds to the

conjunction phase.

4. The code calculates the second derivative of the Fourier series using Equation 2.93 with $m = 2$ to give:

$$I''(\phi) = -\sum_{n=1}^{N} (2\pi n)^2 [a_n \cos(2\pi n\phi) + b_n \sin(2\pi n\phi)] \tag{2.95}$$

At this point, the code uses one of two methods to determine the conjunction phase.

   4a. The code determines the phases of $I''(\phi)$'s minima in $\phi \in [-0.25, 0]$ and $\phi \in [0, 0.25]$. Geometrically, the $I''(\phi)$'s minima should roughly indicate the beginning and end of the partial phase of an eclipse. The reasoning behind this is that the light curve's slope decreases most strongly when the eclipser begins covering the star (causing an increase in the flux decrease) and when it has finished uncovering the star (causing a decrease in the flux increase). Testing this reasoning with Algol-type systems like the non-sample KIC 9479460 confirms this geometric argument. My idea here is that the conjunction phase should occur exactly halfway between the beginning and end of the partial phase, so the code sets the target system's zero phase to the average of the two minima phases.

   4b. The code determines the phases of $I''(\phi)$'s maxima in $\phi \in [-0.25, 0]$ and $\phi \in [0, 0.25]$. Geometrically, the $I''(\phi)$'s maxima should roughly indicate the beginning and end of the total phase of a totally eclipsing system. The reasoning behind this is that the light curve's slope increases most strongly when the eclipser has finished the process of

covering the star (causing the flux change to transition from decreasing to being nearly constant) and when it begins the process of uncovering the star (causing the flux change to transition from being nearly constant to increasing). This argument does not hold for partially eclipsing systems, where $I''(\phi)$'s maximum should occur at the phase of conjunction. In the totally eclipsing case, my idea here is that the conjunction phase should occur exactly halfway between the beginning and end of the total phase, so the code sets the target system's zero phase to the average of the two maxima phases.

5. The code creates a set of 1,001 phase offsets to check in $\phi \in [-0.05, 0.05]$. It then offsets the Fourier curve by each phase offset in this set and mirrors the Fourier curve in $\phi \in [-0.5, 0]$ across the line $\phi = 0$ to lie in $\phi \in [0, 0.5]$. Next, it calculates a $\chi^2$ statistic given by:

$$\chi^2 = \sum_{\phi=0}^{0.5} \frac{[I_{\text{right}}(\phi) - I_{\text{left}}(\phi)]^2}{I_{\text{left}}(\phi)} \tag{2.96}$$

   Finally, the code sets the target system's zero phase to the phase with the lowest $\chi^2$ among the 1,001 checked phase offsets. My idea here is that the minimum $\chi^2$ occurs when each half of the light curve best aligns with the other, which should occur when the conjunction phase lies at $\phi = 0$.

6. This method is identical to Method 2 except that it uses the convolved curve (described in Section 2.1.3.2) to calculate the derivative instead of the Fourier curve. Being derived from the averaged light curve, this method lacks the increased resolution the Fourier curve provides for the other methods. As it turns out, this lack of resolution is a fatal flaw of this method. I could

expand this method trivially to incorporate Methods 4a and 4b, but my tests indicated that such efforts would not be an efficient use of my time due to the aforementioned lack of phase resolution.

Despite the time invested in determining the conjunction phase, no method proved adequate at doing so in all circumstances. For example, only Method 3 could accurately determine the conjunction phase of the core sample systems KIC 622-3646 and KIC 9935311 (Figure 1.3). However, Method 3 is by far the worst method of determining the core sample system KIC 8822555's conjunction phase and is, in fact, worse than simply using the phase of minimum light! Despite the similarity of Methods 2, 4a, and 4b, they routinely give different conjunction phases. My testing shows that a phase shift as small as $\Delta\phi = 0.0003$ in the non-sample system KIC 9479460 (about 53.8 s given the system's 2.074-day period) increases the AP by over an order of magnitude, eliminating any possibility of using Method 6. Most frustratingly, the only known method of definitively determining the correct conjunction phase is through time-consuming binary modeling with a program like PHOEBE (Section 1.3), meaning that I lack a way to determine which method most accurately gives the conjunction phase. In light of these difficulties, I was forced to calculate the AP using the time of minimum light as the zero phase.

The AP itself is defined very similarly to the LCA (Equation 2.13):

$$\mathrm{AP} = \sqrt{\frac{1}{\phi_1}\int_0^{\phi_1}\left(\frac{I(\phi) - I(1-\phi)}{I(\phi)}\right)^2 d\phi} \tag{2.97}$$

In fact, the only changes I made to Equation 2.13 were to set the integral's upper limit to an arbitrary phase $\phi_1$ and introduce a normalization factor to account for different limits of integration between systems. I derived the value of $\phi_1$ from

the first derivative of the Fourier series (Equation 2.94). Like Method 2, I found the phase $\phi_{min}$ at $I'(\phi)$'s minimum in $\phi \in [-0.25, 0]$ and the phase $\phi_{max}$ at $I'(\phi)$'s maximum in $\phi \in [0, 0.25]$. These two phases represent where the ingress and egress of the eclipse are steepest, which can be roughly used to mark where the eclipse "is." I defined $\phi_1$ to be the average of $|\phi_{min}|$ and $\phi_{max}$. The AP's error is also very similar to the LCA's (Equation 2.74):

$$\sigma_{\mathrm{AP}} \approx \frac{\sigma_X}{\mathrm{AP}\phi_1} \int_0^{\phi_1} \left[ \left( \frac{I(\phi) - I(1-\phi)}{I(\phi)} \right)^2 \times \sqrt{\frac{2\left(N + \frac{1}{2}\right)}{\left[I(\phi) - I(1-\phi)\right]^2} + \frac{N + \frac{1}{2}}{I(\phi)^2}} \right] d\phi \tag{2.98}$$

where $\sigma_X$ is given by Equation 2.45.

Ultimately, the AP I have defined suffers from several issues that make it less robust than the SVP. Perhaps the most pressing issue is that it cannot be used to create an objective, well-defined sample. This issue occurs because, regardless of the AP value I choose to define when a system has an asymmetric minimum, there are invariably systems with an AP above that value that lack asymmetric minima. Additionally, unless I set an unrealistically small cutoff value, there will be systems with asymmetric minima and an AP below the cutoff. One cause of this issue is the form of the Fourier series: a function comprised of sinusoids can never accurately represent straight lines or sharp corners, both of which occur in my sample. Therefore, systems with completely flat minima (such as KIC 10544976's short-cadence data) have their Fourier series minima near the beginning and end of the total eclipse phase, significantly increasing the AP. Another cause of this issue is the oscillating systems I discussed at the end of Section 2.2.2.2, such as KIC 8248967 (Figure 2.8; ironically, KIC 8248967 is one of the few systems showing

these oscillations that has a genuinely asymmetric minimum). These oscillations can significantly offset the Fourier series' minimum, artificially inflating the AP. Even if the oscillations do not introduce an offset, they are often asymmetric around the minimum, which has the same effect as introducing an offset. I can remove these oscillations from the Fourier series by reducing the number of terms $N$ (which I have set to 50, as in Section 2.2.2.2) below the oscillation frequency $\nu$, but using $N = 20$ significantly reduces the AP of some asymmetric minima systems (such as KIC 8822555, which has its AP decrease from 0.013 to 0.005). Furthermore, some systems (such as KIC 11498689) have $\nu < 20$, and reducing $N$ further begins to remove the very asymmetry I am trying to characterize. I have yet to formulate a solution for this issue.

The AP's other major issue is that, unlike the SVP with regards to temporal variation, it does not accurately reflect the degree of asymmetry in the minimum. In other words, a larger AP does not necessarily translate to a more asymmetric minimum. I believe the root cause of this issue is that I could not determine the conjunction phase, which results in the symmetry axis being inconsistent between systems. Therefore, the AP measures different quantities in different systems, rendering a comparison of the AP meaningless. One example of this would be a system where the AP's symmetry axis is considerably offset from the conjunction phase, which would cause an eclipse's ingress and egress to not overlap. Equation 2.97 would compute a large AP in such a case, even if the system was perfectly symmetrical about the conjunction phase. By contrast, a system with an asymmetric minimum could have a much smaller AP if the symmetry axis coincided with the conjunction phase. The AP will not be well-defined until a consistent symmetry axis can be identified.

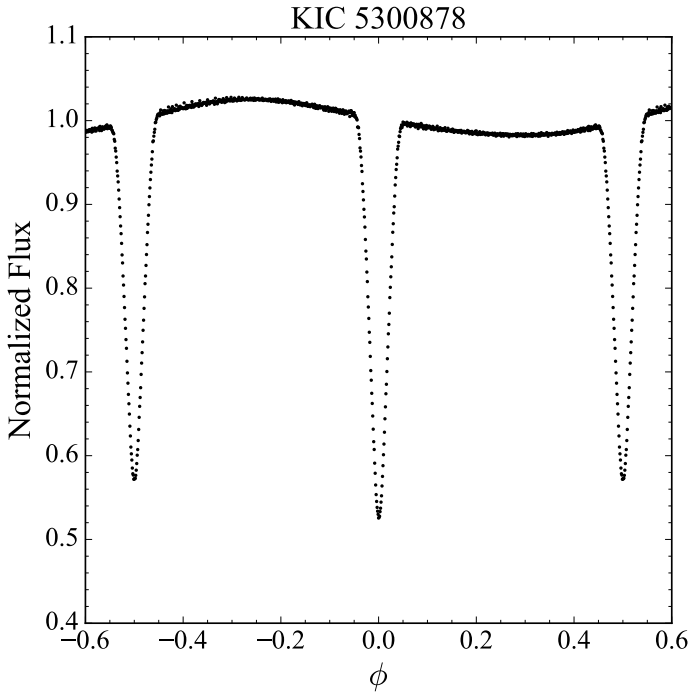


Figure 2.13: Averaged light curve of KIC 5300878 showing its concave-up region following the primary eclipse.

### 2.2.8 Concavity Parameter

Figure 2.13 shows that the region following the core sample system KIC 5300878's primary eclipse is concave-up. I created two concavity parameters to quantify the concave-up regions of a light curve: the left concavity parameter (LCP), covering the region before the primary eclipse, and the right concavity parameter (RCP), covering the region after the primary eclipse. I wrote a dedicated code to calculate the LCP and RCP, which I will now describe.

The code uses the same 1,001-bin averaged light curve Section 2.1.3.2 describes to calculate the Fourier coefficients as Section 2.2.2.2 describes. It scales the light curve such that it spans exactly the flux range $[0, 1]$. It then creates a Fourier series as in Section 2.2.2.2, except that $N = 25$ terms for the concavity parameters

instead of the 50 terms used for OER and LCA. I chose fewer terms to create a smoother representation of the light curve. Since I use the derivatives of the Fourier series to calculate the concavity parameters and differentiating a function exaggerates unsmooth features, having a smooth function creates more well-behaved derivatives. However, the reduced number of terms was still insufficient to produce a smooth enough light curve due to the nature of systems exhibiting a concave-up region.

In particular, as Figure 2.13 shows, KIC 5300878 is an Algol-type system, a trait shared by all systems with a concave-up region. Algol-type systems typically have sharp changes in the slope around the eclipses, which manifests as the Gibbs phenomenon in Fourier analysis (Hewitt & Hewitt 1979). The Gibbs phenomenon creates large oscillations around the sharp edge, with smaller oscillations created further from the edge. While my concavity parameters are not concerned with the region with the largest oscillations, the smaller oscillations interfere with identifying concave-up regions. Therefore, I used the $\sigma$-approximation (Lanczos 1956), which adds a factor of $\operatorname{sinc}\left(\frac{n}{N+1}\right)$ to the sum in the Fourier series' definition (Equation 2.15), giving:

$$I_\sigma(\phi) = \frac{a_0}{2} + \sum_{n=1}^{N} \operatorname{sinc}\left(\frac{n}{N+1}\right)[a_n \cos(2\pi n\phi) + b_n \sin(2\pi n\phi)] \tag{2.99}$$

Here, sinc refers to the normalized cardinal sine function:

$$\operatorname{sinc} x = \frac{\sin \pi x}{\pi x} \tag{2.100}$$

The $\sigma$-approximation almost completely eliminates the Gibbs phenomenon at the cost of decreased accuracy in representing the light curve, particularly during the

eclipses.

The sign of a point's second derivative determines its concavity. As a result, integrating the second derivative over a region $[b, a]$ describes the concavity of that region, giving a generic concavity parameter (CP):

$$\text{CP} = \int_b^a I''_\sigma(\phi)\ d\phi \tag{2.101}$$

However, the antiderivative of the second derivative is just the first derivative. Because the $\sigma$-approximation is $\phi$-independent, it only adds a multiplicative term to the Fourier series's first derivative (Equation 2.94), giving:

$$I'_\sigma(\phi) = -\sum_{n=1}^{N} 2\pi n \,\text{sinc}\left(\frac{n}{N+1}\right)[a_n \sin(2\pi n\phi) - b_n \cos(2\pi n\phi)] \tag{2.102}$$

I integrate $I''_\sigma(\phi)$ over the intervals $\phi \in [-0.4, -0.1]$ for the LCP and $\phi \in [0.1, 0.4]$ for the RCP. These phase ranges exclude the eclipses, focusing on the inter-eclipse region of interest. The code uses Equation 2.102 to calculate the LCP and RCP, which, using the Fundamental Theorem of Calculus, are given by:

$$\text{LCP} = I'_\sigma(-0.1) - I'_\sigma(-0.4) \tag{2.103}$$

and:

$$\text{RCP} = I'_\sigma(0.4) - I'_\sigma(0.1) \tag{2.104}$$

respectively. A given system contains a concave-up region wherever one of these parameters is positive.

The error analysis for the concavity parameters is straightforward. Applying

Equation 2.39's partial derivatives to Equation 2.102 gives:

$$\frac{\partial I'_\sigma(\phi)}{\partial a_0} = 0 \tag{2.105a}$$

$$\frac{\partial I'_\sigma(\phi)}{\partial a_n} = -2\pi n \operatorname{sinc}\left(\frac{n}{N+1}\right) \sin\left(2\pi n\phi\right) \tag{2.105b}$$

$$\frac{\partial I'_\sigma(\phi)}{\partial b_n} = 2\pi n \operatorname{sinc}\left(\frac{n}{N+1}\right) \cos\left(2\pi n\phi\right) \tag{2.105c}$$

Therefore, by Equation 2.39:

$$\sigma_{I'_\sigma(\phi)} = 2\pi\sigma_X \sqrt{\sum_{n=1}^{N} n^2 \operatorname{sinc}^2\left(\frac{n}{N+1}\right)} \tag{2.106}$$

where $\sigma_X$ is given by Equation 2.45. Careful analysis[6] of the sum reveals that:

$$\sum_{n=1}^{N} n^2 \operatorname{sinc}^2\left(\frac{n}{N+1}\right) = \frac{(N+1)^3}{2\pi^2} \tag{2.107}$$

Therefore:

$$\sigma_{I'_\sigma(\phi)} = 2\sigma_X \sqrt{\frac{(N+1)^3}{2}} \tag{2.108}$$

The LCP's error is then:

$$\sigma_{\text{LCP}} = \sqrt{\sigma^2_{I'_\sigma(-0.1)} + \sigma^2_{I'_\sigma(-0.4)}} \tag{2.109}$$

$$= 2\sigma_X \sqrt{(N+1)^3} \tag{2.110}$$

since Equation 2.108's $\phi$-independence implies that $\sigma_{I'_\sigma(-0.1)} = \sigma_{I'_\sigma(-0.4)}$. Equa-

[6] https://www.wolframalpha.com/input?i=sum+n%5E2+*+%28sinc%28pi*%28n+%2F+%28N+%2B+1%29%29%29%29%5E2+from+n%3D1+to+N

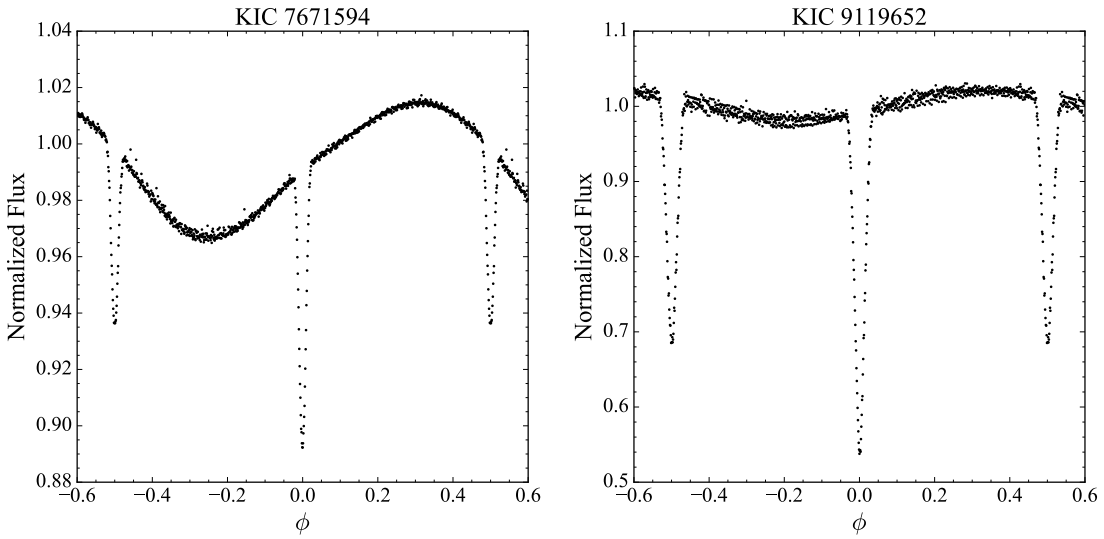


Figure 2.14: Averaged light curves of KIC 7671594 (left) and KIC 9119652 (right), which have an LCP of 2.194 and 0.474, respectively.

tion 2.108's $\phi$-independence further implies that $\sigma_{\mathrm{LCP}} = \sigma_{\mathrm{RCP}}$. The errors are zero at the precision I give the LCP and RCP.

Like the SVP described in Section 2.2.6 does with temporal variation, the LCP and RCP objectively quantify the degree that a region is concave-up in the sense that a larger concavity parameter indicates a more strongly concave-up region. For example, Figure 2.14 compares the light curves of KIC 7671594 (left) and KIC 9119652 (right). KIC 7671594's LCP of 2.194 is significantly larger than KIC 9119652's LCP of 0.474, and KIC 7671594's left inter-eclipse region does indeed appear to be more strongly concave-up relative to the light curve's amplitude than KIC 9119652's. The LCP and RCP also allow for the creation of an objective sample of concave-up systems, as a system where LCP $> 0$ or RCP $> 0$ has a concave-up region by definition. I discuss these concave-up systems and the concavity parameters further in Section 5.4.

## 2.3 Observational Methods

I observed 10 of the 258 systems in my sample over a period of three years using two instruments. I used several tools to reduce and analyze the data I obtained. I also developed several methods to plan these observations and analyze their results. I describe my observations in Chapter 6.

### 2.3.1 Observations

I conducted my observations using two of the three telescopes operated by the Southeastern Association for Research in Astronomy (SARA; Keel et al. 2017): SARA-KP and SARA-RM. FIT, a founding member of SARA, has around 20 guaranteed nights total each observing semester across the three SARA telescopes. I operated these observatories remotely.

SARA-KP is a 0.96-meter telescope at Kitt Peak National Observatory (KP-NO) in Arizona. Mounted on the telescope is a camera custom-built by Astronomical Research Cameras (ARC), Inc. thermoelectrically cooled to −110 °C. The camera contains a 2,048 × 2,048 pixel CCD chip with a 14.98' field of view. The filter wheels include B, V, R, and I filters following the Bessell prescription (Bessell 1990) that closely approximates the Johnson–Cousins filter system (Johnson & Morgan 1953; Cousins 1976). The filter wheel also contains a U filter of unknown prescription. The Kitt Peak site suffers from comparatively poor seeing, and the telescope mirror suffers from astigmatism (Keel et al. 2017). Additionally, telescope tracking and focus have become less reliable since Keel et al. (2017) was published, so I generally reserved SARA-KP for systems brighter than magnitude 14. The National Optical-Infrared Astronomy Research Laboratory (NOIRLab)

completely shuts down KPNO during the Arizona monsoon season from mid-July to early September, which is, unfortunately, the ideal time to observe the *Kepler* field. SARA-KP is equipped with a spectrograph also built by ARC, but even the brightest of my systems are too faint to easily observe with that spectrograph.

SARA-RM is a 1.0-meter telescope at Observatorio del Roque de los Muchachos (ORM) in the Canary Islands. Mounted on the telescope is an Andor Ikon-L camera thermoelectrically cooled to $-50$ °C. SARA upgraded the cooling unit in September 2019, after which the camera was cooled to $-100$ °C. The camera contains a 2,048 $\times$ 2,048 pixel CCD chip with an 11.62' field of view. The filter wheels include standard Johnson–Cousins U, B, V, $R_C$, and $I_C$ filters. ORM has more stable seeing conditions than KPNO, often obtaining a full-width at half-maximum under 1", and clear nights occur more consistently than at KPNO. Furthermore, the SARA-RM telescope has better tracking and focus than SARA-KP. Combined with the slightly larger light-collecting area, this makes SARA-RM ideal for observing some of the fainter systems in my sample. However, the Andor Ikon-L camera on SARA-RM is lower quality than the ARC camera on SARA-KP, and I had to remain vigilant to avoid the bad columns present on the camera. SARA-RM is not equipped with a spectrograph.

The *Kepler* field of view ranges in RA from $18^{\rm h}38^{\rm m}31^{\rm s}$ to $20^{\rm h}7^{\rm m}18^{\rm s}$[7], which means that the optimal time to view the *Kepler* field is from June to September. To plan my observations, I created a code to predict the altitude and orbital phase of a target on a given night. The code required the target's right ascension ($\alpha$) and declination ($\delta$) and the observing site's latitude ($\phi$) and longitude ($\lambda$) to calculate

[7] `https://archive.stsci.edu/search_fields.php?mission=kepler_stellar17`

these values. It first calculated the number of days elapsed since 12:00 UT on January 1, 2000, for a Julian date $T_{JD}$:

$$\Delta T = T_{JD} - 2451545.0 \tag{2.111}$$

It then calculated the local sidereal time on the Prime Meridian using the equation:

$$\mathrm{GMST} = (18.697374558 + 24.065709824\Delta T) \pmod{24} \tag{2.112}$$

I obtained this equation by combining Equations 2.10 and 2.12 from Kaplan (2005), dividing by 3,600 to convert to hours, and keeping only the $\theta$ term. The local sidereal time is then:

$$\mathrm{LST} = \mathrm{GMST} + \lambda \tag{2.113}$$

where $\lambda$ is in units of hours. The hour angle HA of the target at $T_{JD}$ is:

$$\mathrm{HA} = \mathrm{LST} - \alpha \tag{2.114}$$

while the altitude of the target is:

$$a = \arcsin\left(\sin\phi \sin\delta + \cos\phi \cos\delta \cos \mathrm{HA}\right) \tag{2.115}$$

where HA is in hours in Equation 2.114 and degrees in Equation 2.115.

To accurately determine the phase, I need to account for the light travel time delay of up to 16 min caused by Earth's revolution about the Sun. To do this, I calculated the heliocentric Julian date (HJD) using the method outlined in Henden & Kaitchuck (1990). The HJD is a measure of the time at the solar center at

a terrestrial time JD. The Sun is an effectively stationary reference point for measuring time because the maximum delay caused by the Sun's motion around the Solar System's barycenter is a negligible 8 s. The code begins by determining the number of centuries elapsed since 1900 for a Julian date $T_{JD}$:

$$\Delta T = \frac{T_{JD} - 2415020.0}{36525} \tag{2.116}$$

It then calculates the precession of the Sun since 1950:

$$p_{\odot} = [1.396041 + 0.000308(\Delta T + 0.5)](\Delta T - 0.499998) \tag{2.117}$$

Next, it calculates the mean solar longitude at $T_{JD}$:

$$L_{\odot} = 279.696678 + 36000.76892\Delta T + 0.000303(\Delta T)^2 - p_{\odot} \tag{2.118}$$

the mean solar anomaly at $T_{JD}$:

$$g_{\odot} = 358.475833 + 35999.04975\Delta T - 0.000150(\Delta T)^2 \tag{2.119}$$

the rectangular coordinates of the Sun at $T_{JD}$:

$$\begin{aligned} X = {} & 0.999860 \cos L_\odot - 0.025127 \cos (g_\odot - L_\odot) + 0.008374 \cos (g_\odot + L_\odot) \\ & + 0.000105 \cos (2g_\odot + L_\odot) + 0.000063 \Delta T \cos (g_\odot - L_\odot) \\ & + 0.000035 \cos (2g_\odot - L_\odot) \end{aligned} \tag{2.120a}$$

$$\begin{aligned} Y = {} & 0.917308 \sin L_\odot + 0.023053 \sin (g_\odot - L_\odot) + 0.007683 \sin (g_\odot + L_\odot) \\ & + 0.000097 \sin (2g_\odot + L_\odot) - 0.000057 \Delta T \sin (g_\odot - L_\odot) \\ & - 0.000032 \sin (2g_\odot - L_\odot) \end{aligned} \tag{2.120b}$$

and the obliquity of the ecliptic at $T_{JD}$:

$$\epsilon = 23.452294 - 0.013013 \Delta T - 0.000002 (\Delta T)^2 \tag{2.121}$$

Finally, the code calculates the heliocentric correction with:

$$\delta T = -0.0057755 [X \cos \delta \cos \alpha + Y (\tan \epsilon \sin \delta + \cos \delta \sin \alpha)] \tag{2.122}$$

The HJD is then:

$$T_{HJD} = T_{JD} + \delta T \tag{2.123}$$

The code determines the phase at $T_{HJD}$ using Equation 2.8 with the KEBC period and the KEBC $T_0$ (or my observed $T_0$, if available), then saves the phase and altitude of the target in 5-minute intervals to a file.

I conducted most of my observations using the B, V, $R_C$, and $I_C$ filters on each telescope, although I observed using the U filter on a couple of the brighter targets in 2019. To reduce exposure length and readout time, I used $2 \times 2$ binning on all

observations, which reads out a $2 \times 2$ square of pixels as a single pixel. Overall, I collected over ten thousand science images during more than 50 nights of observation, with a comparable number of observing nights lost for reasons ranging from weather and mechanical issues to a moth infestation and the COVID-19 pandemic. This observing campaign resulted in complete, high-quality, multi-wavelength light curves of seven core sample systems and lower quality or incomplete light curves of two more core sample systems and a marginal sample system.

### 2.3.2 Data Reduction

I reduced my images using the `ccdred` package in the Image Reduction and Analysis Facility (IRAF; Tody 1993). `ccdred` used bias, dark, flat field, and fringe correction images as needed to correct instrumental artifacts and produce an image suitable for photometric measurements. Additionally, `ccdred` used the overscan region present on SARA-KP images to further assist in bias correction.

The bias is a positive offset added to each pixel value to prevent negative values (Howell 2006). Biases are zero-second exposures designed to remove and correct any pixel-to-pixel variations in this offset. A zero-second integration time means that the CCD is not exposed but simply read out. I took these exposures in large batches (typically 30) at the beginning and end of each observing session. IRAF's `zerocombine` combines the biases into a single master bias frame and takes the mean value of each pixel. `zerocombine` ignores the highest value for each pixel to reduce the impact of cosmic ray events (Massey 1997).

Overscan regions do not correspond to a physical area of the CCD chip but are instead a set of pixels that are read out a second time after reading out the image (Howell 2006). IRAF uses these pixels to assist bias correction because

they represent the bias level for that image. This does not obviate the need for dedicated bias images, however, because the overscan region does not capture full-frame patterns in the bias. Once IRAF uses the overscan region, it trims those pixels away as they constitute an unusable portion of the image.

Dark current is caused by thermal fluctuations in the CCD chip and is strongly dependent on the chip temperature (Howell 2006). SARA-KP's camera is sufficiently cool that the dark current is negligible (Keel et al. 2017), as is SARA-RM's camera following the September 2019 upgrade to its cooler. Before the upgrade, however, the dark current for SARA-RM was sufficient to require correction. Darks are closed-shutter exposures of length comparable to the longest science images taken in an observing session. The closed shutter means that any signal is due to a combination of the bias and dark current. I took three dark exposures at the beginning and end of each observing session. IRAF's `darkcombine` subtracts the master bias from each dark and then combines the darks into a single master dark frame. It takes the median value of each pixel instead of the mean so that events such as cosmic rays do not affect the master dark.

Flat field effects result from inhomogeneities across the image frame. These inhomogeneities arise from pixel-to-pixel differences in gain and sensitivity as well as artifacts like out-of-focus dust grains on the mirror or filter wheel (Howell 2006). Flats are exposures of a uniformly lit background highlighting these inhomogeneities, allowing for their removal from science images. I typically took five flats per filter near the beginning and end of each observing session. The time frame for taking flats is narrow, as the twilight sky must be bright enough that the exposures are reasonably short but not so bright that the images are saturated. I shift the telescope between each exposure so that stars do not fall on the

same pixel in each image. IRAF's `flatcombine` subtracts the master bias (and the master dark, if applicable) from each flat and then scales the images such that their modes are equal. It then combines the flats into one master flat frame per filter. `flatcombine` takes the median value of each pixel to remove stars in the field, leaving only the intrinsic inhomogeneities.

Fringing arises from interference patterns between reflected light waves within a CCD (Howell 2006). The reflections occur because silicon is partially transparent to infrared radiation, causing reflections to occur at the top and bottom of a silicon wafer. I created a fringe correction image by taking several 100-second exposures of the night sky, with a shift between each exposure to move the stars. I used IRAF's `imcombine` with parameters identical to those used in `flatcombine` to produce a single image from my exposures. IRAF's `mkfringe` creates a master fringe frame by using a "boxcar" smoothing algorithm on the image created by `imcombine`. The boxcar algorithm subtracts the average value of a variably sized box from the pixel at the center of the box. This variably sized box is largest at the image center and smallest at the image edge. `mkfringe` excludes from this smoothing algorithm any pixels that are a specified number of standard deviations above or below the average, allowing the fringing pattern to remain. Fortunately, the fringing pattern is constant on a given instrument for a given wavelength (Howell 2006) and only affects my $I_C$-filter data, so I only needed three fringe correction images: one for SARA-KP and two for SARA-RM, as the fringing pattern at SARA-RM changed slightly after the September 2019 mirror realuminization.

Finally, I used the IRAF command `ccdproc` to reduce my science images. `ccdproc` subtracts the master bias frame, subtracts the master dark frame (if applicable), divides by the appropriate master flat frame, and subtracts the master

fringe frame (if applicable) from each science image taken. If the science image's exposure time differs from the exposure times used to create the master dark or fringe frames, `ccdproc` linearly scales these frame's pixel values to the science image's exposure time. These images are then suitable for photometric measurements.

### 2.3.3 Photometry and Phasing

For photometry, I used AstroImageJ (AIJ; Collins et al. 2017), an image processing program based on ImageJ (Rasband et al. 1997; Schneider et al. 2012) tailored for astronomy. I performed differential photometry using a comparison star and a check star. I carefully selected the comparison star using SIMBAD and VizieR to ensure that it was not a known variable, that it had similar brightness and color to the target, and that it had published Johnson–Cousins B, V, $R_C$, and $I_C$ magnitudes. My primary source for the magnitudes was the AAVSO (American Association of Variable Star Observers) Photometric All-Sky Survey (APASS). I used both APASS Data Release 9 (Henden et al. 2016), accessible from VizieR, and APASS Data Release 10 (Henden 2019), accessible from the AAVSO website[8]. APASS uses Sloan r' and i' filters, but these magnitudes can be transformed to Cousins $R_C$ and $I_C$ magnitudes using the following equations from Table 1 of Jester et al. (2005):

$$R_C = V - [1.09\,(r' - i') + 0.22] \tag{2.124a}$$

$$I_C = V - [2.09\,(r' - i') + 0.43] \tag{2.124b}$$

[8] `https://www.aavso.org/download-apass-data`

I error propagated Equations 2.124a and 2.124b to obtain the errors in these magnitudes:

$$\sigma_{\mathrm{R_C}} = \sqrt{\sigma_{\mathrm{V}}^2 + 1.09^2\left(\sigma_{\mathrm{r'}}^2 + \sigma_{\mathrm{i'}}^2\right)} \tag{2.125a}$$

$$\sigma_{\mathrm{I_C}} = \sqrt{\sigma_{\mathrm{V}}^2 + 2.09^2\left(\sigma_{\mathrm{r'}}^2 + \sigma_{\mathrm{i'}}^2\right)} \tag{2.125b}$$

Ideally, the comparison star should be of similar brightness to the variable so it would have a high signal-to-noise ratio (SNR) and thus provide more precise photometry. I wanted the comparison to be a similar color to minimize the effect of atmospheric reddening on the photometry. Finally, I wanted magnitudes for each filter so I could transform the differential magnitudes into apparent magnitudes. I was more flexible when choosing the check star, as it only needed to be non-variable and of similar brightness to the comparison star. I used the check star only to ensure that the comparison star was non-variable.

AIJ outputs the fluxes and flux errors of the target and check star relative to the comparison star. I developed a code to produce a phased light curve of differential magnitudes. It converts the relative flux $F$ to a differential magnitude $\Delta m$ using Pogson's equation (Pogson 1856):

$$\Delta m = -2.5 \log F \tag{2.126}$$

Error propagating Equation 2.126 gives:

$$\sigma_{\Delta m} = \frac{2.5}{\ln 10} \frac{\sigma_F}{F} \tag{2.127}$$

The code converts the JD of the observations to HJD and then phase using Equa-

tions 2.123 and 2.8, respectively. It then saves the HJD, phase, differential magnitude, and error in differential magnitude to a file. I also included the option to export apparent magnitude rather than differential magnitude for future work on modeling these systems.

I used the method outlined by Kwee & van Woerden (1956) to determine the time of minimum $T_0$. Their method begins by defining the time interval used to find the minimum. They create $2n+1$ magnitudes equally spaced by time $\Delta t$ and recommend letting $2n+1 \approx N$, where $N$ is the total number of observations in the selected interval. To create these $2n+1$ magnitudes, they linearly interpolate the adjacent observed data to the required time. One of these magnitudes at a time $T_e$ represents an initial estimate for the time of minimum. They mirror the points on one side of $T_e$ to the other, creating a set of equidistant times with two magnitude measurements each. They take the difference in each pair $\Delta m_i$ and define:

$$s(T_e) = \sum_{i=1}^{n} (\Delta m_i)^2 \tag{2.128}$$

The idea is that at the exact time of minimum, the sums should all become zero assuming they neglect the observational error. They then shift the symmetry axis to $T_e + \frac{\Delta t}{2}$ and then $T_e - \frac{\Delta t}{2}$, calculating $s\left(T_e + \frac{\Delta t}{2}\right)$ and $s\left(T_e - \frac{\Delta t}{2}\right)$ using Equation 2.128. The smallest of the three sums should be $s(T_e)$ if they correctly picked $T_e$. If another sum is smaller than $s(T_e)$, they change the time offset in the smaller sum by another factor of $\Delta t/2$ and check if that sum is larger than $s(T_e)$. This process continues until a larger sum is found or the method is terminated. The number of points $n$ must be kept the same for all of these sums. These sums

serve to define a parabolic function of the form:

$$s(T) = aT^2 + bT + c \tag{2.129}$$

Any three non-colinear points $(x_1, y_1)$, $(x_2, y_2)$, and $(x_3, y_3)$ with $x_1 \neq x_2 \neq x_3$ define the constants $a$, $b$, and $c$ by[9]:

$$a = \frac{x_3(y_2 - y_1) + x_2(y_1 - y_3) + x_1(y_3 - y_2)}{(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)} \tag{2.130a}$$

$$b = \frac{x_1^2(y_2 - y_3) + x_3^2(y_1 - y_2) + x_2^2(y_3 - y_1)}{(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)} \tag{2.130b}$$

$$c = \frac{x_2x_3y_1(x_2 - x_3) + x_3x_1y_2(x_3 - x_1) + x_1x_2y_3(x_1 - x_2)}{(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)} \tag{2.130c}$$

They find the time of minimum by calculating the minimum of Equation 2.129, giving:

$$T_0 = -\frac{b}{2a} \tag{2.131}$$

with error:

$$\sigma_{T_0} = \sqrt{\frac{4ac - b^2}{4a^2(Z - 1)}} \tag{2.132}$$

Here, $Z$ is the number of independent pairs of observations. Using pairs of observations would imply $Z = N/2$, but the linear interpolation adds another factor of $1/2$ since the interpolation is calculated from two observations. Therefore, unless linear interpolation is unnecessary, $Z = N/4$.

I created my own code implementing their method. My code plots the data in an interactive Python plot, from which I can visually pick the minimum estimate

---

[9] `https://math.stackexchange.com/questions/889569/finding-a-parabola-from-three-points-algebraically`

and time interval. The code differs from the formulation in Kwee & van Woerden (1956) in that I select the minimum before the time interval. Additionally, I only allow the program to change the offset for the sum a maximum of two times. If $s(T_e \pm \frac{3\Delta t}{2})$ is still smaller than $s(T_e)$, the code generates a warning and suggests using a different time interval or minimum estimate. While Kwee & van Woerden (1956) do not explain why $s(T_e)$ needs to be the smallest sum, my tests indicate that time intervals not containing the minimum can still have a time of minimum calculated without this constraint. I choose which calculated time of minimum I want to save to file before exiting the program.

### 2.3.4 O'Connell Effect Size Determination

I used the same technique to find the OES of systems I observed that I used on the *Kepler* data in Section 2.1.3.2. Because of the reduced number of observations, however, I chose to use 101 bins rather than 1,001 to create the averaged light curve. Other than the reduced number of bins, my methodology is identical to the one Section 2.1.3.2 describes. I determined the OES separately for each filter I observed with.

### 2.3.5 Color Index and Temperature Determination

I chose to use the B − V color index to estimate the primary component's temperature, which is a common procedure in eclipsing binary studies (see Terrell & Nelson 2014, Mennickent et al. 2020, and Alton & Wiley 2021 as examples). Determining the B − V color index requires the apparent B and V magnitudes, which my code computed by adding the known apparent B and V magnitudes of the comparison star to the differential B and V magnitudes AstroImageJ produced.

The code found apparent magnitude error by adding the photometric errors in quadrature with the reported errors of the comparison star's magnitudes. The B and V magnitudes were not recorded at the same time, however. To overcome this limitation, my code linearly interpolated the B-filter observations to coincide with the V-filter observations using equations similar in form to Equations 2.37 and 2.38. The code then converted the observations to phase space and produced a B − V curve, showing how the system's color index changes as a function of $\phi$.

The B − V color index is a flux-weighted average of the colors of the system's components. The color index at $\phi = 0.5$ gives the best estimate of the primary component's color because this is the phase at which it blocks the most light of the secondary component. Therefore, my code created a convolved color curve using the same method that I described in Section 2.3.4. I defined the primary component's B − V color index to be the value of the convolved color curve at $\phi = 0.5$. The error of the primary's color index is generally dominated by the error in the comparison star's apparent magnitudes. Using this color index, I then estimated the primary component's temperature using the results of Flower (1996). I did not account for interstellar reddening due to dust, so my estimates may underestimate the primary component's temperature.

# Chapter 3

# Literature Review

I have reviewed the literature surrounding in my sample and the systems therein. In this chapter, I give an overview of the most relevant works that reference the systems in my sample. I present the methods used in each paper, the major conclusions the authors draw, and how they relate to my sample. Because the papers I review cover a multitude of topics, this chapter will necessarily jump from topic to topic. I have organized this chapter by grouping papers with similar themes together in sections, although several papers are too specific to group with any other. I give special attention to Kouzuma (2018) in its own section. At the end of this chapter, I summarize the papers that directly focus on individual systems in the sample in the order of the relevant system's *Kepler* identifier.

## 3.1 Flares

Balona (2015), Gao et al. (2016), and Davenport (2016) are three papers that discuss stellar flares in *Kepler* targets. Balona (2015) defined a flare star as a star

that has at least one observed flare. Like the KEBC, he used the SAP flux over the PDCSAP flux (I defined SAP and PDCSAP in Section 2.1.4). In his case, he used SAP flux because PDCSAP flux could mistake flares as outliers and remove them. He searched for flares in a subsample of 20,810 long-cadence light curves and all 4,758 short-cadence light curves. Fifty-seven of the short-cadence systems were eclipsing binaries, of which five (KICs 6205460, 6836140, 8608490, 9091810, and 9328852) are in the core sample and two (KICs 4660997 and 7885570) are in the marginal sample. He found that flares occur in stars of spectral type A and even two subdwarf B stars, and that larger stars tend to have larger flares, a fact he attributed to larger active regions. The relative number of flares decreases by a factor of four in his sample, but he states that this is a selection effect due to the lower spectral contrast between flares and F- and A-type stars and the dilution of flare intensity by the higher luminosities of F- and A-type stars. In light of this, he claims that the difference in flare rates between cool and hot stars is much smaller or possibly non-existent. Finally, he looked at three eclipsing binaries that show numerous flares and find no phase dependence on the flare rate. This is inconsistent with magnetic recombination between the components being responsible for flares in close binaries, as this recombination would occur most frequently in the region between the two stars.

Gao et al. (2016) focused on flares in *Kepler* eclipsing binaries. They analyzed the light curves of 1,049 KEBC systems. Their search excluded systems with $P < 0.4$ d because it would be hard to disentangle a flare signal from the eclipse signal in very short-period systems. They also excluded heartbeat stars (discussed in Section 2.1.5), non-binaries, and systems without a valid morphology parameter, among other "peculiar" systems. They define each system's morphological

classification using $\mu$ (which I discussed in Section 2.2.3, where I also discuss why I avoid this convention). They needed to detrend the long-cadence PDCSAP data to remove the light curve variation due to orbital motion. They used a median filter for detached systems that creates a baseline curve of variability without flares or eclipses. For all other systems, they fit a four-term Fourier series to the data and iteratively sigma-clipped all points more than $3\sigma$ away from the average ten times, then removed any fitted points more than $1\sigma$ away from the average. They created the baseline curve by taking the mean of the remaining points, excluding any points during an eclipse. After subtracting the baseline curve from the observed curve, they flagged events as potential flares when three consecutive data points had values $3\sigma$ above the detrended flux. They confirmed flare candidates by eye. Gao et al. (2016) found 6,818 total flares in 234 KEBC systems, of which 16 are in my core sample and 15 in the marginal sample. They determined that semi-detached systems have the highest incidence of flares at nearly 33%, followed by detached systems with 22%. Only 12% of overcontact systems exhibit flares. The cumulative flare energy is highest for detached and semi-detached systems and systems with a period of about three days. Like Balona (2015), they did not find a correlation between flare activity and phase for eclipsing binaries. Finally, they discuss two systems that are not in their sample, KIC 5952403 and KIC 11347875. The former system is not a member of my sample. By contrast, KIC 11347875 is one of the most notable members of my core sample, as it both exhibits the largest OES in the sample (Section 4.1.1) and is the archetype of a class of systems displaying a concave-up region in their light curves (Section 5.4). For KIC 11347875, they find that flares are significantly more common at the quadrature following secondary eclipse than at any other phase.

Davenport (2016) presents a catalog of flare stars observed by *Kepler*. He analyzed the available long- and short-cadence light curves for all 207,617 systems *Kepler* observed. Like Balona (2015), he used the SAP flux over the PDCSAP flux due to the latter's tendency to remove flares as outlier data points. He produces a quiescent light curve using a rolling median filter, removing sinusoidal signals with Lomb-Scargle fitting, and removing flares and starspots using spline fitting. He then detects flares by cross-correlating the quiescent curve subtracted *Kepler* data with a flare profile defined in Davenport et al. (2014). He also injected 100 artificial flares into each continuous segment of the raw SAP flux data and attempted to recover them using the procedure outlined above. The percentage of flares recovered as a function of flare energy represents the completeness threshold for that continuous segment. Davenport (2016) found 2,304,930 candidate flares, the vast majority of which, he claims, are spurious. He therefore restricted further analysis to systems with at least 100 candidate flares and at least ten candidate flares with energies above the 68% completeness threshold for a given continuous segment. This resulted in a sample of 4,041 systems, of which 109 are in my core sample and 16 are in the marginal sample. He found that the percentage of stars that flare increases as the stellar mass decreases, but the maximum flare energy increases with increasing mass. Using the stellar rotation periods McQuillan et al. (2014) gives for 402 systems in his sample, he found an inverse relationship between rotation period and flare luminosity.

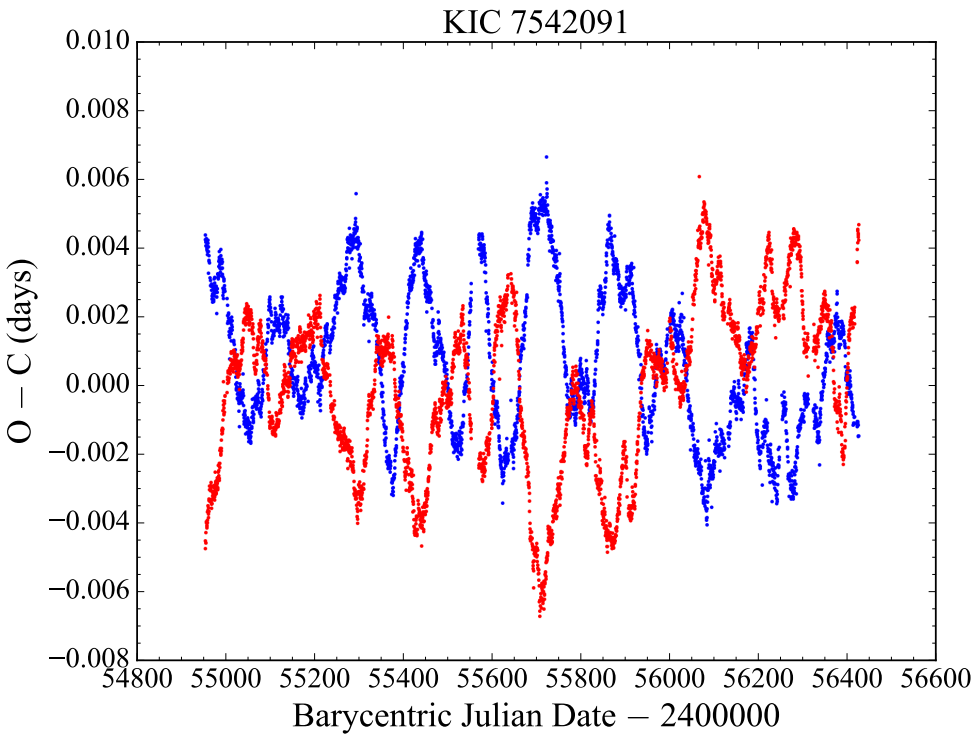


Figure 3.1: Eclipse timing variation of KIC 7542091. The blue points are the eclipse timings for the primary eclipse while the red points are the eclipse timings for the secondary eclipse. The primary and secondary eclipse timings are strongly anticorrelated.

## 3.2 Spots

Tran et al. (2013) and Balaji et al. (2015) discuss starspots in *Kepler* eclipsing binaries. Tran et al. (2013) related the anticorrelated ETV pattern seen in Figure 3.1 and many other overcontact systems to the presence of spots. These systems also have anticorrelated quadrature timing variation (QTV, a term introduced by Borkovits et al. 2016) that are 90° out-of-phase with the ETV. They median-normalized the long-cadence PDCSAP flux for all targets in the then-latest release of the KEBC, then smoothed the resultant light curve by applying a boxcar function. They subtracted this smoothed curve from the unsmoothed curve,

removing most features with a frequency under half the orbital frequency. To accurately determine eclipse times using data with long-cadence time resolution, the fit a parabola to each data point representing a local minimum and the immediately adjacent points and took the time of the parabola's minimum as the eclipse time of minimum. This procedure was generally accurate to within 30 seconds for orbital periods under a day. They calculated the times of minimum for each eclipse assuming a linear ephemeris and subtracted them from the corresponding observed time to produce the ETV. They found 390 systems with an ETV exhibiting quasi-periodic variation and an anticorrelated relation between primary and secondary eclipses, but they focus their analysis on a subset of 32 systems, all with periods between 0.2 and 0.5 d. Six (KICs 2159783, 5033682, 7542091, 7773380, 9071104, and 9832227) of those 32 systems are in the core sample, while one (KIC 5283839) is in the marginal sample. They note that a true change in the period cannot explain this anticorrelated nature because this would produce a correlated ETV, and mass transfer could not produce such rapid period changes in any event. Similarly, slightly eccentric orbits are deemed implausible for such short-period systems. Therefore, they explain the anticorrelation as a consequence of a starspot migrating in longitude. The proposed starspot must be persistently visible across all longitudes and all points in the orbit, requiring the spot to be at a high latitude and the orbital inclination to be low enough that the spot is never occluded. They tested this hypothesis on the non-sample system KIC 3437800 using PHOEBE (Prša & Zwitter 2005) and reproduced the anticorrelated ETV and QTV with a spot at 45° latitude.

Balaji et al. (2015) built upon the work of Tran et al. (2013) by tracking the longitude of the spots proposed in the latter paper. Balaji et al. (2015) looked

at a larger sample of 414 systems with an anticorrelated ETV from the then-latest release of the KEBC. These 414 systems include 89 core sample systems. They introduced a four-term equation to describe their light curves: a constant term equal to the mean flux, a sinusoid with a frequency twice that of the orbital frequency, a sinusoid with a frequency equal to the orbital frequency, and a second sinusoid with a frequency twice that of the orbital frequency representing the starspot. They collapsed this into a three-term equation by incorporating the phase and amplitude of the starspot term into the other two sinusoids, and they fit this equation to their sample light curves. The $\phi_1$ term in this equation is the phase offset term of the first sinusoid, and it is closely associated with the spot longitude. They plotted each term in their equation as a function of time and analyzed these plots for KIC 6431545, a system in the marginal sample. They highlighted the $\phi_1$ plot as striking, as it shows a pronounced linear decrease in $\phi_1$ over time. It makes a full 360° rotation six times over *Kepler*'s primary mission, with a seventh rotation in progress at mission end. They attribute this to the starspot moving around the star, either due to incomplete tidal locking, differential rotation, or another effect such as a time-varying magnetic field. This motion is retrograde (relative to the binary orbit) in KIC 6431545, and they found such retrograde motion in 34.3% of their sample. Prograde motion was seen in 13.0% of their sample, while 7.2% showed erratic behavior. The remaining 45.6% did not see $\phi_1$ make a full rotation during the *Kepler* mission. They end by noting that the phase offset term for their second sinusoid, $\phi_2$, should remain constant in the absence of external perturbations. They used the variation of $\phi_2$ to identify 39 triple systems, and they obtained orbital elements for KIC 5478466 similar to those found in Conroy et al. (2014).

## 3.3 Mass Transfer

Kouzuma (2018) discusses mass transfer in overcontact *Kepler* eclipsing binaries. He considered three cases of mass transfer: mass exchange from the more massive component to the less massive one (MEML), mass exchange from the less massive component to the more massive one (MELM), and mass loss from the system (ML). The rate of change in period $\dot{P}$ is negative for the first case and positive for the latter two. He used the second version of the KEBC presented in Slawson et al. (2011) containing 2,165 eclipsing binaries. He produced eclipse times and ETVs – for the primary eclipse only – using the KEBC long-cadence data and manually selected ETVs that appeared parabolic. He then fit a parabola to the ETV using least-squares and determined $\dot{P}$ from the quadratic coefficient. Determining the mass transfer rate $\dot{M}$ requires the individual stellar masses, which he estimated from the mass-temperature relationship given for main sequence stars in Harmanec (1988). He found that $\dot{M}$ generally decreases for $P < 0.6$ d for all three mass transfer cases, although it increases for periods under 0.4 d for MEML. Aside from a continued negative correlation between 0.6 and 0.8 d for MEML, however, he found no correlations for $P > 0.6$ d. He states that systems with $P > 0.6$ are less likely to be overcontact, and thus it is unsurprising to find different mass transfer properties in such systems. He also calculated the semi-major axis $a$ for each system via Kepler's third law and found a negative correlation between $\dot{M}$ and $a$ for $a < 4$ $R_\odot$. For $a > 4$ $R_\odot$, he found a positive correlation between $\dot{M}$ and $a$. Another quantity he investigated was the mass ratio $q$, which he defines as the mass of the gaining star divided by the mass of the donor star, rather than the more common definition of the secondary component's mass divided by the primary component's mass. The MEML systems positively correlate with $\dot{M}$ over

the range of $q$, while the MELM systems negatively correlate for $q < 2$. For $q > 2$, the MELM systems are nearly uncorrelated with $q$, with a possible slight positive correlation. For the ML systems, $\dot{M}$ generally decreases for $q < 0.5$ and increases for $q > 0.5$. Finally, he explored the relationship between the primary component's temperature $T$ and $\dot{M}$. He found a positive correlation for $T < 6{,}200$ K and a negative correlation for $6{,}200 \text{ K} < T < 7{,}000$ K for the MEML sample. The MELM and ML samples have a negative correlation throughout the range of $T$. He also investigated the relationships with the fillout factor (a quantity that measures the degree of overcontact) and stellar mass, but I do not discuss these as they were less precisely determined. Twenty-two of the 111 systems described in this paper are in the core sample.

## 3.4 Multiplicity

Borkovits et al. (2016) and Furlan et al. (2017) discuss higher-order multiplicity in *Kepler* eclipsing binaries. Borkovits et al. (2016) focused on studying triples using ETV. They introduced a model of the ETV described by the equation:

$$\Delta = \sum_{i=0}^{3} \left( c_i E^i \right) + \left[ \Delta_{\text{LTTE}} + \Delta_{\text{dyn}} + \Delta_{\text{apse}} \right]_0^E \tag{3.1}$$

where $c_i$ is a constant, $E$ is the epoch number ($E^{\text{th}}$ eclipse since some reference eclipse), and $\Delta_{\text{LTTE}}$, $\Delta_{\text{dyn}}$, and $\Delta_{\text{apse}}$ are the ETV contributions from light-travel time effects (LTTE), short-term dynamical perturbations, and apsidal motion effects, respectively. The zeroth- and first-order terms represent the standard linear ephemeris used in variable star studies, the second-order term accounts for a linear change in $P$, and the third-order term accounts for additional ETV of unknown

origin in five systems. They give analytical equations for $\Delta_{\mathrm{LTTE}}$ and $\Delta_{\mathrm{dyn}}$ that I do not rewrite, but the former depends on the semi-major axis of the eclipsing binary and the eccentricity and inclination of the third body, while the latter also depends on orbital periods, component masses, and mutual inclination. $\Delta_{\mathrm{apse}}$ is significant in only a few systems. They analyzed the QTV for systems with significant ellipsoidal variation. They are ambiguous about the version of the KEBC they used, as they state it was the then-current version. Given the paper's year, this would imply the third version, but they do not cite Kirk et al. (2016). Regardless, they used data taken directly from the KEBC to create a phased light curve that they then binned using 1,000 equally-spaced phase bins. They used these binned light curves to create an eclipse template, which they used to search for eclipses in the full data set. Finally, they used a Levenberg–Marquardt (LM) fit to determine the precise eclipse times for each eclipse. Section 4 of their earlier work, Borkovits et al. (2015), describes their procedure in greater detail. This procedure found about 400 ETVs they reviewed further. Given the results of Tran et al. (2013), they averaged the primary and secondary ETVs as well as the two QTVs to reduce the influence of starspots. They smoothed the light curves of detached systems using a low-order polynomial to lessen the effects of intrinsic brightness variations. This allowed them to remove false positives from their sample more effectively. For instance, the KEBC flags the core sample system KIC 11247386 as a multiple, but Borkovits et al. (2016) states that it is a false-positive and that its ETV arises from a different origin. They base this conclusion on the differing amplitudes of KIC 11247386's ETV and QTV, which is atypical for LTTE. They found evidence of a third body in 222 KEBC systems, of which 21 are in the core sample and five are in the marginal sample. Their analysis of the mutual inclination between

the eclipsing binary and third body revealed two peaks: one indicating co-planar orbits and another around $i_m = 40°$. The latter peak is close to the 39.2° cutoff for the Kozai mechanism (Kozai 1962), to which they attribute this peak. They also found a near-complete absence of short period ($P < 200$ d) third bodies around the shortest-period binaries ($P < 1$ d), which cannot be explained as an observational bias. Instead, they explain the dearth of systems in this regime as a consequence of the dynamical processes that created the close binary. Finally, they note that there are "hundreds" of other ETVs that are more-or-less parabolically-shaped, paralleling the conclusion I will draw in Section 4.1.3.

Furlan et al. (2017) describes efforts to image companions to *Kepler* Objects of Interest using high-resolution imaging. They combined observations from nine observatories using four techniques: adaptive optics (Keck, Palomar, Lick, and MMT), speckle interferometry (Gemini North, WIYN, and the Discovery Channel Telescope), lucky imaging (Calar Alto), and *Hubble* observations. They imaged 3,557 *Kepler* systems, four of which (KICs 2569494, 3338660, 5300878, and 10676927) are in the core sample. They found close companions around 1,903 of these systems, and they note that companions closer than 1" away from the primary are likely bound based upon the results of Horch et al. (2014). Of the six systems from my sample that they observed (including KICs 4037163 and 8949316 from the marginal sample), only KIC 2569494 had companions, both farther than 1" away. Nevertheless, such close companions will reduce the transit depth, be it a planetary transit or a stellar eclipse.

## 3.5 High-Energy Observations

Smith et al. (2015) and Pizzocaro et al. (2019) report on high energy (ultraviolet and X-ray) observations conducted in the *Kepler* field. Smith et al. (2015) conducted their observations with the X-Ray Telescope (XRT) and UV/Optical Telescope (UVOT) instruments on the *Swift* space telescope (Burrows et al. 2005). These instruments are aligned such that simultaneous observations image the same field. They observed four of the twenty-one regions covered by one of *Kepler*'s CCD pairs and found 93 X-ray sources with SNR $> 3$. They then identified the nearest KIC source by overplotting the XRT data on the simultaneously obtained UVOT image. The source they labeled KSw 53 corresponds to the marginal sample system KIC 7284688, a temporally varying system (as defined in Section 5.1.1) with an anticorrelated ETV (as discussed in Section 3.2). They obtained optical spectra on 31 sources using the Hale telescope, but KIC 7284688 was not among them. Aside from KIC 7284688, they confirmed 32 other sources as stellar, and they note many of these were G- and K-type stars with broad emission features indicative of highly active chromospheres. Frasca et al. (2016) spectrally classified KIC 7284688 as a G2 V star which, combined with its temporal variation and the results from this paper, indicates that it is likely chromospherically active.

Pizzocaro et al. (2019) analyzed data from 3XMM-DR5 (Rosen et al. 2016), a catalog of sources serendipitously discovered by *XMM-Newton* (Jansen et al. 2001). Pizzocaro et al. (2019) cross-referenced the source positions from Rosen et al. (2016) with the KIC source positions from Brown et al. (2011), noting that $\sigma_{\mathrm{KIC}} \ll \sigma_{\mathrm{XMM}}$. They extracted the photometry for each matched source from the KIC and excluded sources without 2MASS IR photometry, as this would result in unreliable spectral classifications. They also fit each source's spectral energy

distribution with a stellar photospheric model (Allard et al. 2012) to remove non-stellar X-ray sources. This resulted in a sample of 125 *Kepler* targets, two of which (KIC 8842170 and KIC 9532591) are members of my core sample. They drew the temperature, surface gravity, and metallicity values from Huber et al. (2014) and Frasca et al. (2016), while they inverted the parallaxes reported by Gaia Collaboration et al. (2018) to determine distance rather than using the results from Bailer-Jones et al. (2018). They estimated the mass and bolometric magnitude by projecting onto the Dartmouth Stellar Evolution Program (Dotter et al. 2008) isochrone corresponding to their temperature and surface gravity. They explicitly excluded eclipsing binaries from a significant portion of their analyses, including rotational, spectral, and flare analysis. They spectrally classed KIC 8842170 as a G-type star with an effective temperature of $5{,}831 \pm 204$ K, and they found an X-ray luminosity of $1.955 \times 10^{23}$ W for the system. They also spectrally classed KIC 9532591 as a G-type star with an effective temperature of $5{,}502 \pm 103$ K, and they found an X-ray luminosity of $1.062 \times 10^{23}$ W for the system. Their determined temperature is about 200 K hotter than the *Kepler* temperature and 600-700 K hotter than the *Gaia* temperature for both systems.

## 3.6 Miscellaneous

Gaulme et al. (2013) presents a catalog of red giants in *Kepler* eclipsing binaries. They selected their sample by cross-referencing a list[1] of *Kepler* red giants compiled by the *Kepler* team with the then-current version of the KEBC presented

[1] `https://archive.stsci.edu/kepler/red_giant_release.html`

in Slawson et al. (2011) and Matijevič et al. (2012), which resulted in a list of 70 systems. They knew that not all of these systems contained a red giant, however, because the orbital periods of several were too short to accommodate a giant star. They note that the surface gravity and temperature distributions resemble the red giant distribution, indicating that even systems in their sample without a red giant are dominated by a red giant's flux. They considered three possible cases: the red giant was part of the eclipsing binary, the red giant was a gravitationally bound companion to the eclipsing binary, or the red giant was unassociated with the eclipsing binary. To differentiate between these cases, they inspected the pixel map files and ETV. The former allowed them to determine if the eclipsing signal coincided with the red giant source or if it was from a nearby source, while the latter allowed them to determine if there was a third body in the system. They found 13 eclipsing binaries and one heartbeat star where one of the components is a red giant, but none are in my sample. The remaining 56 systems contain five core sample systems (KICs 5820209, 8846978, 9091810, 9489411, and 11671660). Of these, they list KIC 11671660 as a triple system containing a red giant bound to an eclipsing binary while the other four do not contain red giants.

Liakos & Niarchos (2017) presents a catalog of known $\delta$ Scuti variables that are members of an eclipsing binary. $\delta$ Scuti variables are generally short-period ($P <$ 0.25 d), low-amplitude ($A < 0.01$ mag) pulsating variables of spectral type A or F (Breger 2000). The catalog is continuously updated, but it had 199 systems upon publication, none of which are in the core sample. However, the catalog includes one marginal sample system (KIC 11175495). Additionally, one core sample system (KIC 11819135) is listed in a table of ambiguous cases, with the comment that it had an unclear pulsation type. They find that the companion star does not

influence the pulsations on the $\delta$ Scuti component in binaries with $P > 13$ d.

Zoła et al. (2017) investigated the effects of phase-smearing on overcontact binary light curves and provides parameters for 17 overcontact KEBC systems. However, they exclude systems showing signs of magnetic activity, explicitly citing the O'Connell effect as an example. The only system from my sample they mention – the core sample system KIC 2159783 – is used as an example of the type of system they exclude. Regardless, their analysis of phase smearing warrants discussion. Phase smearing refers to the fact that the 30-minute long-cadence exposure time covers a non-negligible phase interval for a given system, particularly systems with short periods. As a result, the flux obtained from a given observation represents the average of a rapidly-changing intensity, leading to the smoothing of sharp light curve features like kinks. To test this, they simulated the light curve of an overcontact system at *Kepler* cadence for a range of periods using an older version of the WD program (Wilson & Devinney 1971) that did not account for phase smearing and attempted to regain the original parameters through modeling. They found that WD had difficulty recovering the inclination for the systems with $P < 1.5$ d and the inclination, potential, and mass ratio were significantly different for the systems with $P < 0.5$ d. The inclination differed by up to 15° for the $P = 0.2$ d system, while the potential and mass ratio differed by up to 22.7% and 5.5%, respectively. They conclude that phase smearing must be taken into account for systems with $P < 1.5$ d to recover accurate parameters.

Kjurkchieva et al. (2017) discusses eccentricity in detached *Kepler* eclipsing binaries. While they cite Kirk et al. (2016), they note that the third version of the KEBC does not include the orbital parameters that the second version presented by Slawson et al. (2011) did, so they drew their sample from the second version. They

selected systems with two eclipses, well-defined boundaries into and out-of eclipses, eclipse depths greater than 0.0001 in normalized flux, and either eclipses that are separated by a value other than 0.5 or eclipse widths that differ by more than 5%. This produced a sample of 529 eclipsing binaries, including ten core sample systems and two marginal sample systems. They used PHOEBE to model each system's orbital parameters (eccentricity $e$ and argument of periastron $\omega$) while fixing other parameters to reduce the fitting time. Systems with near-circular orbits will show correspondingly small deviations from the circular orbit case, while $\omega = 90°,\ 270°$ indicates that the semi-major axis is perfectly radial, meaning eclipses will occur half a phase apart regardless of eccentricity. It is therefore difficult to tell systems with such an $\omega$ apart from systems with circular orbits. All systems in my sample have $e \leq 0.05$ and $\omega$ within 20° of 90° or 270° except KIC 4049124 and KIC 12164751, which have $e = 0.094$ and $\omega = 347.3°$, respectively. However, KIC 4049124 has $\omega = 272.4°$ and KIC 12164751 has $e = 0.006$, offsetting the other element's large value. None of these 12 systems show obvious signs of an eccentric orbit. Kjurkchieva et al. (2017) found only a faint trend of eccentricity increasing with orbital period.

Matson et al. (2017) presents radial velocity data on 41 *Kepler* eclipsing binaries. They conducted spectroscopic observations using KPNO's 4-meter Mayall telescope, Lowell Observatory's 1.8-meter Perkins telescope, and the Dominion Astrophysical Observatory's (DAO) 1.8-meter Plaskett telescope. These observations covered wavelengths between 3,930 Å and 4,600 Å and resolving power between 4,200 and 6,200. The data were reduced, extracted, and wavelength-calibrated with IRAF. The KPNO and DAO spectra were wavelength-calibrated using standard lamps, while the Lowell spectra required standard star observations. They

determined the orbital elements for each system using the fitting routine described in Morbey & Brosterhus (1974). They found the semi-major axis and component masses by using the inclination reported by Slawson et al. (2011) in conjunction with their derived values for $a \sin i$ and $M \sin^3 i$. Two of their 41 systems are in my sample: KIC 2305372 (core sample) and KIC 4660997 (marginal sample), both of which are double-lined spectroscopic binaries. They obtained six spectra of KIC 2305372 and nine of KIC 4660997 from KPNO. Slawson et al. (2011) report $\sin i > 1$ for KIC 2305372, so Matson et al. (2017) sets $i = 90°$, which is incompatible with the lack of flat-bottomed eclipses in KIC 2305372's light curve. Their reported semi-major axis ($a = 6.4 \pm 0.2$ $R_\odot$) and component masses ($M_1 = 1.2 \pm 0.1$ $M_\odot$ and $M_2 = 0.62 \pm 0.04$ $M_\odot$) for KIC 2305372 therefore underestimate the true values. They report a semi-major axis $a = 3.6 \pm 0.1$ $R_\odot$ and component masses $M_1 = 1.16 \pm 0.07$ $M_\odot$ and $M_2 = 0.88 \pm 0.07$ $M_\odot$ for KIC 4660997. Both systems have mass ratios significantly less than 1 ($0.52 \pm 0.03$ and $0.76 \pm 0.05$ for KICs 2305372 and 4660997, respectively), placing them in the intermediate-mass secondary category. There are comparatively few systems in this category, which they attribute to the difficulty of distinguishing the increasingly extreme component flux ratio. Finally, seven systems in their sample were suspected triples based on ETV analysis (Gies et al. 2012, 2015), of which five are double-lined spectroscopic binaries. One of these five is KIC 2305372, which they single out as an ideal candidate for high-resolution spectroscopy to detect all three components. They determine that the third body in KIC 2305372 has a mass between the eclipsing binary components' masses.

Kobulnicky et al. (2022) performs a Bayesian analysis of 10 contact or near-contact *Kepler* eclipsing binaries. They chose systems that had deep eclipses,

$0.74 \leq \mu \leq 0.95$, and $P \approx 1$ d. They used the *Kepler* data provided by Kirk et al. (2016) combined with spectroscopic data obtained from the 2.3‘meter telescope at the Wyoming Infrared Observatory (WIRO) and the 3.5-meter Astrophysical Research Consortium telescope at Apache Point Observatory (APO). They modeled this data using PHOEBE v2.2 (Prša et al. 2016) combined with the Python package `emcee`'s (Foreman-Mackey et al. 2013) implementation of the Markov chain Monte Carlo (MCMC) method to solve the inverse problem. Two of the 10 systems they studied (KIC 9164694 and KIC 9345838) are members of my core sample. KIC 9164694 serves as the exemplar of the asymmetric minima system class (Section 5.2) and was one of 10 systems I observed for this project. They found that both KIC 9164694 and KIC 9345838 are overcontact systems ($f = 0.61$ and 0.89, respectively) comprised of stars of very different mass ($q_{\rm ph} = 0.24$ and $q_{\rm sp} = 0.20$ for both systems). Their model of KIC 9164694 has $i = 89.8°$, $T_2/T_1 = 0.82$, and a third-body light contribution of 22%, while their model of KIC 9345838 has $i = 75.9°$, $T_2/T_1 = 1.00$, and a third-body light contribution of 47%. Both systems required a hot starspot on the primary component. I detail my own observations of KIC 9164694 in Section 6.1.6.

## 3.7 Other Surveys

Hartman et al. (2004) describes the results of the Hungarian-made Automated Telescope network (HATnet) survey of variability in the *Kepler* field. They observed an $8.2° \times 8.2°$ field centered on the J2000 equatorial coordinates $19^{\rm h}44^{\rm m}00^{\rm s}$ $+37°32'00''$, covering a portion of the *Kepler* field. They obtained 788 5-minute exposures on this field and selected 98,000 stars contained therein with a Cousins I

magnitude brighter than 14. However, the density of objects was high, resulting in severe blending between targets. This rendered traditional photometric methods (aperture photometry and point-spread function – or PSF – fitting) impractical, so they used the image subtraction method (Alard & Lupton 1998; Alard 2000). Image subtraction subtracts each image from a reference image, taking advantage of the fact that non-variable sources will disappear in the resultant image, but variable sources will remain. To perform PSF photometry on this image, however, the PSF needs to be known to high-precision. Alard & Lupton (1998) describes a kernel that transforms the PSF of the reference frame into the PSF of the active frame despite the two having different seeing in general. They then performed PSF fitting on the subtracted image to obtain their photometric measurements. They identified 1,617 variable stars with an amplitude above 0.032 magnitudes, of which they classed 157 as eclipsing binaries. Six of these eclipsing binaries (KICs 2305372, 4660997, 5123176, 5282464, 5557368, and 6205460) are in my core sample, while two (KICs 3557421 and 4660997) are in the marginal sample. Additionally, they classified the core sample system KIC 5294739 as a pulsating variable with $P = 0.1886$ d instead of an eclipsing binary.

Hoffman et al. (2008) and Hoffman et al. (2009) discuss variable stars in the Northern Sky Variability Survey (NSVS; Woźniak et al. 2004). The former focuses on Algol- and $\beta$ Lyrae-type eclipsing binaries, while the latter expands to include other variable stars. The NSVS was a survey that produced light curves for ~14,000,000 objects north of $-28°$ aimed at detecting variable stars. Hoffman et al. (2008) was the paper I chose NSVS 7322420 (Section 1.4) from. They analyzed NSVS systems brighter than magnitude 13.5 (unfiltered), a standard deviation greater than 0.1 magnitudes around the median magnitude, and at least 30 ob-

servations. They applied a Fourier transform to the data and analyzed the power spectrum to determine the period. This produced ~10,000 candidate variables, which they reduced to ~3,000 by selecting systems that spent most phases dimmer than the median magnitude. They manually inspected these systems to identify Algol- and $\beta$ Lyrae-type candidates, which they further analyzed with the eclipsing light curve code (Orosz & Hauschildt 2000). They identified 409 new Algol- or $\beta$ Lyrae-type eclipsing binary candidates. The core sample system KIC 7657914 was one of the new candidates, while the core sample system KIC 7272739 and the marginal sample system KIC 4660997 were previously known.

Hoffman et al. (2009) extended the analysis of Hoffman et al. (2008) to other periodic variable stars, including W Ursae Majoris-type eclipsing binaries, short-period pulsators like $\delta$ Scuti- and RR Lyrae-type variables, and long-period pulsators like Cepheid- and Mira-type variables. They analyzed the full ~10,000 systems Hoffman et al. (2008) found. They fit a six-term Fourier cosine series to each light curve using LM fitting, and if the fit failed to converge, they used slightly different initial conditions and integer harmonics of their determined period. They used the $a_4$ coefficient to classify systems: if $a_4 > 0$ and $P < 1.2$ d, they classified the system as an RR Lyrae, while if $P > 1.2$ d, they classified it as a Cepheid or long-period variable. If $a_4 < 0$, they determined the primary and secondary minima depth and analyzed the system further to classify it. They identified 4,659 variable stars, including 551 Algol- or $\beta$ Lyrae-type eclipsing binaries and 2,332 W Ursae Majoris-type eclipsing binaries. Note that the 551 Algols and $\beta$ Lyrae do not include the systems identified in Hoffman et al. (2008). Among the systems identified as eclipsing binaries are seven core sample systems (KICs 5123176, 7272739, 7542091, 7584739, 7871200, 8703528, and 10259530). Additionally, they

classified the core sample system KIC 9345838 as a Cepheid variable with half its correct period.

Devor et al. (2008) identifies and describes 773 eclipsing binaries discovered by the Trans-Atlantic Exoplanet Survey (TrES; Alonso et al. 2004). The data they analyzed consisted of ten fields observed by a 0.1-meter telescope at Palomar Observatory. The telescope has a $6^\circ \times 6^\circ$ FoV, a resolution of 11" per pixel, and a mounted Sloan r' filter. Each of the 185,445 observed systems has ~2,000 photometric measurements taken at a nine-minute cadence, which they obtained by the method of image subtraction (Alard & Lupton 1998; Alard 2000) and then binning five consecutive 90-second exposures. Devor et al. (2008) applied two routines – the Detached Eclipsing Binary Light curve fitter (DEBiL; Devor 2005) and the Method for Eclipsing Component Identification (MECI; Devor & Charbonneau 2006a,b) – designed for coeval detached systems. Their pipeline broadly consisted of fitting the orbital parameters and removing contaminated light curves, which eliminated 97% of the objects. They used box-fitting least squares (Kovács et al. 2002) and ANOVA (Schwarzenberg-Czerny 1989, 1996) to determine the period of each system. Finally, they placed each system into one of seven groups, the last three of which contained systems that could not be fit using MECI. The 773 systems they classified contains 18 core sample systems and two marginal sample systems. Devor et al. (2008) misidentified three core sample systems (KICs 9328852, 11154110, and 11341950) as nearby stars (KICs 9328864, 11154102, and 11394704, respectively), but I examined the *Kepler* data from MAST and found that the eclipsing signal comes from the systems in my sample. An interesting non-sample binary that they observed was KIC 9998899, which they identified as T-Lyr1-13166 but is most commonly known as V361 Lyrae (previously mentioned in Section 1.2.3).

V361 Lyrae is a well-observed system (Andronov & Richter 1987; Shugarov et al. 1990; Kałużny 1990, 1991; Gray et al. 1995; Hilditch et al. 1997) that had – when W09 was published – the largest OES of any known system. Hilditch et al. (1997) found that V361 Lyrae is a rare type of binary consisting of solar-type stars undergoing mass transfer as the binary evolves towards a contact configuration. These properties make the system highly relevant to my project. Unfortunately, while V361 Lyrae nominally falls in *Kepler*'s FoV, its light fell into the area between the CCD chips in *Kepler*'s photometer array.

Pigulski et al. (2009) describes variables in the *Kepler* field observed by the All Sky Automated Survey (ASAS; Pojmanski 1997). ASAS conducted these observations with optical V and I filters and have low spatial – 15" per pixel – and temporal – 2-4 day cadence – resolution. They determined periodicity by using a Fourier amplitude periodogram. This Fourier method, combined with the data quality and quantity, imposes a large detection variability threshold, which increases from ~0.01 mag for $I \leq 11$ to 0.15 mag at $I = 14$, indicating that their sample is strongly biased towards high-amplitude variables. They selected 947 variable stars in the *Kepler* field, 186 of which they classified as eclipsing binaries. These 186 eclipsing binaries include 43 core sample systems and 4 marginal sample systems. They also classified the core sample system KIC 11347875 as a Cepheid variable.

Ramsay et al. (2014) outlines a high-cadence variability survey of the *Kepler* field. Their objective was to discover pulsating compact objects, and to that end, they conducted observations over 42% of the *Kepler* field using the 2.5-meter Isaac Newton Telescope (INT) at ORM and the 1.3-meter McGraw-Hill Telescope at KPNO at a high-cadence (50 seconds at ORM and 85 seconds at KPNO).

They obtained their photometry using SExtractor (Bertin & Arnouts 1996) for sparse fields and an updated version of Difference Image Analysis Package (DIAPL; Woźniak 2000), DIAPL2, for crowded fields. DIAPL2 is based on the method of image subtraction (Alard & Lupton 1998; Alard 2000) other surveys I have covered have used. They removed systematic trends by applying the SYSREM algorithm (Tamuz et al. 2005), resulting in ~710,000 detrended light curves. They searched for periodicity using a Lomb-Scargle periodogram and defined a median absolute deviation (MAD) to identify outliers. They selected sources that obeyed the relation $\log \mathrm{FAP} < n \cdot \mathrm{MAD}_{\log \mathrm{FAP}} + \mathrm{Median}_{\log \mathrm{FAP}}$, where FAP is the false-alarm probability and $n$ is some integer, which they chose to be 18. This resulted in a sample of 65 objects, from which they selected 18. They obtained optical spectra of these 18 targets using the 10.4-meter Gran Telescopio Canarias (GTC) and successfully bid to have them observed as short-cadence targets by *Kepler* during the last several months of *Kepler*'s primary mission. Four of these 18 systems are eclipsing binaries, two of which (KIC 7667885 and KIC 9786165) are core sample systems. They classify both systems with a "mid-G" spectral type based on the GTC spectra, but they do not report an effective temperature. Despite the appearance of KIC 7667885's light curve, they identify both systems as likely overcontact based solely on their short periods. KIC 9786165 was added to the KEBC based on this paper, but KIC 7667885 was not, and the latter system is therefore the only one in my sample that is not in the KEBC.

Clark Cunningham et al. (2019) describes *Kepler* eclipsing binaries observed by the Apache Point Observatory Galactic Evolution Experiment (APOGEE; Majewski et al. 2017). APOGEE was conducted using the 300 fiber, near-infrared APOGEE spectrograph (Wilson et al. 2019) attached to the 2.5-meter Sloan Foun-

dation Telescope at APO. It has provided high-resolution ($R \sim 22{,}000$) spectra of ~150,000 stars, a subset of which *Kepler* observed. For their sample, they selected detached or semi-detached *Kepler* eclipsing binaries brighter than $H = 14$ that were visited by APOGEE at least three times and showed multiple cross-correlation peaks in one epoch. This resulted in 33 promising spectroscopic eclipsing binaries, of which KIC 3848919 is in the core sample and KICs 4077442 and 4660997 are in the marginal sample. The remainder of the paper gives a detailed analysis of seven of the 33 systems, none of which are in my sample.

## 3.8 System-Specific Papers

Debski et al. (2014) describes spot migration on two *Kepler* overcontact systems, KIC 2159783 and KIC 6118779. The former is a member of the core sample while the latter is not in the sample, although it does present an O'Connell effect below the cutoff. Both systems are totally eclipsing (i.e. they have flat-bottomed eclipses), and KIC 6118779 shows an asymmetric minimum (Section 5.2). Both also show temporal variation in their light curves, although not to the level described in Section 5.1. They found evidence of spot migration in both systems, briefly for KIC 2159783 and continuously for KIC 6118779. They also found that the OES is slowly increasing in KIC 2159783 on timescales of order years, while it varies cyclically in KIC 6118779 with a period of about 45 d. Finally, they find that the phase separation between the two maxima varies in both systems. This variation is periodic in KIC 6118779, with a period nearly identical to that of the OES variation. They discuss KIC 6118779 further in Debski et al. (2015), where they find that only a large polar spot can replicate the observed variation.

Yoldaş (2021) describes the starspot and flare activity of the core sample system KIC 6044064, one of only seven systems in my sample displaying a concave-up region (Section 5.4). They used Kwee & van Woerden's (1956) method described in Section 2.3.3 to determine 525 times of minimum from 264 primary eclipses and 261 secondary eclipses. Their ETV analysis showed that starspots affected the secondary minima timing much more strongly than the primary minima timing. They analyzed KIC 6044064's light curve by averaging two consecutive orbital cycles with minimal variation together with a phase interval of 0.005 and modeling the resultant light curve using PHOEBE v0.32. They determined a primary temperature $T_1$ of 5,375 K using de-reddened $(H-K)_0$ and $(J-H)_0$ color indices from 2MASS (Cutri et al. 2003), which they held fixed during modeling. Their model indicates a detached system with $i = 83.04 \pm 0.03°$, $q = 0.948 \pm 0.005$, and $T_2 = 3{,}951 \pm 50$ K. They used calibrations from Tokunaga (2000) to estimate spectral types of G7 V and K9 V and component masses $M_1 = 0.86$ $M_\odot$ and $M_2 = 0.54$ $M_\odot$. While the authors do not remark upon it, their model's near-unity mass ratio is at odds with the determined temperatures if the components are assumed to be main-sequence stars, implying either an inconsistent model or an evolved component. Because the spots evolve rapidly, they split the *Kepler* data into 11 subsets and modeled the subsets with sinusoids, finding two sets of three distinct starspot regions on the secondary component separated by 120° in longitude. The two sets make a complete rotation around the secondary in ~1.67 y and ~3 y, respectively. Finally, they analyzed KIC 6044064's flare activity by combining the PHOEBE light curve and starspot modulations to create a synthetic light curve, which they then compared to the observed data to find flares. They

calculated each flare's equivalent duration using the equation (Gershberg 1972):

$$P = \int \frac{I_{\text{flare}} - I_0}{I_0} \, dt \tag{3.2}$$

where $I_{\text{flare}}$ is the flux during the flare and $I_0$ is the expected flux at the same time using the synthetic light curve. They also defined a value they call *Plateau*, defined by the equation:

$$y = y_0 + (Plateau - y_0)\left(1 - e^{-kx}\right) \tag{3.3}$$

where $y$ is the equivalent duration of a given flare (Equation 3.2), $y_0$ is the theoretical equivalent duration for a flare of minimum duration, $k$ is a constant, and $x$ is the total duration of a given flare. They state that *Plateau* is the maximum equivalent duration for a given star and is related to the maximum flare energy of a star. They find KIC 6044064's *Plateau* to be 3.983 s, which they note is higher than the *Plateau* value of EV Lacertae (3.014 s), one of the most active flare stars known. In fact, they state that KIC 6044064's *Plateau* value is much higher than any other known system's. As a result, they conclude that KIC 6044064 has a "remarkably high level of magnetic activity."

NegmEldin et al. (2019) presents an analysis of three *Kepler* eclipsing binaries: KIC 2715417, KIC 6050116, and KIC 6287172. KIC 2715417 is a near-contact system that displays an O'Connell effect, but its OES is below the cutoff and it is not a member of my sample. KIC 6050116 is a member of the core sample and displays a large, negative O'Connell effect (OES = $-0.036$). Finally, KIC 6287172 is an ellipsoidal variable showing no O'Connell effect. They modeled these systems' long-cadence *Kepler* light curves using PHOEBE v0.31a. One issue I have encountered with v0.31a is that it does not include the *Kepler* bandpass, and there

is no ability to add new bandpasses to the program until v0.32 (which includes the *Kepler* bandpass by default). Presumably, this is why NegmEldin et al. (2019) presents the *Kepler* data as Johnson V-band data throughout the paper, as this is the closest approximation in PHOEBE v0.31a. Their model of KIC 6050116 indicates a marginally overcontact ($f = 0.0704$) system with $i = 69.2 \pm 0.7°$, $q = 0.573 \pm 0.012$, and component temperatures within 100 K of each other. To explain the O'Connell effect, their model includes a cool starspot in the secondary component's mid-latitudes. They determined KIC 6050116's absolute parameters by estimating the mass using three methods: an empirical relation from Harmanec (1988), the total mass-luminosity relationship given by Maceroni & van't Veer (1996), and the period-mass relationship given by Gazeas & Niarchos (2006). Unfortunately, their results imply that KIC 6050116's distance is about 230 pc, but the *Gaia* EDR3 parallaxes give a distance of $448 \pm 3$ pc (Bailer-Jones et al. 2021), indicating that their absolute parameters are inaccurate. They accounted for the effects of interstellar extinction, but neglecting this increases the distance by only ~50 pc.

Blättler & Diethelm (2000a) and Blättler & Diethelm (2000b) are observation reports detailing the observations of five eclipsing binaries. Of the five, GSC 3131.476 Lyrae from Blättler & Diethelm (2000a) and GSC 3547.216 Cygni from Blättler & Diethelm (2000b) are the core sample systems KIC 7871200 and KIC 10727655, respectively. These observations were conducted with an unfiltered 0.15-meter refractor in Switzerland and have a relatively high photometric scatter as a result. The O'Connell effect is not discernible in either light curve.

Kunt & Dal (2017) describes the starspot and flare activity of the marginal sample system KIC 7885570, a totally eclipsing Algol-type binary exhibiting flares,

strong temporal variation (Section 5.1), and – in its short-cadence light curve – asymmetric minima (Section 5.2). *Kepler* observed the system in short-cadence for over six months during four non-consecutive periods. Kunt & Dal (2017) selected a single orbit of short-cadence *Kepler* data exhibiting minimal out-of-eclipse variation to model using PHOEBE v0.32. They determined a primary temperature $T_1$ of 6,530 K from the de-reddened $(H-K)_0$ and $(J-H)_0$ color indices from 2MASS (Cutri et al. 2003), which they held fixed during modeling. Their model of KIC 7885570 indicates a detached system with $i = 80.6 \pm 0.1°$, $q = 0.43 \pm 0.01$, and $T_2 = 5{,}732 \pm 4$ K. They included a cool starspot on the secondary component to account for some out-of-eclipse variation. They then removed eclipses and flares from the light curve and modeled the starspot distribution over 35 time intervals using the SpotModel program (Ribárik 2002; Ribárik et al. 2003). They used two spots on the secondary component, the first with a temperature factor of 0.75 and the second with a temperature factor of 0.80. They found that the spots would remain stationary for several cycles, then some combination of latitude, longitude, and spot radius would abruptly change by several degrees. Finally, they analyzed KIC 7885570's flares by combining the PHOEBE light curve and starspot modulations to create a synthetic light curve, which they then compared to the observed data. They found 113 flares and calculated the equivalent duration using Equation 3.2. They found that the flaring frequency was ~0.004 flares per hour and the total flare duration was ~0.001% of the total observing time. I detail my observations of this system in Section 6.2.2.

KIC 9832227 is a highly referenced system due to a prediction that its component stars will merge in 2022 (Molnar et al. 2017). I will summarize four papers referencing KIC 9832227 in their title: Molnar et al. (2017), Pavlenko et al. (2018),

Socia et al. (2018), and Kovács et al. (2019). Molnar et al. (2017) quotes Eggleton (2012) as describing contact binary evolution as one of the "great unsolved questions of stellar evolution," as it remains unknown how they form or what causes the components to merge. They review the past difficulty in classifying KIC 9832227, as Kinemuchi et al. (2006) described it as an RR Lyrae variable, while Prša et al. (2011) described it as an eclipsing binary. Molnar et al. (2017) observed the system with two identical 0.4-meter telescopes in Grand Rapids, Michigan, and Rehoboth, New Mexico, using Bessell R or Sloan r' filters, with additional filters used on certain nights. They combined this data with photometric measurements of KIC 9832227 taken by the NSVS, ASAS, the Wide Angle Search for Planets (WASP; Pollacco et al. 2006), and *Kepler*. They compared these measurements to the pre-outburst photometry of V1309 Scorpii, which underwent a luminous red nova in 2008 that Tylenda et al. (2011) determined was caused by a contact binary merging. They fit a 10-term Fourier series to the *Kepler* data and attributed the maxima asymmetries (the O'Connell effect, though they do not call it such) and minima asymmetries to starspots, while their analysis of the ETV revealed a tertiary component in a ~590-day orbit. They obtained high-resolution ($R \sim$ 31,500) spectroscopic data from APO and WIRO. They used this data and produced a model of the system using BinaryMaker3 (BM3; Bradstreet & Steelman 2002), indicating a $1.395 \pm 0.011$ $M_\odot$ primary and a $0.318 \pm 0.005$ $M_\odot$ secondary on a low-inclination ($i = 53.19 \pm 0.10°$) orbit. Using a combination of their data and *Kepler*'s, Molnar et al. (2017) found an exponentially decreasing period – as Tylenda et al. (2011) found for V1309 Scorpii before the 2008 outburst – and predicted that KIC 9832227 would merge in early 2022.

Pavlenko et al. (2018) describes a near-infrared spectrum of KIC 9832227.

They obtained it using the 3.0-meter Infra-Red Telescope Facility in Hawaii at intermediate resolution ($R$ = 1,200), revealing that the Ca II triplet at 8,500 Å in KIC 9832227 was weaker than their model spectra (Pavlenko 2003) predicted. They explained the weak Ca II lines by introducing a chromosphere into their model, which provided a better fit. They also found a strong He I absorption feature at 10,830 Å that is too narrow to be connected to the rotating components. They instead attributed it to material outside of the component's photospheres, suggesting the common envelope surrounding both stars as a possible source.

Socia et al. (2018) refutes the merger Molnar et al. (2017) predicted based on a reanalysis of the data and introducing new data. They conducted observations with the ill-named 1.06-meter 40-inch Telescope at Mount Laguna Observatory using a Cousins R filter. They also inspected archival data taken with the 0.1-meter Vulcan Photometer instrument (Caldwell et al. 2004) at Lick Observatory. To confirm that Vulcan produced sufficiently accurate eclipse times, they compared the 2003 Vulcan eclipse times of KIC 9592855 – an eclipsing binary showing no sign of period change – and found that they agreed with the eclipse times predicted by the KEBC's linear ephemeris within $1\sigma$, as did WASP observations from 2008. They determined that a starspot was present in KIC 9832227 with a rotational period of 0.460 days, slightly greater than the 0.458-day orbital period. This reflects the results of Balaji et al. (2015). Their reexamination of the NSVS data revealed a 1.01-hour discrepancy between their calculated eclipse time measurement and that of Molnar et al. (2017). They found the source of this discrepancy to be the definition of MJD Woźniak et al. (2004) used in their summary of the NSVS data (MJD $\equiv$ JD $-$ 2400000.0 instead of the standard MJD $\equiv$ JD $-$ 2400000.5), an issue that also proved irritating during my master's project. This 0.5-day offset

produced the apparent exponential decrease Molnar et al. (2017) reported. While Socia et al. (2018) found that KIC 9832227's period was changing, they found that it increased during some time intervals and decreased in others. They concluded that the system would not merge in 2022. They also constrained the minimum mass of a tertiary component to 0.7 $M_\odot$ – which they found suspicious considering that no tertiary component is seen in spectra – and the minimum period to 7,200 d – which they note is suspiciously close to the range of their observations.

Finally, Kovács et al. (2019) revisited the hypothesized tertiary component in KIC 9832227. They combined data from Vulcan, HATnet, the NSVS, ASAS, WASP, the All-Sky Automated Survey for Supernovae (ASAS-SN; Shappee et al. 2014; Kochanek et al. 2017), the Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010), *Kepler*, and observations taken by members of the American Association of Variable Star Observers (AAVSO) to create a detailed $O-C$ curve. They fit this curve with the ephemeris equation:

$$O - C = c_1 + c_2 t + c_3 t^2 + c_4 \sin\frac{2\pi t}{P_2} + c_5 \cos\frac{2\pi t}{P_2} \tag{3.4}$$

using a Monte Carlo method to investigate periods ranging from 4,000 to 8,000 d. In Equation 3.4, $t$ is the time since HJD 2455000.0 and $P_2$ is the period of the third body. Equation 3.4 accounts for both the linear period change from mass transfer and the LTTV due to the third body. Their best-fit parameters indicated a linear period change of $-1.097 \pm 0.047 \times 10^{-6}$ days per year and a periodic variation with a period of $4{,}925 \pm 142$ d, which they attribute to a third body.

Almenara et al. (2012) and Almeida et al. (2019) both discuss the marginal sample system KIC 10544976. KIC 10544976, which I discuss further in Section 5.3,

is an eclipsing binary comprised of a white dwarf and a red dwarf. Almenara et al. (2012) present details about their discovery and subsequent observations of KIC 10544976. They discovered the system in 2005 while conducting a photometric survey with INT. They performed follow-up photometric observations with the 4.2-meter William Herschel Telescope (WHT) and the 2.5-meter Nordic Optical Telescope and spectroscopic observations with the WHT and the 3.6-meter Galileo National Telescope, all of which are at ORM. From these observations, they determined that it was a post-common envelope binary (PCEB), the only one known in the *Kepler* field prior to *Kepler*'s launch. PCEBs are short-period eclipsing binaries consisting of a main sequence star and a white dwarf. They could not see KIC 10544976's secondary eclipse (where the white dwarf transits the red dwarf) in their data. The red dwarf spectrum best fit the M4 V template, while the white dwarf best fit spectral class DA. They modeled their photometric and spectroscopic data using the BinaRoche code described in Appendix A of Lazaro et al. (2009). BinaRoche was developed independently of WD but appears quite similar conceptually. Almenara et al. (2012) obtained the following parameters from their model: $T_1 = 20{,}470 \pm 1{,}300$ K, $T_2 = 3{,}200 \pm 100$ K, $M_1 = 0.61 \pm 0.04$ $M_\odot$, $M_2 = 0.39 \pm 0.03$ $M_\odot$, and $i = 89.6 \pm 1.5°$.

Almeida et al. (2019) investigated orbital period variations observed in KIC 10544976. They reanalyzed the WHT data from Almenara et al. (2012) and combined it with *Kepler* and more recent WHT and INT data to determine the system's ETV. They fit the eclipses with WD to determine the precise eclipse times and found an ETV signal that appeared periodic. They considered two ideas to explain this signal: an LTTE effect caused by an eccentric third body and the Applegate mechanism (Applegate 1992). The Applegate mechanism is a genuine period

change caused by a magnetically active star changing its shape due to variations in the magnetic quadrupole moment. The period of the Applegate mechanism should correspond to the magnetic activity cycle (MAC) of the active star, which they attempted to determine using two methods. The first method involved analyzing the numerous flares observed in the *Kepler* light curve. They normalized the short-cadence PDCSAP data and removed the reflection effect by subtracted a fitted sinusoidal function. Their program flagged a candidate flare whenever three consecutive data points were more than $3\sigma$ brighter than the baseline flux. They determined that the flare rate varied periodically with a period of $600 \pm 134$ d. They also used the flare power formulation proposed by He et al. (2018) and found a periodicity of $550 \pm 101$ d. Their second method involved analyzing the effect of starspot modulation on KIC 10544976's light curve. They normalized the long-cadence PDCSAP data and applied a multiplicative factor to account for each quarter's different pixel mask. They then removed outliers using a sigma clipping algorithm after removing flares, eclipses, and the 372.5-day trend due to *Kepler*'s orbital period. They determined that the stellar activity – as defined by Basri et al. (2010) – had two frequency peaks at $585 \pm 92$ d and $694 \pm 130$ d, while the method of Montet et al. (2017) using *Kepler* full-frame images indicated a periodicity of $544 \pm 88$ d. Both methods therefore indicated a MAC period of ~600 d, which is inconsistent with the observed ETV. Furthermore, using the results of Völschow et al. (2018), they found that the red dwarf component does not have sufficient energy to cause the observed ETV via the Applegate mechanism. They therefore conclude that the ETV signal is caused by a third body with an orbital period of $16.83 \pm 2.4$ yr, an eccentricity of $0.29 \pm 0.11$, and a minimum mass of 13.4 $M_J$.

Liakos (2017) presents an asteroseismological study of four Algol-type eclipsing

binaries containing a $\delta$ Scuti component. One of the four systems he studied was the marginal sample system KIC 11175495. He obtained low-resolution ($R \sim$ 1,000) spectroscopic data on these four objects with the 2.3-meter Aristarchos telescope in Greece. These spectra covered wavelengths between 4,000 and 7,260 Å and were used for spectral classification. He then modeled the short-cadence *Kepler* data for each system using PHOEBE v0.29d, using the Cousins R filter as a proxy for the *Kepler* bandpass. He found that all four systems were semi-detached, with the secondary component filling its Roche lobe. He assumed the primary component's mass using the spectral type-mass relations given in Cox (2000), with the secondary masses following from $q$. For KIC 11175495, he derived the following parameters: $T_1 = 7{,}550 \pm 100$ K, $T_2 = 3{,}600 \pm 55$ K, $M_1 = 1.70 \pm 0.17$ $M_\odot$, $M_2 = 0.49 \pm 0.05$ $M_\odot$, and $i = 68.6 \pm 0.2°$. He then studied the pulsations in each system using the software PERIOD04 v1.2 (Lenz & Breger 2005), which probes pulsational frequencies using standard Fourier analysis. Liakos (2017) searched for pulsations in the range of 0-80 cycles per day. KIC 11175495 showed two clusters of pulsational frequencies, one between 0-12 $d^{-1}$ and another between 55-70 $d^{-1}$. The existence of pulsational frequencies below 5 $d^{-1}$ in all four systems indicates that they may be hybrid $\delta$ Scuti-$\gamma$ Doradus variables. Such hybrids are discussed further in Balona (2014). Liakos (2017) concludes by noting that KIC 11175495 is the youngest of the four systems he studied, and its pulsations have the highest frequency ever discovered in a binary system.

# Chapter 4

# Results and Analysis

I have determined a wealth of information on my sample using the methods I discussed in Chapter 2. In this chapter, I present this information and discuss the implications of my findings. I first give an overview of the characteristics of the core sample, including those I determined myself and those I found from other sources. Next, I look at the trends, distributions, and correlations of seven characteristics: the OES/|OES|, primary eclipse depth, morphology parameter, distance, temperature/*Gaia* color, luminosity/*Gaia* absolute magnitude, and period. Through these characteristics, I discuss what the trends, distributions, and correlations imply about the core sample. Finally, I present a similar, if more perfunctory, analysis of the marginal sample characteristics, trends, distributions, and correlations.

## 4.1 Sample Characteristics

I now present a detailed summary of the characteristics of the core sample. I focus on the light curve characteristics – such as the distributions of the eclipse depths,

OES, and positive and negative OESs – and the physical characteristics – such as the spatial, luminosity, temperature, and period distributions. Studying these distributions allows me to see trends among them and how they correlate with each other, which provides insight into the physical causes of the O'Connell effect. I also present my results from analyzing the ETV, the OER, and the LCA. I use the comparative subset of the KEBC discussed in Section 2.1.2 in this chapter for creating the histograms, color-magnitude diagram, and Hertzsprung–Russell diagram. The comparative subset excludes four systems: KIC 5217781, a long-period system with severe data corruption, KIC 7667885, due to it not being in the KEBC (as mentioned in Section 2.1.4), and KICs 7950964 and 9137819, due to their removal from the KEBC (as explained in Section 2.1.5).

### 4.1.1 Light Curve Characteristics

I used the *Kepler* photometry from the KEBC to determine several characteristics of each system's light curve. The characteristics I found were the light curve class (using the method described in Section 2.2.2.2), the OES and |OES| (using the method described in Section 2.1.3.2), and the primary eclipse depth (using the method described in Section 2.2.1). Using these characteristics, I also determined the proportion of systems displaying a positive OES and a negative OES along with the ratio |OES|/primary eclipse depth. This section discusses the range and distribution of these characteristics. The core sample contains 54 (25%) Algol-type, 40 (19%) $\beta$ Lyrae-type, and 118 (56%) W Ursae Majoris-type systems. This distribution is markedly different from the KEBC's distribution: 1,990 (68%) Algol-type, 250 (9%) $\beta$ Lyrae-type, and 679 (23%) W Ursae Majoris-type systems. The much larger abundance of Algol-type binaries in the KEBC results from the long-period

($P > 10$ d) systems that are not in the core sample. These systems are largely Algol-type due to the large separation between the components.

The largest OES found in the core sample is KIC 11347875, which is the exemplar of an unusual class of systems I discuss further in Section 5.4. These systems do not have an O'Connell effect in the traditional sense described in Section 1.2 because they lack two clear maxima to calculate a difference between due to displaying a concave-up region in their light curves. My program gives an OES of $-0.265$ for KIC 11347875 in units of normalized flux. The largest OES in systems with a traditional O'Connell effect is KIC 9935311 with a value of 0.115. A clear majority of systems (147, or 69% of the core sample) show a positive O'Connell effect, where the brighter maximum is the one following the primary eclipse. Since a naïve expectation would be that around 50% of the sample would show a positive O'Connell effect, this proportion suggests that the O'Connell preferentially causes the brighter maximum to follow the primary eclipse. Differences in the characteristic trends between positive and negative O'Connell effect systems will be a recurring theme of my analysis.

The primary minimum depth ranges from 0.025 (KIC 8190491) to 0.917 (KIC 9101279) in units of normalized flux. Figure 4.1 shows a histogram of the ratio between |OES| and the primary eclipse depth. I found that the overwhelming majority of systems with a ratio above 0.5 are so-called "heartbeat" stars (Thompson et al. 2012), which are stars in an eccentric orbit that become tidally distorted near periapsis and have tidally-induced pulsations. While tidal distortions at periapsis were the original explanation for the O'Connell effect Roberts (1906) put forth, I do not consider these systems to have an O'Connell effect for the purposes of my study because the flux difference caused by these distortions is much smaller than

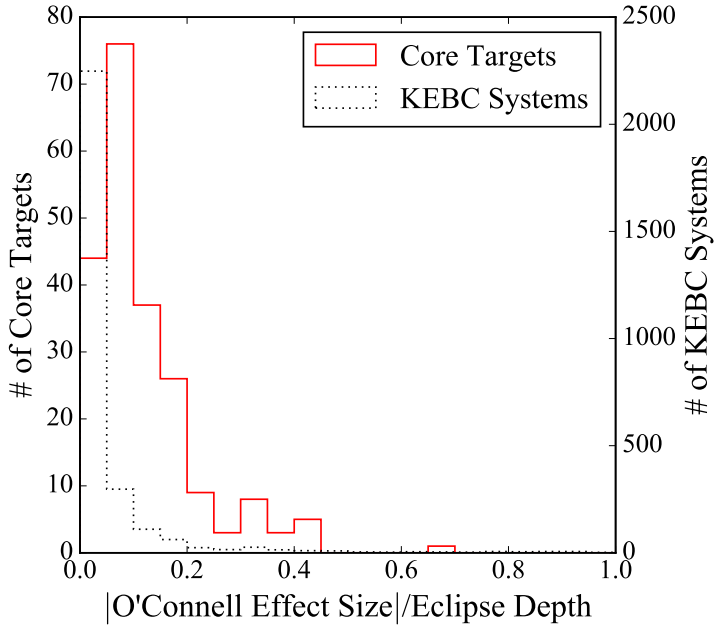


Figure 4.1: Histogram showing the ratio |OES|/primary eclipse depth for the core sample (solid red) and all systems in the KEBC (dotted grey). Systems with a ratio above 1 are not displayed. Core sample systems have a larger ratio on average than other KEBC systems.

the threshold for inclusion in my sample. Additionally, these tidal distortions cannot explain the effect seen in the circular systems comprising my sample. The core sample system among the high ratio systems in Figure 4.1 is the aforementioned non-heartbeat star KIC 11347875, with a ratio of 0.657.

### 4.1.2 Physical Characteristics

I determined several physical characteristics from the published information on the KEBC. These physical characteristics were distance, luminosity, period, temperature, and spectral type. This section focuses on these characteristics' distributions for both the core sample and the KEBC. I also discuss the core sample's color-magnitude and HR diagrams and what they say about the sample's stellar distribution.

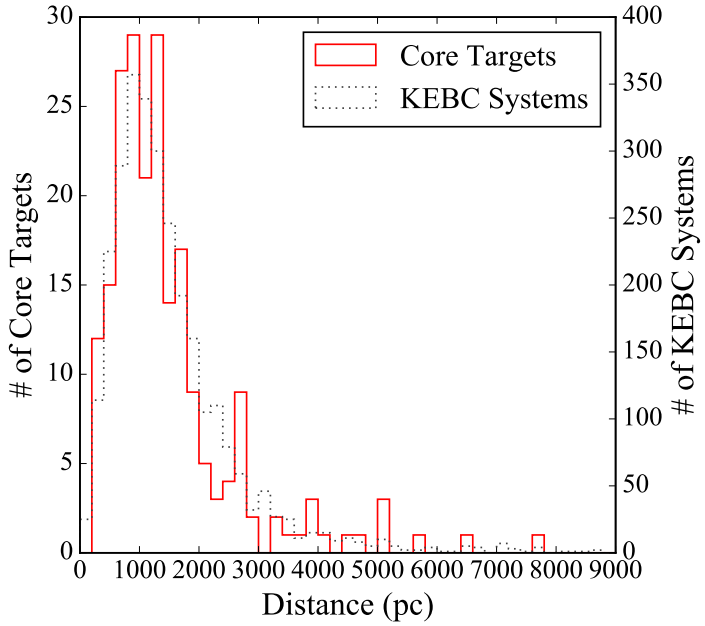


Figure 4.2: Histogram showing the implied distances for the 212 core sample targets (solid red) and 2,860 of 2,862 entries in the KEBC (dotted grey) with *Gaia* EDR3 parallaxes (excluding two systems with $d \geq 9{,}000$ pc). The two distributions are similar.

Figure 4.2 shows a histogram of the core sample's distance distribution determined from *Gaia* EDR3 parallaxes (Gaia Collaboration et al. 2021; Bailer-Jones et al. 2021), along with the distance distribution for 2,860 of the 2,862 KEBC entries with *Gaia* EDR3 parallaxes (excluding two systems with $d \geq 9{,}000$ pc for clarity). The two distributions are similar over the entire range of distances, a fact reinforced by the statistical analysis I discuss in Section 4.2.2. The similarity of the distributions implies that the O'Connell effect is independent of a binary's spatial position. Such a result is expected, assuming the O'Connell effect arises from the binary rather than a phenomenon occurring along the same line of sight. The core sample system's distances range from $210 \pm 1$ pc (KIC 7671594) to $7{,}743^{+1{,}972}_{-1{,}933}$ pc (KIC 4474637).

I used Bailer-Jones et al.'s (2021) distance estimate $d$ and the apparent *Gaia* G magnitude $\mathrm{g}_*$ to calculate the total luminosity $L_*$ in the *Gaia* G-band for each system using the luminosity equation:

$$L_* = 10^{0.4[\mathrm{G}_\odot - \mathrm{g}_* + 5\log(d) - 5]} \tag{4.1}$$

where $L_*$ is in units of solar luminosities ($\mathrm{L}_\odot$) and $\mathrm{G}_\odot$ is the G-band absolute magnitude of the Sun, which I assumed to be $4.68 \pm 0.02$ (Čotar et al. 2019). Error propagating Equation 4.1 using Equation 2.39 gives:

$$\sigma_{L_*} = \sqrt{L_*^2 (0.4 \ln 10)^2 \left(\sigma_{\mathrm{G}_\odot}^2 + \sigma_{\mathrm{g}_*}^2\right) + 4 \cdot 10^{0.8(\mathrm{G}_\odot - \mathrm{g}_* - 5)} d^2 \sigma_d^2} \tag{4.2}$$

The distance errors $\sigma_d$ were calculated by subtracting Bailer-Jones et al.'s (2021) $+1\sigma$ and $-1\sigma$ distances from the distance estimate $d$ and taking the absolute value of this difference. The two errors in distance give two errors in luminosity. Figure 4.3 shows a histogram of the luminosity distribution in the core sample and the 2,862 KEBC entries with *Gaia* EDR3 parallaxes, indicating that systems in the sample are less luminous on average than others in the KEBC. The luminosity I calculated is the combined luminosity of each system, including both components of the eclipsing binary and any potential companions unresolved by *Gaia*. Therefore, one way to explain the core sample's lower average luminosity is that the secondary component is less luminous than average, or that there is a lower rate of unresolved companions in the sample. Another explanation could be that there are fewer evolved stars in the sample than in the KEBC, which would argue against a connection between stellar evolution and a significant O'Connell effect. Finally, the stars in the sample may have a later spectral type on average than

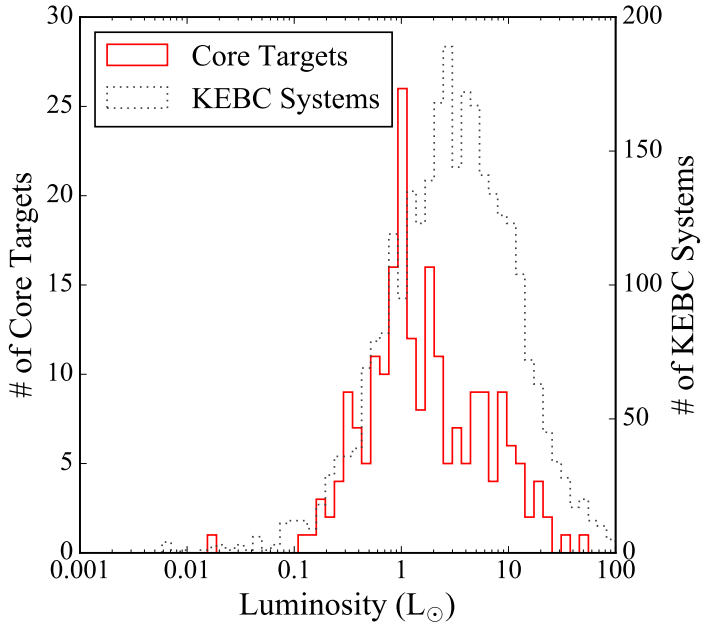


Figure 4.3: Histogram comparing the luminosities for the 212 core sample targets (solid red) and all 2,862 entries in the KEBC (dotted grey) with *Gaia* EDR3 parallaxes. Note the logarithmic scale of the x-axis. Core sample systems are less luminous on average than other KEBC systems.

the KEBC, which would mean lower luminosities for stars on the main sequence. Looking ahead to the temperature histogram in Figure 4.6 reveals that the core sample systems are slightly cooler on average than the KEBC systems, lending support to the last hypothesis. The core sample system's luminosities range from $0.0169 \pm 0.0004$ L$_\odot$ (KIC 7671594) to $50.25^{+5.30}_{-5.41}$ L$_\odot$ (KIC 5820209).

The orbital periods of the systems were taken directly from the KEBC and range from 0.234 d (KIC 6050116) to 9.752 d (KIC 6197038). Figure 4.4 shows a histogram of the period distribution for the core sample and all 2,915 KEBC entries with $P \leq 1{,}000$ d. The distributions between the KEBC and the core sample are similar for $P < 0.5$ d, but the population of systems with longer periods declines much more sharply for the core sample. For instance, the core sample comprises

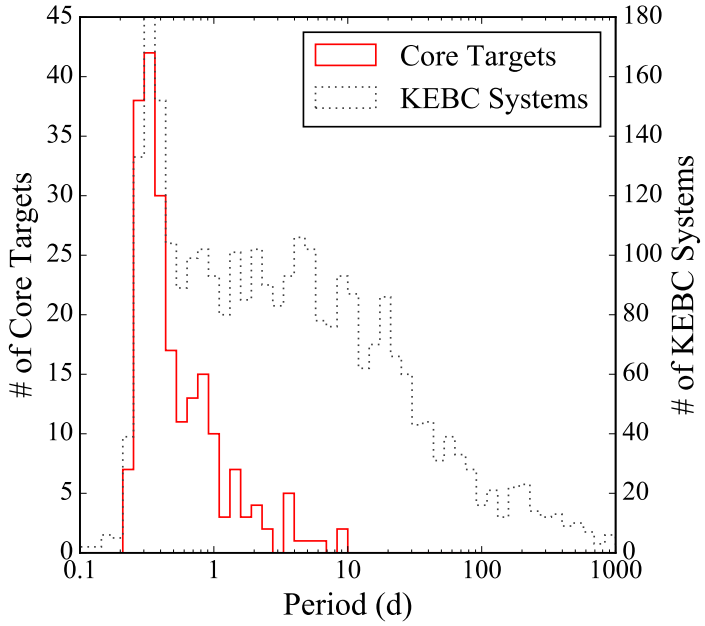


Figure 4.4: Histogram comparing the periods for all 212 core sample targets (solid red) and all 2,915 entries in the KEBC (dotted grey) with $P \leq 1{,}000$ d. Note the logarithmic scale of the x-axis. The distributions are similar under 0.5 d, but the core sample's long-period population decreases much more quickly than the KEBC's.

about a fifth of the KEBC for $P \approx 0.5$ d, but only a ninth for $P \approx 1$ d and less than 3% for $P \approx 10$ d. No core sample system has $P > 10$ d. Therefore, Figure 4.4 implies that the O'Connell effect is strongly associated with short-period systems, which also implies a strong association between the O'Connell effect and stars that are in close proximity to each other.

Both the KEBC and *Gaia* DR2 (Gaia Collaboration et al. 2018) provide temperatures for most of the sample. The core sample system's *Kepler* temperatures range from 3,717 K (KIC 7671594) to 8,540 K (KIC 10857342), while the *Gaia* temperatures range from $3{,}808^{+580}_{-413}$ K (KIC 7671594) to $8{,}540^{+218}_{-229}$ K (KIC 10857342). Nine core sample systems do not have a *Kepler* temperature, while seven do not have a *Gaia* temperature. Figure 4.5 compares the *Kepler* and *Gaia* temperatures

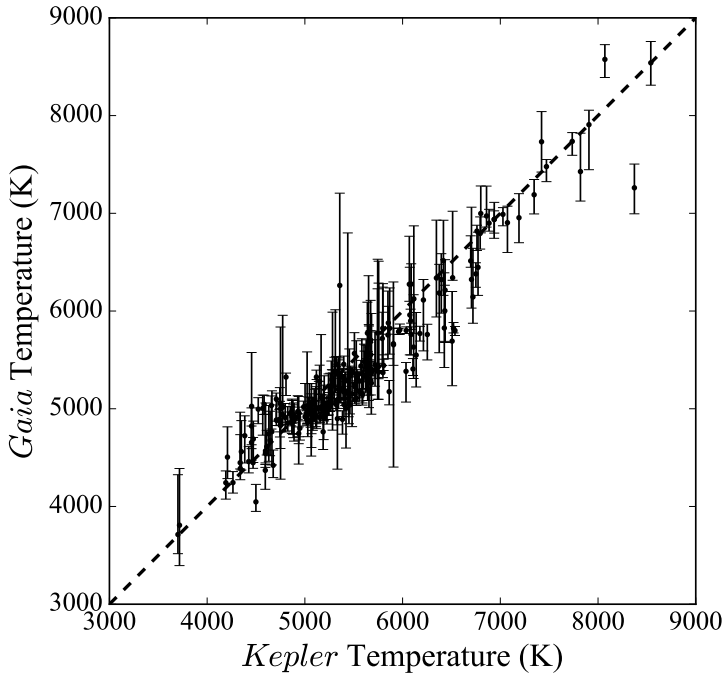


Figure 4.5: Plot comparing the *Kepler* temperatures to the *Gaia* DR2 temperatures for 240 of the 258 complete (core and marginal) sample targets. Error bars are plotted for the *Gaia* temperatures, but no errors were provided for the *Kepler* temperatures. The dashed line indicates where the two temperatures are equal.

for all 240 complete sample systems with both temperatures, demonstrating that most systems in the sample have similar *Kepler* and *Gaia* temperatures. I chose to use *Gaia* as my primary source of temperature because of the more complete coverage *Gaia* provides and the fact that it provides errors. Figure 4.6 shows a histogram of the temperatures for the 205 systems in the core sample and all 2,825 KEBC entries with a *Gaia* temperature. The distributions between the sample and the KEBC are similar below ~5,500 K, but the core sample has fewer hotter systems than the KEBC. Combined with the smaller average luminosity Figure 4.3 shows, this implies that the O'Connell effect is more common in binaries with late-type stars. The discrete peaks present in both populations are known artifacts of

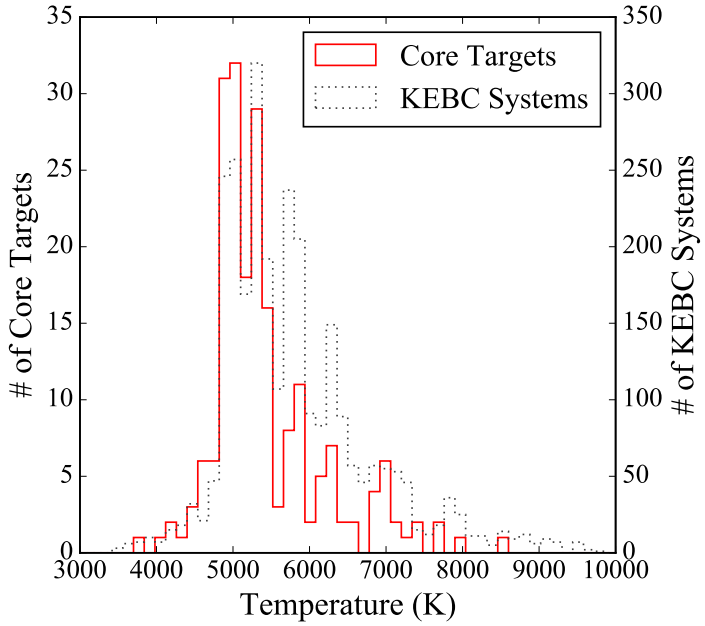


Figure 4.6: Histogram comparing the *Gaia* DR2 temperatures for 205 of the 212 core sample targets (solid red) and all 2,825 entries in the KEBC (dotted grey) with a *Gaia* DR2 temperature. Core sample systems are slightly cooler on average than other KEBC systems.

the *Gaia* DR2 data caused by the inhomogeneous temperature distribution of the training sample used to produce the temperature values (Andrae et al. 2018).

Frasca et al. (2016) provides spectral types for 31 core systems, which the authors determined using optical spectra. They did not establish whether they assumed solitary stars during their analysis, however, so the faster rotation of these close binaries may have broadened the spectral lines sufficiently to misidentify these stars' spectral types. Nevertheless, the 20 reported main sequence spectral types range from K5 (KIC 12109575) to A2 (KIC 8904448). The 11 remaining systems have an explicit non-main sequence spectral type, as listed in Table A.1. Ramsay et al. (2014) spectrally classified KIC 7667885 and KIC 9786165 as "mid

G" systems. I used the online table[1] of photometric colors periodically updated by Dr. Eric Mamajek (Pecaut & Mamajek 2013) in combination with the *Gaia* BP – RP colors provided by EDR3 to expand the spectral classification to the entire target sample. I did this by assigning a given system the spectral type listed in Dr. Mamajek's table that most closely corresponded to its EDR3 color. While this process does not account for interstellar reddening, evolutionary stage (Dr. Mamajek's table lists only luminosity class V spectral types), or color blending of the binary components, it provides an estimate of the spectral class for all systems in the sample, which Table A.1 also lists. By this estimate, the spectral types of the sample range from M2.5 V (KIC 7671594) to A7 V (KIC 10857342).

Figures 4.7 and 4.8 show the core sample's color-magnitude and Hertzsprung-Russell (HR) diagrams, respectively, using *Gaia* EDR3 data. The marker color and shape in Figures 4.7 and 4.8 indicate the light curve classification for each system determined by the criteria given in Section 2.2.2.2. I calculated the absolute magnitude $G_*$ for each system using:

$$G_* = g_* - 5\log(d) + 5 \tag{4.3}$$

with the error given by:

$$\sigma_{G_*} = \sqrt{\sigma_{g_*}^2 + \left(\frac{5}{\ln 10}\frac{\sigma_d}{d}\right)^2} \tag{4.4}$$

Figures 4.7 and 4.8 show the *Kepler* target selection bias discussed in Batalha

[1] `http://www.pas.rochester.edu/~emamajek/EEM_dwarf_UBVIJHK_colors_Teff.txt`

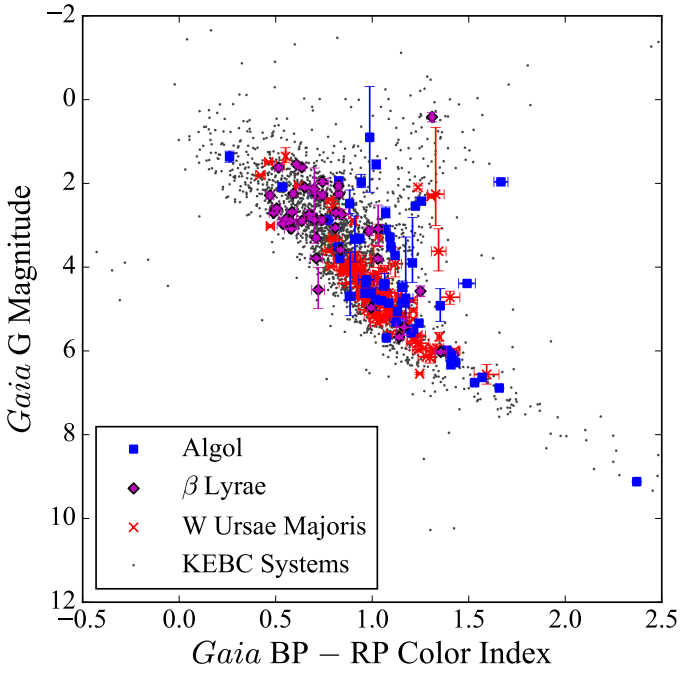


Figure 4.7: Color-magnitude diagram showing the 212 systems in the core sample overplotting the rest of the KEBC systems. The *Gaia* color is plotted against the absolute magnitude in *Gaia*'s G band. The systems are color coded according to their light curve class.

et al. (2010) and Section 2.1.1, as there are few systems with BP − RP $\geq$ 1.5 (or $T$ < 4,000 K) or BP − RP $\leq$ 0.0 (or $T$ > 9,000 K). Indeed, the core sample only has one clear red dwarf system (KIC 7671594) and one clear A-type system (KIC 10857342), and the paucity of both stellar types is likely a consequence of *Kepler*'s selection function. The fact that there are only two clear giants in the sample (KIC 5820209 and KIC 9489411) is likewise a consequence of the selection function[2]. However, the sample's distributions in Figures 4.7 and 4.8 are similar

[2]It should be noted that KIC 5820209 has a period of only 0.656 d. Two 1 $M_\odot$ stars with this period imply a maximum star size of ~2.5 $R_\odot$, while a star of KIC 5820209's effective temperature (4,828 K) would need to have a radius of ~10.1 $R_\odot$ to match its estimated luminosity of 50.25 $L_\odot$.

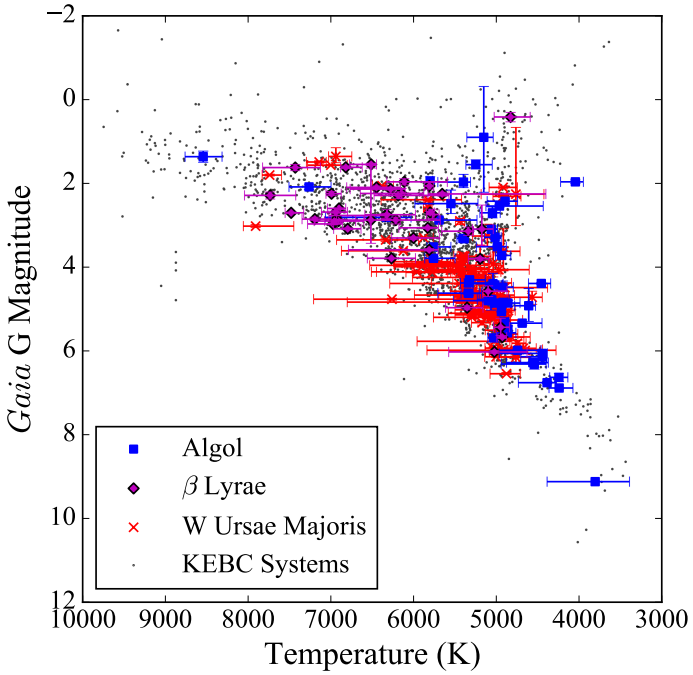


Figure 4.8: HR diagram showing 205 of the 212 systems in the core sample overplotting the rest of the KEBC systems. The *Gaia* temperature is plotted against the absolute magnitude in *Gaia*'s G band. The systems are color coded according to their light curve class.

to the KEBC's along most of the main sequence, suggesting that the sample is not strongly biased beyond *Kepler*'s selection function.

The systems' morphology parameters were taken directly from the KEBC and range from 0.34 (KIC 7671594) to 0.97 (KICs 6264091, 7199183, and 9527167). Figure 4.9 shows a histogram of the morphology parameter distribution for 211 core sample systems (excluding KIC 7667885, as discussed in Section 2.1.4) and all 2,745 KEBC entries with $\mu \neq -1$ (recall from Section 2.2.3 that systems with $\mu = -1$ could not have a valid $\mu$ assigned to them). Figure 4.9 clearly shows three distinct peaks in the distribution at $\mu \approx 0.60, 0.80$, and 0.95. According to Matijevič et al. (2012), the first two peaks occur near the transitions between detached and semi-

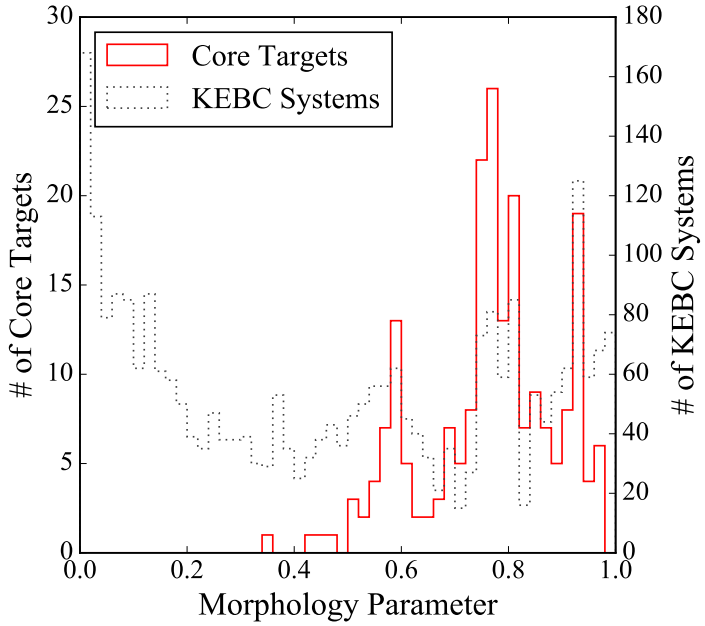


Figure 4.9: Histogram showing the morphology parameters for 211 of the 212 core sample targets (solid red) and all 2,745 entries in the KEBC (dotted grey) that have $\mu \neq -1$. KIC 7667885 is not included because it lacks a value for $\mu$ (see discussion in Section 2.1.4). The two distributions differ significantly, particularly for $\mu < 0.50$. The core sample also has few systems with $0.60 \leq \mu \leq 0.70$.

detached systems (at $\mu = 0.50$) and between overcontact systems and ellipsoidal variables (at $\mu = 0.80$), respectively. The first two peaks also occur near the transitions between Algol- and $\beta$ Lyrae-type systems (at $\mu = 0.60$) and between $\beta$ Lyrae- and W Ursae Majoris-type systems (at $\mu = 0.75$), respectively. These facts suggest that something related to the transition between these different system types may also be related to the O'Connell effect. The third peak in Figure 4.9 does not occur near any transition.

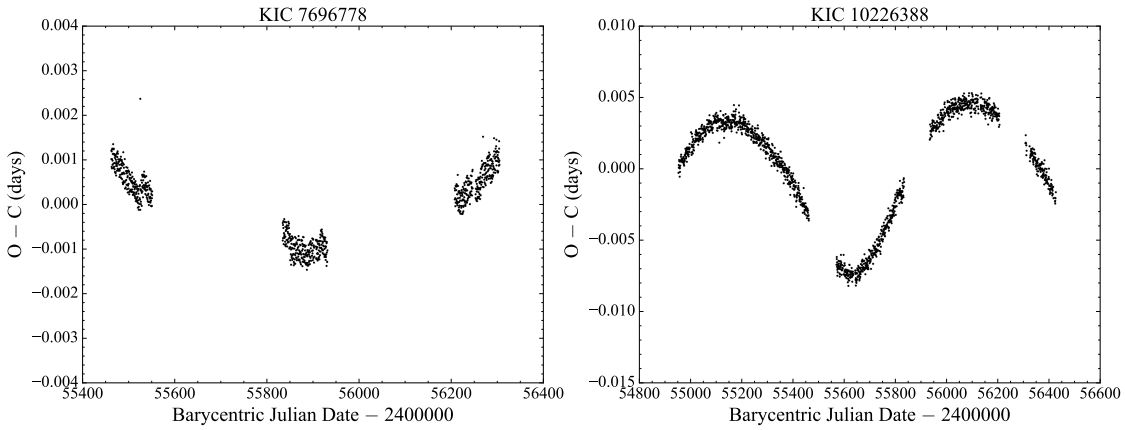


Figure 4.10: Eclipse timing variation of KIC 7696778 (left) and KIC 10226388 (right). KIC 7696778's ETV has a parabolic shape, consistent with a linear change in period caused by mass transfer between components, while KIC 10226388's ETV has a sinusoidal shape, consistent with an LTTE caused by a third body orbiting the binary.

### 4.1.3 Eclipse Timing Variation

I looked at the ETV of systems flagged with one of the three flags defined in Section 2.2.4 (TM, TT, and TI) to study the systems' evolution with time. In this section, I discuss the proportion of systems with each type of ETV. I focus particularly on systems with ETVs indicative of mass transfer due to ETVs being the only method of testing the mass transfer hypothesis with the data I have. To that end, I study the ETVs of the rest of the core sample to determine how many systems have potentially parabolic ETVs.

Only three systems (KICs 5020034, 6044064, and 7696778; 1.4% of the core sample) show the parabolic ETV signal associated with mass transfer. Forty systems (19.0%) show a sinusoidal signal, while forty-seven systems (22.7%) show an ETV classified as "interesting." These proportions seem to imply that mass transfer occurs in only a very small proportion of O'Connell effect binaries, and therefore cannot account for the O'Connell effect in most systems. Figure 4.10's

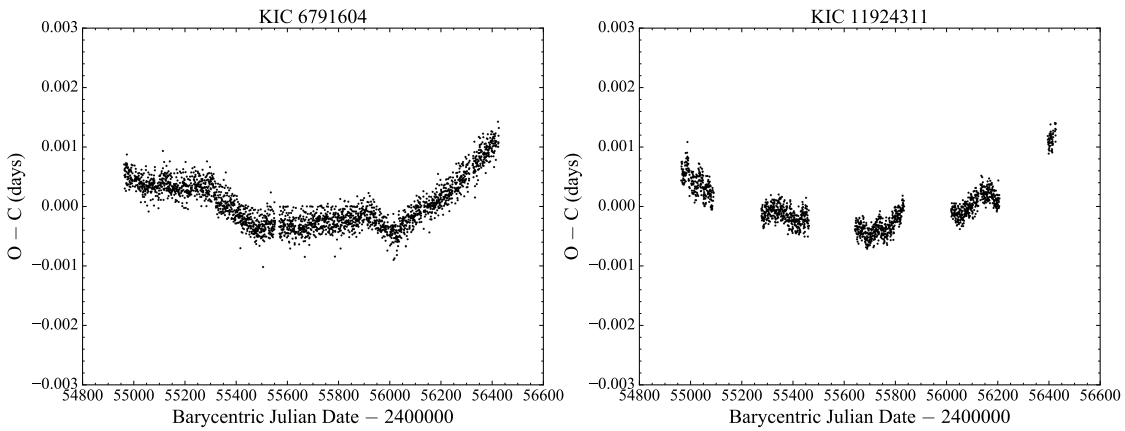


Figure 4.11: Eclipse timing variation of KIC 6791604 (left) and KIC 11924311 (right). KIC 6791604 was flagged with an interesting ETV while KIC 11924311 was flagged as a sinusoidal ETV, but both are consistent with a roughly parabolic ETV as well.

left panel shows KIC 7696778's ETV, demonstrating its parabolic nature, while Figure 4.10's right panel shows KIC 10226388's sinusoidal ETV.

The KEBC may have underestimated the proportion of mass transfer in O'Connell effect binaries, however. Kouzuma (2018; see Section 3.3 for further discussion) cited twenty-two core sample systems as undergoing mass transfer. While these twenty-two systems do not include the three systems flagged with a parabolic ETV, the KEBC flagged sixteen of the twenty-two as having a sinusoidal or otherwise interesting ETV. I visually inspected these systems' ETVs and found that, aside from KIC 2437038, their signals could also be described as roughly parabolic. Figure 4.11 shows the ETVs of two of these sixteen systems, KICs 6791604 and 11924311. The KEBC flagged these systems as having an interesting and sinusoidal ETV, respectively. KIC 6791604's ETV is erratic, consisting of linear regions with several discontinuous slope changes, but it vaguely resembles a parabola. KIC 11924311's ETV more closely resembles a parabola, but the BIC Conroy et al. (2014) used indicated that a sinusoid gave a better fit. Figures 4.12, 4.13,

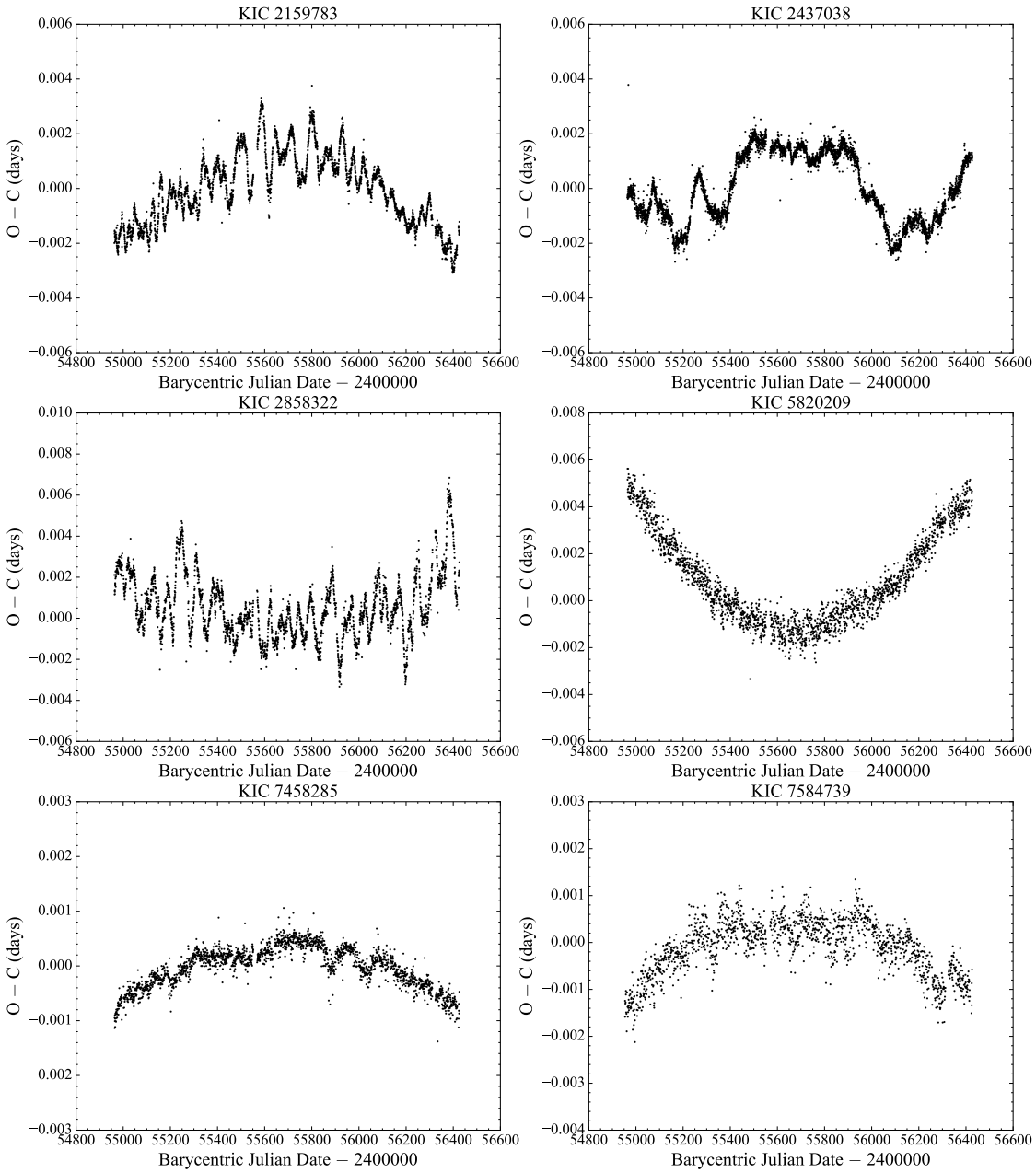


Figure 4.12: Eclipse timing variation of KICs 2159783, 2437038, 2858322, 5820209, 7458285, and 7584739. KICs 2159783, 5820209, and 7584739 were flagged with a sinusoidal ETV while KICs 2437038, 2858322, and 7458285 were flagged as interesting, but all except KIC 2437038 are consistent with a roughly parabolic ETV as well.

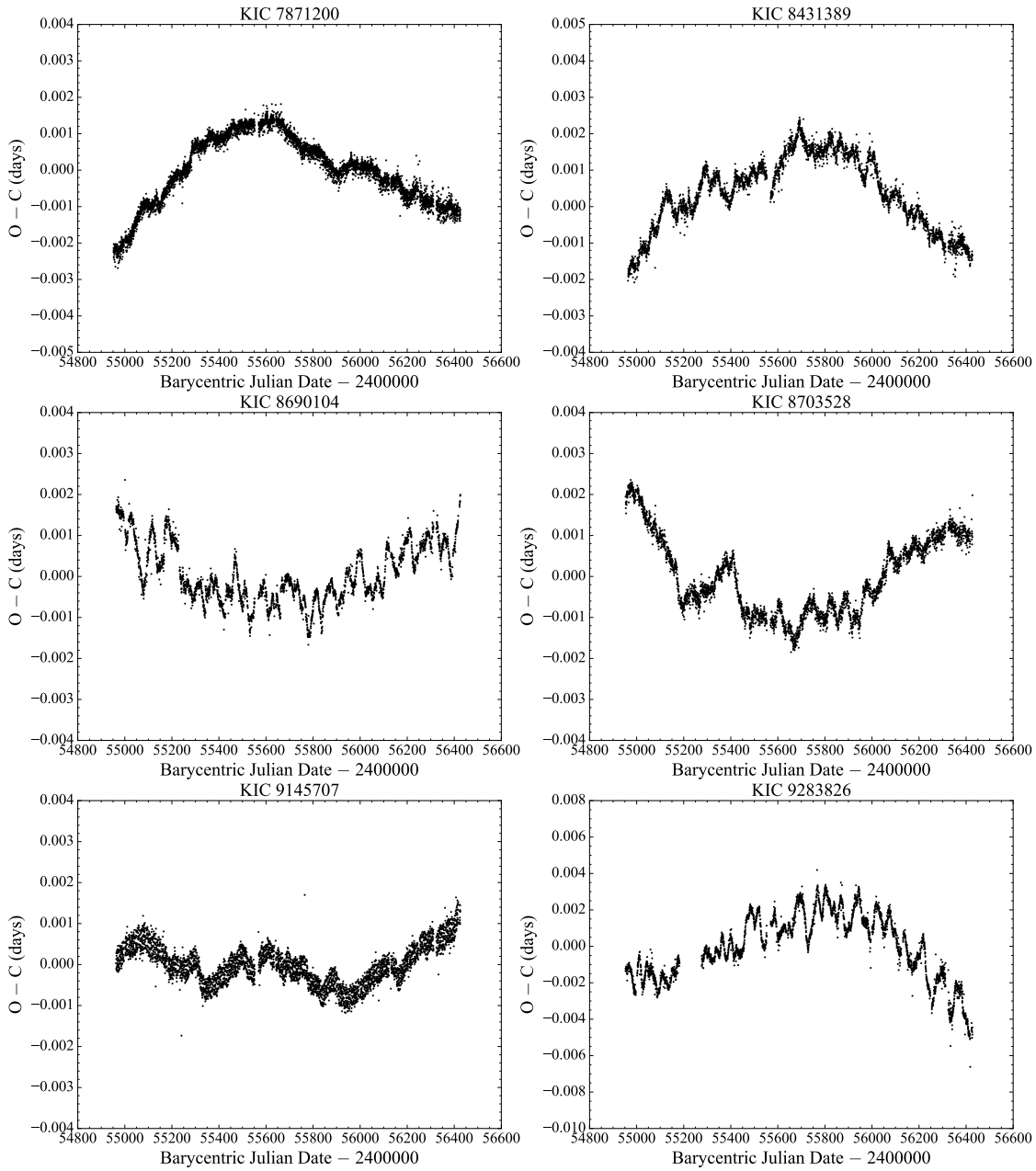


Figure 4.13: Eclipse timing variation of KICs 7871200, 8703528, 8690104, 9145707, and 9283826. KICs 7871200, 8690104, 8703528, and 9283826 were flagged with a sinusoidal ETV while KICs 8431389 and 9145707 was flagged as interesting, but all are consistent with a roughly parabolic ETV as well.

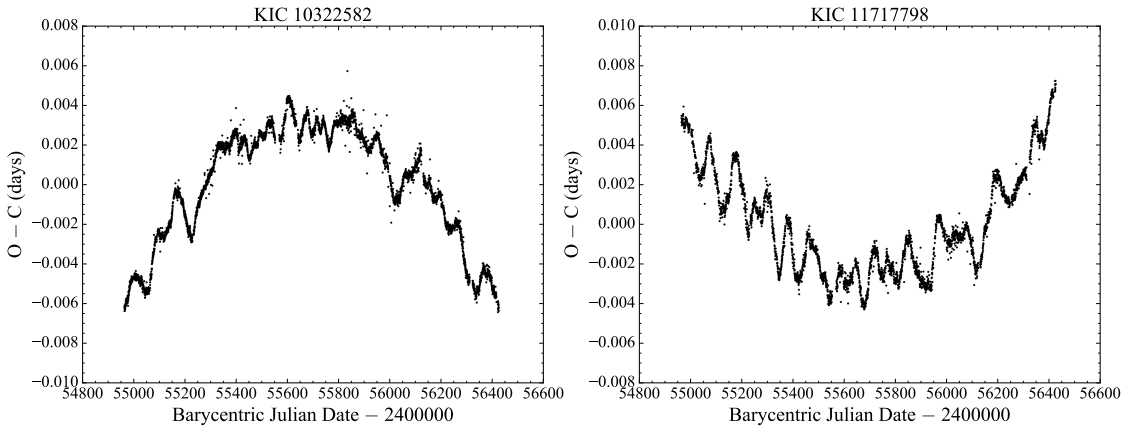


Figure 4.14: Eclipse timing variation of KICs 10322582 and 11717798. Both were flagged with a sinusoidal ETV but are consistent with a roughly parabolic ETV as well.

and 4.14 show the ETVs of the other fourteen systems cited by Kouzuma (2018) and flagged with either a sinusoidal or interesting ETV signal.

I note that Kouzuma (2018) only contained ~10% of the core sample. To better understand what fraction of the core sample had possibly parabolic ETVs, I investigated the core sample systems' ETVs using the Butterworth filter as I described in Section 4.1.3. This investigation showed that 13 systems (11% of the 122 investigated systems) have $R^2 \geq 0.9$, while 25 systems (20%) have a less-stringent $R^2 \geq 0.8$ (for reference, KIC 6791604's ETV – shown in Figure 4.11's left panel – has $R^2 = 0.895$). I regard these 25 systems as having an ETV that is not incompatible with being parabolic, although resolving whether these ETVs are truly parabolic is beyond the scope of my dissertation, as is resolving whether mass transfer causes these ETVs. Because these 122 systems form an unbiased subset of the core sample, I expect the same percentage of core sample systems to show a possibly parabolic ETV. Therefore, I estimate that up to 43 core sample systems have evidence of mass transfer in their ETV, making mass transfer a plausible, if

rare, phenomenon associated with a significant O'Connell effect.

### 4.1.4 O'Connell Effect Ratio and Light Curve Asymmetry

Figures 4.15 and 4.16 compare the OES to the OER (Equation 2.12) and the LCA (Equation 2.13), respectively. The data labeled "Non-Core Systems" include all KEBC systems not in the core sample, excluding only KIC 5217781. My initial analysis found many systems with an OES of 0 but OER $\neq 1$ or LCA $\neq 0$, producing significant vertical scatter in Figures 4.15 and 4.16. Systems in this scatter region of Figure 4.15 are low-amplitude (total change in flux $\Delta F \lesssim 0.03$) binaries with an OES that is large relative to their eclipse depth but small on an absolute scale. In Figure 4.16, the scatter region also includes long-period eccentric binaries. I removed a single data point from KICs 9701423 and 10614158 to avoid a non-physical negative OER, although the low-amplitude system KIC 6948480 retains a negative OER. I excluded the two systems with $|\mathrm{OES}| > 0.1$ (KICs 9935311 and 11347875) from Figures 4.15 and 4.16 for clarity. The figures do not plot errors because they are too small to see at this scale for most systems.

Figure 4.15 shows a strong correlation between the OER and OES. This correlation is unsurprising because both measures take into account the different amount of light under both maxima, albeit in different ways. McCartney (1999) notes that the presence of a constant, third source of light contaminating the data will reduce the OES, but the OER is not affected in such a manner due to subtracting the minimum flux value of the light curve away. Therefore, the OER provides a more consistent measure of the O'Connell effect in this respect. The OER calculation method also means that the OER will be significant for systems like KICs 8912911 and 10905824 with O'Connell effects that are large relative to their eclipse depth

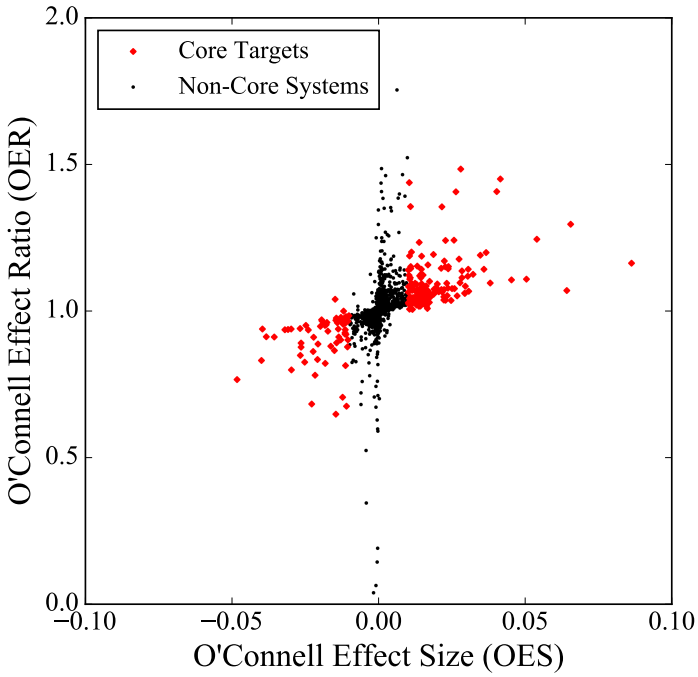


Figure 4.15: Plot comparing the OER to the OES. Core sample targets are shown in red diamonds and non-core sample systems in black circles. KICs 9935311 and 11347875 are excluded for clarity.

but small on an absolute scale. Based on this, I conclude that the OER is ill-suited to describe the O'Connell effect in a general population of eclipsing binaries but is applicable to subsets consisting of binaries with amplitudes larger than ~0.01 in normalized flux.

Figure 4.16 shows that most of the sample lies along lines with slopes of about $\pm 0.5$ that meet at the origin. Many systems lie above these lines, but none lie significantly below them, indicating that the O'Connell effect always produces an LCA at least half as large as the OES. The LCA is also clearly sensitive to asymmetries aside from the O'Connell effect, as evidenced by the large number of systems with a large LCA but small OES. Many of these systems, such as KIC 10909274, are long-period systems with eccentric orbits, implying that the LCA is sensitive

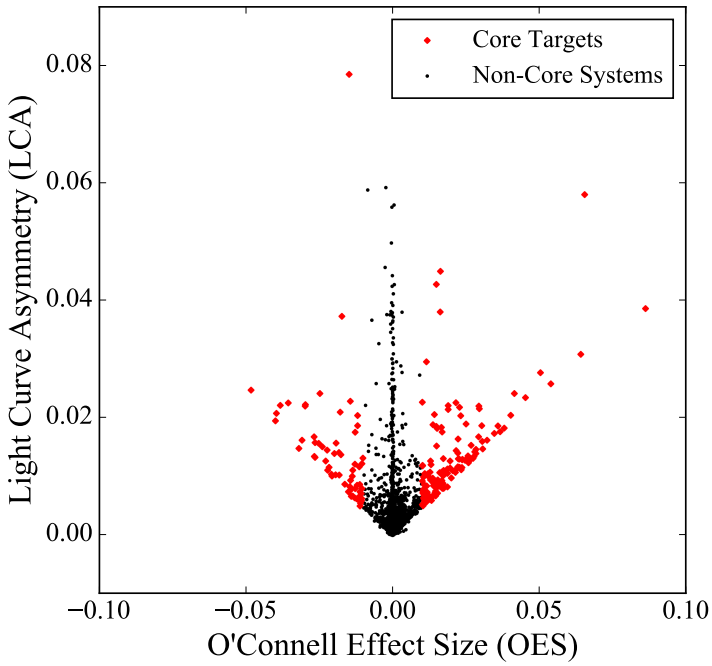


Figure 4.16: Plot comparing the LCA to the OES. Core sample targets are shown in red diamonds and non-core sample systems in black circles. KICs 9935311 and 11347875 are excluded for clarity.

to unevenly-spaced eclipses. The LCA therefore serves as a good indicator of how asymmetric a light curve is, but its sensitivity to other asymmetries means that it should not be relied on as an indicator of the O'Connell effect. Figure 4.16 indicates that the O'Connell effect (assuming that it causes an LCA equal to half the OES) is the most prominent asymmetry for most systems in the sample.

## 4.2 Statistical Analysis

In addition to studying the distributions of the sample's characteristics, I studied how those distributions compared to the KEBC's and how they correlated with each other. In this section, I present a graphical representation of the correlations

between the characteristics I studied (except distance and luminosity). I also discuss the results of my study of the characteristic distributions and correlations. Finally, I further discuss several correlations and their implications on the study of the O'Connell effect, focusing on correlations involving OES, eclipse depth, the morphology parameter, and period.

### 4.2.1 Characteristic Trends

Figures 4.17, 4.18, and 4.19 show corner plots comparing period, OES, *Gaia* temperature, the *Gaia* BP – RP color index, morphology parameter, absolute *Gaia* G magnitude, and primary eclipse depth. I chose the limits of the plots for clarity and, as such, the O'Connell effect size plots exclude KICs 9777984, 9935311, and 11347875. Several trends are visually apparent in Figures 4.17-4.19, most prominently the lower abundance of core sample systems with $0.6 \leq \mu \leq 0.7$ in the morphology parameter panels and the clustering of systems along the right edge of Figure 4.18's center panel. The difference in temperature distribution between positive and negative OES shown in Figure 4.17's center panel is also striking. I discuss these and other trends further in Section 4.2.3. The banding seen in the temperature plots (most prominently in Figure 4.19's right column) is due to the same issue with *Gaia* DR2 data that I discussed regarding Figure 4.6. The banding seen in the morphology parameter plots (most prominently in Figure 4.19's top-center panel) is a quantization effect due to the KEBC giving $\mu$ to only two decimal places.

Figures 4.17-4.19 and 4.20-4.22 use the trend subset of the KEBC discussed in Section 2.1.2, which excludes systems for two reasons: to analyze the core sample without raising Python errors and not have results be affected by invalid parameter

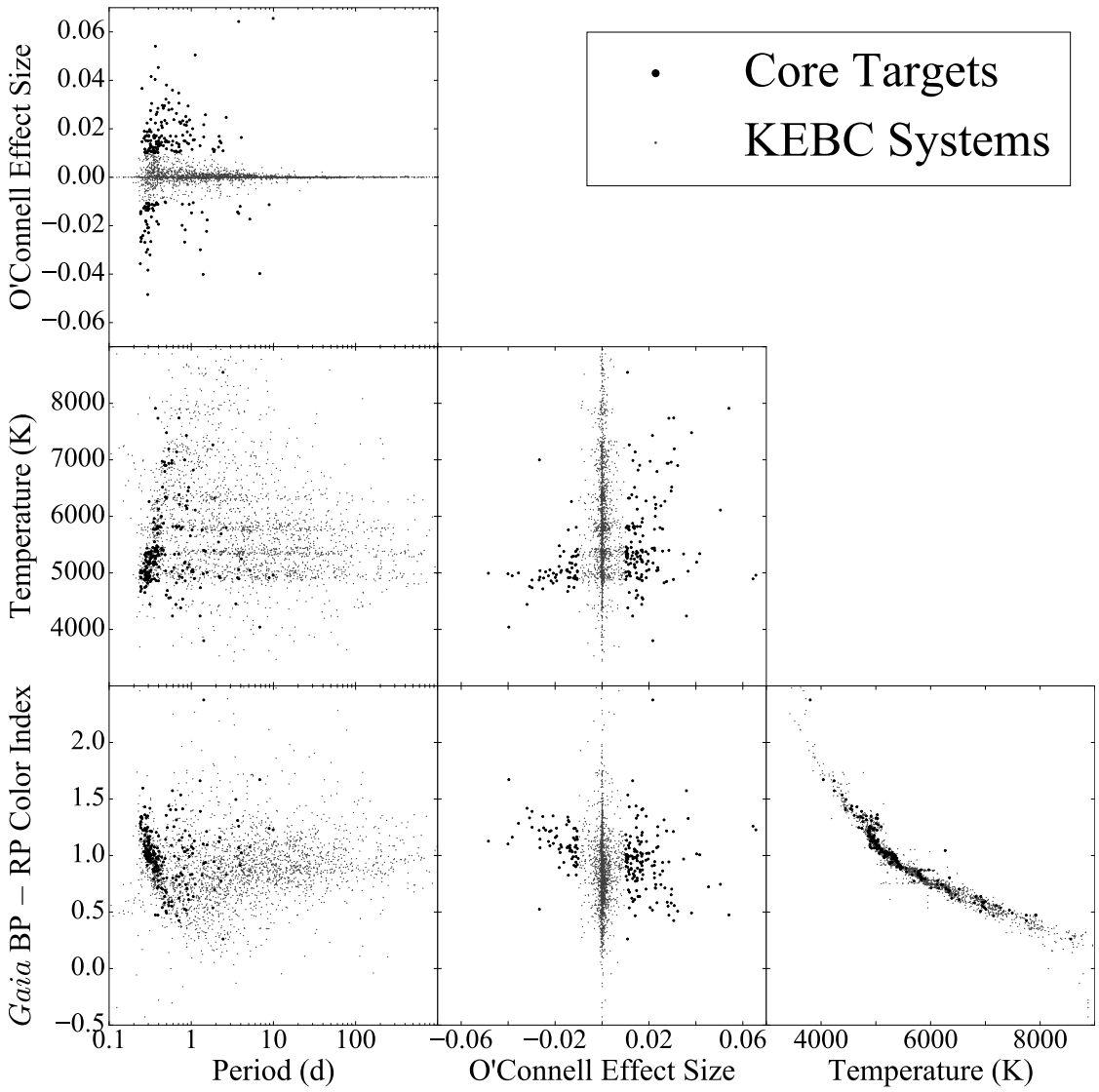


Figure 4.17: Corner plot showing the core sample correlations between three characteristics of interest (period, OES, and temperature) and OES, temperature, and color. Core sample targets are shown in black while non-core KEBC systems are shown in grey. A few outlier target systems have been removed from the OES plots for clarity (see discussion in text). Note the logarithmic x-axis of the period plots.

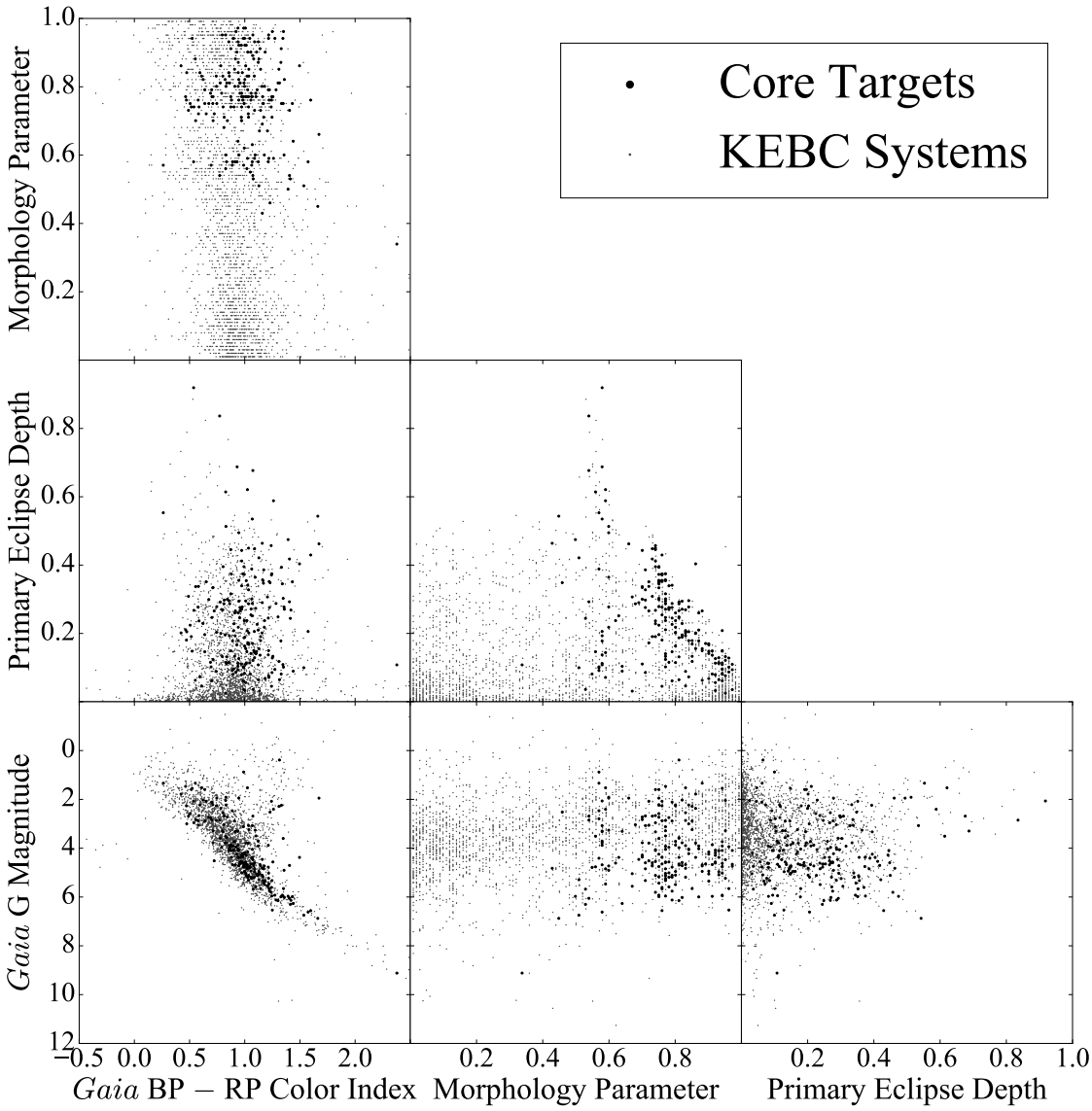


Figure 4.18: Corner plot showing the core sample correlations between three characteristics of interest (color, morphology parameter, and primary eclipse depth) and morphology parameter, primary eclipse depth, and absolute magnitude. Core sample targets are shown in black while non-core KEBC systems are shown in grey.

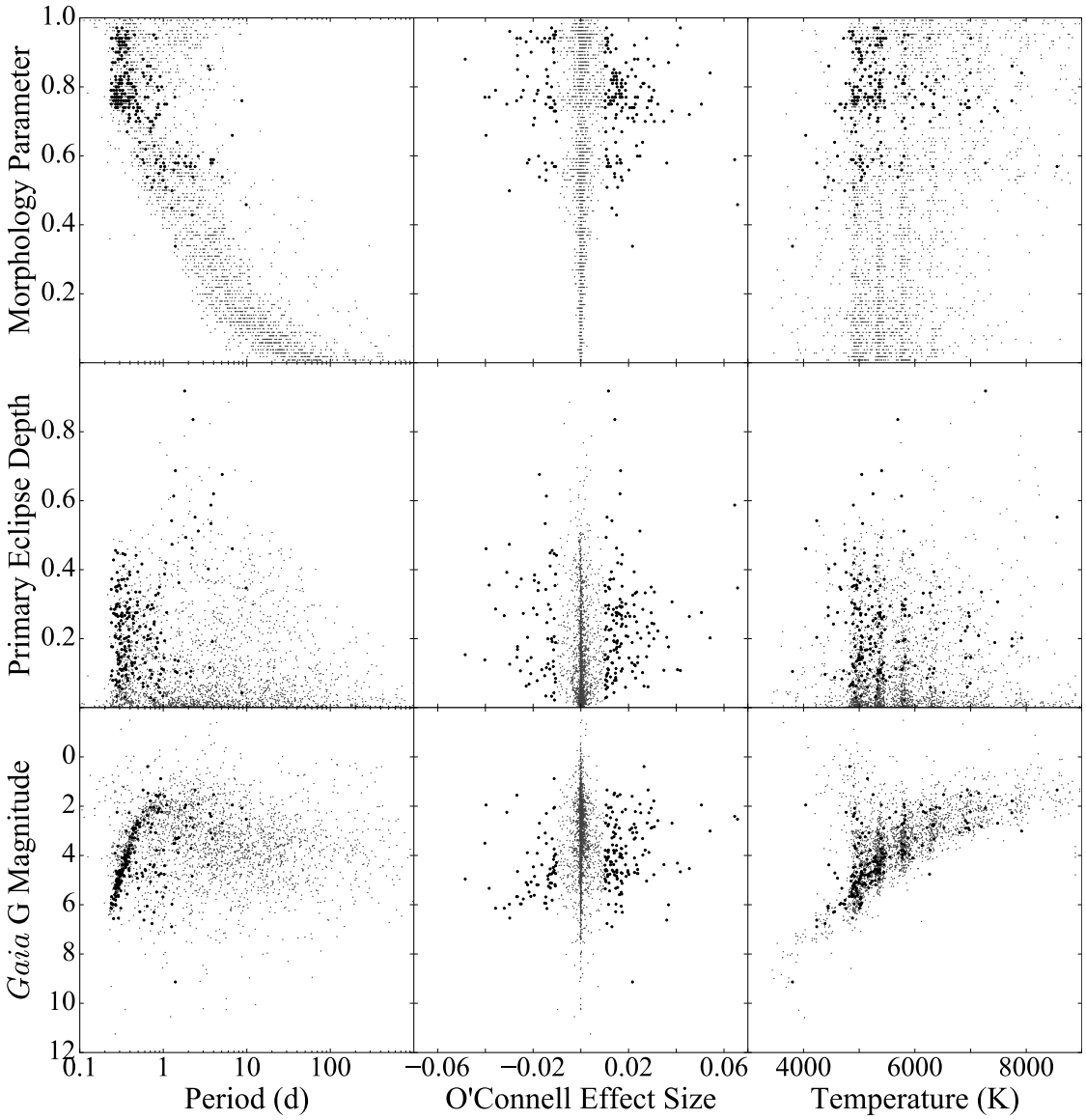


Figure 4.19: Corner plot showing the core sample correlations between three characteristics of interest (period, OES, and temperature) and morphology parameter, primary eclipse depth, and absolute magnitude. Core sample targets are shown in black while non-core KEBC systems are shown in grey. A few outlier target systems have been removed from the OES plots for clarity (see discussion in text). Note the logarithmic x-axis of the period plots.

Table 4.1: Results of analyzing the characteristic distributions between the core sample and the entire KEBC using the Kolomgorov–Smirnov test.

| Characteristic | K–S Statistic | $p$-Value |
|---|---|---|
| O'Connell Effect Size | 0.604 | $< 0.001$ |
| \|O'Connell Effect Size\| | 0.871 | $< 0.001$ |
| Primary Eclipse Depth | 0.398 | $< 0.001$ |
| Morphology Parameter | 0.242 | $< 0.001$ |
| Temperature | 0.277 | $< 0.001$ |
| Distance | 0.092 | 0.080 |
| Absolute Magnitude | 0.285 | $< 0.001$ |
| Period | 0.274 | $< 0.001$ |

values. I exclude a given system because either it had no *Gaia* parallax data (43 systems, plus 1 system without a known *Gaia* EDR3 identifier), it had multiple entries in the KEBC (11 systems, 24 total entries), or it had $\mu = -1$, indicating that the KEBC team could not assign a valid morphology parameter to the system (174 systems). I also explicitly exclude the same systems the comparative subset excludes: KICs 5217781, 7667885, 7950964, and 9137819. In total, these removals reduce the KEBC's size to 2,678 systems. While I remove these 228 systems (241 entries) from the KEBC analysis in Section 4.2.3, I do not remove any from the core sample analysis.

### 4.2.2 Sample Distributions

Table 4.1 gives the results of the K–S test I described in Section 2.2.5. The K–S test rejects the null hypothesis that the two samples were drawn from the same population at a level of 0.001 for every characteristic except distance. I expected that the OES and |OES| distributions would significantly differ because the core sample is looking at the wings of the distribution shown in Figure 2.2. Indeed, the OES differentiates the core sample from the KEBC to begin with. I also expected

the similarity in the distance distribution because there is no reason to suspect a distance dependence on the O'Connell effect. Instead, any biases based on spatial distribution should be due to the inherent biases present in the *Kepler* sample. The remaining characteristics in Table 4.1 are of greater interest as the results point to possible underlying connections between those characteristics and the O'Connell effect.

Table 4.1's results are consistent with the histograms presented in Section 4.1.2. Figure 4.2 shows similar distributions for both samples, strongly suggesting that the O'Connell effect is indeed distance-independent. Figure 4.3 shows that sample systems are less luminous on average than the KEBC systems, and based on that histogram and the one in Figure 4.6, the most likely explanation is that the core sample has a later spectral type on average. Additionally, cooler stars are more chromospherically active than hotter stars (Vaughan & Preston 1980), so the temperature distribution difference shown in Figure 4.6 and backed up by Table 4.1 is consistent with a link between the O'Connell effect and chromospheric activity. Figure 4.4 shows that the sample systems have a shorter period than the KEBC systems, indicating that binary interaction is an important factor in causing the O'Connell effect. Finally, Figure 4.9 shows that a significant O'Connell effect is not found in low morphology systems, further supporting the idea that binary interaction is important in causing the O'Connell effect. I discuss the morphology distribution differences further in Section 4.2.3.3.

### 4.2.3 Characteristic Correlations

My initial analysis of the KEBC found strong correlations between the OES and both period and the morphology parameter. The inclusion of many long-period

($P \geq 10$ d) and well-detached ($\mu \leq 0.3$) systems that are fundamentally different from the core sample's systems strengthened these correlations. These systems have a negligible OES, so there is a strong correlation between a range of periods and morphology parameters and an OES of zero. This correlation is prominently seen as a horizontal line in Figure 4.17's top-left panel for period and as a vertical line in Figure 4.19's top-center panel for morphology. These correlations reinforce the idea that long-period and well-detached systems do not display a significant O'Connell effect, but it does not give insight into the characteristics of systems that lack an O'Connell effect but are otherwise similar to those in the sample. As a result, I considered these initial correlations less relevant to my study. Therefore, the KEBC analysis uses the analysis subset of the KEBC discussed in Section 2.1.2. This subset excludes systems with $P \geq 10$ d (690 systems) or $\mu \leq 0.3$ (1,018 systems) and those excluded by earlier subsets, leaving only 1,639 KEBC systems. As in Sections 4.1 and 4.2.1, I remove no systems from the core sample analysis, and the KEBC population includes the core sample (again, except KIC 7667885). My analysis considers both OES and |OES|, as the former characteristic differentiates the positive and negative O'Connell effects while the latter characteristic does not.

Table 4.2 gives the results of my analysis using Spearman's $\rho$ coefficients, while Table 4.3 gives the results using Kendall's $\tau$ coefficients. Kendall's $\tau$ coefficients are uniformly smaller than Spearman's $\rho$ coefficients and indicate the same correlations. The coefficients indicate the correlation's strength, while the $p$-value indicates how compatible the results of the Spearman test are with the null hypothesis that the characteristics are uncorrelated. I consider two characteristics to be correlated when $|\rho| \geq 0.1$ – keeping consistent with D84 – and strongly correlated when $|\rho| \geq 0.2$. Several correlations found by my analysis, such as the one

Table 4.2: Results of analyzing the correlations between the core sample's characteristics of interest using Spearman's rank correlation coefficient. Bolded entries indicate correlated characteristics, i.e. $|\rho| \geq 0.1$.

| Characteristic | Characteristic | Core Sample | | | KEBC | | |
|---|---|---|---|---|---|---|---|
| One | Two | Coeff. | $p$-Value | Sur.[a] | Coeff. | $p$-Value | Sur.[a] |
| OES | Eclipse Depth | 0.033 | 0.628 | 13 | **0.124** | **$<0.001$** | **12** |
| OES | Morphology | **$-0.148$** | **0.032** | **16** | 0.024 | 0.326 | 0 |
| OES | Temperature | **0.385** | **$<0.001$** | **20** | 0.020 | 0.423 | 0 |
| OES | Distance | **0.218** | **0.001** | **17** | $-0.012$ | 0.641 | 0 |
| OES | Absolute Mag. | **$-0.286$** | **$<0.001$** | **19** | 0.033 | 0.183 | 2 |
| OES | Period | **0.212** | **0.002** | **17** | $-0.069$ | 0.005 | 4 |
| \|OES\| | Eclipse Depth | **0.136** | **0.048** | **10** | **0.573** | **$<0.001$** | **20** |
| \|OES\| | Morphology | $-0.055$ | 0.428 | 10 | **0.210** | **$<0.001$** | **20** |
| \|OES\| | Temperature | $-0.025$ | 0.726 | 9 | **$-0.213$** | **$<0.001$** | **19** |
| \|OES\| | Distance | 0.044 | 0.524 | 9 | 0.010 | 0.687 | 0 |
| \|OES\| | Absolute Mag. | $-0.027$ | 0.696 | 11 | **0.236** | **$<0.001$** | **20** |
| \|OES\| | Period | 0.046 | 0.506 | 8 | **$-0.366$** | **$<0.001$** | **20** |
| Eclipse Depth | Morphology | **$-0.489$** | **$<0.001$** | **20** | **$-0.158$** | **$<0.001$** | **18** |
| Eclipse Depth | Temperature | $-0.039$ | 0.575 | 9 | $-0.091$ | $< 0.001$ | 10 |
| Eclipse Depth | Distance | 0.057 | 0.405 | 10 | 0.071 | 0.004 | 6 |
| Eclipse Depth | Absolute Mag. | $-0.004$ | 0.955 | 12 | **0.110** | **$<0.001$** | **13** |
| Eclipse Depth | Period | 0.054 | 0.432 | 12 | $-0.084$ | $< 0.001$ | 10 |
| Morphology | Temperature | **0.113** | **0.109** | **12** | **0.196** | **$<0.001$** | **20** |
| Morphology | Distance | **$-0.148$** | **0.032** | **13** | 0.025 | 0.317 | 1 |
| Morphology | Absolute Mag. | 0.065 | 0.350 | 12 | **$-0.156$** | **$<0.001$** | **19** |
| Morphology | Period | **$-0.591$** | **$<0.001$** | **20** | **$-0.562$** | **$<0.001$** | **20** |
| Temperature | Distance | **0.245** | **$<0.001$** | **17** | **0.139** | **$<0.001$** | **19** |
| Temperature | Absolute Mag. | **$-0.705$** | **$<0.001$** | **20** | **$-0.695$** | **$<0.001$** | **20** |
| Temperature | Period | **0.248** | **$<0.001$** | **16** | **0.236** | **$<0.001$** | **20** |
| Distance | Absolute Mag. | **$-0.512$** | **$<0.001$** | **20** | **$-0.364$** | **$<0.001$** | **20** |
| Distance | Period | **0.348** | **$<0.001$** | **20** | **0.100** | **$<0.001$** | **8** |
| Absolute Mag. | Period | **$-0.609$** | **$<0.001$** | **20** | **$-0.370$** | **$<0.001$** | **20** |

[a] Number of "surviving" (i.e. having $|\rho| \geq 0.1$) random subsets out of 20; see Section 2.2.5

between distance and absolute magnitude, are expected and serve as validations of my correlation analysis.

Table 4.3: Results of analyzing the correlations between the core sample's characteristics of interest using Kendall's rank correlation coefficient. Bolded entries indicate correlated characteristics, i.e. $|\rho| \geq 0.1$.

| Characteristic | Characteristic | Core Sample | | KEBC | |
|---|---|---|---|---|---|
| One | Two | Coeff. | $p$-Value | Coeff. | $p$-Value |
| OES | Eclipse Depth | 0.026 | 0.567 | **0.090** | **< 0.001** |
| OES | Morphology | **−0.101** | **0.032** | 0.015 | 0.365 |
| OES | Temperature | **0.281** | **< 0.001** | 0.011 | 0.513 |
| OES | Distance | **0.148** | **0.001** | −0.007 | 0.656 |
| OES | Absolute Mag. | **−0.207** | **< 0.001** | 0.026 | 0.114 |
| OES | Period | **0.152** | **< 0.001** | −0.051 | 0.002 |
| \|OES\| | Eclipse Depth | **0.094** | **0.042** | **0.408** | **< 0.001** |
| \|OES\| | Morphology | −0.038 | 0.423 | **0.141** | **< 0.001** |
| \|OES\| | Temperature | −0.014 | 0.767 | **−0.146** | **< 0.001** |
| \|OES\| | Distance | 0.027 | 0.563 | 0.006 | 0.708 |
| \|OES\| | Absolute Mag. | −0.018 | 0.698 | **0.162** | **< 0.001** |
| \|OES\| | Period | 0.033 | 0.481 | **−0.254** | **< 0.001** |
| Eclipse Depth | Morphology | **−0.387** | **< 0.001** | **−0.125** | **< 0.001** |
| Eclipse Depth | Temperature | −0.023 | 0.631 | −0.062 | < 0.001 |
| Eclipse Depth | Distance | 0.040 | 0.391 | 0.048 | 0.003 |
| Eclipse Depth | Absolute Mag. | −0.002 | 0.971 | **0.076** | **< 0.001** |
| Eclipse Depth | Period | 0.032 | 0.490 | −0.057 | < 0.001 |
| Morphology | Temperature | **0.076** | **0.113** | **0.131** | **< 0.001** |
| Morphology | Distance | **−0.105** | **0.025** | 0.017 | 0.303 |
| Morphology | Absolute Mag. | 0.046 | 0.330 | **−0.106** | **< 0.001** |
| Morphology | Period | **−0.421** | **< 0.001** | **−0.412** | **< 0.001** |
| Temperature | Distance | **0.169** | **< 0.001** | **0.095** | **< 0.001** |
| Temperature | Absolute Mag. | **−0.556** | **< 0.001** | **−0.541** | **< 0.001** |
| Temperature | Period | **0.200** | **< 0.001** | **0.168** | **< 0.001** |
| Distance | Absolute Mag. | **−0.365** | **< 0.001** | **−0.256** | **< 0.001** |
| Distance | Period | **0.252** | **< 0.001** | **0.067** | **< 0.001** |
| Absolute Mag. | Period | **−0.493** | **< 0.001** | **−0.266** | **< 0.001** |

#### 4.2.3.1 O'Connell Effect Size

The OES (as defined in Section 2.1.3.2) reflects the difference in brightness at the two maxima, which in turn depends on the brightness difference between the leading and trailing hemispheres of each star. I gave particular focus to correlations and trends involving the OES because it defined my sample and is the characteristic most directly related to the O'Connell effect. This section discusses OES and |OES| correlations with temperature, eclipse depth, and period. It also discusses the differences between the OES and |OES| correlations.

Figure 4.17's center panel comparing OES and temperature shows that a negative O'Connell effect, wherein the brighter maximum occurs before the primary minimum, mainly occurs in systems with $T < 6{,}000$ K. Thirty-four (23% of 143 systems with a *Gaia* temperature) positive O'Connell effect systems have a temperature above 6,000 K, while only two (3% of 62 systems) negative O'Connell effect systems do. Furthermore, one of the two hot negative O'Connell effect systems, KIC 7773380, has a *Kepler* temperature and *Gaia* color implying that it is likely significantly cooler than its *Gaia* temperature indicates. The other system, KIC 7950962, has primary and secondary eclipse depths that differ by only ~0.001 in normalized flux. Furthermore, my light curve of the system has a phase offset of 0.5 from the KEBC light curve. These two facts indicate an ambiguity regarding which of KIC 7950962's eclipses is the primary, and thus if its O'Connell effect is positive or negative. The dearth of hot systems displaying a negative O'Connell effect was unknown before now and is consistent with the idea that starspots are the predominant cause of a negative O'Connell effect. Starspots are expected to exist in the convective envelopes of cooler stars but not in the radiative envelopes of hotter stars.

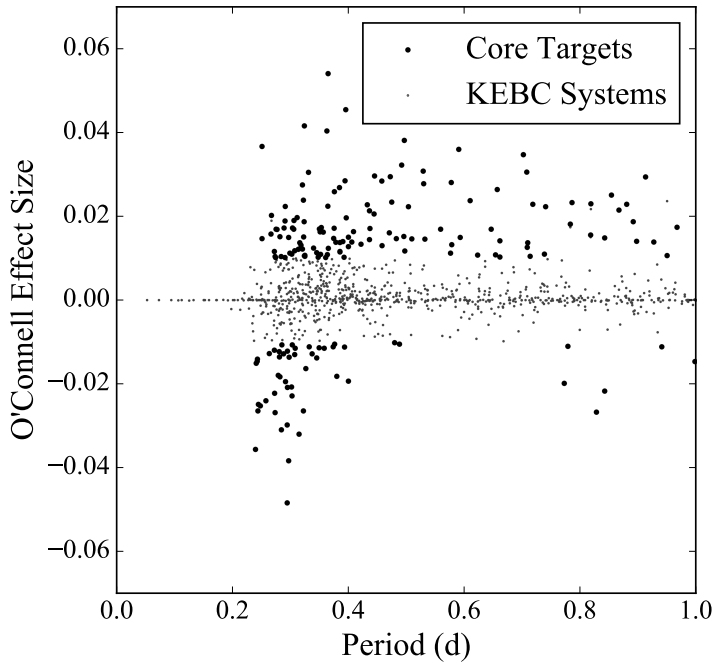


Figure 4.20: Plot comparing the OES to the period. Core sample targets are shown in black while non-target KEBC systems are shown in grey. The y-axis is in units of normalized flux. KICs 9777984, 9935311, and 11347875 are excluded for clarity. Note the positive correlation in core sample systems with $P < 0.5$ d.

Figure 4.17's top-left panel comparing OES and period shows that the OES tends toward zero at longer periods. This trend is consistent with the idea that binary interaction ultimately causes the O'Connell effect. The lack of any systems in the sample with a period greater than ten days strengthens this idea. Table 4.2 shows that this correlation is not robust, however. Of greater concern is that Table 4.2 indicates that the correlation between OES and period is positive in the sample, contradicting the trend observed in Figure 4.17. My explanation for this contradiction is a positive trend between OES and period for the shortest period systems in my sample.

To better display this short-period system trend, Figure 4.20 shows a rescaled

view of Figure 4.17's top-left panel focusing on systems with periods under 1 d. Systems with $P \leq 0.5$ d show this positive correlation clearly. Since Figure 4.4 shows that most systems in the core sample have $P \leq 0.5$ d, this positive correlation dominates the $\rho$ coefficient. I found that the $\rho$ coefficient for the correlation between OES and period more than doubled ($\rho = 0.425$) when analyzing only core sample systems with $P \leq 0.5$ d rather than the entire core sample, supporting my explanation. Furthermore, the short-period system correlation is robust (20 of 20 subsets show a correlation). When I analyzed the core sample systems with $P > 0.5$ d, I found no correlation ($\rho = -0.069$). I interpret the short-period system trend as a result of the following: hotter systems ($T \geq 6{,}000$ K) almost always have OES $> 0$ (as discussed in the previous paragraph), temperature is positively correlated with luminosity, and luminosity is positively correlated with orbital period (as discussed in Section 4.2.3.4). Therefore, the shortest period systems are more likely to have OES $< 0$ than longer period systems, producing the observed trend.

I expected strong correlations between |OES| and other characteristics under the premise that the positive and negative O'Connell effects are fundamentally similar. This premise implies that the positive and negative OES correlations should strengthen when neglecting the OES's sign. Therefore, I was surprised by the *lack* of such correlations in the core sample. The only |OES| correlation I found was between |OES| and the primary eclipse depth, and that correlation is both weak and not robust. Furthermore, Figure 4.21's left panel shows no visible correlation between these two characteristics. The lack of |OES| correlations implies that my premise that the positive and negative O'Connell effects are fundamentally similar is incorrect.

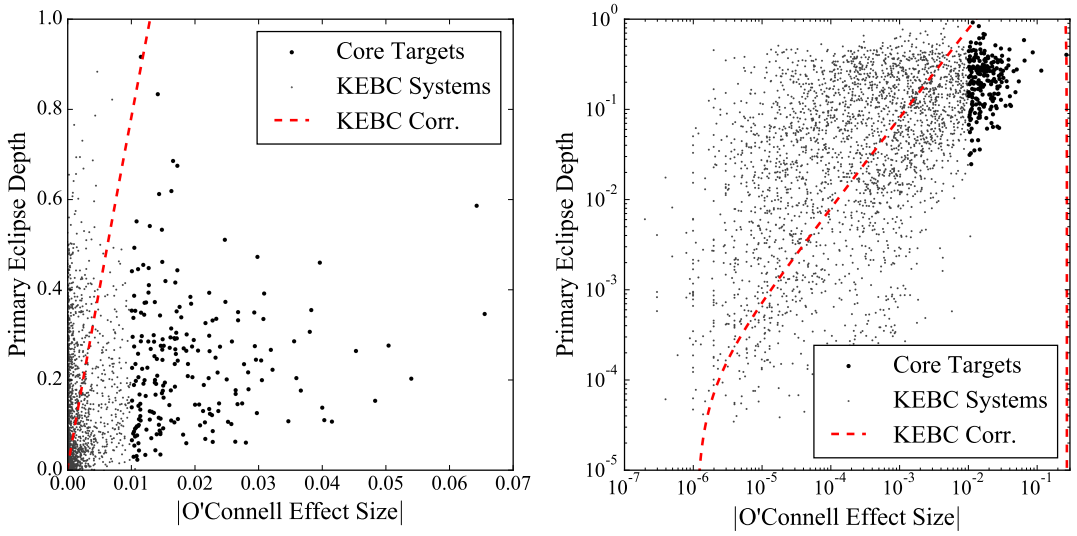


Figure 4.21: Plots comparing the primary eclipse depth to |OES|. Core sample targets are shown in black while non-target KEBC systems are shown in grey. The red dashed lines show the ODR correlation for the KEBC. The left panel's axes are scaled linearly, while the right panel's axes are scaled logarithmically. Both axes are in units of normalized flux. KICs 9777984, 9935311, and 11347875 are excluded from the left panel for clarity. No correlation is apparent in the left panel, but a weak correlation can be seen in the right panel.

Note that every characteristic is correlated with |OES| in the KEBC, the strongest of which is the correlation with primary eclipse depth. I determined a quadratic fit of $D = -298(19)|\mathrm{OES}|^2 + 81(4)|\mathrm{OES}|$ for this correlation using ODR (Section 2.2.5), where $D$ is the primary eclipse depth and the parentheticals are the uncertainties. While Figure 4.21's left panel does not show this correlation, either, Figure 4.21's right panel showing a log-log plot of these characteristics displays a visible, if weak, correlation between them. I suspect that the $\rho$ coefficient detected this trend, resulting in the large coefficient in Table 4.2. Figure 4.21's right panel indicates that this trend extends from the core sample down to much smaller values of |OES| and eclipse depth. Therefore, some systems that are not in the sample may be fundamentally similar to the sample's systems, only with

smaller values for |OES| and eclipse depth. As such, this correlation may be more related to the O'Connell effect than its strength in the core sample would indicate.

Both panels of Figure 4.21 also show a lack of systems displaying a significant O'Connell effect and a small primary eclipse depth. My selection criterion ($|\mathrm{OES}| \geq 0.01$) does not exclude such systems, and their absence is conspicuous since most KEBC systems have small eclipse depths. A bias in the *Kepler* selection function against these systems may explain their absence, which would imply a fundamental difference between such systems and the systems *Kepler* observed. Their absence may also be a true representation of O'Connell effect binaries, although I cannot identify a plausible reason why systems with a significant O'Connell effect cannot have a small primary eclipse depth. As a final note, Figure 4.21's left panel shows that all four KEBC systems with an eclipse depth above 0.8 have a non-negligible |OES|. However, this fact may result from the rarity of KEBC systems with such deep eclipses rather than a genuine link between deep eclipses and the O'Connell effect.

I found different correlations for the OES (the morphology parameter, temperature, distance, absolute magnitude, and period) as compared to |OES| (eclipse depth). Since the OES distinguishes between the positive and negative O'Connell effects but |OES| does not, this suggests a more fundamental difference between the positive and negative O'Connell effects than previously thought. Assuming that spots are the O'Connell effect's primary cause, a larger OES implies that spots cover a greater area, have a more extreme temperature factor, or are further offset from the line connecting the stellar centers. Kouzuma (2019) found a weak positive correlation between stellar temperature and spot temperature factor for cool starspots in W-type overcontact systems, wherein the smaller star is hotter (in

contrast to A-type overcontact systems; McCartney 1999). The same correlation is much stronger in semi-detached systems. Kouzuma (2019) found weak positive correlations in overcontact systems between stellar temperature and spot size and between orbital period and spot size. The correlations for cool spots are similar but generally stronger than for hot spots. His results imply positive correlations between temperature and |OES| and between orbital period and |OES|. Table 4.2 shows both correlations for OES but neither for |OES|.

To further test this correlation, I determined Spearman's $\rho$ coefficient for the positive O'Connell effect systems and for the negative O'Connell effect systems. I found that the correlations for the positive O'Connell effect systems ($\rho_{\text{temp}} = 0.122$ and $\rho_{\text{per}} = 0.083$) are weak while the correlations are much stronger for the negative O'Connell effect systems ($\rho_{\text{temp}} = 0.402$ and $\rho_{\text{per}} = 0.142$), the former has the wrong sign (i.e. the OES gets closer to zero as the temperature increases) while the latter is weak. These correlations are therefore inconsistent with the results of Kouzuma's (2019) starspot study. However, such inconsistencies may not indicate that starspots do not cause the O'Connell effect. Instead, they may result from the core sample's mixture of systems with different O'Connell effect causes, or perhaps from differences between my sample and his. Determining the cause of these inconsistencies is beyond the scope of my project.

The three other robust correlations that I have not discussed are between OES and the morphology parameter, distance, and absolute magnitude. I interpret these three correlations as arising from unrelated correlations with other characteristics. For instance, the correlation between the OES and absolute magnitude arises from the discussed correlation between OES and temperature combined with the strong correlation between temperature and absolute magnitude that is a well-

known feature of main-sequence stars. I do not discuss these secondary correlations further due to their dependence on other correlations that have a more fundamental explanation.

The different correlations I found among positive and negative O'Connell effect systems, combined with the fact that a much larger number of systems display a positive O'Connell effect, leads me to conclude that the preference for a brighter maximum following the primary eclipse is not an observational bias. It is instead fundamental to the O'Connell effect. Additionally, I conclude that the positive and negative O'Connell effects have different causes, or that one cause is common in one case and rare in the other. I consider these findings important results of my project.

#### 4.2.3.2 Eclipse Depth

The primary eclipse depth (as defined in Section 2.2.1) is influenced by four parameters: temperature ratio, relative radii, component shapes, and orbital inclination. This section discusses the correlation between the eclipse depth and the morphology parameter. It also discusses a couple of trends between these two characteristics seen in Figure 4.18's center panel.

The eclipse depth strongly correlates with the morphology parameter in the sample, while the KEBC shows a weaker correlation. I expect such a correlation because, as Wilson (2006) states, the eclipse depth is a monotonic function of the stellar radii ratio for overcontact systems (i.e. systems with a large $\mu$). Meanwhile, Matijevič et al. (2012) notes that the primary eclipse width increases with $\mu$, and the eclipse widths measure the sum of the relative stellar radii. Thus, in overcontact systems, the primary eclipse depth decreases as the morphol-

ogy parameter increases. The best fit to this correlation is the linear function $\mu = -2.13(13)D + 1.26(9)$. The KEBC's correlation is weaker as it includes more small $\mu$ systems, for which the eclipse depth and stellar radii ratio relation does not hold.

Figure 4.18's center panel comparing the primary eclipse depth to the morphology parameter shows a sharp edge toward the panel's right side in both the sample and the KEBC. This trend indicates that eclipses get shallower as $\mu$ increases above 0.7, a consequence of the relation discussed in the previous paragraph. The systems in the sample appear to be clustered along this edge in Figure 4.18's center panel. This clustering indicates that O'Connell effect systems tend to have the deepest eclipses of systems with similar light curves (as measured by $\mu$), hinting at a connection between the OES and the eclipse depth. The clustering around the right edge is seen with the non-core systems, although it is not as pronounced.

Figure 4.18's center panel shows another trend: all KEBC systems with $D \geq 0.6$ have $\mu \approx 0.6$. Söderhjelm & Dischler (2005) states that the maximum eclipse depth for two main-sequence stars is 0.75 magnitudes (equivalent to a relative flux of 0.5), implying that systems with deeper eclipses must have an evolved component. Matijevič et al. (2012) says that a $\mu$ of 0.6 indicates a semi-detached system, which occurs when an evolving star fills its Roche lobe. This trend therefore make sense because close binaries with evolved components would be expected to have $\mu \approx 0.6$ by Matijevič et al. (2012). However, I cannot discount the possibility that this trend is a statistical artifact caused by the rarity of systems with such deep eclipses in the KEBC. The eclipse depth determination method discussed in Section 2.2.1 may also influence this trend, as it can underestimate the primary eclipse depth of Algol-type systems.

#### 4.2.3.3 Morphology Parameter

The morphology parameter (described in Section 2.2.3) is primarily a measure of eclipse widths (Matijevič et al. 2012), but it correlates well with a given system's morphology class. I want to reemphasize the point from Matijevič et al. (2012) that the morphology parameter only provides a "best-guess" estimate of the morphology class. This section focuses on the morphology parameter distribution shown in Figure 4.9. It also discusses the initially surprising correlation between the morphology parameter and temperature.

Figure 4.9 shows a lower abundance of systems with $0.6 \leq \mu \leq 0.7$ in the core sample. The same region is depleted in the KEBC, but not to the extent of the core sample, suggesting that this underabundance is not wholly due to the parent population. Matijevič et al. (2012) states that systems with $\mu$ in this range are semi-detached. My interpretation of systems in this range is that the accreting star grows larger with increasing morphology parameter, with the system becoming contact around $\mu = 0.7$. The larger the accretor is, the less deeply a matter stream will penetrate into its potential well. The matter stream will therefore impact the surface with less energy and cause less dramatic heating. Assuming the accretor is on the main sequence, a larger star will be hotter, further reducing the degree of heating. These factors could cause the paucity of systems in this region of parameter space showing a significant O'Connell effect.

Figure 4.9 also shows that systems with $\mu \lesssim 0.5$ rarely show a significant O'Connell effect, and none with $\mu \leq 0.3$ do. This is significant because the selection criterion should not be biased with respect to $\mu$. Stars in systems with a small $\mu$ are generally farther apart and are less likely to significantly affect each other, so this result reinforces the link between binary interaction and the O'Connell effect

discussed in Section 4.2.3.1.

I was surprised by the morphology parameter and temperature correlation because I assumed that they would be uncorrelated. I now believe that this correlation results from a selection effect. Two processes work in tandem to cause this selection effect: the influence of geometry on the morphology parameter and the influence of the initial mass function (IMF) on temperature. Regarding the first process, eclipses only occur if $\Delta < R_1 + R_2$ at $\phi = 0$ (conjunction), where $\Delta = a \cos i$ (Prša 2006, Equations 3.37 and 3.38) and the $R_n$ are the component radii. Therefore, closer binaries (smaller $a$) exhibit eclipses over a broader range of inclinations. Since the components of large $\mu$ systems are generally closer than in small $\mu$ systems, it is more likely to observe a given system eclipse in the former case. The IMF, meanwhile, makes hot stars rarer than cool stars, a fact compounded by *Kepler*'s selection biases discussed in Section 2.1.1. Acting together, these processes mean that most hot eclipsing binaries observed by *Kepler* should have large $\mu$. The resulting dearth of hot, small $\mu$ systems creates an apparent positive correlation between the characteristics.

#### 4.2.3.4 Period

The period is a function of component masses and orbital separation due to Kepler's third law. This section discusses the correlation between the period and the absolute magnitude/luminosity in the core sample. Figure 4.19's lower-left panel demonstrates this correlation in the dense clustering seen near the panel's left side.

A correlation between period and absolute magnitude initially seems odd. Low luminosity stars can have very wide orbits, after all, while very luminous stars can be short-period contact binaries. However, there is a lower period limit for stars

of a given luminosity. For example, on the main sequence, a star's size and surface temperature (and thus absolute magnitude) are a function of its mass. If binary components are too close, however, they will merge into a single star. Kepler's third law states that $P \propto a^{3/2}$, while the Stefan–Boltzmann law states that $L_* \propto R^2$ for a sphere. Therefore, a lower limit on component separation based on stellar radius implies a lower limit on the period for a given absolute magnitude for main-sequence stars. This limit increases for evolved stars since they become cooler and larger as they evolve. The correlation between period and absolute magnitude is weaker in the KEBC because of the KEBC's numerous long-period systems, which do not have a relationship between component separation and stellar radius.

Both Figure 4.19's lower-left panel and Figure 4.22 show a tight clustering of systems along their left edge. My interpretation of this clustering is that it represents the lower limit on the period for a given absolute magnitude or luminosity. Systems to the left of this edge are likely subdwarf systems or have poorly determined distances resulting in large luminosity errors. The fact that such a large percentage of the core sample lies on or near this edge – particularly systems with $L_* \geq 1$ $L_\odot$ in Figure 4.22 – suggests that the O'Connell effect is more common in contact or near-contact binaries. This trend is therefore strong evidence that binary interaction is a critical factor in causing the O'Connell effect.

## 4.3 The Marginal Sample

My analysis of the marginal sample is more perfunctory than the core sample's. I chose to do a less-detailed analysis because, unlike the core sample, the marginal sample is not well-defined: a system's inclusion in the marginal sample is arbitrary

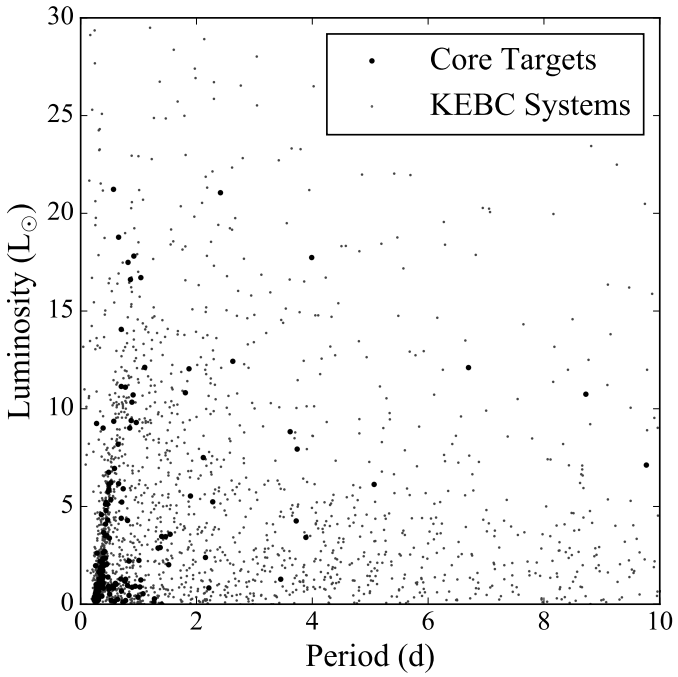


Figure 4.22: Plot comparing the luminosity to the period. Core sample targets are shown in black while non-target KEBC systems are shown in grey. KICs 3662635 and 5820209 are excluded for clarity. The sharp edge near the lower left corner denotes the short-period limit for a given luminosity. Most of the core sample is clustered near or along this edge.

beyond it not having a large enough OES to be in the core sample. Since the marginal sample is therefore not complete, my analysis of its characteristics is less compelling than the core sample's. With this caveat in mind, I now present an overview of my analysis of the marginal sample.

### 4.3.1 Sample Characteristics

I now present a detailed summary of the characteristics of the marginal sample. Like Section 4.1, I focus on the light curve and physical characteristics, along with my results from analyzing the marginal sample's ETV. I omit further discussion

of the OER and the LCA, however, because Section 4.1.4 covered the discussion for all samples I consider. As with Section 4.1, I use the comparative subset of the KEBC discussed in Section 2.1.2 for creating the histograms, color-magnitude diagram, and Hertzsprung–Russell diagram presented here.

#### 4.3.1.1 Light Curve Characteristics

The marginal sample contains 30 (65%) Algol-type, 7 (15%) $\beta$ Lyrae-type, and 9 (20%) W Ursae Majoris-type systems, which differs greatly from the core sample's distribution (25% Algol-type, 19% $\beta$ Lyrae-type, and 56% W Ursae Majoris-type; Section 4.1.1). Indeed, the marginal sample's distribution is much closer to the KEBC's (68% Algol-type, 9% $\beta$ Lyrae-type, and 23% W Ursae Majoris-type). I note that temporal variation is more common in the marginal sample than in the core sample (see Section 5.1.1). Therefore, the larger proportion of Algol-type systems in the marginal sample may indicate that the temporal variation plays a more prominent role in causing the O'Connell effect in Algol-type systems.

The largest OES found in the marginal sample is 0.0099 (KIC 7035139), while the smallest is 0.0005 (KIC 4049124), both in units of normalized flux. As a reminder, while these average OES values are below the threshold, around half of the marginal sample shows significant temporal variation in their light curves. As a result, the OES is occasionally above the threshold for much of the marginal sample. As with the core sample, a clear majority of marginal sample systems (32, or 70% of the sample) show a positive O'Connell effect. This ratio is very similar to the core sample's (69%; Section 4.1.1), supporting the idea that this ratio is intrinsic to the O'Connell effect and not a selection effect caused by the $|\mathrm{OES}| \geq 0.1$ cutoff. Finally, the primary minimum depth ranges from 0.012 (KIC 5615528) to

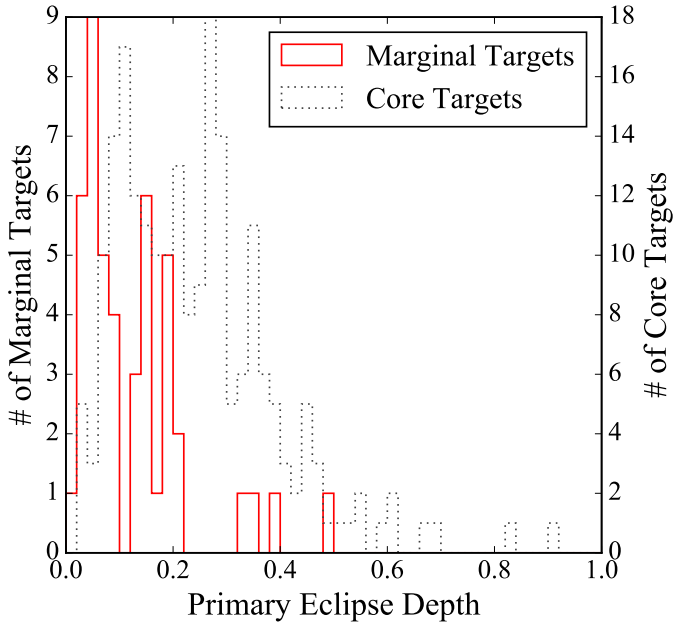


Figure 4.23: Histogram showing the primary eclipse depths for the marginal sample (solid red) and the core sample (dotted grey). Marginal sample systems have shallower eclipses on average than core sample systems.

0.492 (KIC 4660997). Figure 4.23 shows a histogram of the primary eclipse depths for the core and marginal samples, indicating that marginal sample systems have shallower eclipses than core sample systems.

#### 4.3.1.2 Physical Characteristics

Figure 4.24 shows a histogram of the marginal sample's distance distribution determined from *Gaia* EDR3 parallaxes (Gaia Collaboration et al. 2021; Bailer-Jones et al. 2021), along with the distance distribution for 2,860 of the 2,862 KEBC entries with *Gaia* EDR3 parallaxes (excluding two systems with $d \geq 9{,}000$ pc for clarity). While small-sample statistics may skew the distribution, it seems that marginal sample systems are slightly nearer on average than the KEBC. The statistical analysis I present in Section 4.3.2 indicates that the marginal sample's

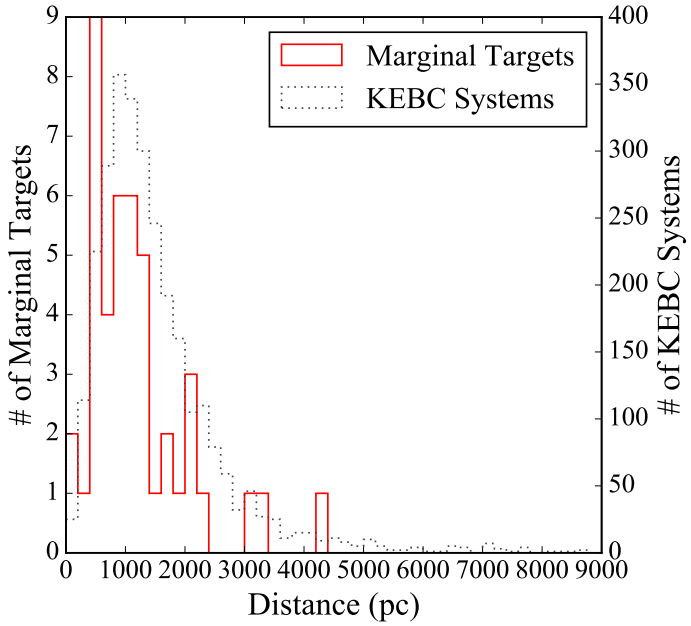


Figure 4.24: Histogram comparing the implied distances for 44 of the 46 marginal sample targets (solid red) and 2,860 of 2,862 entries in the KEBC (dotted grey) with *Gaia* EDR3 parallaxes (excluding two systems with $d > 9{,}000$ pc). Two marginal systems (KICs 7284688 and 11198068) lack *Gaia* parallax data and so are not included. Marginal sample systems are slightly closer on average than other KEBC systems.

distance population differs from the KEBC's more strongly than the core sample's. The marginal sample system's distances range from $94.4 \pm 0.1$ pc (KIC 3557421) to $4{,}387^{+2{,}537}_{-1{,}454}$ pc (KIC 8694926).

Figure 4.25 shows a histogram of the luminosities in the marginal sample and the 2,862 KEBC entries with *Gaia* EDR3 parallaxes. As with the core sample, Figure 4.25 indicates that the marginal sample systems are less luminous on average than others in the KEBC. The marginal sample sample system's luminosities range from $0.0059^{+0.0010}_{-0.0008}$ $L_\odot$ (KIC 10544976) to $16.78^{+1.98}_{-1.94}$ $L_\odot$ (KIC 5215999).

Figure 4.26 shows a histogram of the periods for the marginal sample and all 2,915 KEBC entries with $P \leq 1{,}000$ d. The small sample size makes it difficult

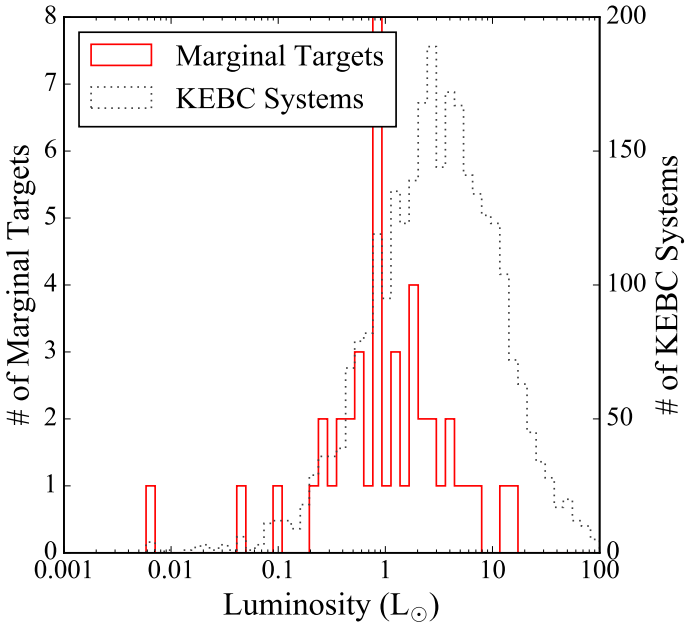


Figure 4.25: Histogram comparing the luminosities for 44 of the 46 marginal sample targets (solid red) and all 2,862 entries in the KEBC (dotted grey) with *Gaia* EDR3 parallaxes. Two marginal systems (KICs 7284688 and 11198068) lack *Gaia* parallax data and so are not included. Note the logarithmic scale of the x-axis. Marginal sample systems are less luminous on average than other KEBC systems.

to draw many conclusions from Figure 4.26, but comparison with Figure 4.4 indicates that marginal sample systems have longer periods than the core sample on average. In the 0.7-0.9 d range, the marginal sample alone makes up ~10% of the KEBC. The marginal sample system's periods range from 0.251 d (KIC 12602985) to 4.804 d (KIC 4049124).

Figure 4.27 shows a histogram of the temperatures for the 42 systems in the marginal sample (excluding KICs 4037163, 7284688, 10544976, and 11920266) and all 2,825 KEBC entries with a *Gaia* temperature. Figure 4.27 indicates that the marginal sample system's temperatures are strongly concentrated between 4,500 and 6,000 K. Additionally, marginal sample systems are cooler on average than core sample systems (compare Figure 4.27 with Figure 4.6), and only two

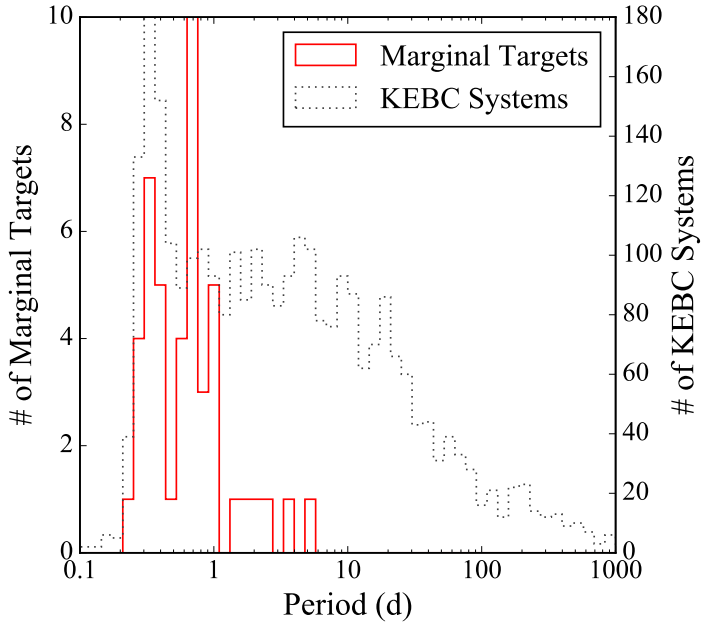


Figure 4.26: Histogram comparing the periods for the 46 marginal sample targets (solid red) and all 2,915 entries in the KEBC (dotted grey) with $P \leq 1{,}000$ d. Note the logarithmic scale of the x-axis. Marginal sample systems have shorter periods on average than other KEBC systems.

(KICs 8285349 and 11175495) have a temperature above 6,000 K. One possible explanation for this is that the O'Connell effect in marginal systems is predominately caused by starspots, which are not expected to exist in stars significantly hotter than ~6,000 K. The marginal sample's *Kepler* temperatures range from 3,701 K (KIC 8949316) to 8,070 K (KIC 11175495), while their *Gaia* temperatures range from 3,713 K (KIC 8949316) to 8,575 K (KIC 11175495).

Frasca et al. (2016) provide spectral types for nine marginal systems, and the main sequence spectral types range from G3 (shared by KICs 6431545, 7284688, and 11244501) to A4 (KIC 11175495). Only KICs 7885570 (G0 IV) and 11920266 (F7 IV) have an explicit non-main sequence spectral type given by Frasca et al. (2016). Almenara et al. (2012) spectrally classified KIC 10544976 as consisting

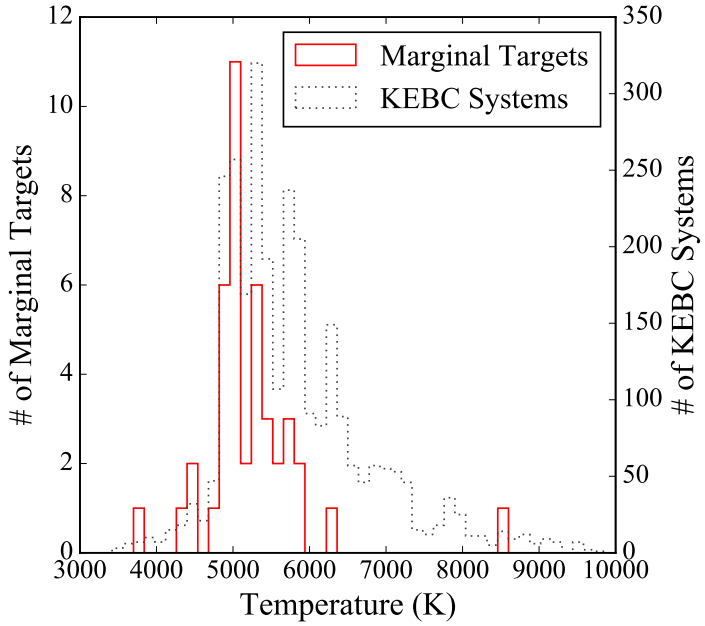


Figure 4.27: Histogram comparing the *Gaia* DR2 temperatures for 42 of the 46 marginal sample targets (solid red) and all 2,819 systems in the KEBC (dotted grey) with a *Gaia* DR2 temperature. Marginal sample systems are cooler on average than other KEBC systems.

of a DA white dwarf primary and an M4 V secondary. The spectral types of the marginal sample estimated using Dr. Mamajek's table range from K5.5 V (KIC 4037163) to A7 V (KIC 11175495). Table A.1 lists both spectral classifications.

Figures 4.28 and 4.29 show the marginal sample's color-magnitude and Hertzsprung-Russell (HR) diagrams, respectively. I excluded the two systems lacking *Gaia* parallaxes (KICs 7284688 and 11193447) from Figure 4.28. I also excluded the three additional systems (KIC 4037163, 10544976, and 11920266) lacking *Gaia* temperatures from Figure 4.29 (note that KIC 7284688 also lacks a temperature and is not counted among the three additional systems). The marker color and shape indicate the light curve classification for each system determined by the

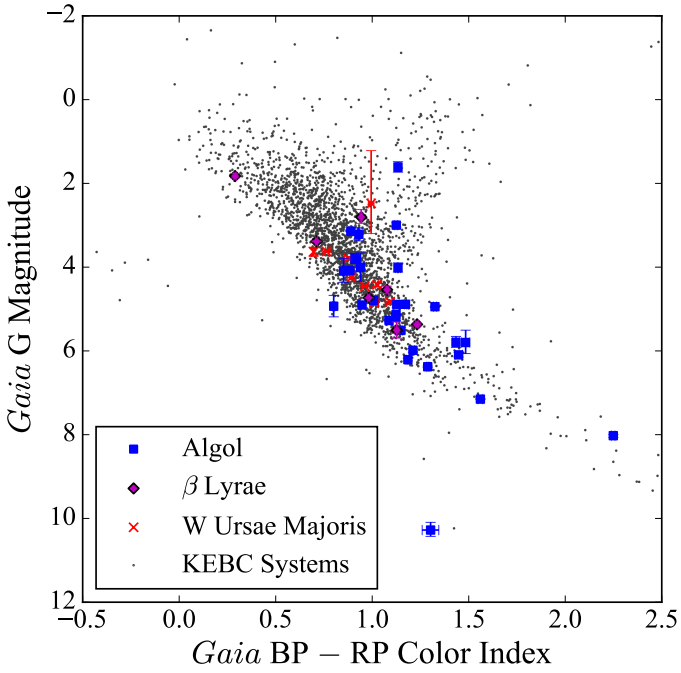


Figure 4.28: Color-magnitude diagram showing the 44 of the 46 systems in the marginal sample overplotting the rest of the KEBC systems. The *Gaia* color is plotted against the absolute magnitude in *Gaia*'s G band. The systems are color coded according to their light curve class. Two marginal sample systems (KICs 7284688 and 11198068) lack *Gaia* parallax data and so are not included.

criteria given in Section 2.2.2.2. Unlike the core sample, the marginal sample differs from the KEBC distribution by its near-total lack of high-temperature ($T \geq$ 6,000 K) systems (the two exceptions being KICs 8285349 and 11175495). My interpretation of this fact is that if the O'Connell effect manifests in high-temperature systems, it is usually significant in the sense I have defined (i.e. $|\mathrm{OES}| \geq 0.1$), which is not true for cooler systems. Note that Figure 4.28 clearly shows the white dwarf system KIC 10544976.

Figure 4.30 shows a histogram of the marginal sample's morphology parameter distribution. Figure 4.30 shows a prominent peak at $\mu \approx 0.60$ that is only weakly

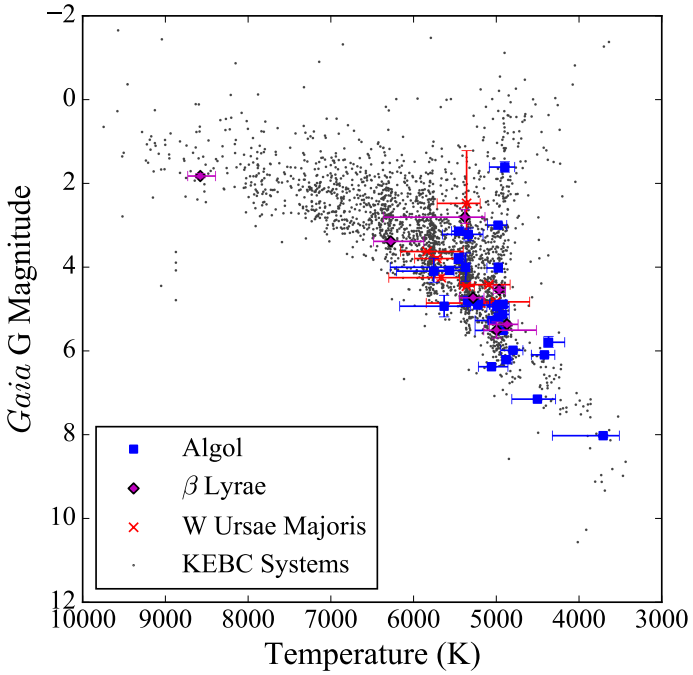


Figure 4.29: HR diagram showing 41 of the 46 systems in the marginal sample overplotting the rest of the KEBC systems. The *Gaia* temperature is plotted against the absolute magnitude in *Gaia*'s G band. The systems are color coded according to their light curve class. Five marginal sample systems lack either *Gaia* parallax data or a *Gaia* temperature and so are not included (see discussion in text).

reflected in the KEBC population and which, while present in the core sample, is less prominent than in the marginal sample. According to Matijevič et al. (2012), $\mu = 0.60$ corresponds to the onset of semi-detached systems, indicating a possible connection between near-contact or semi-detached systems and a slight O'Connell effect.

#### 4.3.1.3 Eclipse Timing Variation

Unlike the core sample, no marginal sample systems show the parabolic ETV signal expected for mass transfer. By contrast, seven systems (15.2% of the marginal

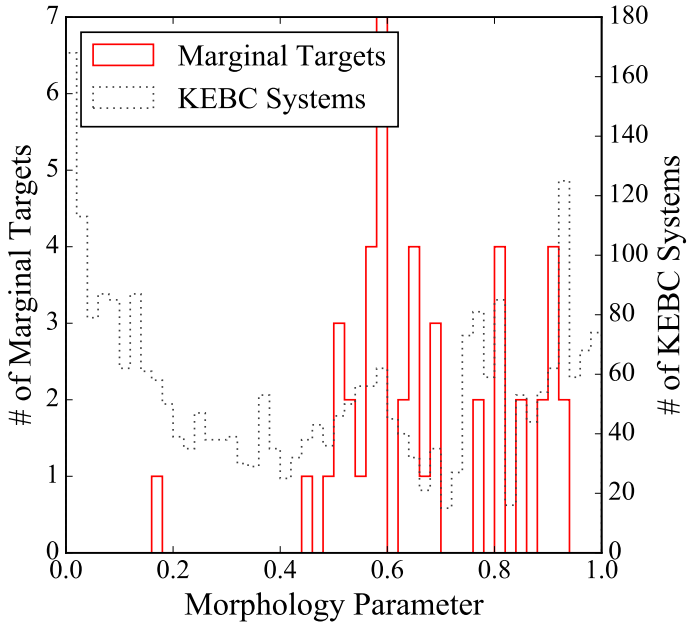


Figure 4.30: Histogram showing the morphology parameters for all 46 marginal sample targets (solid red) and all 2,745 entries in the KEBC (dotted grey) that have $\mu \neq -1$. The two distributions differ significantly, particularly for $\mu < 0.50$. The marginal sample also has few systems with $0.70 \leq \mu \leq 0.80$.

sample) show a sinusoidal signal, and a further seven (15.2%) show an "interesting" ETV. Therefore, all three ETV signals occur less frequently in the marginal sample than in the core sample (1.4%, 19.0%, and 22.7%). Kouzuma (2018) cited no systems in the marginal sample as undergoing mass transfer. However, as with the core sample, I investigated all marginal sample systems' ETVs using the Butterworth filter. I found that four systems (20% of the 20 investigated systems) had $R^2 \geq 0.8$, which is the same fraction that I found for the core sample (Section 4.1.3. Keeping in mind the caveat that the marginal sample is incomplete, I conclude that mass transfer is as common in marginal sample systems as it is in core sample systems.

### 4.3.2 Statistical Analysis

This section closely mirrors Section 4.2. I present a graphical representation of the characteristic correlations, discuss the results of my study of the characteristic distributions and correlations, and further discuss several insightful correlations. The marginal sample's smaller size means that the $p$-values of the statistical tests are larger than a similar result would be in the core sample. In practice, this means that it is more difficult to reject the null hypothesis for a given test.

#### 4.3.2.1 Characteristic Trends

Figures 4.31, 4.32, and 4.33 show corner plots comparing period, OES, *Gaia* temperature, *Gaia* BP – RP color index, morphology parameter, absolute *Gaia* G magnitude, and primary eclipse depth. One trend immediately apparent upon comparing these figures with Figures 4.17-4.19 is that marginal sample systems have shallower eclipses on average than core sample systems, as Figure 4.23 indicated. Additionally, the clustering seen on the right side of Figure 4.18's center panel is absent in Figure 4.32's center panel. There is a lower abundance of systems with $\mu \approx 0.7$ similar to the underabundance found in the core sample, which Figure 4.30 confirms is real. Finally, Figure 4.31's center panel lacks the temperature distribution difference present in Figure 4.17's center panel. These differing trends imply that there are fundamental differences between the marginal sample and the core sample beyond the OES. As with Section 4.2.1, I used the comparative subset for these figures.

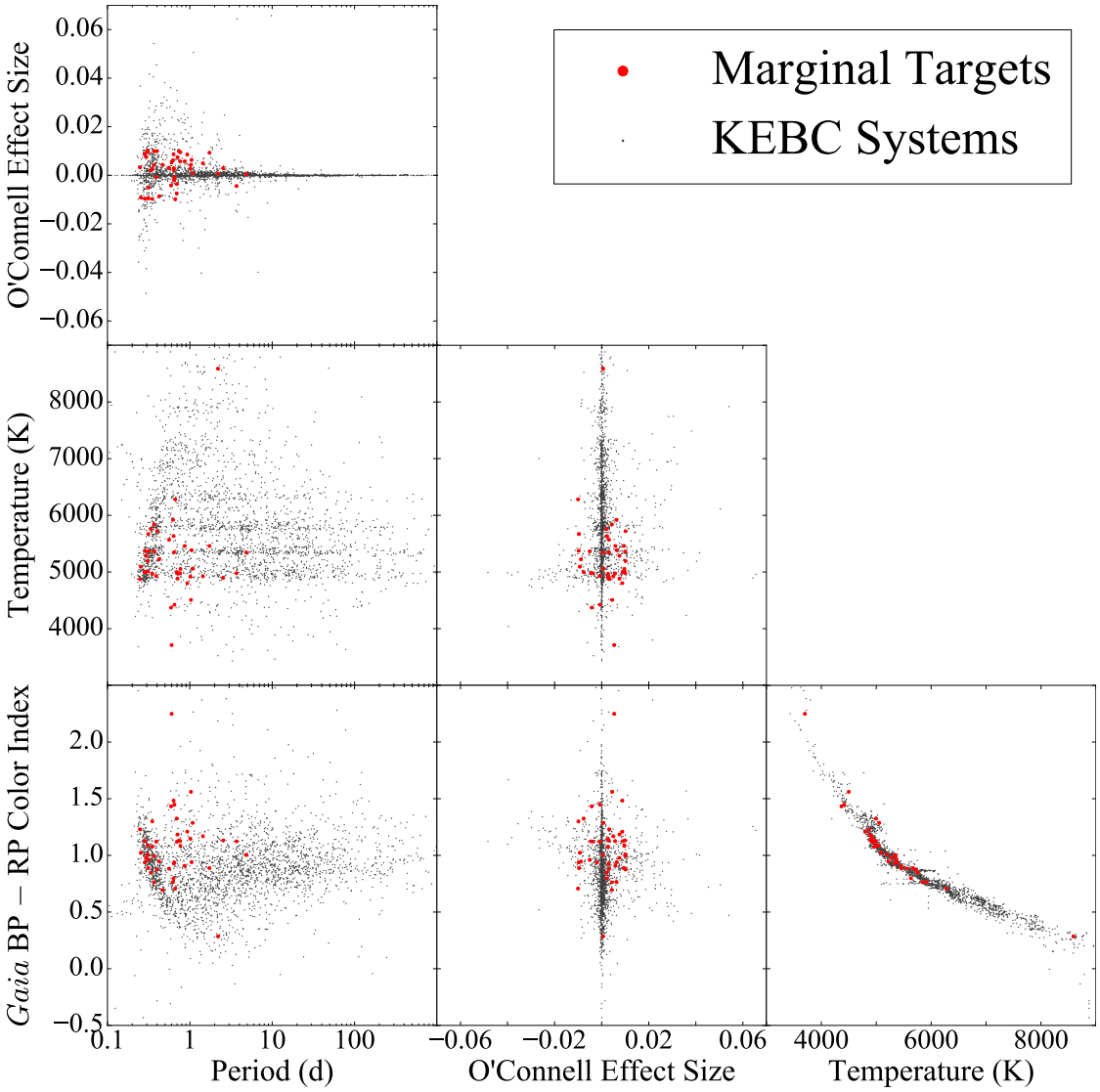


Figure 4.31: Corner plot showing the marginal sample correlations between three characteristics of interest (period, OES, and temperature) and OES, temperature, and color. Marginal sample targets are shown in red while non-marginal KEBC systems are shown in grey.

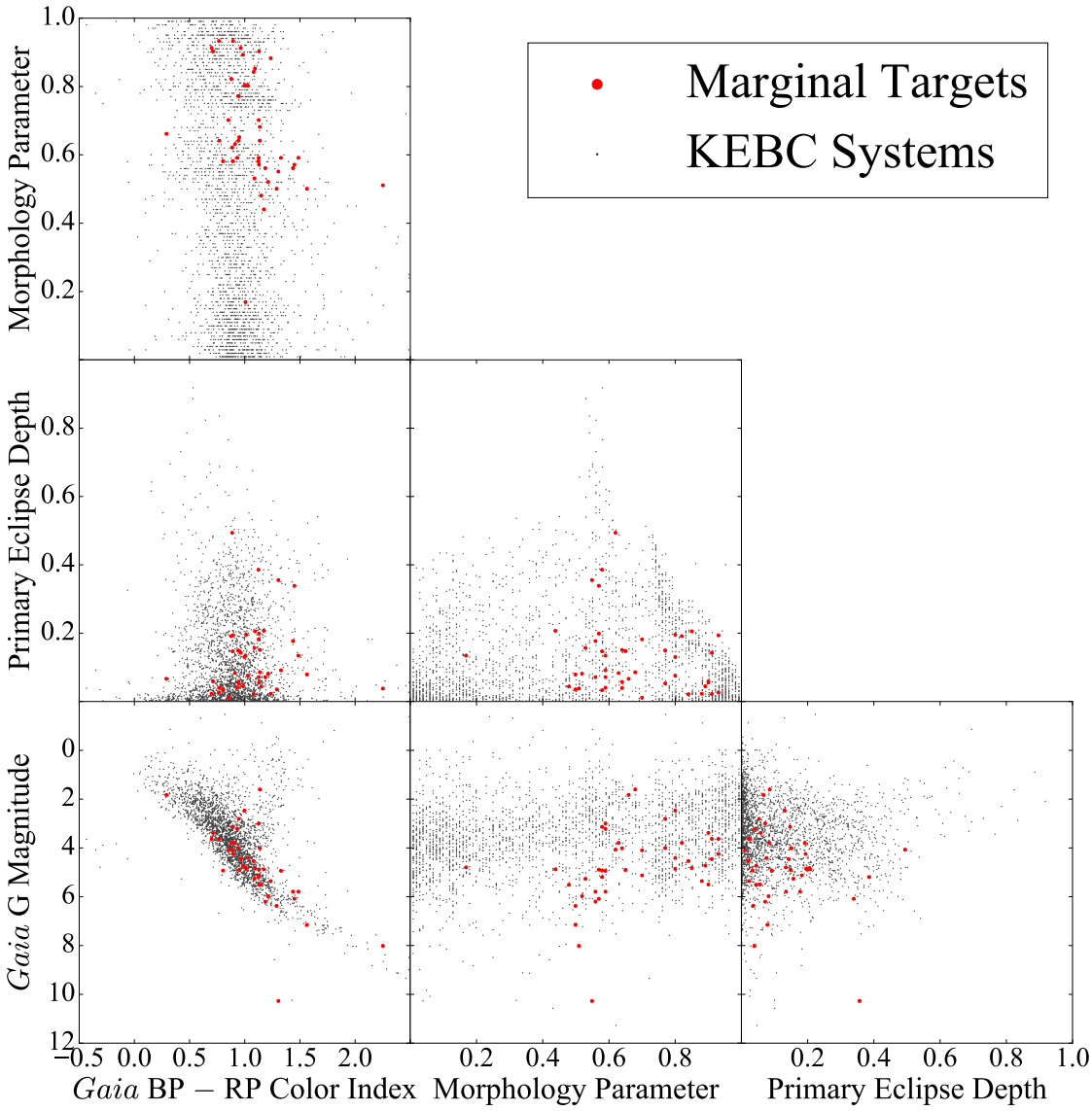


Figure 4.32: Corner plot showing the marginal sample correlations between three characteristics of interest (color, morphology parameter, and primary eclipse depth) and morphology parameter, primary eclipse depth, and absolute magnitude. Marginal sample targets are shown in red while non-marginal KEBC systems are shown in grey.

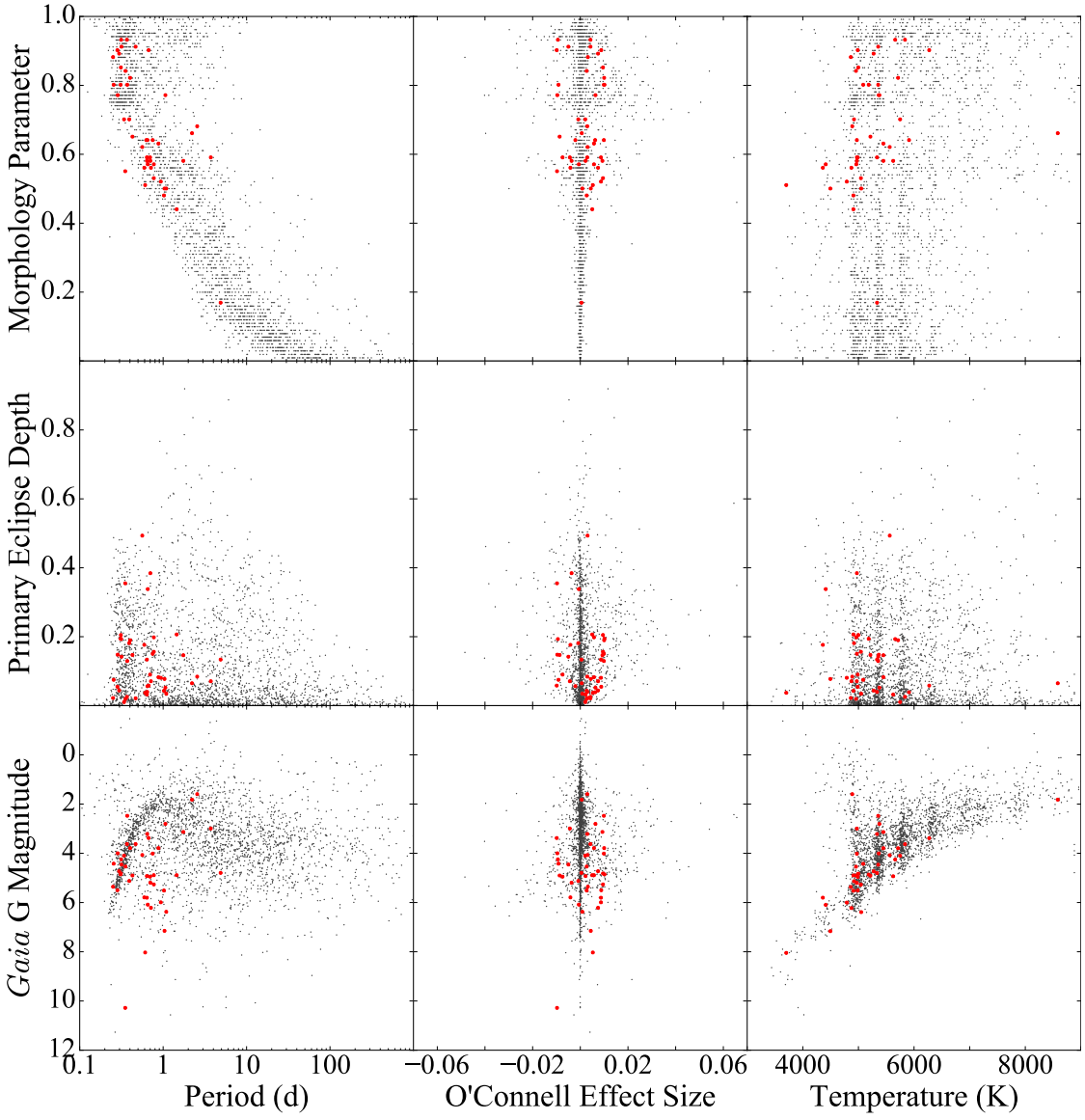


Figure 4.33: Corner plot showing the marginal sample correlations between three characteristics of interest (period, OES, and temperature) and morphology parameter, primary eclipse depth, and absolute magnitude. Marginal sample targets are shown in red while non-marginal KEBC systems are shown in grey.

Table 4.4: Results of analyzing the characteristic distributions between the marginal sample and the entire KEBC using the Kolomgorov–Smirnov test.

| Characteristic | K–S Statistic | $p$-Value |
|---|---|---|
| O'Connell Effect Size | 0.371 | $< 0.001$ |
| \|O'Connell Effect Size\| | 0.485 | $< 0.001$ |
| Primary Eclipse Depth | 0.252 | 0.005 |
| Morphology Parameter | 0.189 | 0.071 |
| Temperature | 0.348 | $< 0.001$ |
| Distance | 0.193 | 0.063 |
| Absolute Magnitude | 0.415 | $< 0.001$ |
| Period | 0.272 | 0.002 |

#### 4.3.2.2 Sample Distributions

Table 4.4 gives the results of the K–S test for the marginal sample. The K–S test rejects the null hypothesis that the two samples were drawn from the same population at a level of 0.001 for the OES, |OES|, temperature, and absolute magnitude. However, it cannot reject the null hypothesis for the primary eclipse depth, morphology parameter, distance, or period. The lower $p$-values are due in part to the smaller sample size used. For instance, comparing Tables 4.1 and 4.4 shows that the period's K–S statistic is nearly the same for the core and marginal samples (0.274 versus 0.272, respectively), but only in the former can the null hypothesis be rejected. Compared to the core sample, the marginal sample's K–S statistic is smaller for the OES, |OES|, primary eclipse depth, and morphology parameter, larger for the temperature, distance, and absolute magnitude, and about the same for the period.

The K–S statistic for the absolute magnitude and the morphology parameter differ particularly significantly from the core sample's. Figure 4.34 shows a histogram comparing the marginal and core sample's absolute magnitude distributions. The primary difference between the two is the lower proportion of luminous

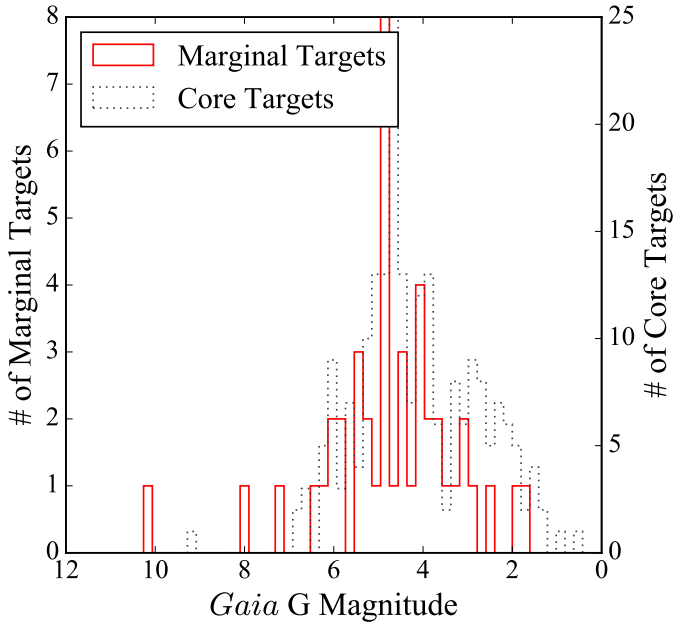


Figure 4.34: Histogram comparing the absolute magnitudes for 44 of the 46 marginal sample targets (solid red) and all 212 core sample targets (dotted grey). Two marginal systems (KICs 7284688 and 11198068) lack *Gaia* parallax data and so are not included. The marginal sample has a lower proportion of bright ($G \leq 3$) systems than the core sample.

systems in the marginal sample, particularly systems with $G \leq 3$. Figure 4.25 indicates that many KEBC systems have $L \approx 5$ $L_{\odot}$, thus explaining the larger K–S statistic for the marginal sample.

#### 4.3.2.3 Characteristic Correlations

Table 4.5 gives the results of my marginal sample analysis using both Spearman's $\rho$ and Kendall's $\tau$ coefficients (the KEBC's $\rho$ and $\tau$ coefficients are given in Tables 4.2 and 4.3, respectively). I immediately noticed several differences between the core and marginal sample correlations. I will discuss the differences in the OES/|OES|, morphology, and period correlations.

The marginal sample has only a single, weak correlation (distance) for the OES,

Table 4.5: Results of analyzing the correlations between the marginal sample's characteristics of interest using Spearman's and Kendall's rank correlation coefficients. Bolded entries indicate correlated characteristics, i.e. $|\rho| \geq 0.1$.

| Characteristic | Characteristic | Spearman's | | Kendall's | |
|---|---|---|---|---|---|
| One | Two | Coeff. | $p$-Value | Coeff. | $p$-Value |
| OES | Eclipse Depth | −0.013 | 0.932 | 0.009 | 0.932 |
| OES | Morphology | 0.003 | 0.983 | 0.012 | 0.909 |
| OES | Temperature | −0.065 | 0.685 | −0.060 | 0.573 |
| OES | Distance | **0.187** | **0.223** | **0.129** | **0.217** |
| OES | Absolute Mag. | −0.040 | 0.790 | −0.026 | 0.798 |
| OES | Period | 0.061 | 0.689 | 0.040 | 0.698 |
| \|OES\| | Eclipse Depth | **0.347** | **0.018** | **0.262** | **0.010** |
| \|OES\| | Morphology | **0.288** | **0.053** | **0.197** | **0.056** |
| \|OES\| | Temperature | **0.180** | **0.255** | **0.126** | **0.242** |
| \|OES\| | Distance | **0.124** | **0.422** | **0.085** | **0.418** |
| \|OES\| | Absolute Mag. | **−0.126** | **0.405** | **−0.094** | **0.358** |
| \|OES\| | Period | **−0.318** | **0.031** | **−0.217** | **0.033** |
| Eclipse Depth | Morphology | **−0.195** | **0.195** | **−0.117** | **0.255** |
| Eclipse Depth | Temperature | **−0.139** | **0.381** | **−0.093** | **0.386** |
| Eclipse Depth | Distance | 0.027 | 0.864 | −0.008 | 0.935 |
| Eclipse Depth | Absolute Mag. | 0.049 | 0.745 | 0.038 | 0.712 |
| Eclipse Depth | Period | 0.027 | 0.857 | 0.026 | 0.798 |
| Morphology | Temperature | **0.470** | **0.002** | **0.313** | **0.004** |
| Morphology | Distance | **0.193** | **0.210** | **0.124** | **0.240** |
| Morphology | Absolute Mag. | **−0.496** | **<0.001** | **−0.320** | **0.002** |
| Morphology | Period | **−0.663** | **<0.001** | **−0.489** | **<0.001** |
| Temperature | Distance | **0.174** | **0.277** | **0.101** | **0.351** |
| Temperature | Absolute Mag. | **−0.572** | **<0.001** | **−0.453** | **<0.001** |
| Temperature | Period | **−0.137** | **0.385** | **−0.077** | **0.474** |
| Distance | Absolute Mag. | **−0.440** | **0.003** | **−0.326** | **0.002** |
| Distance | Period | 0.021 | 0.890 | 0.025 | 0.808 |
| Absolute Mag. | Period | −0.069 | 0.650 | −0.051 | 0.616 |

but every characteristic is correlated with |OES|. This arrangement is nearly the converse of the core sample's. The correlation between OES and distance is similar in both samples, but the correlation between |OES| and eclipse depth is more than twice as strong in the marginal sample. However, Figure 4.35 shows that the ODR correlation ($D = -48.7(11.5)|\text{OES}| + 0.3(1)$) is not readily visible.

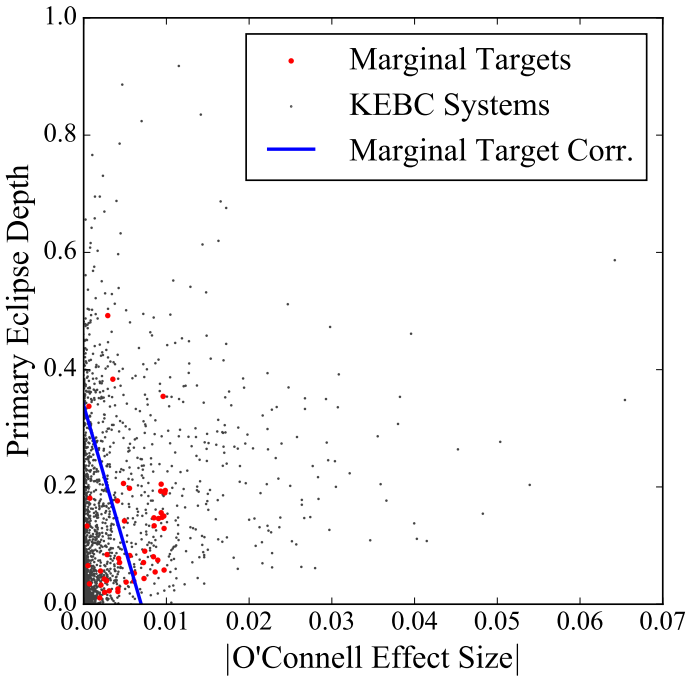


Figure 4.35: Plot comparing the primary eclipse depth to |OES|. Marginal sample targets are shown in black while non-target KEBC systems are shown in grey. The solid blue line shows the ODR correlation for the marginal sample. Both axes are in units of normalized flux. The best fit to the trend is not readily visible.

The correlation between |OES| and period is over an order of magnitude stronger in the marginal sample than in the core sample. This correlation is negative, unlike the core sample's correlation between OES and period. In spite of this, Figure 4.36 shows that ODR's best fit to the marginal sample's correlation ($P = 35.4(5.5)|\text{OES}| + 0.3(1)$) is positive. Figure 4.36 also displays the best fit line to the KEBC's correlation ($P = -3{,}021(781)|\text{OES}| + 7(2)$).

The correlation between |OES| and morphology parameter is also nearly an order of magnitude stronger in the marginal sample than in the core sample, and is similar in strength to the KEBC's correlation. Figure 4.37 plots the two characteristics against each other and shows the ODR best fit for the marginal sample

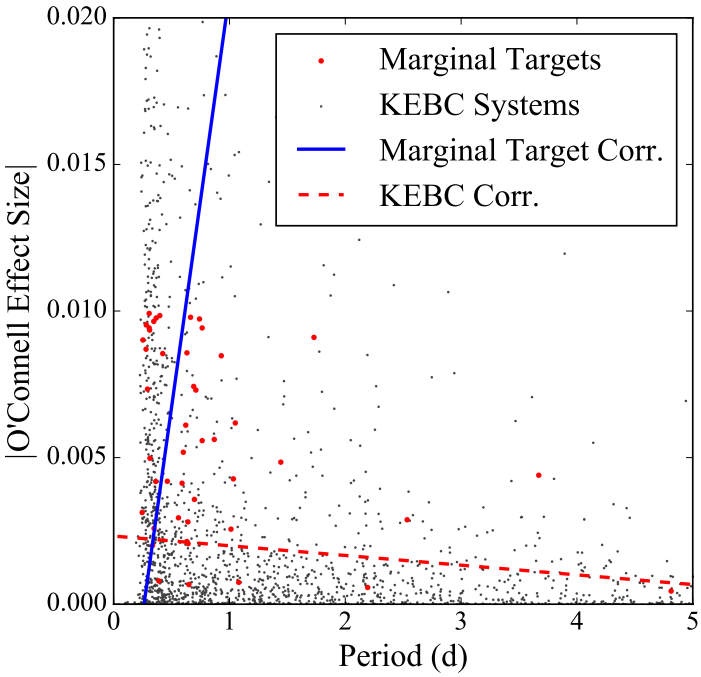


Figure 4.36: Plot comparing |OES| to the period. Marginal sample targets are shown in black while non-target KEBC systems are shown in grey. The solid blue and red dashed lines show the ODR correlations for the marginal sample and the KEBC, respectively. The y-axis is in units of normalized flux. The marginal sample's correlation trend has the opposite sign to what its $\rho$ and $\tau$ values imply.

and the KEBC ($\mu = 68.7(10.6)|\text{OES}| + 0.3(1)$ and $\mu = 74.2(1.8)|\text{OES}| + 0.37(3)$, respectively). The two lines are similar and show that |OES| increases with $\mu$. Because $\mu$ partially measures the distortion of one component by the other's gravity, this correlation is very strong evidence that binary interaction is an important factor to the O'Connell effect's cause. However, what I find odd is that the core sample does not display this correlation. My explanation for this lies in one difference between Figures 4.9 and 4.30: the marginal sample is concentrated between $\mu$ of 0.5 and 0.7, while the core sample is more widely distributed over a slightly larger range of $\mu$ (0.6 to 0.9). In particular, the three distant peaks separated by

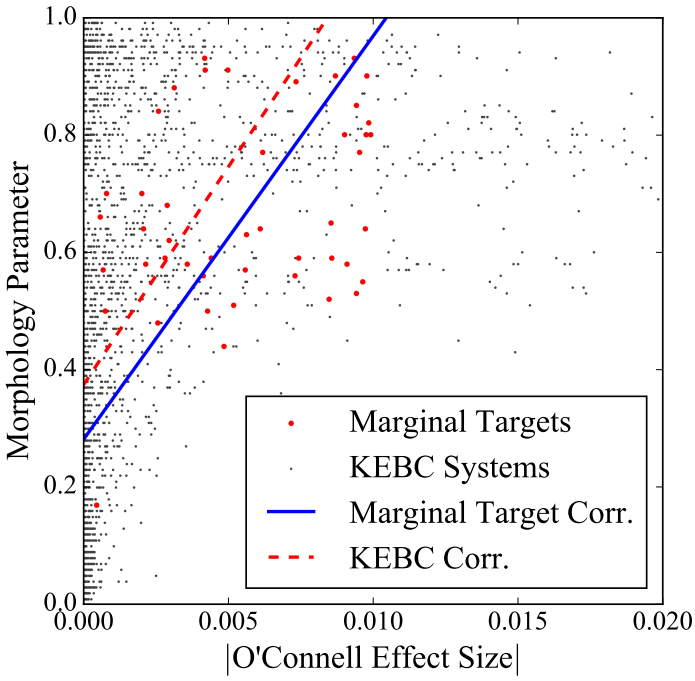


Figure 4.37: Plot comparing the morphology parameter to |OES|. Marginal sample targets are shown in black while non-target KEBC systems are shown in grey. The solid blue and red dashed lines show the ODR correlations for the marginal sample and the KEBC, respectively. The x-axis is in units of normalized flux. The marginal sample and the KEBC have similar trends.

$\mu \approx 0.2$ are quite different than the marginal sample's single prominent peak at $\mu = 0.6$. This distribution spreads the OES across a range of $\mu$, as Figure 4.19's top-center panel shows. I conclude from this difference that there may be several prominent causes for a significant O'Connell effect that causes this spread in $\mu$, while a smaller O'Connell effect may be due to a single cause. This single cause also occurs in systems with a significant O'Connell effect, as a peak around $\mu = 0.6$ occurs in both samples.

Table 4.5 indicates that the correlation between period and absolute magnitude discussed in Section 4.2.3.4 is absent in the marginal sample. This absence is most

likely a result of the lower proportion of marginal sample systems with periods at the minimum period cutoff mentioned in Section 4.2.3.4 (see Figure 4.26). Figure 4.33's lower-left panel confirms that fewer marginal sample systems lie on the cutoff line near the panel's left edge.

# Chapter 5

# Peculiar System Classes

There are four system classes containing peculiar features in their light curve present in the sample. These four classes are those displaying significant temporal variation, those with an asymmetric minimum, those with a concave-up region in their light curves, and a white dwarf. Any given system can belong to multiple classes. I have selected an exemplar system to serve as an identifier for each class for the purposes of categorization. In this chapter, I discuss each of the four classes and identify characteristics systems in each class share. I also statistically analyze three classes to discover any trends or correlations that may exist. In this chapter, I compare the peculiar system classes to the rest of the complete sample (excluding that class's systems) rather than the KEBC.

## 5.1 KIC 7433513: Temporally Varying Systems

Many systems in my sample exhibit dramatic changes in their light curves over multiple orbits. I chose KIC 7433513 as the exemplar of this class, as it was the first

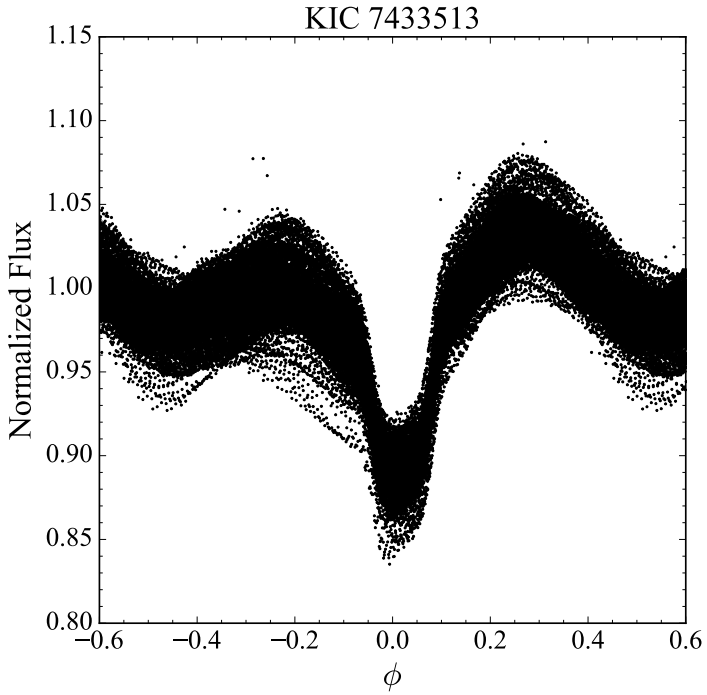


Figure 5.1: *Kepler* light curve of KIC 7433513 showing significant scatter caused by temporal variation.

such system I noticed and the first one I observed (Section 6.2.1). My light curve averaging removes the temporal variation present in such systems, so the variation must be studied separately. I created the variance parameter (Section 2.2.6) to quantify the amount of variation, and I also produced diagrams showing how these systems vary over time. The most likely explanation for this variation is the movement and evolution of chromospheric starspots across one or both components.

### 5.1.1 Class Description

KIC 7433513 and systems like it (labeled with the flag TV in Table A.1) display strong temporal variation in their light curves. Figure 5.1 shows KIC 7433513's phased light curve using all *Kepler* data, clearly showing the significant data scat-

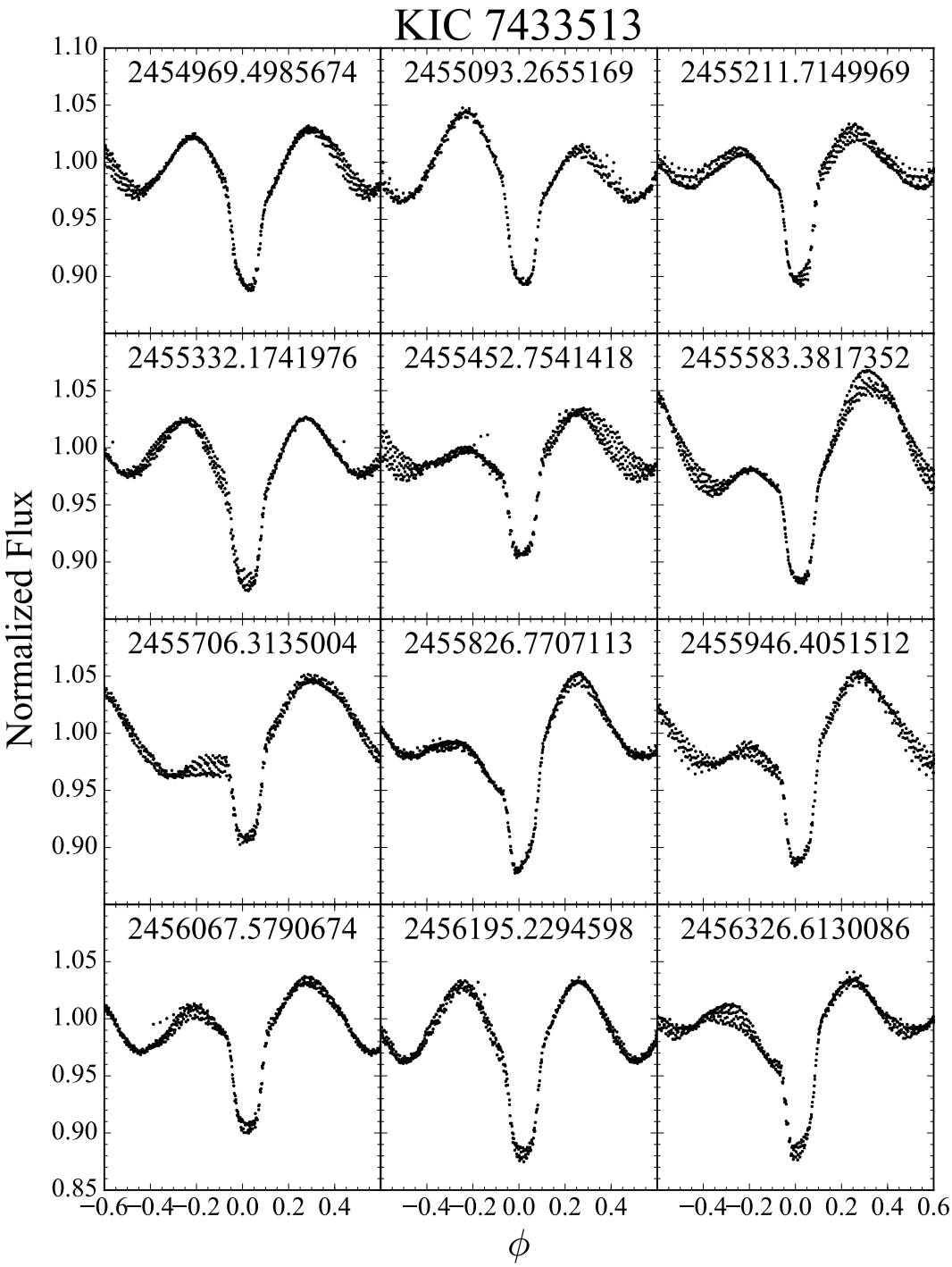


Figure 5.2: Ten-day slices of KIC 7433513's phased *Kepler* data separated by several months, showing the significant changes in the light curve over time. The number in each time slice's subplot is the BJD for the midpoint of that time slice.

Table 5.1: List of temporally varying systems.

| *Kepler* ID | Unscaled Variance | Scaled Variance | Sample |
|---|---|---|---|
| 2569494 | 0.023 | 0.155 | C |
| 2577756 | 0.012 | 0.117 | M |
| 2695740 | 0.012 | 0.079 | C |
| 3124420 | 0.015 | 0.063 | C |
| 3339563 | 0.019 | 0.087 | C |
| 3342425 | 0.014 | 0.054 | C |
| 3351945 | 0.013 | 0.398 | M |
| 3656322 | 0.021 | 0.399 | M |
| 3659940 | 0.016 | 0.186 | C |
| 3662635 | 0.011 | 0.096 | C |
| 3954227 | 0.029 | 0.160 | M |
| 3965242 | 0.017 | 0.519 | C |
| 4037163 | 0.009 | 0.065 | M |
| 4077442 | 0.012 | 0.125 | M |
| 4349483 | 0.014 | 0.136 | C |
| 4474193 | 0.014 | 0.065 | C |
| 4474637 | 0.013 | 0.067 | C |
| 5215999 | 0.006 | 0.056 | M |
| 5615528 | 0.003 | 0.218 | M |
| 5774375 | 0.024 | 0.058 | C |
| 6034812 | 0.002 | 0.058 | M |
| 6191574 | 0.031 | 0.467 | M |
| 6197038 | 0.051 | 0.144 | C |
| 6264091 | 0.021 | 0.172 | C |
| 6267702 | 0.006 | 0.078 | M |
| 6421483 | 0.011 | 0.092 | C |
| 6431545 | 0.015 | 0.083 | M |
| 6436038 | 0.018 | 0.160 | C |
| 6836140 | 0.014 | 0.058 | C |
| 7119757 | 0.011 | 0.065 | M |
| 7199183 | 0.010 | 0.061 | C |
| 7284688 | 0.007 | 0.116 | M |
| 7433513 | 0.012 | 0.080 | C |
| 7584826 | 0.014 | 0.050 | C |
| 7671594 | 0.006 | 0.056 | C |

Table 5.1 *continued on next page*

Table 5.1 *(continued)*

| *Kepler* ID | Unscaled Variance | Scaled Variance | Sample |
|---|---|---|---|
| 7885570 | 0.014 | 0.088 | M |
| 8112324 | 0.017 | 0.110 | C |
| 8211824 | 0.014 | 0.178 | C |
| 8285349 | 0.004 | 0.052 | M |
| 8294484 | 0.009 | 0.070 | C |
| 8298344 | 0.013 | 0.149 | C |
| 8479107 | 0.013 | 0.078 | M |
| 8846978 | 0.018 | 0.099 | C |
| 9137992 | 0.020 | 0.166 | C |
| 9145846 | 0.022 | 0.127 | C |
| 9328852 | 0.016 | 0.053 | C |
| 9419603 | 0.005 | 0.064 | M |
| 9450883 | 0.014 | 0.115 | C |
| 9466316 | 0.003 | 0.117 | M |
| 9527167 | 0.003 | 0.061 | C |
| 9551200 | 0.021 | 0.299 | M |
| 9767437 | 0.004 | 0.118 | C |
| 9935245 | 0.025 | 0.084 | C |
| 9935311 | 0.020 | 0.053 | C |
| 10189523 | 0.013 | 0.329 | M |
| 10322582 | 0.018 | 0.053 | C |
| 10350225 | 0.022 | 0.108 | C |
| 10351767 | 0.002 | 0.062 | M |
| 10676927 | 0.021 | 0.129 | C |
| 10711646 | 0.013 | 0.129 | C |
| 11076176 | 0.022 | 0.055 | M |
| 11127048 | 0.005 | 0.057 | C |
| 11193447 | 0.007 | 0.145 | M |
| 11244501 | 0.004 | 0.059 | M |
| 11347875 | 0.035 | 0.076 | C |
| 11404698 | 0.023 | 0.103 | C |
| 11671660 | 0.006 | 0.096 | C |
| 11704155 | 0.007 | 0.061 | C |
| 11920266 | 0.002 | 0.072 | M |
| 12055517 | 0.013 | 0.073 | C |
| 12109575 | 0.011 | 0.115 | C |

Table 5.1 *continued on next page*

Table 5.1 *(continued)*

| *Kepler* ID | Unscaled Variance | Scaled Variance | Sample |
|---|---|---|---|
| 12218858 | 0.022 | 0.126 | M |
| 12400729 | 0.017 | 0.213 | M |
| 12458797 | 0.006 | 0.051 | C |
| 12602985 | 0.004 | 0.109 | M |

ter caused by this temporal variation. Figure 5.2 shows 10-day time slices of KIC 7433513's *Kepler* data separated by a few months, along with the Barycentric Julian Date (BJD) for the midpoint of each time slice. Figure 5.2 could be understood as individual time intervals of Figure 2.10, albeit with a 10-day interval instead of a 4.087-day interval. Figure 5.2 demonstrates that KIC 7433513 varies significantly even on these short timescales.

I used the SVP to quantify the degree of variation for each system in my sample. Based on visual inspection of the light curves, I chose to use the criterion SVP $\geq$ 0.05 to define the temporally varying system class. Figure 5.3 shows the complete sample's SVP distribution, and SVP = 0.05 roughly corresponds to the $75^{\text{th}}$ percentile. Based on this criterion, 46 core sample systems (22% of the core sample) and 29 marginal sample systems (63% of the marginal sample) are members of this class. Table 5.1 lists all 75 temporally varying systems, along with each system's UVP and SVP and which sample they belong to. I should stress that criterion SVP $\geq$ 0.05 is arbitrary, and a continuum exists between temporally varying systems and temporally stable systems.

### 5.1.2 Class Analysis

In this section, I will summarize the characteristic distributions and trends and look at any correlations between the characteristics. I also discuss the results of

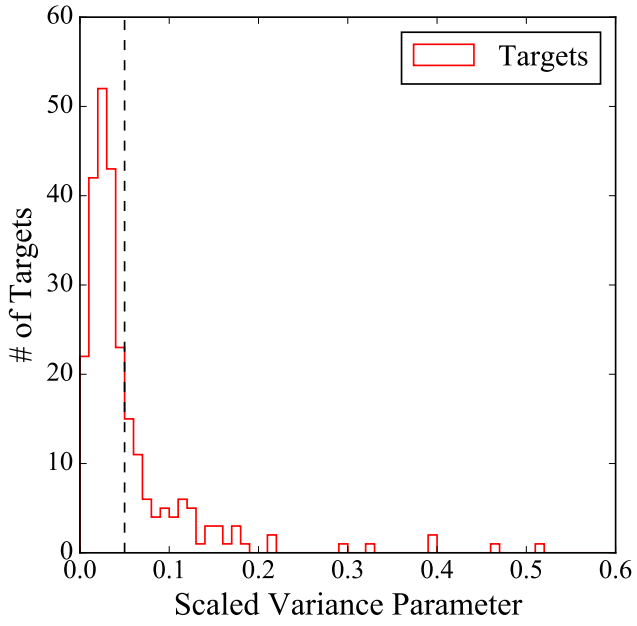


Figure 5.3: Histogram showing the SVPs of all 258 complete sample systems (red). The dashed line indicates where SVP = 0.05, which defines the cutoff for temporally varying systems.

my analysis of this class of systems. For the temporally varying systems, I have added the SVP into the list of characteristics I analyze.

#### 5.1.2.1 Sample Characteristics

The temporally varying system class is comprised of 46 (61%) Algol-type, 10 (13%) $\beta$ Lyrae-type, and 19 (25%) W Ursae Majoris-type systems. These proportions are nearly identical to the marginal sample (65% Algol-type, 15% $\beta$ Lyrae-type, and 20% W Ursae Majoris-type; Section 4.3.1.1). I found this surprising because the majority of temporally varying systems (46; 61%) are core sample systems, which has a very different proportion of light curve classes (25% Algol-type, 19% $\beta$ Lyrae-type, and 56% W Ursae Majoris-type; Section 4.1.1).

The largest OES found in this class is –0.265 (KIC 11347875), while the small-

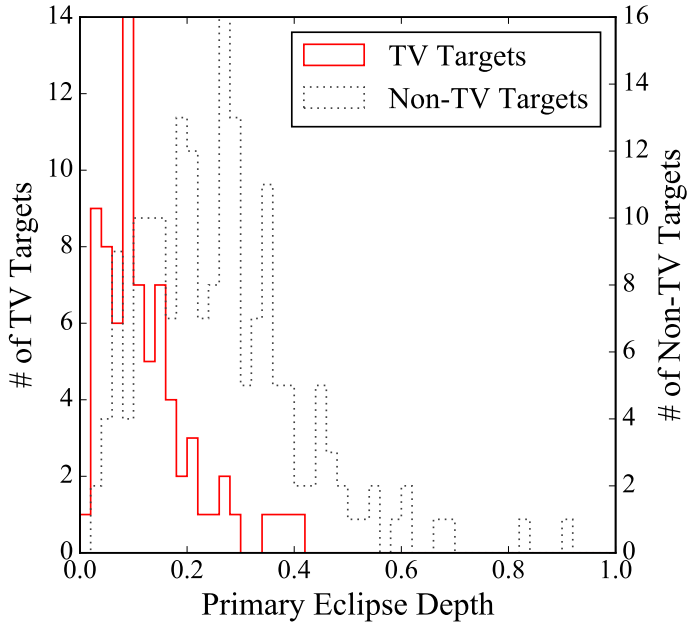


Figure 5.4: Histogram comparing the primary eclipse depths for all 75 temporally varying targets (solid red) and the remaining 183 complete sample targets (dotted grey). Temporally varying systems have shallower eclipses on average than other complete sample systems.

est is 0.0008 (KIC 3351945), both in units of normalized flux. Most temporally varying systems show a positive O'Connell effect (49 systems, or 65% of this class), similar to the core (69%) and marginal (70%) samples. The temporally varying systems' SVP ranges from 0.050 (KIC 7584826) to 0.519 (KIC 3965242, shown in Figure 2.11's right panel). By contrast, the rest of the complete sample's SVP ranges from 0.002 (KIC 10857342, shown in Figure 2.11's left panel) to 0.049 (KIC 3557421).

The primary eclipse depth ranges from 0.012 (KIC 5615528) to 0.403 (KIC 1134-7875). Figure 5.4 shows a histogram comparing the temporally varying systems' eclipse depth distribution to the remaining 183 complete sample systems' distribution. The stark difference between the two distributions demonstrates that tem-

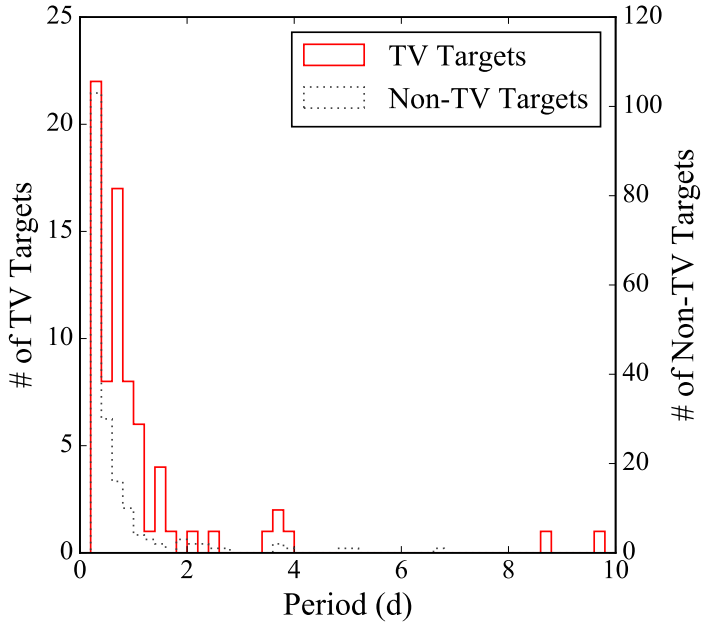


Figure 5.5: Histogram comparing the periods for all 75 temporally varying systems (solid red) and the remaining 183 complete sample targets (dotted grey). Temporally varying systems have longer on average than other complete sample systems.

porally varying systems have much shallower eclipses than non-varying systems in the sample.

The temporally varying systems' distances range from 105.1±0.1 pc (KIC 1124-4501) to $6{,}474^{+3{,}601}_{-3{,}897}$ pc (KIC 3662635). The class's luminosities range from 0.0169± 0.0004 $L_\odot$ (KIC 7671594) to $32.24^{+35.87}_{-38.82}$ $L_\odot$ (KIC 3662635). The orbital periods of this class range from 0.241 d (KIC 12458797) to 9.752 d (KIC 6197038). Figure 5.5 shows a histogram comparing the temporally varying systems' period distribution to the remaining 183 complete sample systems' distribution. The temporally varying systems have a much higher proportion of systems with 0.5 d $\leq P \leq$ 1 d than the rest of the sample. This fact suggests that temporal variation is somewhat less dependent on binary interaction than the O'Connell effect is.

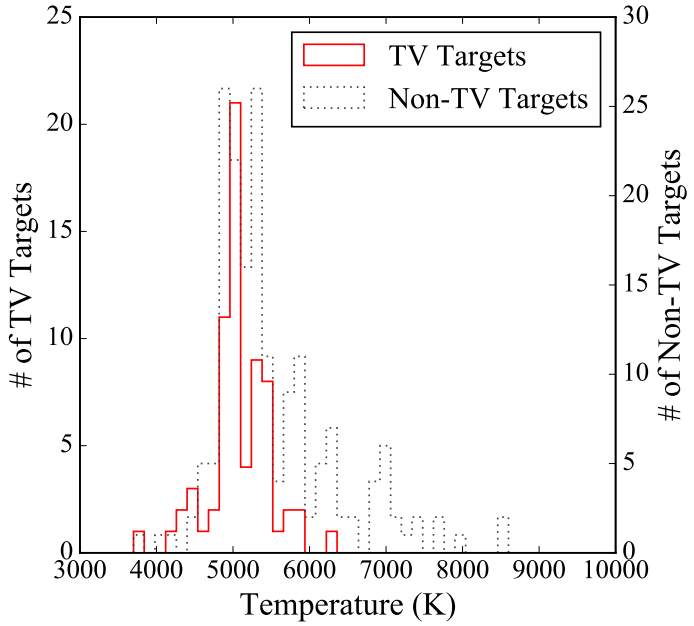


Figure 5.6: Histogram comparing the *Gaia* DR2 temperatures for 69 of the 75 temporally varying systems (solid red) and 178 of the remaining 183 complete sample targets (dotted grey) with a *Gaia* DR2 temperature. Temporally varying systems are cooler on average than other complete sample systems.

Figure 5.6 shows a temperature histogram for the temporally varying systems compared to the rest of the complete sample. It indicates that temporally varying systems are among the cooler systems of my sample. The hottest temporally varying system, KIC 8285349, is estimated to be $6{,}276^{+208}_{-400}$ K. I interpret the complete lack of high-temperature temporally varying systems as a suggestion that starspots cause the temporal variation. Starspots, like sunspots on our parent star, are caused by convective cells in the star's upper layers. Stars hotter than ~7,000 K, however, have radiative upper layers, and so convection (and thus starspots) is not expected for such stars. This fact does not preclude starspots at temperatures between 6,000 K and 7,000 K, however. Looking at the SVPs of these systems, only two systems have SVP $\geq$ 0.025 and $T \geq$ 6,000 K: KIC 8285349 (SVP =

0.052, $T = 6{,}276^{+208}_{-400}$ K) and KIC 11175495 (SVP = 0.040, $T = 8{,}575^{+151}_{-185}$ K). Both systems have the KEBC flag SPD, which means that, while they appear to be overcontact systems, they are short-period detached systems. Being detached, it is plausible that starspots could occur on the cooler component rather than the hotter one. Based on the near lack of significant variance above 6,000 K, I conclude that stellar activity sharply drops above this temperature, explaining the lack of hot temporally varying systems. This drop in activity may be related to the thinning convective layer in the star's upper layers as the temperature increases.

Frasca et al. (2016) reported spectral types for temporally varying systems ranging from K5 Ve (KIC 12109575) to F3/F5 V (KIC 8285349). Among temporally varying systems, there are two subgiants (KICs 7885570 and 11920266), one giant (KIC 3339563), and one bright giant (KIC 9137992). Except for KIC 7885570, these four systems' orbital periods and estimated luminosities are too small to support Frasca et al.'s (2016) spectral types, making it more likely that these spectral types are misidentified due to the stars' rapid rotation. The spectral types estimated using Dr. Mamajek's table range from M2.5 V (KIC 7671594) to F8 V (KIC 11920266).

The temporally varying systems' morphology parameters range from 0.34 (KIC 7671594) to 0.97 (KICs 6264091, 7199183, 9527167). Figure 5.7 shows a histogram of the temporally varying systems' $\mu$ distribution compared to the rest of the complete sample's. The histogram clearly shows a lower abundance of temporally varying systems in that region, which contrasts with the rest of the complete sample's most prominent peak occurring there. On the other hand, Figure 5.7's largest peak coincides with the leftmost of Figure 4.9's three prominent peaks. Indeed, temporally varying systems make up a majority of complete sample systems

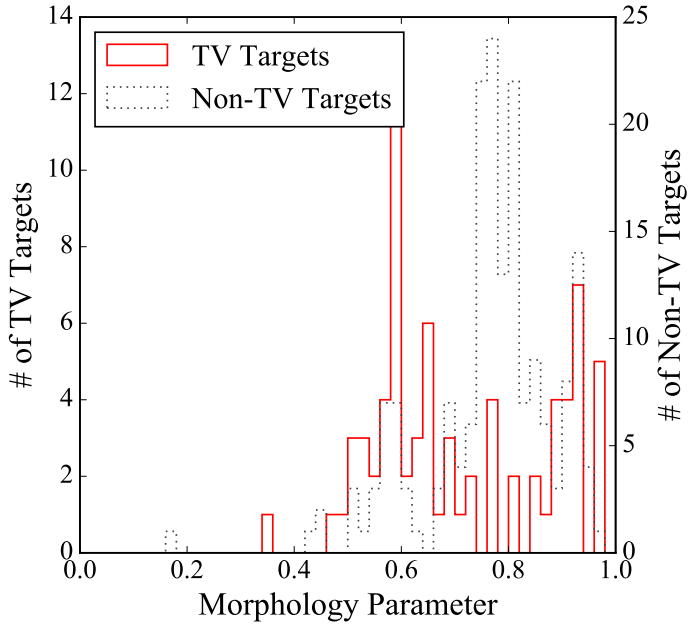


Figure 5.7: Histogram comparing the morphology parameters for all 75 temporally varying systems (solid red) and 182 of the remaining 183 complete sample targets (dotted grey) that have $\mu \neq -1$. Temporally varying systems have two peaks at $\mu \approx 0.60$ and $\mu \approx 0.95$, and relatively few temporally varying systems lie where the other complete sample systems peak near $\mu \approx 0.80$.

with $\mu \approx 0.55$-$0.60$. The temporally varying systems have a secondary peak near $\mu \approx 0.95$, coinciding with the rightmost of Figure 4.9's three prominent peaks.

#### 5.1.2.2 Characteristic Trends

Figures 5.8-5.10 show the temporally varying systems' corner plot comparing period, OES, *Gaia* temperature, the *Gaia* BP – RP color index, morphology parameter, absolute *Gaia* G magnitude, primary eclipse depth, and scaled variance parameter. One readily apparent trend is the tendency for these systems to have shallow eclipses, agreeing with the distribution Figure 5.4 shows. Like the marginal sample, temporally varying systems lack the core sample's temperature contrast between positive and negative O'Connell effect systems. The largest

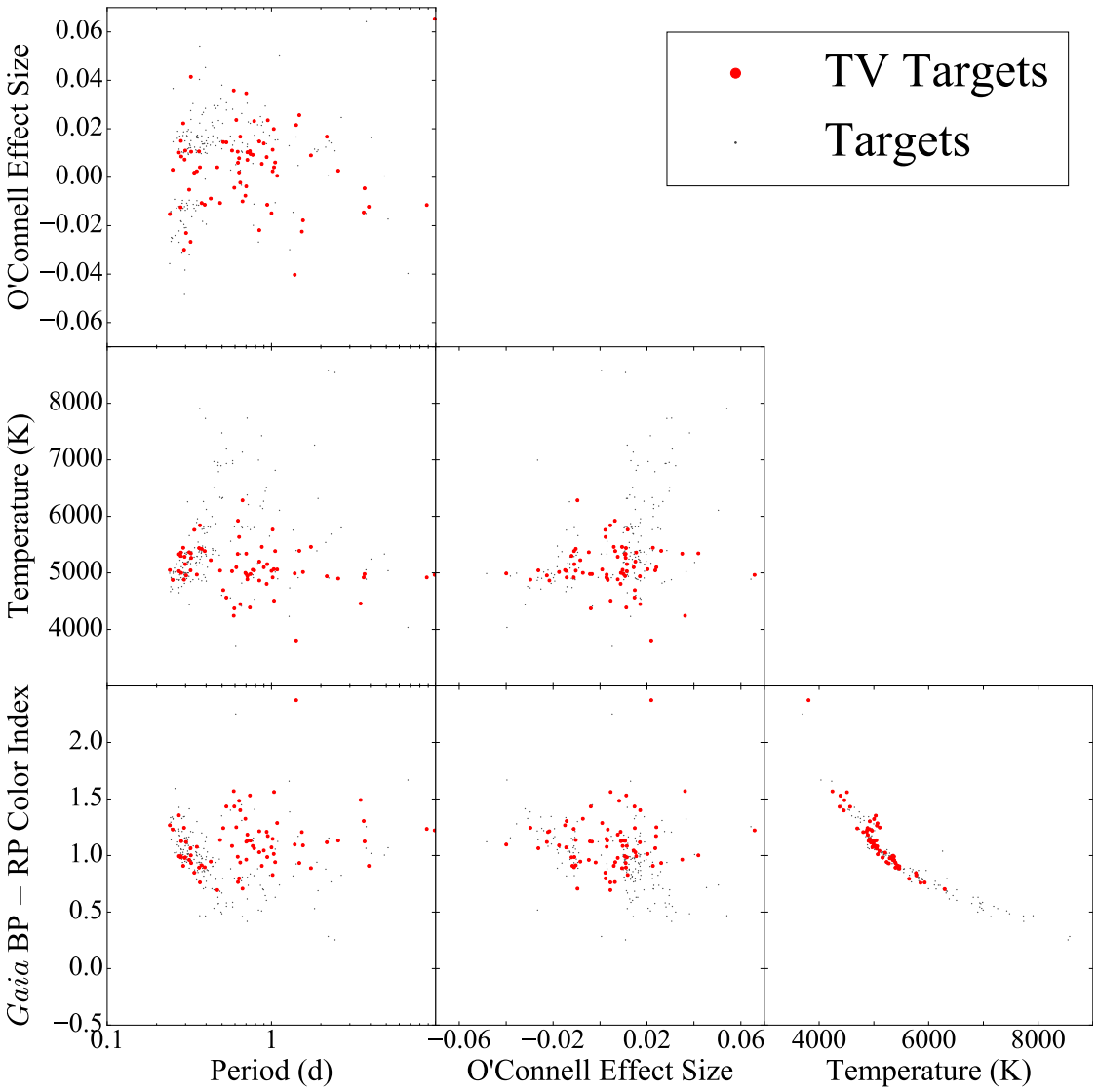


Figure 5.8: Corner plot showing the temporally varying systems' correlations between three characteristics of interest (period, OES, and temperature) and OES, temperature, and color. Temporally varying targets are shown in red while the other complete sample systems are shown in grey. KICs 9935311 and 11347875 have been removed from the OES plots for clarity (see discussion in Section 4.2.1). Note the logarithmic x-axis of the period plots.

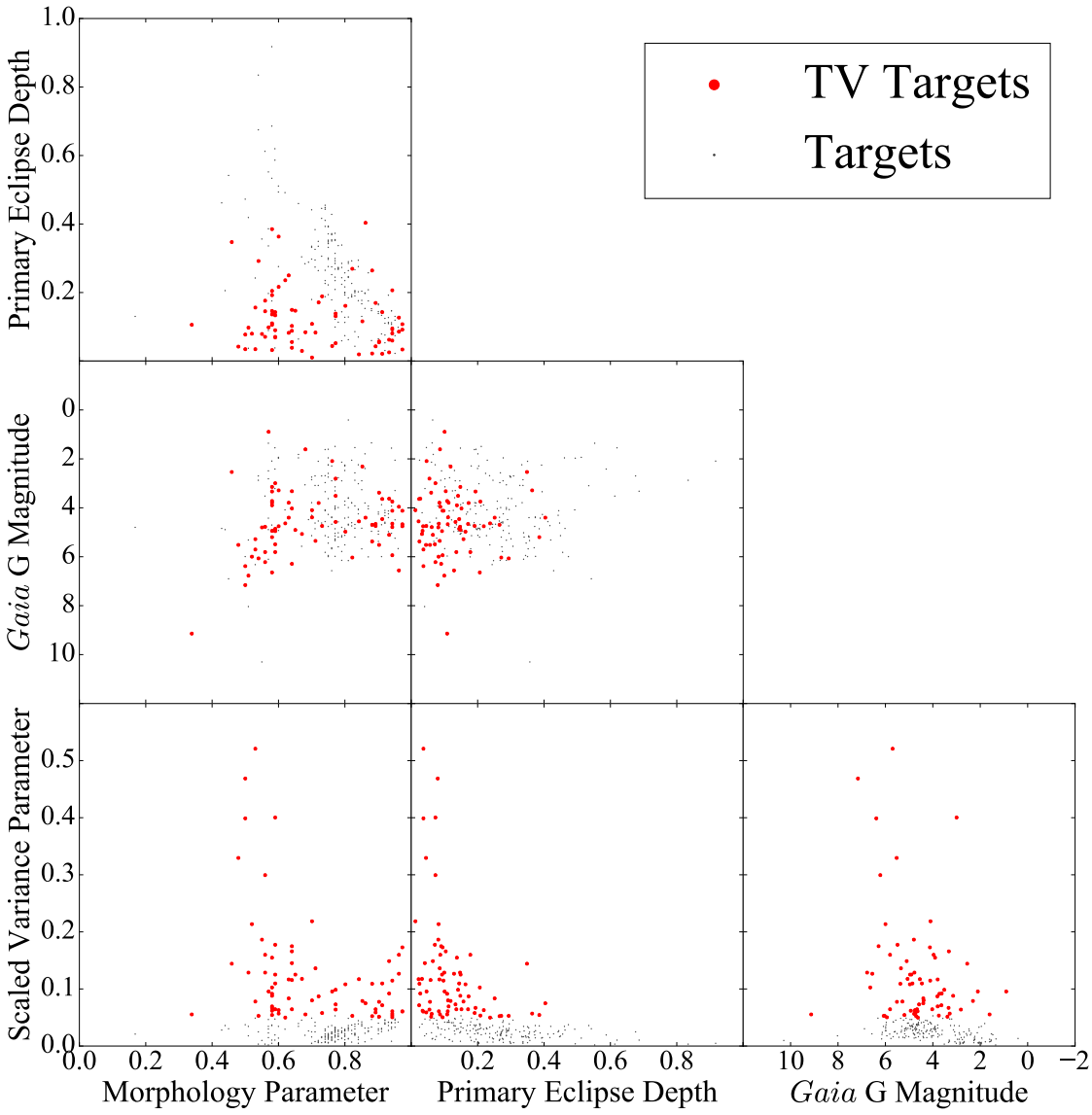


Figure 5.9: Corner plot showing the temporally varying systems' correlations between three characteristics of interest (morphology parameter, primary eclipse depth, and absolute magnitude) and primary eclipse depth, absolute magnitude, and scaled variance. Temporally varying targets are shown in red while the other complete sample systems are shown in grey.

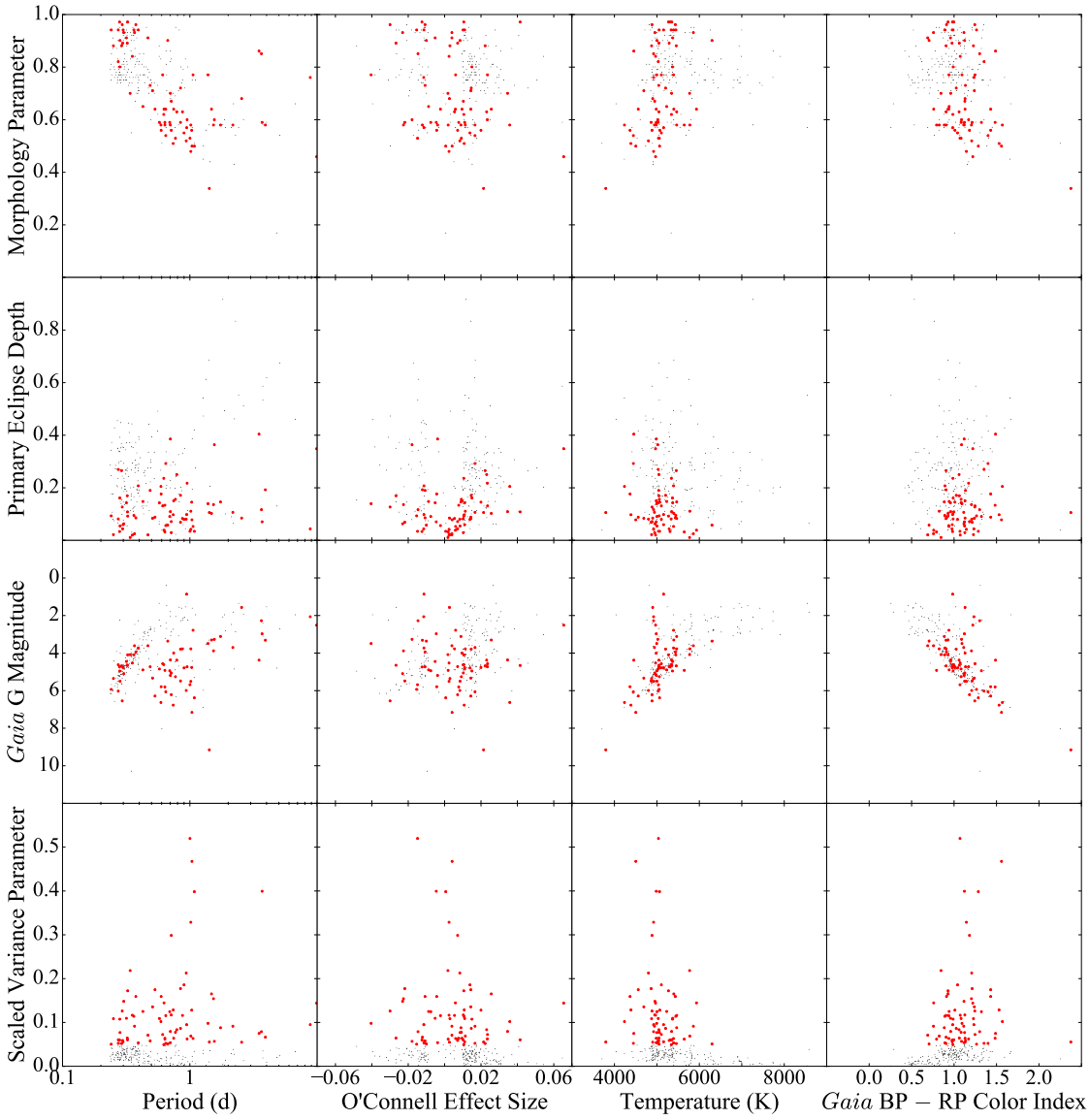


Figure 5.10: Corner plot showing the temporally varying systems' correlations between four characteristics of interest (period, OES, temperature, and color) and morphology parameter, primary eclipse depth, absolute magnitude, and scaled variance. Temporally varying targets are shown in red while the other complete sample systems are shown in grey. KICs 9935311 and 11347875 have been removed from the OES plots for clarity (see discussion in Section 4.2.1). Note the logarithmic x-axis of the period plots.

SVP values (i.e. SVP $\geq 0.3$) are concentrated in relatively narrow ranges of OES ($-0.015 \leq$ OES $\leq 0.005$), temperature ($4{,}500 \leq T \leq 5{,}100$ K), morphology parameter ($0.48 \leq \mu \leq 0.59$), and primary eclipse depth ($D \leq 0.01$). The primary eclipse depth trend is particularly strong, and the largest SVP for a system with $D > 0.1$ is 0.166 (KIC 9137992). However, the primary eclipse depth trend may be influenced by the fact that most temporally varying systems have shallow eclipses. Regardless, these SVP concentrations imply that the most extreme variation only occurs under rather specific conditions.

Another subtle trend is that Figures 5.8-5.10 do not clearly show a lack of systems with $\mu \approx 0.65$-$0.75$ like Figures 4.17-4.19 and Figures 4.31-4.33 do, although the abundance is somewhat reduced. Figure 5.7 clearly shows an underabundance of temporally varying systems in this range of $\mu$. As noted previously, Figure 5.7's largest peak coincides with a peak in the core sample distribution (Figure 4.9), and temporally varying systems make up the majority of the sample around $\mu = 0.60$.

#### 5.1.2.3 Sample Distributions

Table 5.2 gives the results of the K–S test for the temporally varying systems. The K–S test rejects the null hypothesis for all but three characteristics: temperature, distance, and absolute magnitude. The similarity in the distance is unsurprising because the presence of temporal variation should be distance-independent, assuming it is due to the physical conditions of the system rather than an effect of intervening matter. The temperature distribution's similarity is unexpected, however, particularly in light of the stark difference Figure 5.6 shows. I suspect that the small K–S statistic is due to the distributions' similarity in the low-temperature regime, which coincides with the peak of the remaining systems' distribution. Therefore,

Table 5.2: Results of analyzing characteristic distributions between the temporally varying systems and the rest of the complete sample using the Kolomgorov–Smirnov test.

| Characteristic | K–S Statistic | $p$-Value |
|---|---|---|
| O'Connell Effect Size | 0.307 | $< 0.001$ |
| \|O'Connell Effect Size\| | 0.341 | $< 0.001$ |
| Primary Eclipse Depth | 0.524 | $< 0.001$ |
| Morphology Parameter | 0.366 | $< 0.001$ |
| Temperature | 0.221 | 0.009 |
| Distance | 0.137 | 0.249 |
| Absolute Magnitude | 0.202 | 0.022 |
| Period | 0.355 | $< 0.001$ |
| Scaled Variance Parameter | 1.000 | $< 0.001$ |

the high-temperature systems constitute a small number of outliers from an otherwise similar distribution, causing the K–S statistic's small value. The large K–S statistic for primary eclipse depth reflects the distribution difference seen in Figure 5.4. Finally, the K–S statistic of 1.0 for the SVP – indicating that the SVP distributions of the temporally varying systems and the rest of the complete sample are disjoint – is required by the definition of the temporally varying class: any system with an SVP large enough to overlap with the temporally varying systems would be included in the class to begin with.

#### 5.1.2.4 Characteristic Correlations

Table 5.3 gives the results of my analysis of the temporally varying systems using Spearman's $\rho$ and Kendall's $\tau$ coefficients. Four correlations are particularly strong: |OES| with eclipse depth, morphology with period, and both temperature and distance with absolute magnitude. I expected the last three based on underlying physical principles. The correlation between |OES| and eclipse depth, however, has been consistently found in each sample I have analyzed. I will focus

Table 5.3: Results of analyzing the correlations between the temporally varying systems' characteristics of interest using Spearman's and Kendall's rank correlation coefficients.

| Characteristic One | Characteristic Two | Spearman's Coeff. | Spearman's $p$-Value | Kendall's Coeff. | Kendall's $p$-Value |
|---|---|---|---|---|---|
| OES | Eclipse Depth | **0.203** | **0.081** | **0.158** | **0.045** |
| OES | Morphology | **−0.181** | **0.120** | **−0.120** | **0.132** |
| OES | Temperature | 0.064 | 0.604 | 0.036 | 0.660 |
| OES | Distance | −0.033 | 0.782 | −0.025 | 0.753 |
| OES | Absolute Mag. | **0.127** | **0.278** | **0.077** | **0.330** |
| OES | Period | −0.054 | 0.648 | −0.031 | 0.691 |
| OES | SVP | **−0.100** | **0.393** | **−0.069** | **0.382** |
| \|OES\| | Eclipse Depth | **0.550** | **< 0.001** | **0.404** | **< 0.001** |
| \|OES\| | Morphology | **0.109** | **0.352** | **0.084** | **0.290** |
| \|OES\| | Temperature | **−0.154** | **0.207** | **−0.110** | **0.180** |
| \|OES\| | Distance | 0.083 | 0.484 | 0.044 | 0.581 |
| \|OES\| | Absolute Mag. | −0.074 | 0.526 | −0.055 | 0.484 |
| \|OES\| | Period | 0.066 | 0.572 | 0.043 | 0.586 |
| \|OES\| | SVP | **−0.159** | **0.173** | **−0.110** | **0.163** |
| Eclipse Depth | Morphology | **−0.149** | **0.203** | **−0.090** | **0.262** |
| Eclipse Depth | Temperature | **−0.113** | **0.355** | **−0.089** | **0.279** |
| Eclipse Depth | Distance | **0.203** | **0.085** | **0.131** | **0.101** |
| Eclipse Depth | Absolute Mag. | **−0.106** | **0.363** | **−0.065** | **0.413** |
| Eclipse Depth | Period | **0.145** | **0.215** | **0.103** | **0.192** |
| Eclipse Depth | SVP | **−0.345** | **0.002** | **−0.256** | **0.001** |
| Morphology | Temperature | **0.374** | **0.002** | **0.254** | **0.002** |
| Morphology | Distance | 0.069 | 0.562 | 0.044 | 0.590 |
| Morphology | Absolute Mag. | **−0.251** | **0.030** | **−0.166** | **0.037** |
| Morphology | Period | **−0.654** | **< 0.001** | **−0.467** | **< 0.001** |
| Morphology | SVP | **−0.357** | **0.002** | **−0.259** | **0.001** |
| Temperature | Distance | **0.268** | **0.027** | **0.183** | **0.027** |
| Temperature | Absolute Mag. | **−0.435** | **< 0.001** | **−0.332** | **< 0.001** |
| Temperature | Period | **−0.246** | **0.041** | **−0.161** | **0.051** |
| Temperature | SVP | **−0.231** | **0.056** | **−0.158** | **0.055** |
| Distance | Absolute Mag. | **−0.528** | **< 0.001** | **−0.383** | **< 0.001** |
| Distance | Period | **0.230** | **0.050** | **0.171** | **0.032** |
| Distance | SVP | −0.076 | 0.521 | −0.051 | 0.523 |
| Absolute Mag. | Period | **−0.361** | **0.001** | **−0.245** | **0.002** |
| Absolute Mag. | SVP | **0.207** | **0.075** | **0.134** | **0.090** |
| Period | SVP | **0.240** | **0.038** | **0.165** | **0.036** |

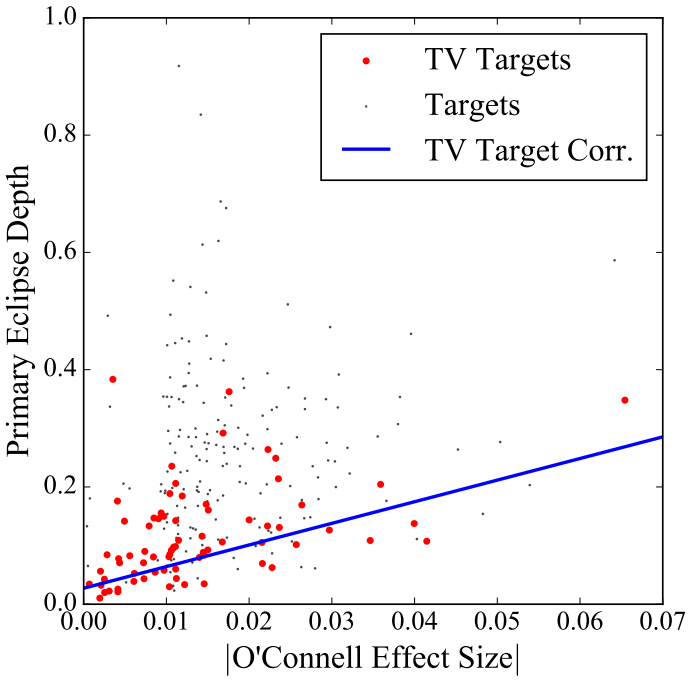


Figure 5.11: Plot comparing the primary eclipse depth to |OES|. Temporally varying systems are shown in red while other complete sample targets are shown in grey. The solid blue line shows the ODR correlation for the temporally varying systems. Both axes are in units of normalized flux. KICs 9935311 and 11347875 are excluded for clarity. The trend is readily visible.

on discussing this correlation in the context of temporally varying systems.

The correlation between |OES| and the eclipse depth is much stronger among temporally varying systems than in either the core or marginal samples. Figure 5.11 plots these characteristics against each other along with ODR's best fit to the data ($D = 3.68(47)|\mathrm{OES}| + 0.03(1)$). Unlike with the core and marginal sample plots (Figures 4.21 and 4.35, respectively), however, the best fit relation is readily discernible. While there are several outliers (including KICs 9935311 and 11347875, which I excluded from Figure 5.11 for clarity), it should be noted that three of the most prominent outliers (KICs 5774375, 9328852, and 11076176) have SVP $< 0.06$,

only slightly above the cutoff for inclusion in this class. There appears to be a weak but genuine relationship between |OES| and primary eclipse depth for temporally varying systems, wherein |OES| increases as the primary eclipse gets deeper. However, it is also possible that this is merely a consequence of the lack of deep eclipses among temporally varying systems, as Figure 5.4 indicates.

Finally, while tangential to the temporally varying systems, the complete sample's correlation between the SVP and the primary eclipse depth shown in Figure 5.9's lower-center panel is important to discuss. The two characteristics appear to be inversely related to each other when considering both the red and grey points. The scaling applied in Section 2.2.6 to produce the SVP would be expected to produce precisely such a trend if the data scatter was an artifact of the scaling rather than a physical process. This possibility appears unlikely considering the differences I have found between the temporally varying systems and non-temporally varying systems. Therefore, this trend suggests that there is some relationship between the eclipse depth and temporal variability. In other words, the shallower a primary eclipse is, the more temporal variation we tend to observe.

### 5.1.3 Potential Explanations

In Section 3.8, I discussed Kunt & Dal's (2017) analysis of KIC 7885570, a temporally varying marginal sample system. Kunt & Dal (2017) described KIC 7885570 as an RS Canum Venaticorum variable. These binary systems have at least one star with an active chromosphere that produces large spots (Rodonò et al. 1995), and the migration of such spots is a plausible source for the temporal variations observed. Indeed, Figure 5.12, displaying KIC 3339563's variance diagram, shows that the variance evolves in a manner invoking the image of a feature moving in

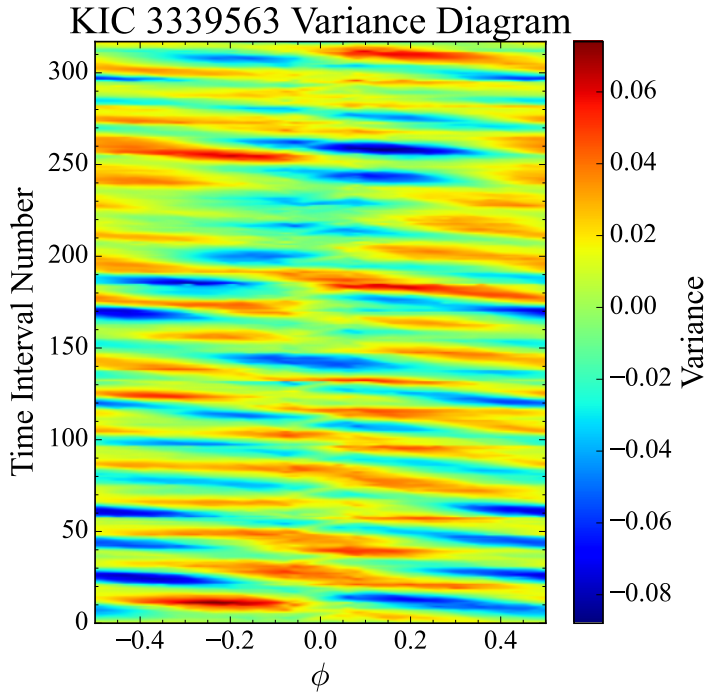


Figure 5.12: Variance diagram of KIC 3339563 showing a variance pattern similar to what migrating starspots would produce. KIC 3339563's time intervals are 4.087 d long. The striped pattern resembles a feature moving in longitude.

longitude across one of the components.

I hypothesize that chromospheric starspots are responsible for the temporal variation observed in temporally varying systems. To test this hypothesis, I looked at the findings of Tran et al. (2013) and Balaji et al. (2015), which I discussed in Section 3.2. Those two papers listed 27 temporally varying systems (36% of the class) that exhibit an anticorrelated ETV. Both Tran et al. (2013) and Balaji et al. (2015) ascribe anticorrelated ETVs to chromospheric starspots migrating in longitude, so at first glance, over a third of temporally varying systems have clear evidence of starspot evolution from these papers alone. However, these papers focused exclusively on overcontact or near-contact systems, but the majority of temporally varying systems are Algol-type, which should not fall under the scope

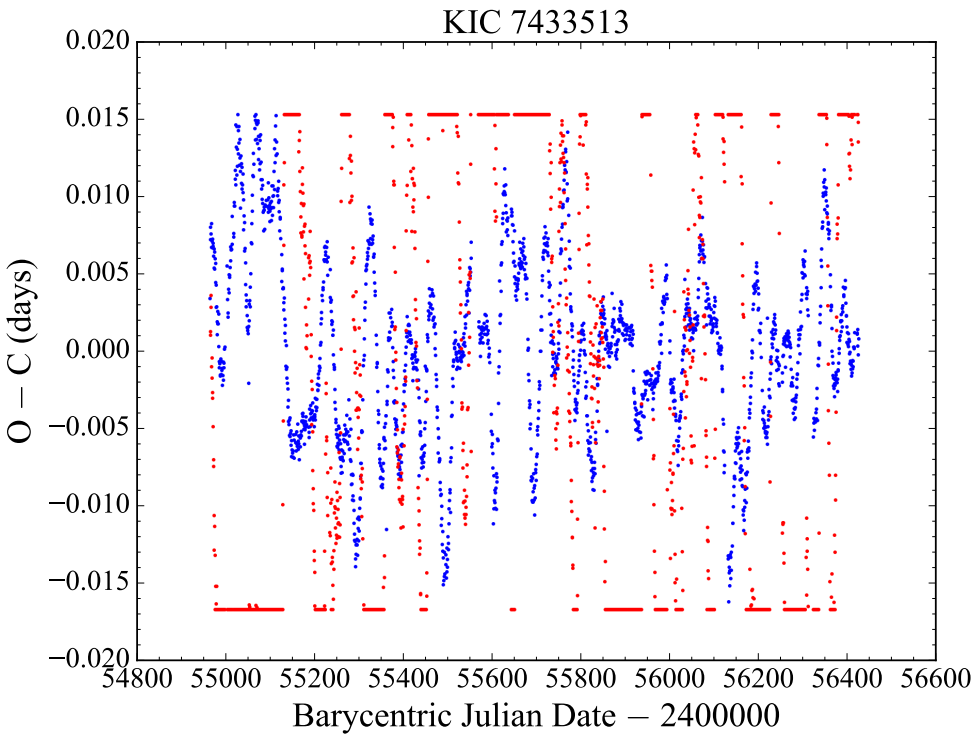


Figure 5.13: Eclipse timing variation of KIC 7433513. The blue points are the eclipse timings for the primary eclipse while the red points are the eclipse timings for the secondary eclipse. The ETV data are capped at +22 min ($0.01527$ d) and −24 min ($0.01\bar{6}$ d) by Conroy et al. (2014). I cannot determine if KIC 7433513 has an anticorrelated ETV or not due to the ETV cap.

of Tran et al. (2013) or Balaji et al. (2015). Nevertheless, I classified 9 of the 27 systems from these papers as Algol-type, with the remaining 18 systems split between 6 $\beta$ Lyrae-type systems and 12 W Ursae Majoris-type systems. These 18 systems account for 62% of the 29 non-Algol temporally varying systems. Therefore, most temporally varying systems fall outside the scope of Tran et al. (2013) and Balaji et al. (2015), and a clear majority of systems falling under those papers' scopes show an anticorrelated ETV.

Since Tran et al. (2013) and Balaji et al. (2015) did not consider several systems in my sample, I manually inspected every temporally varying system's ETV. I

found an additional 20 systems that showed a clear anticorrelation pattern in their ETV, while 8 systems lacked such a pattern in their ETV. I could not determine whether 16 of the remaining 20 systems had an anticorrelated ETV or not, however. Figure 5.13 displays the ETV of one of these 16 systems, the class exemplar KIC 7433513. The secondary eclipse timings vary enough to hit the +22 min and −24 min limits imposed by Conroy et al. (2014) on the ETV data, and therefore, it is impossible to see whether the ETV is anticorrelated during those times. As a result, while the remaining data does not suggest anticorrelation, I cannot rule out anticorrelation when the ETV data is capped. The remaining four systems did not have ETV data from Conroy et al. (2014) because their $\mu \leq 0.5$.

The additional systems I found show that, of the 55 temporally varying systems that I can identify as having an anticorrelated ETV or not, 47 (85%) show an anticorrelated ETV. Such a large percentage of systems strongly suggests that starspots migrating in longitude caused the temporal variation *Kepler* observed. Therefore, I conclude that chromospheric starspot migration primarily causes the temporal variation that distinguishes this class of peculiar systems.

## 5.2 KIC 9164694: Asymmetric Minima Systems

Several systems in my sample display a primary or secondary minimum that is appreciably asymmetric in the same way as NSVS 7322420. I chose KIC 9164694 as the exemplar of this class, as it had the most overtly asymmetric minimum in my initial training sample and was the system I spent the most time observing (Section 6.1.6). My attempts to create an asymmetry parameter (AP; Section 2.2.7) created something less versatile than the SVP, but the AP did flag some asym-

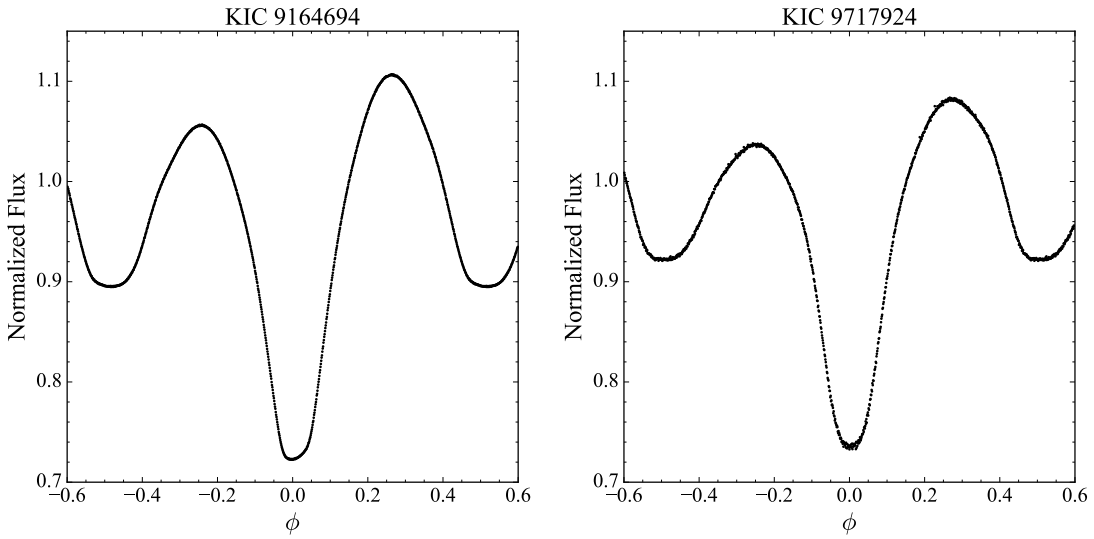


Figure 5.14: Averaged light curves of KIC 9164694 (left) and KIC 9717924 (right) showing an asymmetric primary and secondary minimum, respectively.

metric minima systems I had missed when I manually selected systems for this class. Like the temporally varying systems, I consider the most likely explanation for this asymmetry to be due to spots. However, while migrating chromospheric spots likely explain this asymmetry in some systems, other systems may have their asymmetry explained better by stationary hotspots caused by mass transfer.

### 5.2.1 Class Description

KIC 9164694 and systems like it (labeled with the flag AM in Table A.1) contain an asymmetric minimum in addition to the O'Connell effect, as seen in Figure 5.14's left panel showing KIC 9164694's light curve. KIC 9164694's asymmetry is rather subtle, and is best described as looking like someone has taken their finger and "pushed" the light curve upward and leftward in the region just to the right of minimum light. Asymmetries can appear in the primary minimum (as in KIC 9164694) or the secondary minimum (as in KIC 9717924; see Figure 5.14's right panel).

Asymmetric minima are a rarely discussed feature seen in some eclipsing binary systems, including NSVS 7322420 (Section 1.4). Table 5.4 lists the 35 systems I have found in the literature that contain an asymmetric minimum, along with each system's first reference showing the minimum asymmetry. Older literature, such as Brownlee (1957) and van't Veer (1973), also refers to an asymmetric minimum. However, the asymmetry they describe seems to refer to differing slopes on ingress and egress to the eclipse, as seen in Figure 1.3. This asymmetry occurs closer to the beginning and end of the eclipse. The asymmetry I am describing, by contrast, occurs around the center of the eclipse.

I manually inspected the light curve of every system with AP $\geq 0.008$ to identify potential asymmetric minima systems. However, recall the AP's issues I discussed in Section 2.2.7. As a result of these issues, several systems with AP $\geq 0.008$ turned out to be false positives, which I will discuss shortly. Additionally, I identified several systems as having asymmetric minima before creating the AP, and some of these systems have AP $< 0.008$. That being said, 34 systems (13% of my sample) display an asymmetric minimum in either eclipse. There appear to be two subsets of these systems: those that are temporally varying, such as KIC 7433513, and those that are temporally stable, such as KICs 9164694 and 9717924. My analysis shows several fundamental differences between these two populations, suggesting that the asymmetric minimum has different causes in these two subsets. Table 5.5 lists the systems from and differentiates between the temporally stable and temporally varying systems. It further lists each system's AP, SVP, and which sample they are in. Note that KICs 7885570, 9328852, and 9786165 only show an asymmetric minimum in their short-cadence data, but Table 5.5 lists the values obtained from their long-cadence data.

Table 5.4: List of 35 published asymmetric minima systems.

| Variable Star Designation | Reference |
|---|---|
| $\beta$ Lyrae | Kuiper (1941) |
| VV Orionis | Duerbeck (1975) |
| AK Herculis | Woodward & Wilson (1977) |
| AW Ursae Majoris | Hrivnak (1982) |
| BB Pegasi | Leung et al. (1985) |
| V361 Lyrae | Kałużny (1990) |
| EH Hydrae | Samec et al. (1991) |
| W Crucis | Zoła (1996) |
| V600 Persei | Campos-Cucarella et al. (1997) |
| XX Leonis | Stark et al. (2000) |
| AB Andromedae | Djurašević et al. (2000) |
| GR Tauri | Zhang et al. (2002) |
| CZ Canis Minoris | Yang et al. (2005) |
| RR Centauri | Yang et al. (2005) |
| RY Scuti | Djurašević et al. (2008) |
| UV Monocerotis | Wilsey & Beaky (2009) |
| V963 Persei | Samec et al. (2010) |
| V396 Monocerotis | Liu et al. (2011) |
| AG Virginis | Pribulla et al. (2011) |
| TZ Boötis | Christopoulou et al. (2011) |
| TT Herculis | Terrell & Nelson (2014) |
| BX Pegasi | Li et al. (2015) |
| NSVS 5066754 | Gardner et al. (2015) |
| GSC 02049-01164 | Fox-Machado et al. (2016) |
| ROTSE1 J164341.65+251748.1 | Michel et al. (2016) |
| UCAC4 479-113711 | El-Sadek et al. (2019) |
| NSVS 7322420 | Knote et al. (2019) |
| OGLE-BLG-ECL-157529 | Mennickent et al. (2020) |
| IO Cancri | Liao et al. (2021) |
| CRTS J082700.8+462851 | Li et al. (2021) |
| 1SWASP J132829.37+555246.1 | Li et al. (2021) |
| QZ Carinae | Harmanec et al. (2022) |
| WISE J182934.18-395010.8 | Zasche et al. (2022) |
| TYC 4002-2628-1 | Guo et al. (2022) |
| KN Persei | Gao et al. (2022) |

Table 5.5: List of asymmetric minima systems.

| *Kepler* ID | Asymmetry Parameter | Scaled Variance | Sample |
|---|---|---|---|
| Temporally Stable Systems | | | |
| 2159783 | 0.018 | 0.026 | C |
| 2449084 | 0.016 | 0.006 | C |
| 5283839 | 0.016 | 0.028 | M |
| 6205460 | 0.174 | 0.036 | C |
| 8248967 | 0.055 | 0.012 | C |
| 8696327 | 0.011 | 0.017 | C |
| 8822555 | 0.013 | 0.006 | C |
| 8842170 | 0.006 | 0.022 | C |
| 9164694 | 0.039 | 0.007 | C |
| 9283826 | 0.072 | 0.040 | C |
| 9717924 | 0.018 | 0.010 | C |
| 9786165 | 0.010 | 0.017 | C |
| 10528299 | 0.038 | 0.028 | C |
| 10861842 | 0.019 | 0.013 | C |
| 11395645 | 0.016 | 0.006 | C |
| 11924311 | 0.021 | 0.007 | C |
| Temporally Varying Systems | | | |
| 2569494 | 0.008 | 0.155 | C |
| 2577756 | 0.003 | 0.117 | M |
| 2695740 | 0.002 | 0.079 | C |
| 3124420 | 0.011 | 0.063 | C |
| 3339563 | 0.025 | 0.087 | C |
| 3342425 | 0.013 | 0.054 | C |
| 4474637 | 0.055 | 0.067 | C |
| 6421483 | 0.007 | 0.092 | C |
| 7433513 | 0.019 | 0.080 | C |
| 7885570 | 0.002 | 0.088 | M |
| 8294484 | 0.009 | 0.070 | C |
| 9328852 | 0.001 | 0.053 | C |
| 9419603 | 0.005 | 0.064 | M |
| 9935245 | 0.009 | 0.084 | C |
| 10350225 | 0.012 | 0.108 | C |
| 10676927 | 0.004 | 0.129 | C |
| 11076176 | 0.016 | 0.055 | M |
| 11347875 | 0.044 | 0.076 | C |

Due to the AP's aforementioned issues, I found that 14 systems with AP $\geq$ 0.008 were false positive systems with no noticeably asymmetric minima. All 14 systems are in the core sample, and none are temporally varying. The vast majority of these false positives (12 systems, or 86%) were due to a near-resonance between the system's orbital period and the *Kepler* long-cadence integration time (as discussed in Sections 2.2.2.2 and 2.2.7). The remaining two systems, KICs 9533706 and 9777984, do not have this near-resonance, nor do they have any evidence for an asymmetric minimum. I conceived of two possible explanations: either the AP is sensitive to a genuine, slight asymmetry that is undetectable to the eye, or the form of the Fourier series has introduced a slight offset that inflates the AP. The fact that both systems' AP drops to under 0.008 if I change the number of Fourier series terms $N$ to 30 supports the latter explanation, but I cannot discount the possibility that the reduction of $N$ masks a genuine asymmetry. Table 5.6 lists the 14 false-positive systems along with their AP, SVP, and the ratio of their period to the *Kepler* long-cadence integration time.

### 5.2.2 Class Analysis

In this section, I will summarize the characteristic distributions and trends and look at any correlations between the characteristics. I also discuss the results of my analysis of this class of systems. I emphasize the differences between the temporally stable and temporally varying subsets throughout this section. I do not analyze the AP due to its previously discussed issues.

Table 5.6: List of false positive asymmetric minima systems.

| *Kepler* ID | Asymmetry Parameter | Scaled Variance | Cadence Ratio |
|---|---|---|---|
| 2854752 | 0.020 | 0.015 | 23.022 |
| 5425950 | 0.060 | 0.032 | 18.507 |
| 8386048 | 0.021 | 0.021 | 40.003 |
| 8608490 | 0.010 | 0.035 | 52.991 |
| 8690104 | 0.093 | 0.013 | 20.005 |
| 9345163 | 0.076 | 0.008 | 11.996 |
| 9533706 | 0.009 | 0.017 | 13.904 |
| 9777984 | 0.010 | 0.022 | 12.650 |
| 10128961 | 0.022 | 0.018 | 17.018 |
| 10395202 | 0.011 | 0.020 | 14.972 |
| 10743600 | 0.009 | 0.032 | 39.983 |
| 10802917 | 0.038 | 0.018 | 13.333 |
| 11246163 | 0.036 | 0.024 | 13.665 |
| 11498689 | 0.034 | 0.014 | 14.990 |

#### 5.2.2.1 Sample Characteristics

The asymmetric minima system class as a whole is comprised of 14 (41%) Algol-type, 11 (32%) $\beta$ Lyrae-type, and 9 (26%) W Ursae Majoris-type systems. These proportions are not similar to any of the other samples I have analyzed thus far. Looking at each class subset, I found that the temporally varying subset is comprised of 13 (72%) Algol-type, 3 (17%) $\beta$ Lyrae-type, and 2 (11%) W Ursae Majoris-type systems, which is similar to the marginal sample's proportions (65% Algol-type, 15% $\beta$ Lyrae-type, and 20% W Ursae Majoris-type; Section 4.3.1.1) and, unsurprisingly, the temporally varying systems' proportions (61% Algol-type, 13% $\beta$ Lyrae-type, and 25% W Ursae Majoris-type; Section 5.1.2.1). On the other hand, the temporally stable subset is comprised of 1 (6%) Algol-type, 8 (50%) $\beta$ Lyrae-type, and 7 (44%) W Ursae Majoris-type systems, which is unlike any other

sample's proportions. The temporally stable subset's light curve class proportion, then, is the first piece of evidence that these form a fundamentally different class of systems.

The largest OES found in the temporally stable subset is 0.050 (KIC 9164694), while the smallest is –0.009 (KIC 5283839), both in units of normalized flux. For the temporally varying subset, the largest and smallest OES values are –0.265 (KIC 11347875) and –0.004 (KIC 11076176), respectively. The asymmetric minima systems overall show a similar preference for a positive O'Connell effect (68%) as the other samples I have studied, as do the temporally stable (69%) and varying (67%) subsets individually.

The primary eclipse depth ranges from 0.169 (KIC 8842170) to 0.534 (KIC 620-5460) for the temporally stable subset and from 0.054 (KIC 9419603) to 0.403 (KIC 11347875) for the temporally varying subset. These ranges suggest a difference in the eclipse depth distributions between the two subsets. Figure 5.15 shows a histogram comparing the temporally stable asymmetric minima subset's eclipse depth distribution to the temporally varying subset's and the rest of the complete sample's. Figure 5.15 confirms the distribution difference between the two subsets: the temporally stable subset has deeper eclipses than the temporally varying subset. In particular, 13 of 16 systems (81%) in the temporally stable subset have $D \geq 0.24$, while only 4 of 18 systems (22%) in the temporally varying subset do. This fact is unsurprising in light of the temporally varying systems having shallower eclipses in general, as Figure 5.4 shows. The distribution difference provides further evidence that the two subsets contain fundamentally different objects.

The temporally stable subset's distances range from $646^{+5}_{-4}$ pc (KIC 9283826) to $3{,}826^{+925}_{-800}$ pc (KIC 9717924), while the temporally varying subset's distances

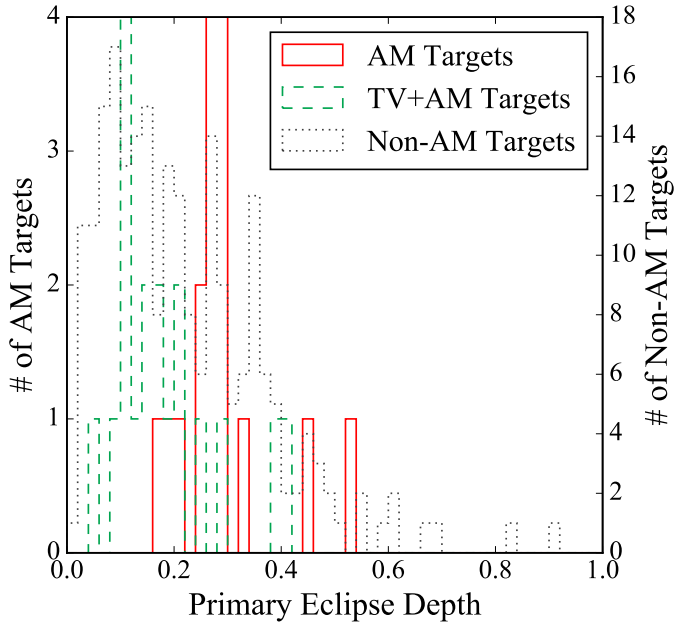


Figure 5.15: Histogram comparing the primary eclipse depths for all 16 temporally stable asymmetric minima targets (solid red), all 18 temporally varying asymmetric minima targets (dashed green), and the remaining 224 complete sample targets (dotted grey). The temporally stable subset have deeper eclipses on average than the temporally varying subset.

range from $511^{+3}_{-2}$ pc (KIC 7885570) to $7{,}743^{+1{,}972}_{-1{,}933}$ pc (KIC 4474637). The temporally stable subset's luminosities range from $1.06 \pm 0.34$ $L_\odot$ (KIC 9786165) to $12.12^{+0.88}_{-0.79}$ $L_\odot$ (KIC 9164694), while the temporally varying subset's luminosities range from $0.28 \pm 0.01$ $L_\odot$ (KIC 9328852) to $8.84^{+0.27}_{-0.26}$ $L_\odot$ (KIC 2695740). Figure 5.16 shows a histogram comparing the temporally stable asymmetric minima subset's luminosity distribution to the temporally varying subset's and the rest of the complete sample's. Both subsets are slightly brighter on average than the rest of the complete sample, but the temporally varying subset is more spread out and has slightly lower luminosity. By contrast, the temporally stable subset is more concentrated among higher luminosities, with no system having a luminosity less than the Sun's.

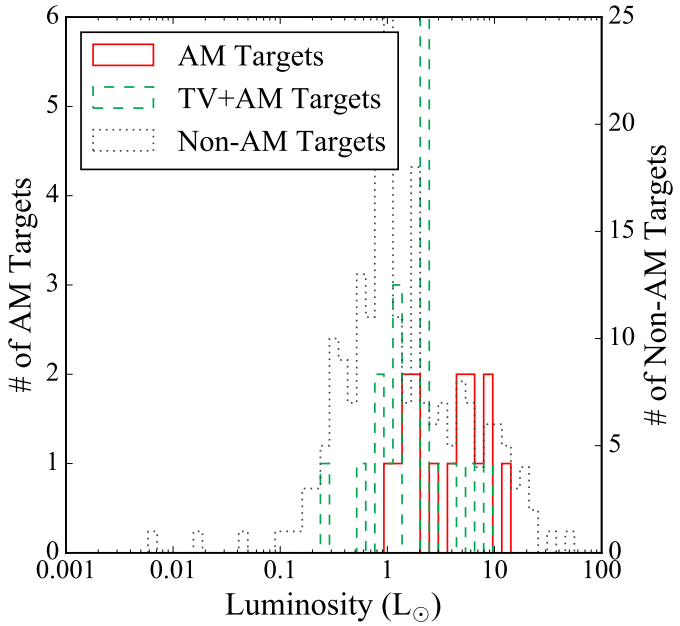


Figure 5.16: Histogram comparing the luminosities for all 16 temporally stable asymmetric minima targets (solid red), all 18 temporally varying asymmetric minima targets (dashed green), and 222 of the remaining 224 complete sample targets (dotted grey) with *Gaia* EDR3 parallaxes. Two non-asymmetric minima systems (KICs 7284688 and 11198068) lack *Gaia* parallax data and so are not included. Note the logarithmic scale of the x-axis. Both subsets are brighter on average than other complete sample systems, with the temporally stable subset being brighter.

The temporally stable subset's periods range from 0.315 d (KIC 5283839) to 3.723 d (KIC 6205460), while the temporally varying subset's periods range from 0.282 d (KIC 10350225) to 3.887 d (KIC 4474637). Figure 5.17 shows a histogram comparing the temporally stable asymmetric minima subset's period distribution to the temporally varying subset's and the rest of the complete sample's. The temporally stable subset's distribution is similar to the rest of the sample's, while the temporally varying subset has a larger proportion of systems with 0.5 d $\leq P \leq$ 1 d. The temporally varying subset's distribution reflects the rest of the temporally varying systems', as shown in Figure 5.5.

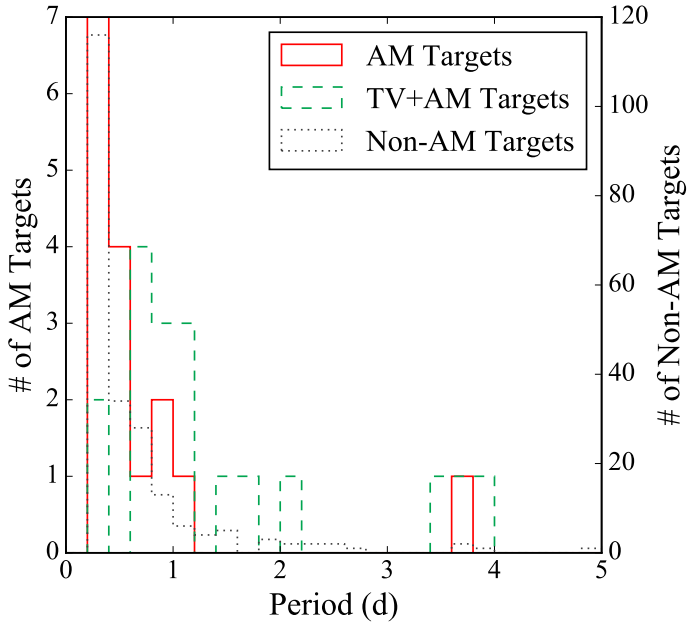


Figure 5.17: Histogram comparing the periods for all 16 temporally stable asymmetric minima targets (solid red), all 18 temporally varying asymmetric minima targets (dashed green), and 179 of the remaining 183 complete sample targets (dotted grey) with $P \leq 5$ d. The temporally stable subset has similar periods to other complete sample systems, while the temporally varying subset has longer periods.

The temporally stable subset's temperatures range from $5{,}063^{+293}_{-76}$ K (KIC 620-5460) to $6{,}955^{+246}_{-256}$ K (KIC 10861842), while the temporally varying subset's temperatures range from $4{,}448^{+516}_{-64}$ K (KIC 9328852) to $5{,}760^{+105}_{-260}$ K (KIC 8294484). Figure 5.18 shows a histogram comparing the temporally stable asymmetric minima subset's temperature distribution to the temporally varying subset's and the rest of the complete sample's. The temporally varying subset's distribution is broadly similar to the rest of the temporally varying systems' (Figure 5.6), while the temporally stable subset's temperatures are significantly hotter on average. Indeed, the temporally stable subset systems are among the hotter systems in my sample, with half of the systems with a *Gaia* DR2 temperature having $T \geq 6{,}000$ K. By contrast, none of the temporally varying subset systems have $T \geq 6{,}000$ K.

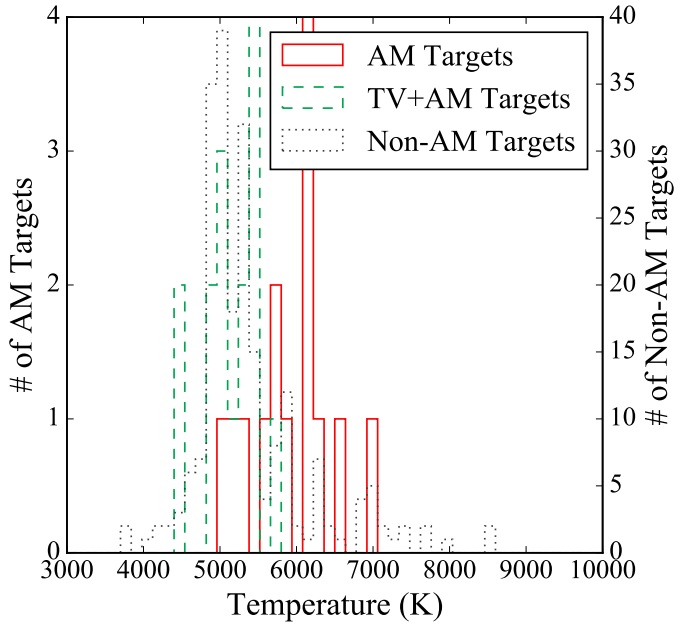


Figure 5.18: Histogram comparing the *Gaia* DR2 temperatures for 14 of the 16 temporally stable asymmetric minima targets (solid red), 15 of the 18 temporally varying asymmetric minima targets (dashed green), and 218 of the remaining 224 complete sample targets (dotted grey) with a *Gaia* DR2 temperature. The temporally stable subset is hotter on average than other complete sample systems, while the temporally varying subset is cooler on average.

Frasca et al. (2016) reported spectral types for four systems in the temporally stable subset ranging from G0 V (KIC 10528299) to F1 V (KIC 10861842). Additionally, Ramsay et al. (2014) reported a spectral type of "mid-G" for KIC 9786165. For the temporally varying subset, Frasca et al. (2016) reported spectral types for two systems: KIC 3339563 (G0 IIIp) and KIC 7885570 (G0 IV). The spectral types estimated using Dr. Mamajek's table range from K2 V (KIC 9786165) to F5 V (KIC 10861842) for the temporally stable subset and from K5.5 V (KIC 11347875) to G3 V (KIC 8294484) for the temporally varying subset.

The temporally stable subset's morphology parameters range from 0.58 (KIC 6205460) to 0.93 (KIC 5283839), while the temporally varying subset's range from

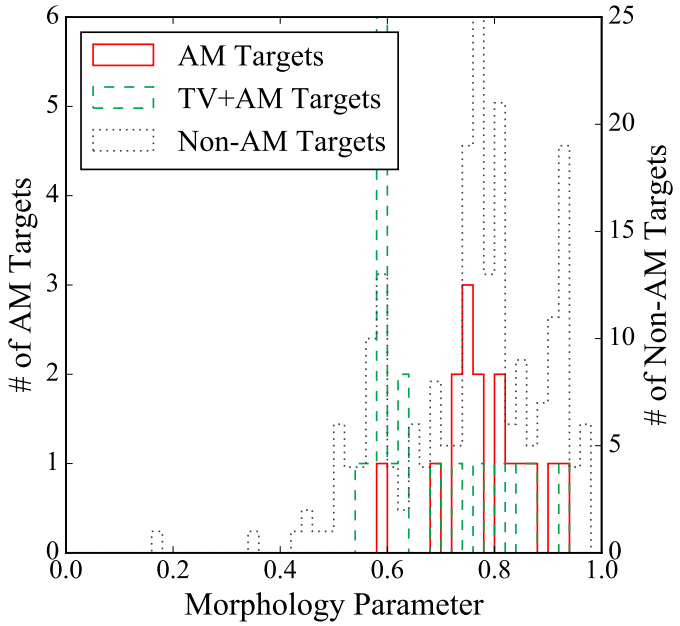


Figure 5.19: Histogram comparing the morphology parameters for all 16 temporally stable asymmetric minima targets (solid red), all 18 temporally varying asymmetric minima targets (dashed green), and 182 of the remaining 183 complete sample targets (dotted grey) that have $\mu \neq -1$. The temporally stable subset peaks at $\mu \approx 0.75$, while the temporally varying subset peaks at $\mu \approx 0.60$.

0.54 (KIC 9328852) to 0.94 (KIC 3342425). Thus, the subsets' ranges of $\mu$ are nearly identical. However, Figure 5.19, showing a histogram of the temporally stable subset's $\mu$ distribution to the temporally varying subset's and the rest of the complete sample's, demonstrates that the subsets' distribution differs radically. The temporally stable subset's distribution peaks around $\mu \approx 0.75$, near the largest peak of the complete sample. The temporally varying subset's distribution, on the other hand, peaks around $\mu \approx 0.60$, around where the temporally varying systems' distribution peaks. Since the temporally varying subset is also a subset of the temporally varying systems, it is not surprising to find that they peak together. However, the fact that the temporally stable systems have a peak at much higher $\mu$ further cements the difference between the two subsets.

#### 5.2.2.2 Characteristic Trends

Figures 5.20-5.22 show the asymmetric minima systems' corner plot comparing period, OES, *Gaia* temperature, the *Gaia* BP – RP color index, morphology parameter, absolute *Gaia* G magnitude, and primary eclipse depth. These corner plots distinguishes between the temporally stable (red points) and the temporally varying (green points) subsets. They also plot temporally varying systems without asymmetric minima (blue points) for comparison. These corner plots highlight two illuminating trends. The first trend is seen in Figure 5.21's lower-left panel showing that 15 of the 16 systems in the temporally stable subset (excepting KIC 6205460) lie along the line of minimum period Section 4.2.3.4 discusses. By contrast, only 2 of the 18 systems in the temporally varying subset (KICs 3342425 and 10350225) lie along this line. Figure 5.21's center panel shows the other trend: systems in the temporally stable subset lie – without exception – along the sharp edge Section 4.2.3.2 discusses. By contrast, systems in the temporally varying subset lie nearly exclusively away from this sharp edge. In a more general sense, Figures 5.20-5.22 show that the temporally stable subset and the temporally varying subset occupy distinct regions of parameter space having limited overlap. This fact is particularly noticeable in Figure 5.20's lower-left panel, Figure 5.21's center, center-left, and lower-center panels, and Figure 5.22's upper-left, center-left, and lower-left panels. This separation in parameter space, combined with the two trends discussed earlier, further implies that the two subsets contain fundamentally different types of systems.

One of the temporally stable subset's systems, KIC 6205460, is notable for being an outlier in many panels in Figures 5.20-5.22. With $P = 3.723$ d, $D = 0.534$, and $\mu = 0.58$, it appears as an outlier in the panels involving period, eclipse

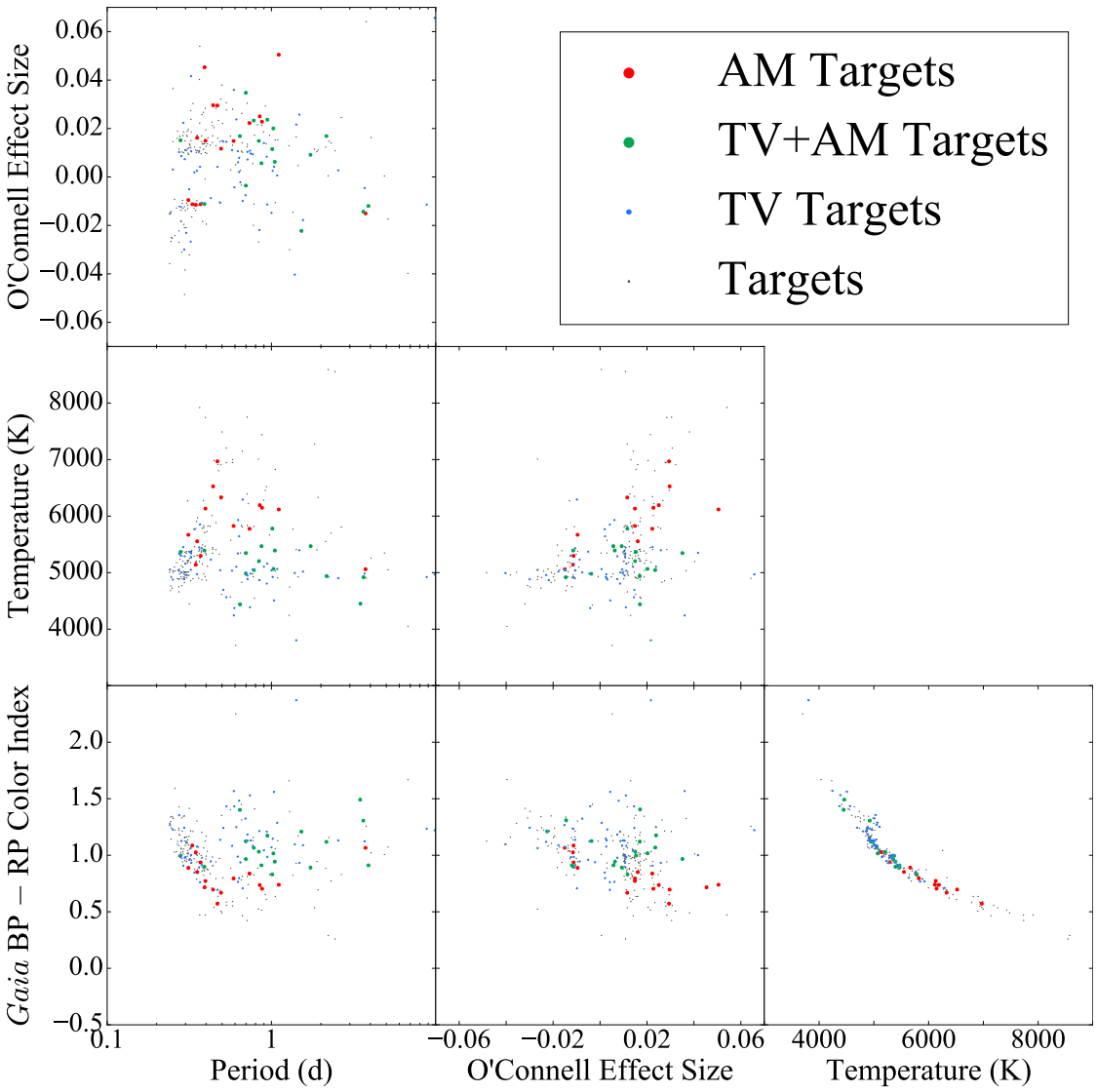


Figure 5.20: Corner plot showing the asymmetric minima system correlations between three characteristics of interest (period, OES, and temperature) and OES, temperature, and color. Systems in the temporally stable subset are shown in red, systems in the temporally varying subset are shown in green, other temporally varying systems are shown in blue, and the other complete sample systems are shown in grey. KIC 11347875 has been removed from the OES plots for clarity (see discussion in Section 4.2.1). Note the logarithmic x-axis of the period plots.

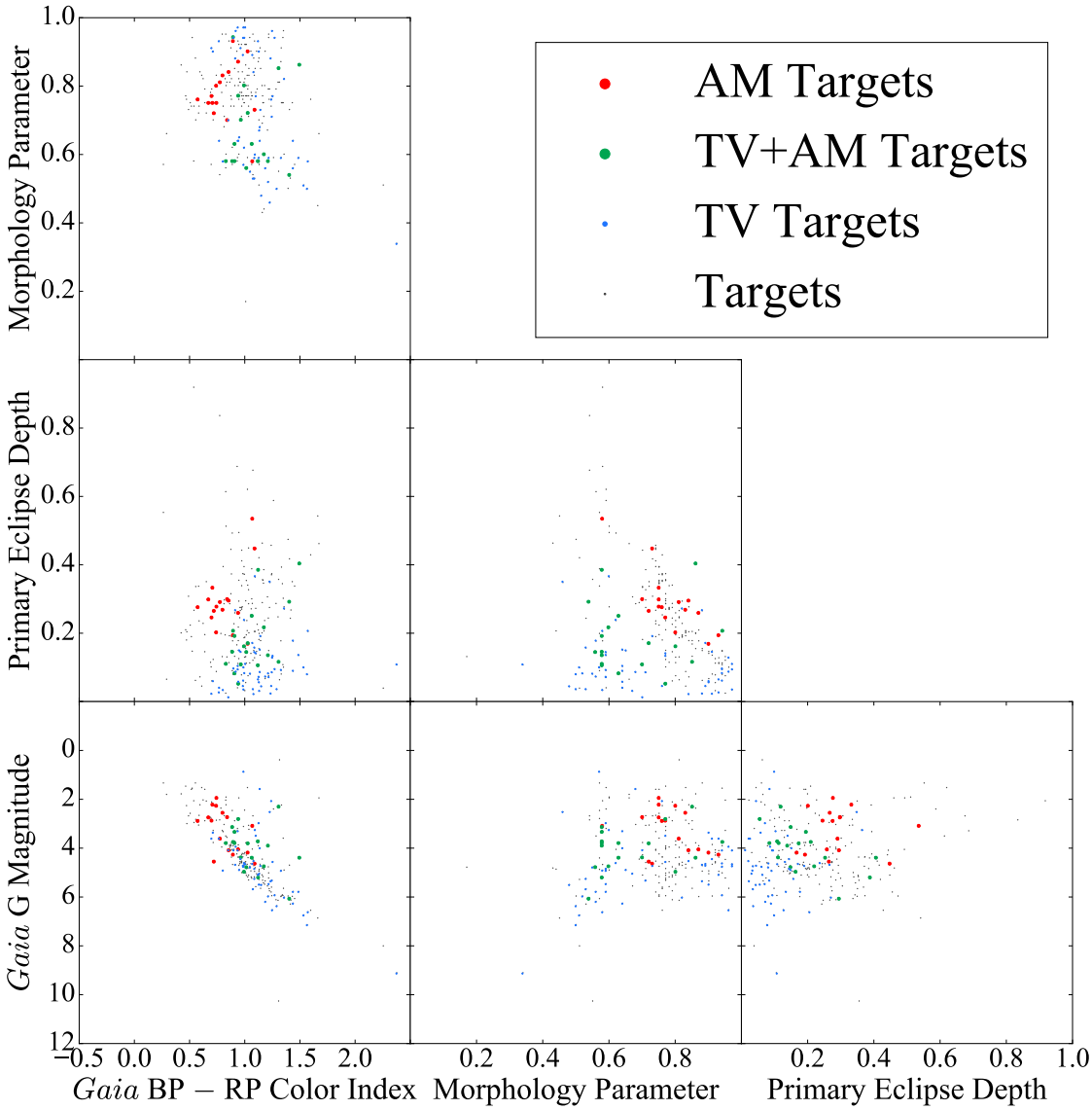


Figure 5.21: Corner plot showing the asymmetric minima system correlations between three characteristics of interest (color, morphology parameter, and primary eclipse depth) and morphology parameter, primary eclipse depth, and absolute magnitude. Systems in the temporally stable subset are shown in red, systems in the temporally varying subset are shown in green, other temporally varying systems are shown in blue, and the other complete sample systems are shown in grey.

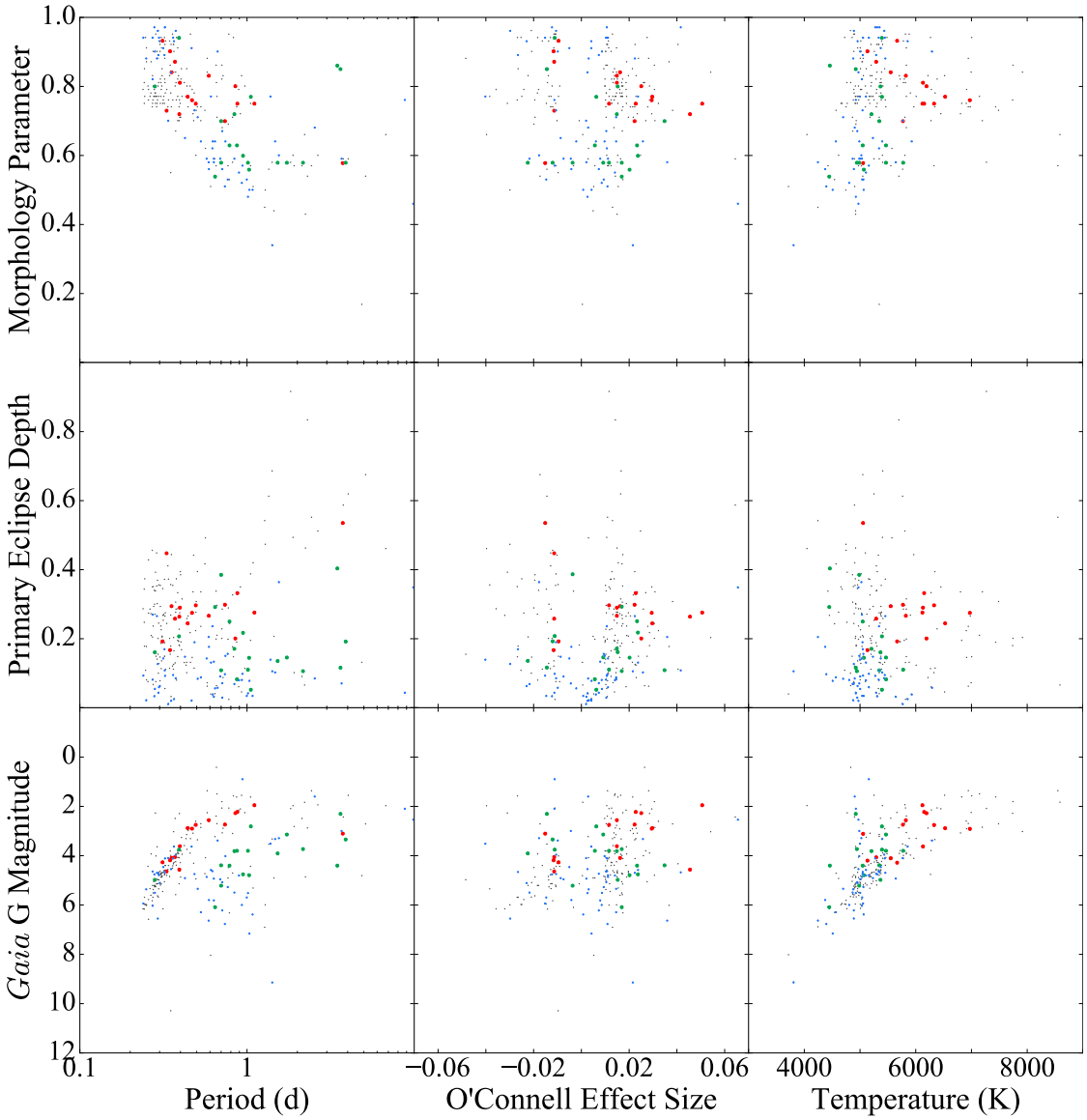


Figure 5.22: Corner plot showing the asymmetric minima system correlations between three characteristics of interest (period, OES, and temperature) and morphology parameter, primary eclipse depth, and absolute magnitude. Systems in the temporally stable subset are shown in red, systems in the temporally varying subset are shown in green, other temporally varying systems are shown in blue, and the other complete sample systems are shown in grey. KIC 11347875 has been removed from the OES plots for clarity (see discussion in Section 4.2.1). Note the logarithmic x-axis of the period plots.

Table 5.7: Results of analyzing characteristic distributions between the asymmetric minima systems and the rest of the complete sample using the Kolomgorov–Smirnov test.

| Characteristic | K–S Statistic | $p$-Value |
|---|---|---|
| O'Connell Effect Size | 0.217 | 0.107 |
| \|O'Connell Effect Size\| | 0.167 | 0.353 |
| Primary Eclipse Depth | 0.209 | 0.132 |
| Morphology Parameter | 0.155 | 0.443 |
| Temperature | 0.274 | 0.019 |
| Distance | 0.348 | 0.001 |
| Absolute Magnitude | 0.346 | 0.001 |
| Period | 0.325 | 0.003 |

depth, and morphology parameter particularly often. KIC 6205460 often appears in the same areas of parameter space that systems from the temporally varying subset do, and its SVP (0.036) is the second largest of the temporally stable subset (only KIC 9283826's SVP of 0.040 is larger). Therefore, KIC 6205460 may be fundamentally more similar to the temporally varying subset's systems than the rest of the temporally stable subset's. Regardless, I consider it a member of the temporally stable subset for the purposes of my analysis.

#### 5.2.2.3 Sample Distributions

Table 5.7 gives the results of the K–S test for the asymmetric minima systems. The K–S test does not reject the null hypothesis for *any* characteristic, although it nearly does so for distance, absolute magnitude, and period. The difference in distance distribution is very surprising to me, as it is difficult to interpret the presence of an asymmetric minimum in a way that would depend on the distance to the system. Instead, I interpret this as a consequence of the fact that asymmetric minima systems are more luminous on average than the rest of the sample (as

Table 5.8: Results of analyzing characteristic distributions between the temporally stable asymmetric minima systems and the temporally varying asymmetric minima systems using the Kolomgorov–Smirnov test.

| Characteristic | K–S Statistic | $p$-Value |
|---|---|---|
| O'Connell Effect Size | 0.271 | 0.495 |
| \|O'Connell Effect Size\| | 0.222 | 0.741 |
| Primary Eclipse Depth | 0.604 | 0.002 |
| Morphology Parameter | 0.549 | 0.007 |
| Temperature | 0.590 | 0.003 |
| Distance | 0.222 | 0.741 |
| Absolute Magnitude | 0.451 | 0.043 |
| Period | 0.576 | 0.004 |

Figure 5.16 shows, supported by the absolute magnitude K–S statistic shown in Table 5.7). As a result, asymmetric minima systems are more easily observed from a greater distance than the other, less luminous members of my sample, which accounts for the distance distribution difference.

Because the evidence strongly suggests that the temporally stable subset and the temporally varying subset represent two fundamentally different populations, I thought it would be illuminating to compute the K–S statistics between the two subsets. Table 5.8 gives the results of this K–S test. While the K–S test does not reject the null hypothesis for any of the characteristics, it does indicate that the two subsets have significantly different distributions of eclipse depth, morphology parameter, temperature, absolute magnitude, and period. These results are supported by Figures 5.15, 5.19, 5.18, 5.16, and 5.17, respectively.

#### 5.2.2.4 Characteristic Correlations

By this point, the idea that the asymmetric minima systems are split into two subsets of fundamentally different binaries has substantial supporting evidence.

Table 5.9: Results of analyzing the correlations between the asymmetric minima system's temporally stable and temporally varying subsets' characteristics of interest using Spearman's rank correlation coefficient.

| Characteristic One | Characteristic Two | Stable Subset Coeff. | Stable Subset $p$-Value | Varying Subset Coeff. | Varying Subset $p$-Value |
|---|---|---|---|---|---|
| OES | Eclipse Depth | **−0.144** | **0.594** | −0.067 | 0.791 |
| OES | Morphology | **−0.221** | **0.410** | **−0.290** | **0.243** |
| OES | Temperature | **0.736** | **0.003** | −0.036 | 0.899 |
| OES | Distance | **0.241** | **0.368** | −0.026 | 0.919 |
| OES | Absolute Mag. | **−0.459** | **0.074** | **0.422** | **0.081** |
| OES | Period | **0.338** | **0.200** | **−0.449** | **0.062** |
| \|OES\| | Eclipse Depth | 0.009 | 0.974 | **0.267** | **0.284** |
| \|OES\| | Morphology | **−0.407** | **0.118** | 0.020 | 0.937 |
| \|OES\| | Temperature | **0.631** | **0.016** | **−0.525** | **0.044** |
| \|OES\| | Distance | **0.150** | **0.579** | 0.063 | 0.804 |
| \|OES\| | Absolute Mag. | **−0.497** | **0.050** | **0.420** | **0.083** |
| \|OES\| | Period | **0.521** | **0.039** | 0.053 | 0.836 |
| Eclipse Depth | Morphology | **−0.714** | **0.002** | −0.012 | 0.964 |
| Eclipse Depth | Temperature | −0.011 | 0.970 | **−0.490** | **0.064** |
| Eclipse Depth | Distance | **0.171** | **0.528** | **−0.164** | **0.515** |
| Eclipse Depth | Absolute Mag. | **−0.150** | **0.579** | **0.577** | **0.012** |
| Eclipse Depth | Period | **0.403** | **0.122** | **−0.259** | **0.299** |
| Morphology | Temperature | **−0.285** | **0.324** | 0.069 | 0.808 |
| Morphology | Distance | **−0.416** | **0.109** | −0.011 | 0.967 |
| Morphology | Absolute Mag. | **0.206** | **0.443** | **−0.278** | **0.265** |
| Morphology | Period | **−0.560** | **0.024** | **−0.106** | **0.675** |
| Temperature | Distance | **0.349** | **0.221** | **0.331** | **0.229** |
| Temperature | Absolute Mag. | **−0.516** | **0.059** | **−0.372** | **0.172** |
| Temperature | Period | **0.196** | **0.503** | **−0.161** | **0.567** |
| Distance | Absolute Mag. | **−0.138** | **0.610** | −0.098 | 0.699 |
| Distance | Period | **0.115** | **0.672** | **0.146** | **0.565** |
| Absolute Mag. | Period | **−0.853** | **$<$0.001** | **−0.622** | **0.006** |

Therefore, I chose to look at each subset's correlations separately rather than the correlations of the system class as a whole. Table 5.9 gives the results of my analysis of the temporally stable and temporally varying subsets using Spearman's $\rho$ coefficient, while Table 5.10 gives the results of my analysis using Kendall's $\tau$ coefficient. Tables 5.9 and 5.10 highlight several differences between the subsets.

Table 5.10: Results of analyzing the correlations between the asymmetric minima system's temporally stable and temporally varying subsets' characteristics of interest using Kendall's rank correlation coefficient.

| Characteristic One | Characteristic Two | Stable Subset Coeff. | Stable Subset $p$-Value | Varying Subset Coeff. | Varying Subset $p$-Value |
|---|---|---|---|---|---|
| OES | Eclipse Depth | **−0.067** | **0.757** | −0.046 | 0.823 |
| OES | Morphology | **−0.194** | **0.299** | **−0.200** | **0.262** |
| OES | Temperature | **0.604** | **0.002** | −0.019 | 0.921 |
| OES | Distance | **0.167** | **0.398** | −0.046 | 0.823 |
| OES | Absolute Mag. | **−0.333** | **0.079** | **0.268** | **0.131** |
| OES | Period | **0.250** | **0.195** | **−0.294** | **0.096** |
| \|OES\| | Eclipse Depth | 0.017 | 0.965 | **0.176** | **0.330** |
| \|OES\| | Morphology | **−0.312** | **0.095** | 0.035 | 0.847 |
| \|OES\| | Temperature | **0.451** | **0.026** | **−0.364** | **0.060** |
| \|OES\| | Distance | **0.050** | **0.825** | −0.007 | 1.000 |
| \|OES\| | Absolute Mag. | **−0.383** | **0.041** | **0.281** | **0.112** |
| \|OES\| | Period | **0.367** | **0.052** | 0.033 | 0.881 |
| Eclipse Depth | Morphology | **−0.565** | **0.002** | −0.021 | 0.908 |
| Eclipse Depth | Temperature | 0.033 | 0.914 | **−0.364** | **0.060** |
| Eclipse Depth | Distance | **0.167** | **0.398** | **−0.124** | **0.501** |
| Eclipse Depth | Absolute Mag. | **−0.100** | **0.626** | **0.399** | **0.021** |
| Eclipse Depth | Period | **0.317** | **0.096** | **−0.163** | **0.369** |
| Morphology | Temperature | **−0.246** | **0.226** | 0.050 | 0.802 |
| Morphology | Distance | **−0.329** | **0.078** | −0.048 | 0.786 |
| Morphology | Absolute Mag. | **0.228** | **0.222** | **−0.214** | **0.230** |
| Morphology | Period | **−0.481** | **0.010** | **−0.104** | **0.562** |
| Temperature | Distance | **0.231** | **0.279** | **0.249** | **0.198** |
| Temperature | Absolute Mag. | **−0.275** | **0.193** | **−0.268** | **0.165** |
| Temperature | Period | **0.099** | **0.667** | **−0.096** | **0.620** |
| Distance | Absolute Mag. | **−0.133** | **0.506** | −0.046 | 0.823 |
| Distance | Period | **0.083** | **0.690** | **0.176** | **0.330** |
| Absolute Mag. | Period | **−0.750** | **$<$0.001** | **−0.451** | **0.009** |

These differences will be my primary focus in this section.

The correlation between OES/|OES| and absolute magnitude are anticorrelated in the temporally stable subset but correlated in the temporally varying subset. The same holds true for the correlation between eclipse depth and absolute magnitude. The opposite (correlated in the temporally stable subset but anticorrelated

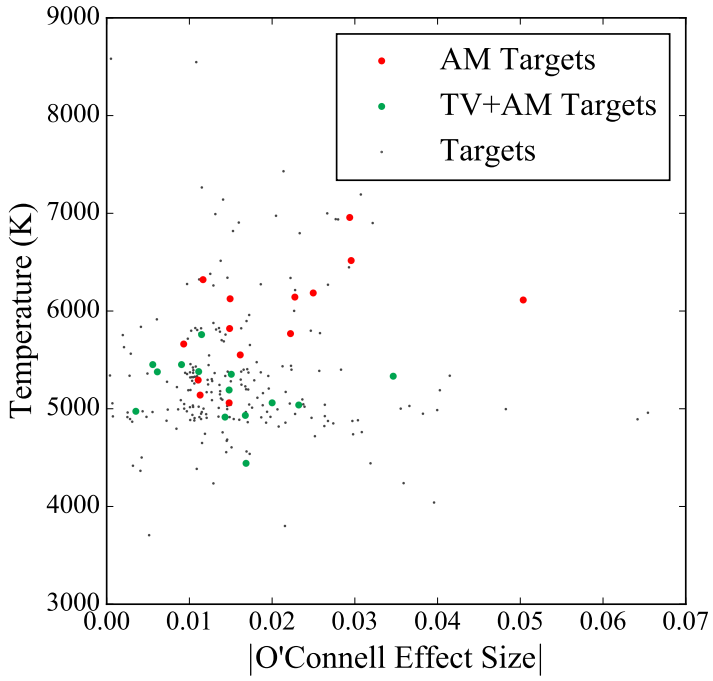


Figure 5.23: Plot comparing the temperature to |OES|. Temporally stable asymmetric minima systems are shown in red, temporally varying asymmetric minima systems are shown in green, while the other complete sample systems are shown in grey. KIC 11347875 is excluded for clarity. Both subsets show a weak correlation.

in the temporally varying subset) holds true for the characteristic pairs |OES| and temperature, eclipse depth and distance, eclipse depth and period, morphology and absolute magnitude, and temperature and period.

Figure 5.23 plots the strongest correlation of these characteristic pairs, namely the one between |OES| and temperature. Figure 5.23 shows a clearly positive correlation among the temporally stable systems and, less clearly, a negative correlation among the temporally varying systems. These correlations may reflect a differing O'Connell effect cause between the subsets. In this case, the temporally stable subset's O'Connell effect would be caused by something that strengthens

with increasing temperature, while the temporally varying subset's O'Connell effect would be caused by something that weakens with increasing temperature. As I concluded in Section 5.1.2.1, stellar activity drops as temperature increases, which explains the temporally varying subset's negative correlation. The temporally stable subset's correlation still requires an explanation.

The temporally stable subset shows a very strong correlation between the primary eclipse depth and morphology that is entirely absent in the temporally varying subset. Figure 5.21's center panel for these shows this correlation as the sharp edge at the panel's right side, just as in the core sample. As I mentioned in Section 5.2.2.2, all temporally stable systems lie along this edge, creating the detected correlation. As I discussed in Section 4.2.3.2, this trend indicates that systems in the temporally stable subset have the deepest eclipses of systems with similar light curves.

The correlations between OES and temperature and between |OES| and period are also very strong in the temporally stable subset but absent in the temporally varying subset. Figure 5.20's center panel shows that temporally stable systems with a negative O'Connell effect are cooler than those with a positive O'Connell effect. The negative O'Connell effect systems differ in another aspect in this subset: they have a larger SVP. Specifically, four of the five negative O'Connell effect systems (KICs 2159783, 5283839, 6205460, and 8842170, but *not* KIC 9786165) have larger SVPs than nine of the eleven positive O'Connell effect systems (the exceptions being KIC 9283826 and KIC 10528299). Therefore, temporally stable asymmetric minima systems with a positive O'Connell effect are hotter and temporally more stable than their negative O'Connell effect counterparts. These systems may therefore represent a further subdivision of the class. However, the

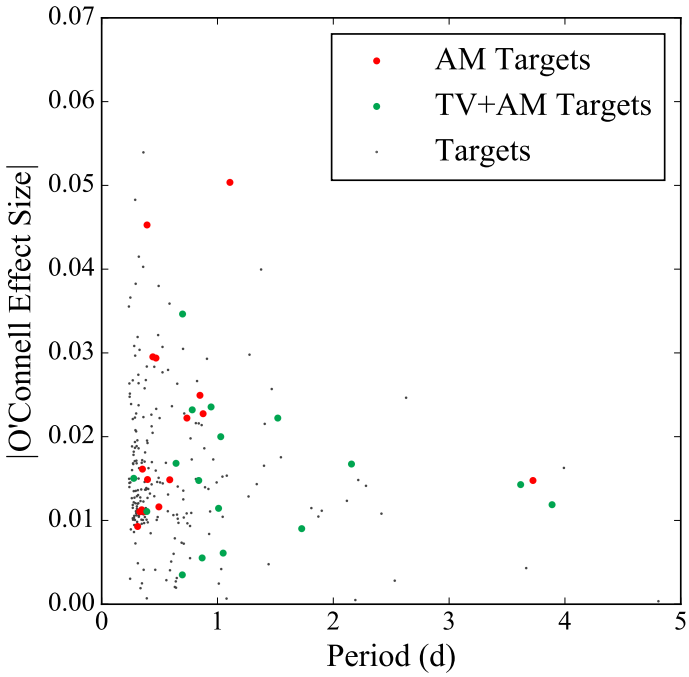


Figure 5.24: Plot comparing |OES| to the period. Temporally stable asymmetric minima systems are shown in red, temporally varying asymmetric minima systems are shown in green, while the other complete sample systems are shown in grey. KIC 11347875 and complete samples systems with $P > 5$ d are excluded for clarity. The temporally stable subset shows a roughly linear correlation.

other characteristics of the temporally stable subset's systems are generally more similar to each other than to the temporally varying subset's.

Figure 5.24 plots |OES| against period, showing that the temporally varying systems are uncorrelated, as Table 5.9 implies. On the other hand, Figure 5.24 also shows that the temporally stable systems have a roughly linear relation between the two characteristics, aside from one outlier system (KIC 6205460 again). If mass transfer through a matter stream drives the O'Connell effect in the temporally stable subset's systems, such a correlation makes physical sense. In such systems, the longer orbital period would imply greater separation between the components,

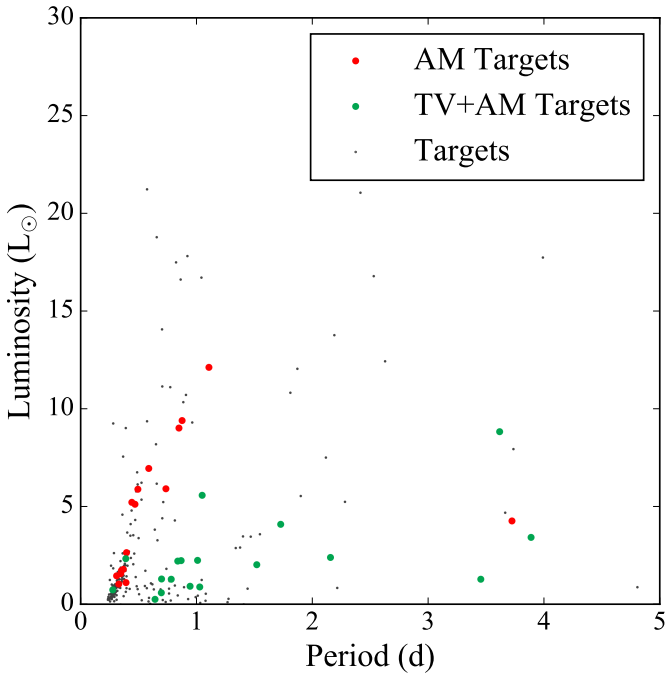


Figure 5.25: Plot comparing the luminosity to the period. Temporally stable asymmetric minima systems are shown in red, temporally varying asymmetric minima systems are shown in green, while the other complete sample systems are shown in grey. Complete samples systems with $P > 5$ d are excluded for clarity. The temporally stable subset shows a very strong linear correlation, while the temporally varying subset shows a weaker correlation.

which would allow the matter stream to acquire more energy before hitting some surface. The extra energy released would increase the disparity between maxima brightness, increasing |OES|.

The final correlation I will discuss is between the absolute magnitude/luminosity and period. The characteristics are correlated in both subsets, but the temporally stable subset's correlation is exceptionally strong. Figure 5.25, which plots the period versus the luminosity, shows why: with a single exception (KIC 6205460 yet again), the temporally stable systems have a nearly linear relationship between the two characteristics. This line is the same as the edge discussed in Section 4.2.3.4

representing the lower limit on period for a given luminosity, meaning the components are in or nearly in contact with each other. The fact that 15 of the 16 temporally stable systems lie along this lower limit strongly implies that the temporally stable subset represents a class nearly exclusively comprised of near-contact or overcontact binaries. The temporally varying systems, by contrast, lie largely away from this line, implying that the temporally varying subset consists largely of detached binaries. The morphology parameter distributions shown in Figure 5.19 support both conclusions.

### 5.2.3 Potential Explanations

The results given throughout Section 5.2.2 paint a clear picture of a class of systems subdivided into two fundamentally different types of binaries. The results also give a clear idea of what systems in each subset look like. Temporally stable asymmetric minima systems are relatively hot, luminous systems with deep eclipses that are near-contact or overcontact binaries, and the systems with the most stable light curves overwhelmingly favor a positive O'Connell effect. Temporally varying asymmetric minima systems are cool, solar-type luminosity detached binaries with relatively shallow eclipses that show a similar preference for a positive O'Connell effect as other O'Connell effect binaries. I explained the temporally varying system class as a result of chromospheric starspot migration in Section 5.1.3, and the temporally varying subset of asymmetric minima systems seems to align well with that explanation. With a viable explanation for the temporally varying subset, I will focus my attention in this section on the temporally stable subset of the asymmetric minima systems.

The most striking commonality between asymmetric minima systems is that

all known examples, including those from both subsets and the literature, exhibit total eclipses. Total eclipses occur when one star is completely occluded during one of the eclipses and are distinguished by flat-bottomed or nearly flat-bottomed minima. The fact that asymmetric minima only occur in totally eclipsing systems offers an important clue as to the phenomenon's cause insofar as the phenomenon is probably related to the presence of total eclipses. However, asymmetric minima cannot be described simply as a consequence of total eclipses in O'Connell effect binaries, as asymmetric minima occur in totally eclipsing systems like KIC 8265951 (Figure 2.12) that lack a significant O'Connell effect. Likewise, some systems with total eclipses and a significant O'Connell effect do not display an asymmetric minimum, such as KIC 8386048.

Based on the connection between asymmetric minima and total eclipses, I focused my hypotheses on the differences between partial and total eclipses. One such difference is that areas near the visible pole of each star are persistently visible over the entire orbit for partially eclipsing systems. By contrast, one of the stars is completely occluded in totally eclipsing systems, making it impossible for any region on that star to be persistently visible. Therefore, my first hypothesis was that a polar feature on one star caused the asymmetric minimum, as it would be occluded in totally eclipsing systems but not in partially eclipsing ones. I tested my hypothesis with BinaryMaker3 (BM3; Bradstreet & Steelman 2002), a program that allows me to create a binary system and see the light curve it would produce. My testing confirmed that a polar spot could create an asymmetric minimum. However, it also showed that equatorial spots could also produce an asymmetric minimum. Since equatorial spots are occluded in both partially and totally eclipsing systems, this hypothesis could not explain why we see asymmetric

Table 5.11: System parameters of the asymmetric minimum toy model used to create Figure 5.26.

| Parameter | Value(s) |
|---|---|
| Orbital Inclination (°) | 74, 75, 76, 77 |
| Mass Ratio | 0.2 |
| $T_{eff}$ of Primary Component (K) | 7,800 |
| $T_{eff}$ of Secondary Component (K) | 4,700 |
| Surface Potential of Primary Component | 2.3511 |
| Surface Potential of Secondary Component | 2.3511 |
| Spot Temperature (K) | 9,400 |
| Spot Latitude (°) | 0 |
| Spot Longitude (°) | 165 |
| Spot Radius (°) | 30 |

minima only in the latter type of binary.

Further testing showed that asymmetric minima are strongly dependent on the presence of total eclipses. I created a toy model of an eclipsing binary with BM3 using the parameters listed in Table 5.11. While these parameters are unrealistic regarding spot size and temperature, it is sufficient to illustrate my point, and I used data from this model to produce Figure 5.26. Figure 5.26 shows that a one degree difference in inclination is sufficient to transform a largely symmetrical eclipse into a significantly asymmetric one. This testing indicates that total eclipses are a requirement for asymmetric minima rather than an observational bias, casting further doubt on the polar hypothesis discussed in the previous paragraph. I required another hypothesis to explain this system class.

Another difference between partially and totally eclipsing systems is in the nature of the total eclipse itself: as the change in the visible area of each star is zero during the total phase, the flux change is also zero, modulo effects such as limb darkening and gravity brightening. This fact allows for the possibility that the asymmetry's cause is a relatively small change in flux that the eclipse's partial

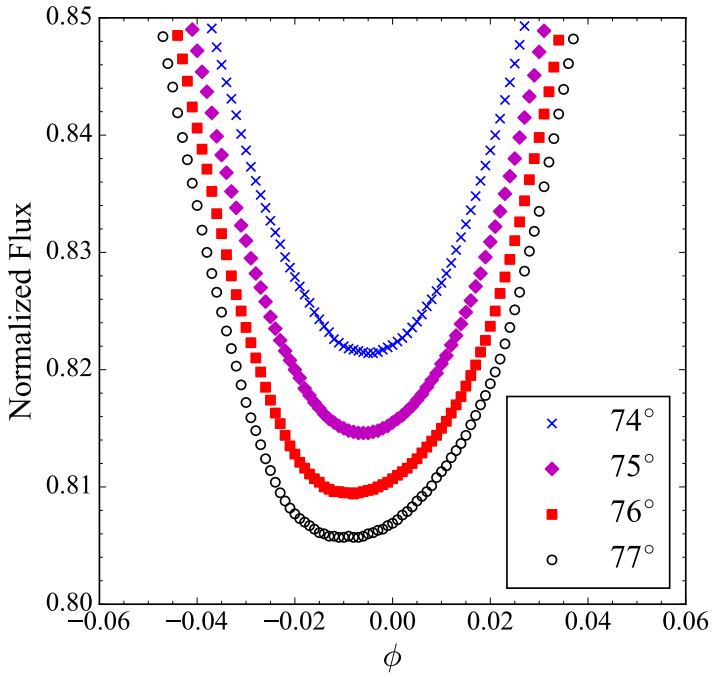


Figure 5.26: Light curves of four eclipsing binary models detailing the area around the primary minimum. Parameters for all four models are identical except for inclination. The asymmetry grows significantly more pronounced as the system changes from partially eclipsing ($i = 74,\ 75°$) to totally eclipsing ($i = 76,\ 77°$), demonstrating the asymmetry's strong dependence on total eclipses.

phase would drown out. A starspot's changing aspect (i.e. viewing angle) would create such a small flux change in a similar manner to limb darkening. With limb darkening, the change in flux from the center of a stellar disk to its limb is due to seeing to higher optical depth (and therefore higher surface brightness) at the center than at the limb (Gray 2008). Similarly, we see to a higher optical depth as a spot rotates from the star's limb to its center. Considering the $T^4$ temperature dependence of surface brightness, the differing temperature of the spot magnifies this effect compared to limb darkening. If the spot is azimuthally offset from the plane both 1.) perpendicular to the orbital plane and 2.) containing both

stellar centers, we will observe its highest optical depth before or after the binary's conjunction phase. The flux change this causes is still considerably less than the flux change produced by the eclipsed star being covered or uncovered during an eclipse's partial phase, preventing the spot from creating a significant asymmetry during the partial phase. During the essentially flux-constant total eclipse phase, however, the small flux change caused by the spot's changing aspect has no other flux change to compete with, allowing for a noticeable asymmetry.

The origin of the spot causing the asymmetry is clear for the temporally varying subset but not for the temporally stable subset. Due to the implication from Section 5.2.2 that such systems have components near or in contact with each other, it is reasonable to suspect that these systems are undergoing mass transfer. The spot, then, can be interpreted as an effect of such mass transfer, perhaps as the impact site of a matter stream on the gainer's surface. Since the impact site should not change position significantly on short timescales, this idea explains the stability of these systems' light curves compared to the temporally varying systems. I would also expect the matter stream to preferentially impact on the trailing hemisphere of the gainer's surface due to conservation of angular momentum. If the gainer is cooler than the donor, the matter stream impact on the gainer will preferentially produce a positive O'Connell effect. However, the gainer can only be cooler in the relatively rare "reverse Algol" class of binary, wherein the more massive star fills its Roche lobe (Leung 1989). To test my mass transfer idea, I looked at the temporally stable subset's ETVs. Exactly half of the 16 systems had a possibly parabolic ETV. While this percentage is smaller than what I would expect if my idea is correct, it is larger than what I found in my sample as a whole.

Therefore, I conclude that the changing aspect of a spot or spots during an

eclipse's total phase causes the asymmetric minima that distinguish this class of peculiar systems. In the temporally varying subset, the spots are likely migrating chromospheric starspots on stars that are not in contact with each other. In the temporally stable subset, this spot is more likely caused by mass transfer between stars that are near or in contact with each other.

## 5.3 KIC 10544976: The White Dwarf

KIC 10544976 (previously discussed in Section 3.8 and labeled with the flag WD in Table A.1) is unique within my sample as it is the only binary to contain a degenerate component. My sample's relative lack of white dwarfs is not surprising considering that *Kepler*'s primary mission was to detect habitable exoplanets, which are unlikely to exist in white dwarf systems. KIC 10544976 – being the only member of this class – serves as its exemplar. Unfortunately, I cannot apply the same statistical tests to a single system that I did with the other system classes, limiting the potential discussion of the white dwarf class.

Figure 5.27's left panel shows KIC 10544976's long-cadence averaged light curve, its right panel shows KIC 10544976's short-cadence averaged light curve, and its bottom panel shows KIC 10544976's light curve using all of *Kepler*'s short-cadence data of the system. The system has an exceptionally sharply defined, flat-bottomed primary eclipse and no noticeable secondary eclipse. The primary eclipse's sharp nature makes the long-cadence data unsuitable for analyzing KIC 10544976 because the time between the beginning of the eclipse to the beginning of totality is only ~8% of a single long-cadence exposure. The system also shows numerous flares that increase the system's luminosity by up to a factor

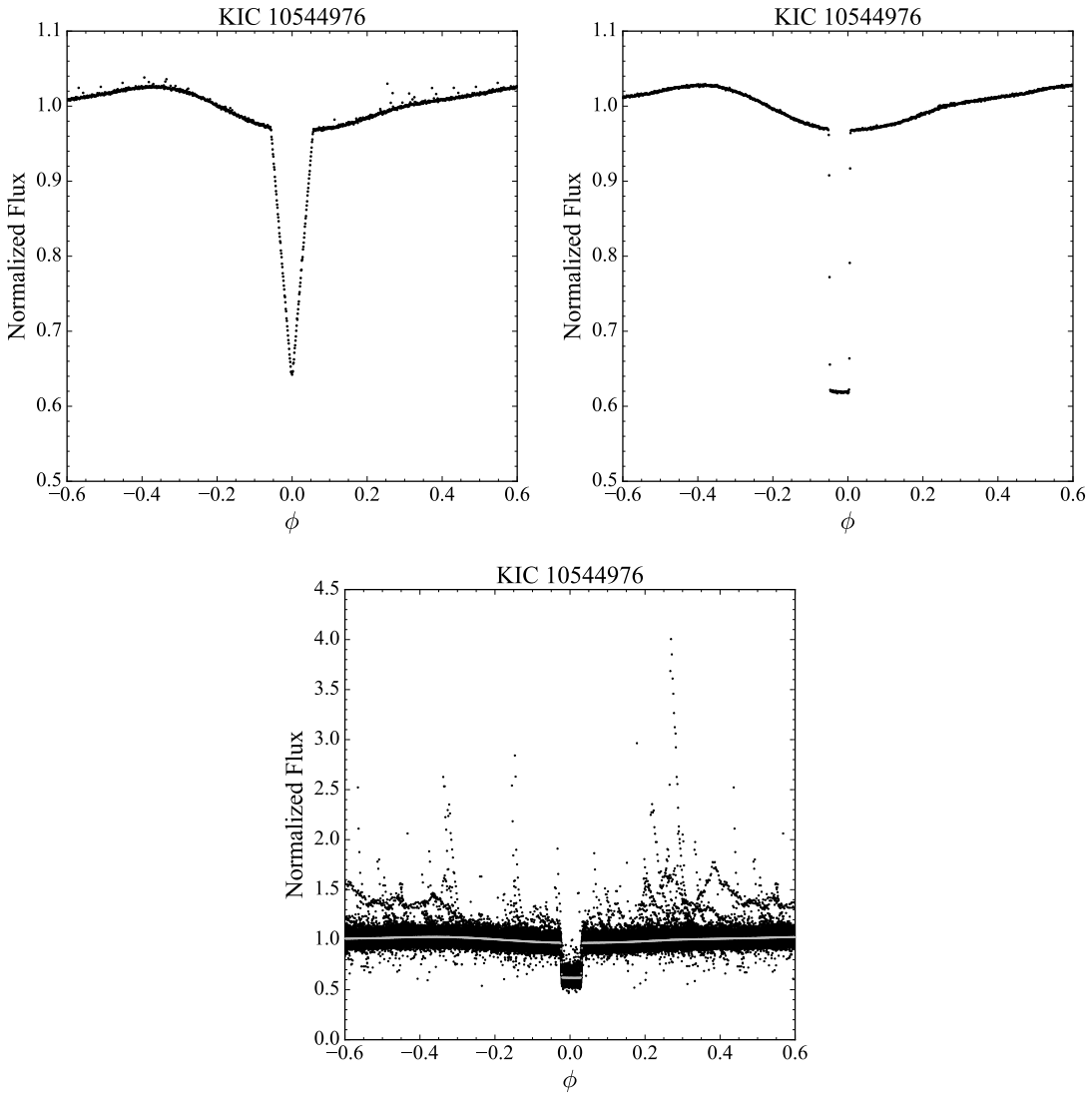


Figure 5.27: Averaged long-cadence (left), averaged short-cadence (right), and *Kepler* short-cadence (bottom) light curves of KIC 10544976 showing important features, including the sharp primary eclipse and numerous flares. The bottom panel also plots the averaged short-cadence light curve in grey for comparison.

of four. Because the flux increases monotonically after the primary eclipse until $\phi \approx -0.35$, there are not two inter-eclipse maxima to measure a difference between. As a result, I consider KIC 10544976 to lack a traditional O'Connell effect in the sense defined in Section 1.2.

My code classifies KIC 10544976 as an Algol-type binary with an OES of $-0.0097 \pm 0.0010$ and an eclipse depth of $0.357 \pm 0.001$ using the long-cadence data (note that the OES and eclipse depth errors are significant enough to quote due to KIC 10544976's apparent magnitude, $K_p = 18.679$). As noted previously, however, the long-cadence data is ill-suited for studying the system. Using the short-cadence data, my code classifies KIC 10577976 as an Algol-type binary with an OES of $-0.0077 \pm 0.0010$ and an eclipse depth of $0.395 \pm 0.001$.

Almenara et al. (2012) identify the system as having a DA white dwarf primary and an M4 V secondary with an 8.4-hour (0.350 d) orbit. Its *Gaia* EDR3 data indicates a distance of $521^{+45}_{-35}$ pc and a luminosity of $0.0059^{+0.0010}_{-0.0008}$ $L_\odot$. It lacks a temperature from either *Kepler* or *Gaia* DR2, but its *Gaia* color implies a spectral type of K4 V using Dr. Mamajek's table. The system has $\mu = 0.55$, and its ETV shows no variation. Almeida et al. (2019) ascribes the numerous flares seen in Figure 5.27's bottom panel to the red dwarf secondary's active chromosphere. KIC 10544976's SVP is 0.011 with the long- and short-cadence data sets, indicating only a modest amount of temporal variation.

## 5.4 KIC 11347875: Concave-Up Systems

A few systems in my sample display an inter-eclipse region that is concave-up rather than concave-down. I chose KIC 11347875 as the exemplar of this class because it has the most extreme example of this concave-up region. I created the left and right concavity parameters (LCP, RCP; Section 2.2.8) to quantify the degree to which the region before and after the primary eclipse were concave-up, respectively. I consider the most likely explanation of this phenomenon to be an

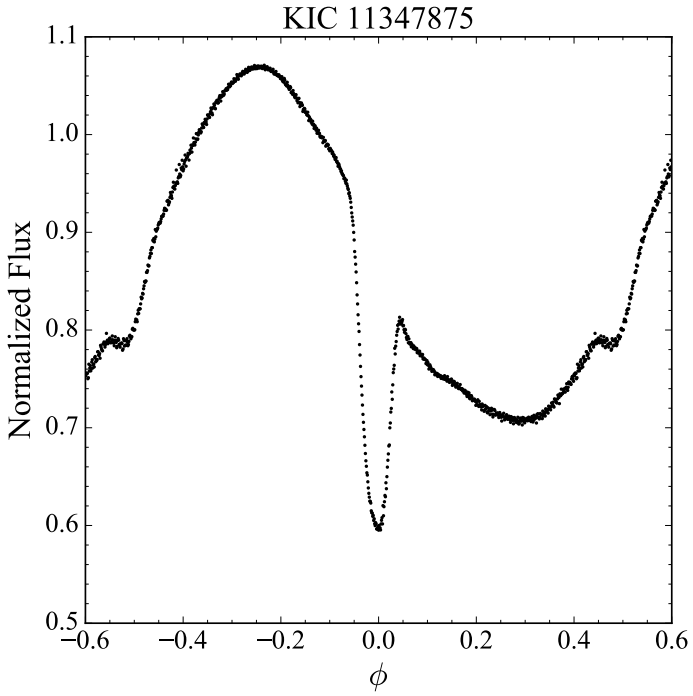


Figure 5.28: Averaged light curve of KIC 11347875 showing the system's concave-up region following primary eclipse.

enormous starspot covering most of one star's hemisphere.

### 5.4.1 Class Description

KIC 11347875 and systems like it (labeled with the flag CU in Table A.1) show a concave-up region before or after their primary eclipse, often causing the system's light curve to resemble a rough sinusoid. Figure 5.28 shows KIC 11347875's concave-up region following the primary eclipse, and Figure 2.7 shows an enlarged view of this region. Because KIC 11347875 lacks two inter-eclipse maxima to measure a difference between, I do not consider it and other concave-up systems to have traditional O'Connell effects in the sense defined in Section 1.2. Six other systems in the complete sample (KICs 5300878, 6044064, 6197038, 7671594, and 9119652

Table 5.12: List of concave-up systems.

| *Kepler* ID | Left Concavity | Right Concavity | Scaled Variance | Sample |
|---|---|---|---|---|
| 5300878 | −0.498 | 0.350 | 0.038 | C |
| 6044064 | −0.160 | 0.798 | 0.025 | C |
| 6197038 | 1.849 | −1.498 | 0.144 | C |
| 6697716 | 0.063 | −0.215 | 0.036 | M |
| 7671594 | 2.194 | −1.562 | 0.056 | C |
| 9119652 | 0.474 | −0.262 | 0.029 | C |
| 11347875 | −4.121 | 2.933 | 0.076 | C |

of the core sample and KIC 6697716 of the marginal sample) have a concave-up region. At least one non-sample system (KIC 5802285) also shows a concave-up region, and some temporally varying systems like KIC 8479107 show such a region during some time intervals. Additionally, at least one system outside the KEBC (EPIC 206036749; Miyakawa et al. 2021) displays this phenomenon. However, KIC 11347875 is by far the most extreme known example.

I used the LCP and RCP to quantify the degree that a system's light curve is concave-up. If a system has LCP $> 0$, the region before the primary eclipse is concave-up, and similarly for the region after the primary eclipse if RCP $> 0$. Based on this definition, six core sample systems (3% of the core sample) and one marginal sample system (2% of the marginal sample) exhibit a concave-up region. Table 5.12 lists all seven concave-up systems, along with each system's LCP, RCP, SVP, and which sample they belong to. Note that all seven systems either have LCP $> 0$ or RCP $> 0$, but not both. Unlike the SVP with respect to the temporally varying systems, the criteria LCP $> 0$ and RCP $> 0$ are objective, and a given system either does or does not have a concave-up region.

### 5.4.2 Class Analysis

In this section, I will summarize the characteristic distributions and trends and look at any correlations between the characteristics. I also discuss the results of my analysis of this class of systems. Unfortunately, my analysis largely finds that small sample statistics inherent to a sample size of seven makes it difficult to infer much about the concave-up systems.

#### 5.4.2.1 Sample Characteristics

The concave-up system class is comprised of seven Algol-type systems, with no $\beta$ Lyrae-type or W Ursae Majoris-type systems exhibiting a concave-up region. This proportion of system classes strongly differs from every other sample I have studied, implying that the concave-up systems are fundamentally different from other O'Connell effect binaries. The lack of $\beta$ Lyrae- or W Ursae Majoris-type systems implies that one or both binary components are well-detached from their Roche lobe.

The largest OES found in this class is –0.265 (KIC 11347875), while the smallest is 0.0049 (KIC 3351945), both in units of normalized flux. More concave-up systems show a positive O'Connell effect (four systems, or 57% of this class). The class's SVPs range from 0.025 (KIC 6044064) to 0.144 (KIC 6197038). While only three concave-up systems are also temporally varying systems (i.e. SVP $\geq$ 0.050), the class's SVP range indicates that all concave-up systems show a non-negligible degree of temporal variation. The positive concavity parameters range from 0.063 (KIC 6697716) to 2.933 (KIC 11347875).

The primary eclipse depth ranges from 0.107 (KIC 7671594) to 0.675 (KIC 6044064). Concave-up systems, then, do not seem to have shallow eclipses. However,

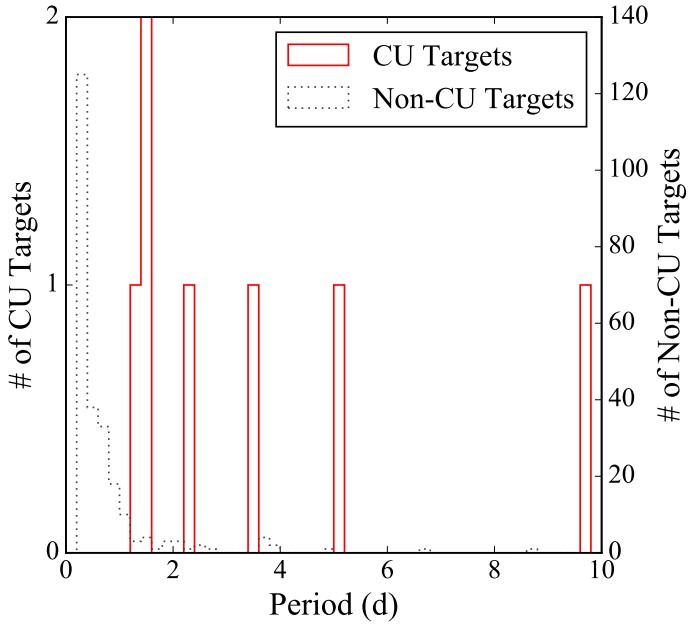


Figure 5.29: Histogram comparing the periods for all 7 concave-up systems (solid red) and the remaining 251 complete sample targets (dotted grey). Concave-up systems have much longer periods than other complete sample systems.

the non-sample concave-up system KIC 5802285 has an eclipse depth of 0.010. Since the OES of concave-up systems is the difference between the single maximum and the point directly before or after an eclipse, my sample may exclude KIC 5802285 due to a relationship between OES and system amplitude (which is related to the eclipse depth). In other words, concave-up systems with shallow eclipses may lie outside of my sample because such systems also have $|\mathrm{OES}| < 0.01$ due to a relation between OES and eclipse depth.

The concave-up systems' distances range from $210 \pm 1$ pc (KIC 7671594) to $2{,}982^{+191}_{-167}$ pc (KIC 6044064). The class's luminosities range from $0.0169 \pm 0.0004$ $L_\odot$ (KIC 7671594) to $7.14^{+0.26}_{-0.27}$ $L_\odot$ (KIC 6197038). The orbital periods of this class range from 1.279 d (KIC 5300878) to 9.752 d (KIC 6197038). This range implies a significant difference between the period distributions of the concave-up systems

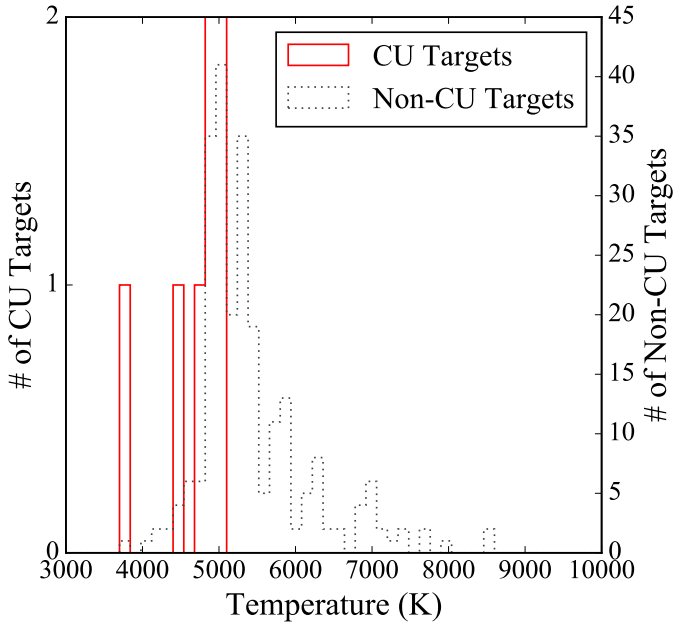


Figure 5.30: Histogram comparing the *Gaia* DR2 temperatures for all 7 concave-up systems (solid red) and 240 of the remaining 251 complete sample targets (dotted grey) with a *Gaia* DR2 temperature. Concave-up systems are much cooler than other complete sample systems.

and the rest of the sample, which Figure 5.29, showing a histogram comparing these distributions, confirms. The long periods of concave-up systems imply that they are likely detached binaries. The fact that all concave-up systems are Algol-type supports this idea.

The concave-up systems' temperatures range from $3{,}808^{+580}_{-413}$ K (KIC 7671594) to $5{,}048^{+152}_{-94}$ K (KIC 6044064). Figure 5.30 shows a temperature histogram for the concave-up systems compared to the rest of the complete sample. It indicates that concave-up systems are some of the coolest in my sample. As with the temporally varying systems, the lack of high-temperature systems suggests that starspots cause this feature.

Frasca et al. (2016) does not report the spectral type of any concave-up sys-

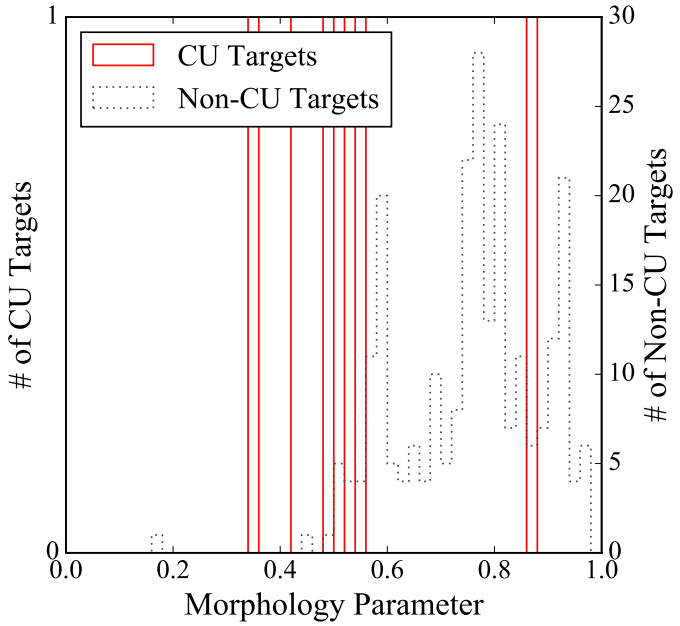


Figure 5.31: Histogram comparing the morphology parameters for all 7 concave-up systems (solid red) and 250 of the remaining 251 complete sample targets (dotted grey) that have $\mu \neq -1$. Concave-up systems have smaller $\mu$ on average than other complete sample systems.

tem. The spectral types estimated using Dr. Mamajek's table range from M2.5 V (KIC 7671594) to K1.5 V (KIC 6044064). Gao et al. (2016) identifies KIC 11347875 as containing two late-type red giants based on the effective temperature and surface gravity estimates, although my luminosity estimate using *Gaia* EDR3 data indicates a total luminosity of only $1.31 \pm 0.04$ $L_{\odot}$. Gao et al. (2016) also lists KICs 6044064, 6197038, 7671594, and 11347875 as exhibiting flares. However, inspecting all seven concave-up systems' *Kepler* light curves shows that all except KIC 9119652 (which *Kepler* only observed for a single quarter) show at least one flare. I discussed Yoldaş's (2021) discussion of KIC 6044064's flare activity in Section 3.8.

The concave-up systems' morphology parameters range from 0.34 (KIC 767-

1594) to 0.86 (KIC 11347875). Figure 5.31 shows a histogram of the concave-up systems' $\mu$ distribution compared to the rest of the complete sample's. The concave-up systems have a smaller $\mu$ on average, which agrees with their longer periods and Algol classification. The lone exception is KIC 11347875 with $\mu = 0.86$. Despite this, Figure 5.28 shows that KIC 11347875 is clearly an Algol-type system, albeit with prominent ellipsoidal variation more characteristic of $\beta$ Lyrae- or W Ursae Majoris-type systems.

#### 5.4.2.2 Characteristic Trends

Figures 5.32-5.34 show the concave-up systems' corner plot comparing period, OES, *Gaia* temperature, the *Gaia* BP – RP color index, morphology parameter, absolute *Gaia* G magnitude, and primary eclipse depth. Unlike with the other samples, there are no readily apparent trends in Figures 5.32-5.34 aside from the characteristic distributions noted in Section 5.4.2.1. The lack of trends may be a symptom of the small number of concave-up systems.

#### 5.4.2.3 Sample Distributions

Table 5.13 gives the results of the K–S test for the concave-up systems. The K–S test rejects the null hypothesis only for the morphology parameter and period. Recall, however, Equation 2.78, indicating that rejecting the null hypothesis depends on the sample size. Since the concave-up sample is comprised of only seven systems, it is more challenging to reject the null hypothesis. For instance, the temperature K–S statistic of 0.649 is insufficient to reject the null hypothesis, whereas the temporally varying systems' OES K–S statistic of 0.307 (Table 5.2) is sufficient for that 75 system sample. I conclude that the K–S statistic for the concave-up

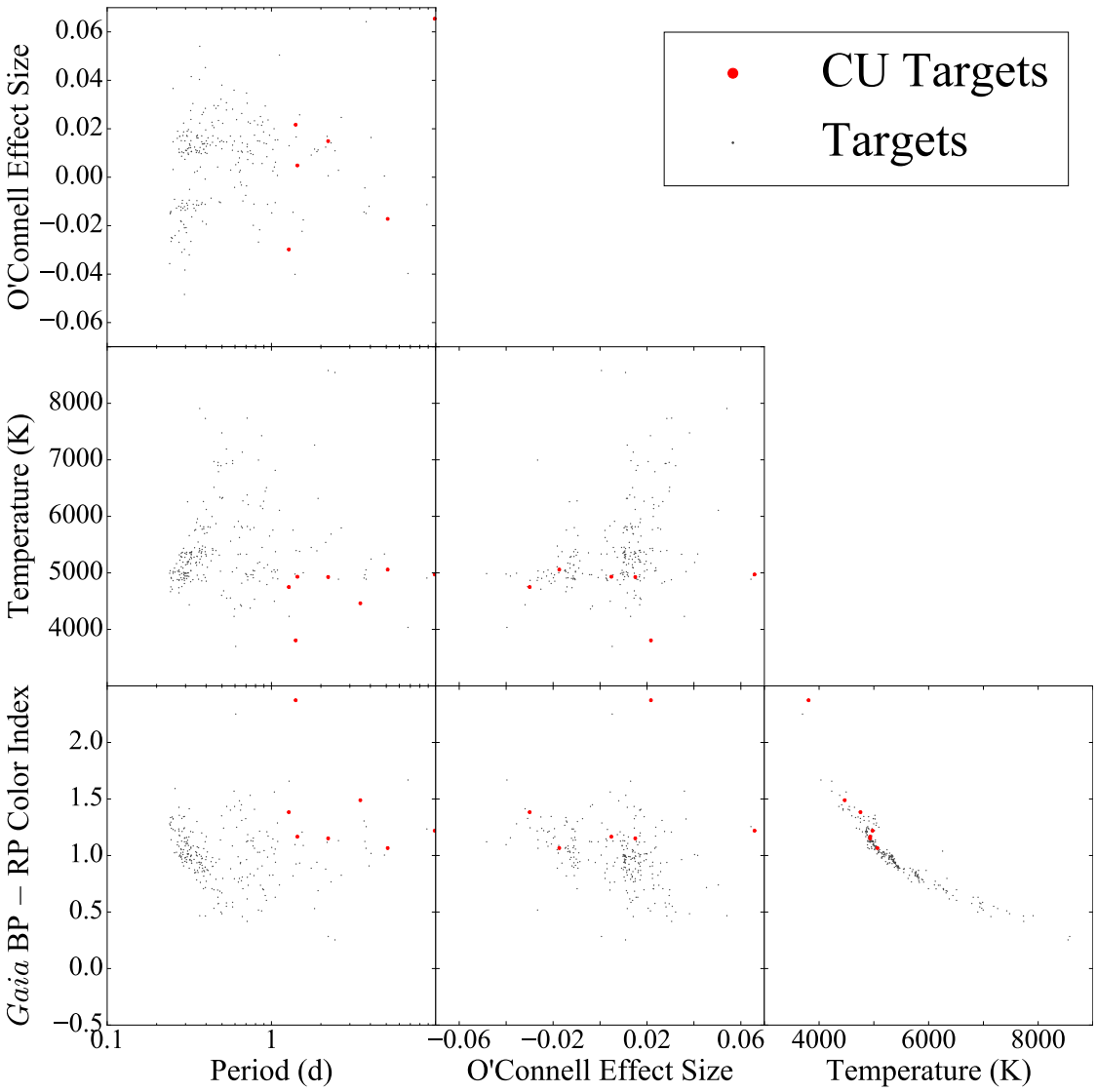


Figure 5.32: Corner plot showing the concave-up system correlations between three characteristics of interest (period, OES, and temperature) and OES, temperature, and color. Concave-up targets are shown in red while the other complete sample systems are shown in grey. KIC 11347875 has been removed from the OES plots for clarity (see discussion in Section 4.2.1). Note the logarithmic x-axis of the period plots.

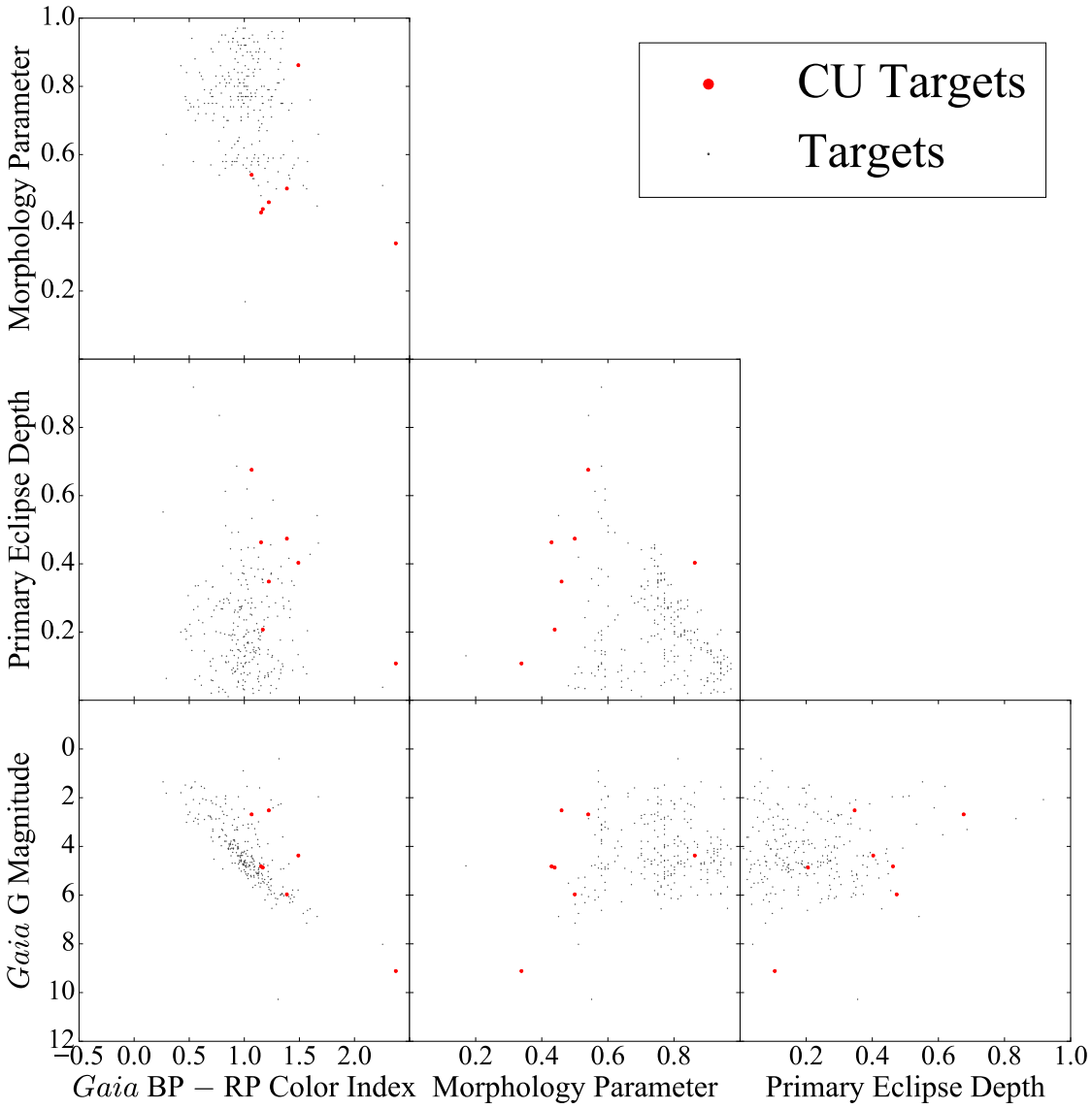


Figure 5.33: Corner plot showing the concave-up system correlations between three characteristics of interest (color, morphology parameter, and primary eclipse depth) and morphology parameter, primary eclipse depth, and absolute magnitude. Concave-up targets are shown in red while the other complete sample systems are shown in grey.

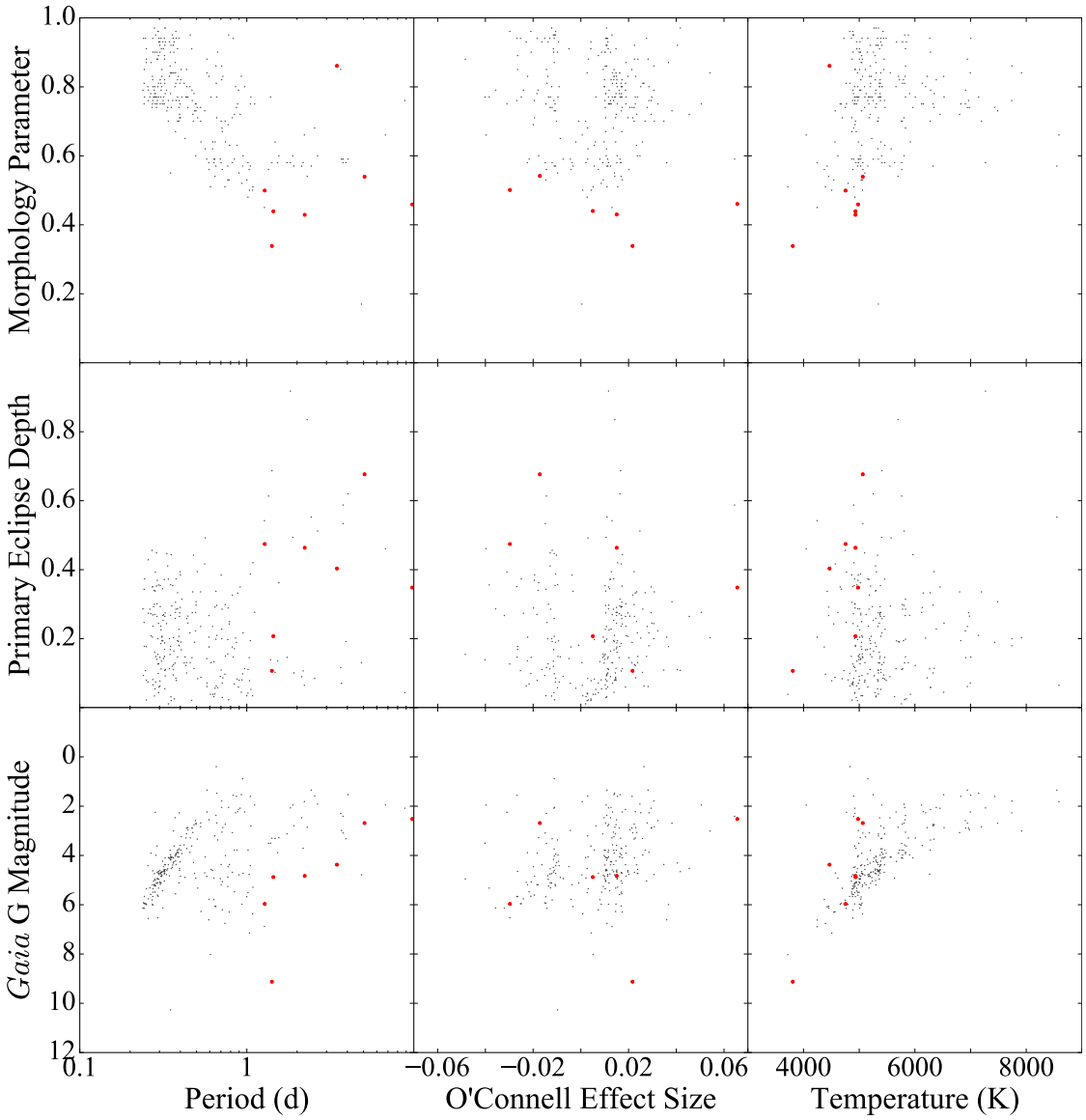


Figure 5.34: Corner plot showing the concave-up system correlations between three characteristics of interest (period, OES, and temperature) and morphology parameter, primary eclipse depth, and absolute magnitude. Concave-up targets are shown in red while the other complete sample systems are shown in grey. KIC 11347875 has been removed from the OES plots for clarity (see discussion in Section 4.2.1). Note the logarithmic x-axis of the period plots.

Table 5.13: Results of analyzing characteristic distributions between the concave-up systems and the rest of the complete sample using the Kolomgorov–Smirnov test.

| Characteristic | K–S Statistic | $p$-Value |
|---|---|---|
| O'Connell Effect Size | 0.317 | 0.423 |
| \|O'Connell Effect Size\| | 0.403 | 0.165 |
| Primary Eclipse Depth | 0.547 | 0.020 |
| Morphology Parameter | 0.801 | $< 0.001$ |
| Temperature | 0.649 | 0.003 |
| Distance | 0.400 | 0.172 |
| Absolute Magnitude | 0.245 | 0.747 |
| Period | 0.896 | $< 0.001$ |

systems is heavily influenced by small sample statistics. Nevertheless, the statistic shows which distributions are more dissimilar than others, so it retains some value. In that light, Table 5.13 shows that the concave-up systems' morphology parameter, temperature, and period distributions differ the most from the rest of the sample's, echoing the results from Section 5.4.2.1.

#### 5.4.2.4 Characteristic Correlations

Table 5.14 gives the results of my analysis of the concave-up systems using Spearman's $\rho$ and Kendall's $\tau$ coefficients. As with K–S test, small sample statistics reduces the utility of these tests, as shown by the fact that – aside from the correlations between OES and absolute magnitude and between |OES| and eclipse depth – *every* characteristic pair is correlated among the concave-up systems. Furthermore, only the correlation between absolute magnitude and period is strong enough to reject the null hypothesis, and even then only using the Spearman's test. Since it is the only correlation of note, I shall focus my discussion on it.

Figure 5.34's lower-left panel shows that, except for the pair KIC 5300878 and

Table 5.14: Results of analyzing the correlations between the concave-up systems' characteristics of interest using Spearman's and Kendall's rank correlation coefficients.

| Characteristic One | Characteristic Two | Spearman's Coeff. | Spearman's $p$-Value | Kendall's Coeff. | Kendall's $p$-Value |
|---|---|---|---|---|---|
| OES | Eclipse Depth | **−0.536** | **0.215** | **−0.333** | **0.381** |
| OES | Morphology | **−0.750** | **0.052** | **−0.619** | **0.069** |
| OES | Temperature | **0.179** | **0.702** | **0.143** | **0.773** |
| OES | Distance | **−0.107** | **0.819** | **−0.143** | **0.773** |
| OES | Absolute Mag. | −0.071 | 0.879 | −0.048 | 1.000 |
| OES | Period | **0.214** | **0.644** | **0.143** | **0.773** |
| \|OES\| | Eclipse Depth | 0.036 | 0.939 | 0.048 | 1.000 |
| \|OES\| | Morphology | **0.536** | **0.215** | **0.333** | **0.381** |
| \|OES\| | Temperature | **−0.321** | **0.482** | **−0.238** | **0.562** |
| \|OES\| | Distance | **0.143** | **0.760** | **0.048** | **1.000** |
| \|OES\| | Absolute Mag. | **−0.286** | **0.534** | **−0.238** | **0.562** |
| \|OES\| | Period | **0.250** | **0.589** | **0.143** | **0.773** |
| Eclipse Depth | Morphology | **0.607** | **0.148** | **0.524** | **0.136** |
| Eclipse Depth | Temperature | **0.429** | **0.337** | **0.333** | **0.381** |
| Eclipse Depth | Distance | **0.393** | **0.383** | **0.238** | **0.562** |
| Eclipse Depth | Absolute Mag. | **−0.357** | **0.432** | **−0.238** | **0.562** |
| Eclipse Depth | Period | **0.179** | **0.702** | **0.143** | **0.773** |
| Morphology | Temperature | **0.250** | **0.589** | **0.238** | **0.562** |
| Morphology | Distance | **0.643** | **0.119** | **0.524** | **0.136** |
| Morphology | Absolute Mag. | **−0.536** | **0.215** | **−0.333** | **0.381** |
| Morphology | Period | **0.393** | **0.383** | **0.238** | **0.562** |
| Temperature | Distance | **0.786** | **0.036** | **0.714** | **0.030** |
| Temperature | Absolute Mag. | **−0.714** | **0.071** | **−0.524** | **0.136** |
| Temperature | Period | **0.643** | **0.119** | **0.429** | **0.239** |
| Distance | Absolute Mag. | **−0.929** | **0.003** | **−0.810** | **0.011** |
| Distance | Period | **0.893** | **0.007** | **0.714** | **0.030** |
| Absolute Mag. | Period | **−0.964** | **$<$0.001** | **−0.905** | **0.003** |

KIC 7671594 (the leftmost two systems in the panel), the relationship between luminosity and period is perfectly monotonic. Considering that there are only seven systems, however, it is difficult to infer whether this relationship is due to a genuine correlation or random chance. Still, the correlation may make physical sense like the same correlation does amongst the temporally stable asymmetric

minima systems: these stars may have the shortest period possible for a given luminosity. If that is true, the fact that these systems lie far away from the main sequence line that the temporally stable asymmetric minima systems lie on implies that concave-up systems contain at least one evolved component. It is nearly certain that KICs 6044064 and 6197038 (the two rightmost systems in Figure 5.34's lower-left panel) contain evolved stars due to having $T < 5{,}100$ K but $L > 5$ $L_\odot$.

### 5.4.3 Potential Explanations

Some process must be either reduce or obscure a binary's light when its light curve is concave-up. Since the concave-up region's position in the light curve seems fixed, I can rule out any source that rotates asynchronously with the binary, suggesting a feature on one of the stars. Since all but one concave-up system has been observed flaring, at least one star must be chromospherically active. The "remarkably high level of magnetic activity" seen in the concave-up system KIC 6044064 (Yoldaş 2021) – which I discussed in Section 3.8 – supports this idea of chromospheric activity. Therefore, starspots are a plausible cause of the concave-up feature. Such starspots would differ significantly from sunspots or the starspots causing temporal variation, however, because small spots could not cause such large dips in flux over nearly half of the light curve. If starspots cause the concave-up region, they must be *enormous*, with sizes on the order of the star itself.

I tested my hypothesis using BM3 to recreate KIC 11347875's light curve using an enormous starspot. I created a toy eclipsing binary using the parameters given in Table 5.15. The toy system is a detached binary wherein the primary star's leading hemisphere is almost entirely covered by a cool starspot. Figure 5.35

Table 5.15: System parameters of the concave-up toy model used to create Figure 5.35.

| Parameter | Value(s) |
|---|---|
| Orbital Inclination (°) | 80 |
| Mass Ratio | 0.7 |
| $T_{eff}$ of Primary Component (K) | 4,800 |
| $T_{eff}$ of Secondary Component (K) | 3,800 |
| Surface Potential of Primary Component | 3.6197 |
| Surface Potential of Secondary Component | 3.6197 |
| Spot Temperature (K) | 4,080 |
| Spot Latitude (°) | 0 |
| Spot Longitude (°) | 270 |
| Spot Radius (°) | 80 |

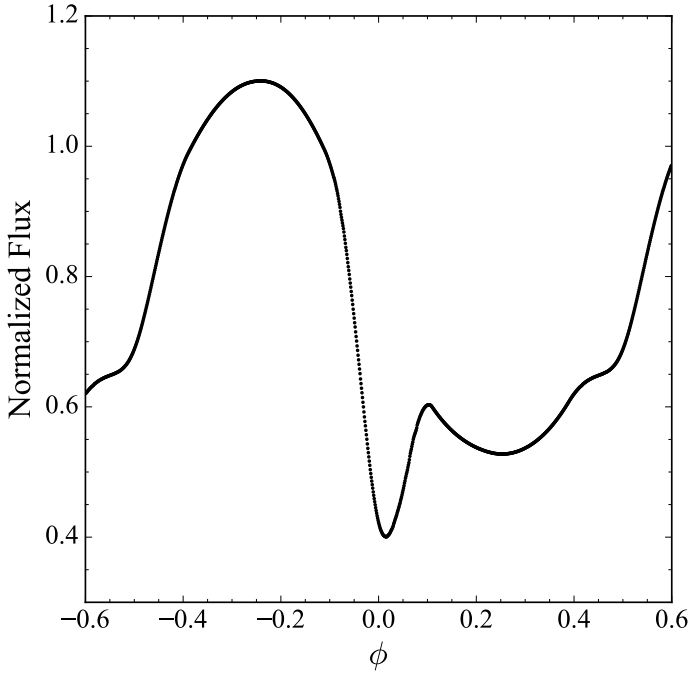


Figure 5.35: Light curve of an eclipsing binary model that looks similar to KIC 11347875's light curve. The concave-up region was created using a large starspot on the primary star covering most of the leading hemisphere, demonstrating how such spots can cause this phenomenon.

shows the light curve of this toy system, and it bears a striking resemblance to KIC 11347875's light curve shown in Figure 5.28. I reason that if an enormous starspot can recreate the most extreme member of the concave-up class, it can recreate every member. Therefore, I conclude that an enormous starspot covering most of one component's hemisphere causes the concave-up region that distinguishes this class of peculiar systems. I do not speculate as to the cause of the starspot.

# Chapter 6

# Observations

I conducted observations of my sample using the methods I discussed in Section 2.3. I obtained multi-wavelength data on ten systems with SARA-KP and SARA-RM: KICs 5195137, 5282464, 6223646, 7433513, 7885570, 8696327, 8822555, 9164694, 10861842, and 11924311. Table 6.1 lists the observation dates for each target, along with the telescopes and filters used on each night. Table 6.2 lists the coordinates and Johnson–Cousins $BVR_CI_C$ magnitudes for each target, comparison, and check star. Seven systems I observed (KICs 5282464, 6223646, 8696327, 8822555, 9164694, 10861842, and 11924311) have complete or nearly complete phase coverage with high-quality data, while the other three systems have incomplete phase coverage (KICs 5195137 and 7885570) or low-quality data (KICs 5195137 and 7433513). I prioritized nine of these systems for either the presence of an asymmetric minimum (Section 5.2; KICs 8696327, 8822555, 9164694, 10861842, and 11924311), significant temporal variation (Section 5.1; KICs 7433513 and 7885570), or an otherwise interesting light curve (KICs 5282464 and 6223646). KIC 5195137 was a target of opportunity, as it happens to lie in the same field of view as

Table 6.1: Dates and locations of observation for each target.

| Target | Observation Date | Location | Filters |
|---|---|---|---|
| KIC 5195137 | 11/11/2019 | ORM | U, B, V, $R_C$, $I_C$ |
| KIC 5282464 | 10/25/2018 | KPNO | B, V, $R_C$, $I_C$ |
| | 06/22/2019 | KPNO | B, V, $R_C$, $I_C$ |
| | 07/01/2019 | KPNO | B, V, $R_C$, $I_C$ |
| | | ORM | B, V, $R_C$, $I_C$ |
| | 11/11/2019 | ORM | U, B, V, $R_C$, $I_C$ |
| KIC 6223646 | 10/14/2018 | ORM | B, V, $R_C$, $I_C$ |
| | 06/01/2019 | KPNO | B, V, $R_C$, $I_C$ |
| | | ORM | B, V, $R_C$, $I_C$ |
| | 06/02/2019 | KPNO | B, V, $R_C$, $I_C$ |
| | 06/03/2019 | KPNO | B, V, $R_C$, $I_C$ |
| | 09/30/2019 | ORM | B, V, $R_C$, $I_C$ |
| | 10/08/2019 | KPNO | B, V, $R_C$, $I_C$ |
| | 05/19/2020 | ORM | B, V, $R_C$, $I_C$ |
| KIC 7433513 | 09/14/2018 | KPNO | B, V, $R_C$, $I_C$ |
| | 09/18/2018 | ORM | B, V, $R_C$, $I_C$ |
| | 09/19/2018 | ORM | B, V, $R_C$, $I_C$ |
| | 09/20/2018 | KPNO | B, V, $R_C$, $I_C$ |
| | | ORM | B, V, $R_C$, $I_C$ |
| KIC 7885570 | 09/01/2019 | KPNO | U, B, V, $R_C$, $I_C$ |
| | | ORM | U, B, V, $R_C$, $I_C$ |
| | 09/02/2019 | KPNO | U, B, V, $R_C$, $I_C$ |
| | | ORM | U, B, V, $R_C$, $I_C$ |
| | 09/04/2019 | KPNO | U, B, V, $R_C$, $I_C$ |
| | 09/05/2019 | KPNO | U, B, V, $R_C$, $I_C$ |
| | | ORM | U, B, V, $R_C$, $I_C$ |
| | 09/06/2019 | KPNO | U, B, V, $R_C$, $I_C$ |
| KIC 8696327 | 06/18/2020 | ORM | B, V, $R_C$, $I_C$ |
| | 06/23/2020 | ORM | B, V, $R_C$, $I_C$ |
| | 06/28/2020 | ORM | B, V, $R_C$, $I_C$ |
| KIC 8822555 | 10/02/2018 | ORM | B, V, $R_C$, $I_C$ |
| | 06/02/2019 | ORM | B, V, $R_C$, $I_C$ |
| | 06/03/2019 | ORM | B, V, $R_C$, $I_C$ |
| | 06/11/2019 | ORM | B, V, $R_C$, $I_C$ |
| | 06/17/2020 | ORM | B, V, $R_C$, $I_C$ |
| | 06/21/2020 | ORM | B, V, $R_C$, $I_C$ |

Table 6.1 *continued on next page*

Table 6.1 *(continued)*

| Target | Observation Date | Location | Filters |
|---|---|---|---|
| KIC 9164694 | 07/10/2017 | ORM | B, V, $R_C$, $I_C$ |
| | 07/11/2017 | ORM | B, V, $R_C$, $I_C$ |
| | 10/12/2017 | KPNO | B, V, $R_C$, $I_C$ |
| | 08/07/2018 | ORM | B, V, $R_C$, $I_C$ |
| | 08/27/2018 | ORM | B, V, $R_C$, $I_C$ |
| | 08/28/2018 | ORM | B, V, $R_C$, $I_C$ |
| | 08/31/2018 | ORM | B, V, $R_C$, $I_C$ |
| | 05/30/2019 | KPNO | B, V, $R_C$, $I_C$ |
| | | ORM | B, V, $R_C$, $I_C$ |
| | 05/31/2019 | ORM | B, V, $R_C$, $I_C$ |
| | 06/04/2019 | ORM | B, V, $R_C$, $I_C$ |
| | 06/08/2019 | ORM | B, V, $R_C$, $I_C$ |
| KIC 10861842 | 04/15/2018 | KPNO | B, V, $R_C$, $I_C$ |
| | 06/08/2018 | KPNO | B, V, $R_C$, $I_C$ |
| | 08/02/2018 | ORM | B, V, $R_C$, $I_C$ |
| | 09/01/2018 | ORM | B, V, $R_C$, $I_C$ |
| | 05/05/2019 | ORM | B, V, $R_C$, $I_C$ |
| | 05/30/2019 | KPNO | B, V, $R_C$, $I_C$ |
| | 08/30/2019 | ORM | B, V, $R_C$, $I_C$ |
| KIC 11924311 | 08/19/2017 | ORM | B, V, $R_C$, $I_C$ |
| | 09/06/2017 | ORM | B, V, $R_C$, $I_C$ |
| | 06/04/2018 | ORM | B, V, $R_C$, $I_C$ |
| | 05/10/2019 | KPNO | B, V, $R_C$, $I_C$ |
| | 05/28/2019 | KPNO | B, V, $R_C$, $I_C$ |
| | 06/07/2019 | KPNO | B, V, $R_C$, $I_C$ |

KIC 5282464 and is of similar brightness. The temporal variation KICs 7433513 and 7885570 exhibit required more coordinated observations, so I will discuss those two systems separately from the other eight. Table 6.3 lists the OES and SVP I calculated from the *Kepler* data and the *Gaia* temperature for each system I observed, along with the OESs and estimated primary component temperature I calculated from my observations. The OESs are in magnitudes, as I converted the *Kepler* OESs from normalized flux to magnitude. Table 6.3 omits OES errors for space.

Table 6.2: Coordinates and magnitudes of target and comparison stars.

| Star | R.A. (J2000) | Dec. (J2000) | B | V | $R_C$ | $I_C$ | Type |
|---|---|---|---|---|---|---|---|
| KIC 5195137 | 19:36:25.2 | +40:18:05.89 | – | – | – | – | Var. |
| KIC 5282464 | 19:35:47.0 | +40:26:58.34 | 13.076 | 12.735 | 12.536 | 12.345 | Var. |
| KIC 5282836 | 19:36:08.1 | +40:24:19.78 | 13.239 | 12.790 | 12.549 | 12.320 | Comp. |
| KIC 5283073 | 19:36:24.0 | +40:27:03.52 | 12.847 | 12.332 | 12.086 | 11.852 | Check |
| KIC 6223646 | 19:47:32.0 | +41:31:07.90 | 13.913 | 13.487 | 13.364 | 13.243 | Var. |
| KIC 6223262 | 19:47:11.3 | +41:32:43.50 | 13.876 | 13.164 | 12.886 | 12.623 | Comp. |
| KIC 6223262 | 19:47:19.1 | +41:34:33.91 | 14.705 | 13.999 | 13.702 | 13.421 | Check |
| KIC 7433513 | 19:12:27.9 | +43:05:39.41 | 15.693 | 14.994 | 14.548 | 14.131 | Var. |
| KIC 7433560 | 19:12:32.7 | +43:03:46.27 | 16.161 | 15.482 | 15.222 | 14.975 | Comp. |
| KIC 7516242 | 19:12:06.1 | +43:07:34.68 | 16.133 | 15.421 | 14.979 | 14.565 | Check |
| KIC 7885570 | 19:19:53.7 | +43:39:13.77 | 12.609 | 11.851 | 11.485 | 11.141 | Var. |
| KIC 7817045 | 19:20:11.4 | +43:35:55.23 | 11.279 | 11.279 | 11.086 | 10.901 | Comp. |
| KIC 7885455 | 19:19:41.9 | +43:36:27.94 | 13.833 | 12.627 | 12.120 | 11.647 | Check |
| KIC 8696327 | 19:38:02.7 | +44:48:24.22 | 15.236 | 14.621 | 14.239 | 13.880 | Var. |
| KIC 8631850 | 19:37:59.4 | +44:47:18.47 | 15.648 | 14.872 | 14.442 | 14.039 | Comp. |
| KIC 8696271 | 19:37:58.1 | +44:48:40.84 | – | – | – | – | Check |
| KIC 8822555 | 19:33:10.4 | +45:02:31.20 | 14.987 | 14.439 | 14.117 | 13.813 | Var. |
| KIC 8822660 | 19:33:19.6 | +45:04:57.70 | 15.132 | 14.602 | 14.226 | 13.873 | Comp. |
| KIC 8822408 | 19:32:58.2 | +45:05:44.95 | 15.384 | 14.797 | 14.385 | 13.999 | Check |
| KIC 9164694 | 19:42:39.8 | +45:30:24.70 | 14.778 | 14.381 | 14.091 | 13.817 | Var. |
| KIC 9099963 | 19:42:27.2 | +45:28:41.97 | 14.464 | 14.259 | 13.877 | 13.518 | Comp. |
| KIC 9164765 | 19:42:45.0 | +45:31:15.25 | 14.809 | 14.347 | 14.093 | 13.852 | Check |
| KIC 10861842 | 19:28:17.2 | +48:14:17.90 | 14.695 | 14.102 | 13.777 | 13.471 | Var. |
| KIC 10861591 | 19:27:51.3 | +48:17:46.49 | 13.980 | 13.647 | 13.202 | 12.786 | Comp. |
| KIC 10861754 | 19:28:08.1 | +48:15:10.62 | 14.480 | 13.759 | 13.340 | 12.947 | Check |
| KIC 11924311 | 19:46:39.9 | +50:14:17.96 | 13.435 | 13.061 | 12.673 | 12.309 | Var. |
| KIC 11924234 | 19:46:33.4 | +50:14:19.27 | 12.920 | 12.615 | 12.310 | 12.022 | Comp. |
| KIC 11924366 | 19:46:44.1 | +50:16:04.20 | 13.473 | 13.175 | 12.923 | 12.684 | Check |

Table 6.3: Selected characteristics and results of my observed targets.

| Target | OES $K_p$ | OES B | OES V | OES $R_C$ | OES $I_C$ | *Gaia* Temp. (K) | Temp. Est. (K) | SVP |
|---|---|---|---|---|---|---|---|---|
| KIC 5195137 | 0.011 | – | – | – | – | $5{,}326^{+87}_{-132}$ | $5{,}386^{+155}_{-173}$ | 0.047 |
| KIC 5282464 | 0.042 | 0.053 | 0.034 | 0.027 | 0.044 | $7{,}478^{+72}_{-154}$ | $7{,}063^{+101}_{-99}$ | 0.006 |
| KIC 6223646 | 0.060 | 0.087 | 0.057 | 0.071 | 0.042 | $7{,}907^{+148}_{-461}$ | $6{,}453^{+225}_{-213}$ | 0.005 |
| KIC 7433513 | 0.038 | −0.072 | −0.077 | −0.100 | −0.079 | $5{,}337^{+99}_{-305}$ | $5{,}470^{+281}_{-239}$ | 0.080 |
| KIC 7885570 | 0.010 | – | – | – | – | $5{,}455^{+84}_{-57}$ | $5{,}470^{+29}_{-28}$ | 0.088 |
| KIC 8696327 | 0.025 | 0.011 | 0.018 | 0.012 | 0.019 | $6{,}144^{+499}_{-269}$ | $5{,}855^{+138}_{-147}$ | 0.017 |
| KIC 8822555 | 0.027 | 0.022 | 0.023 | 0.016 | 0.025 | $6{,}185^{+293}_{-612}$ | $6{,}496^{+91}_{-87}$ | 0.006 |
| KIC 9164694 | 0.056 | 0.032 | 0.025 | 0.027 | 0.021 | $6{,}113^{+210}_{-520}$ | $7{,}825^{+61}_{-59}$ | 0.007 |
| KIC 10861842 | 0.032 | 0.024 | 0.024 | 0.020 | 0.027 | $6{,}955^{+246}_{-256}$ | $6{,}632^{+794}_{-668}$ | 0.013 |
| KIC 11924311 | 0.033 | 0.039 | 0.035 | 0.027 | 0.027 | $6{,}515^{+414}_{-355}$ | $6{,}867^{+400}_{-371}$ | 0.007 |

## 6.1 The Stables: KICs 5195137, 5282464, 622-3646, 8696327, 8822555, 9164694, 10861842, and 11924311

These eight systems do not show significant temporal variation. Indeed, with the exception of KIC 5195137, their light curves are remarkably stable with time. This stability allowed me to observe a given system for over two years without fear that I may be unable to combine the data.

### 6.1.1 KIC 5195137

KIC 5195137 has an orbital period of 0.324 d and a $\mu$ of 0.92, suggesting a strongly overcontact system or ellipsoidal variable. The KEBC has a single flag for the system – UNC – indicating that it is uncertain that the system is an eclipsing binary. The system's *Kepler* data shows an OES of 0.011, a primary eclipse depth of 0.056, and an SVP of 0.047. *Gaia* reports a temperature of $5{,}326^{+87}_{-132}$ K for the system, while the *Gaia* color indicates a spectral type of K1 V. KIC 5195137's *Gaia* EDR3 data implies a luminosity of $1.10 \pm 0.03$ $L_{\odot}$. Balaji et al. (2015) and Davenport (2016) cite the system as having starspots and flares, respectively.

I observed KIC 5195137 – which lies close enough to KIC 5282464 to allow for simultaneous observations of both systems – only on the night of November 11, 2019, with SARA-RM. My previous observations of KIC 5282464 did not include KIC 5195137 because I had not yet added the latter system to the sample, and so I made no effort to include it in the field of view. The data collected on the system covers only the secondary eclipse. The data is also low-quality because

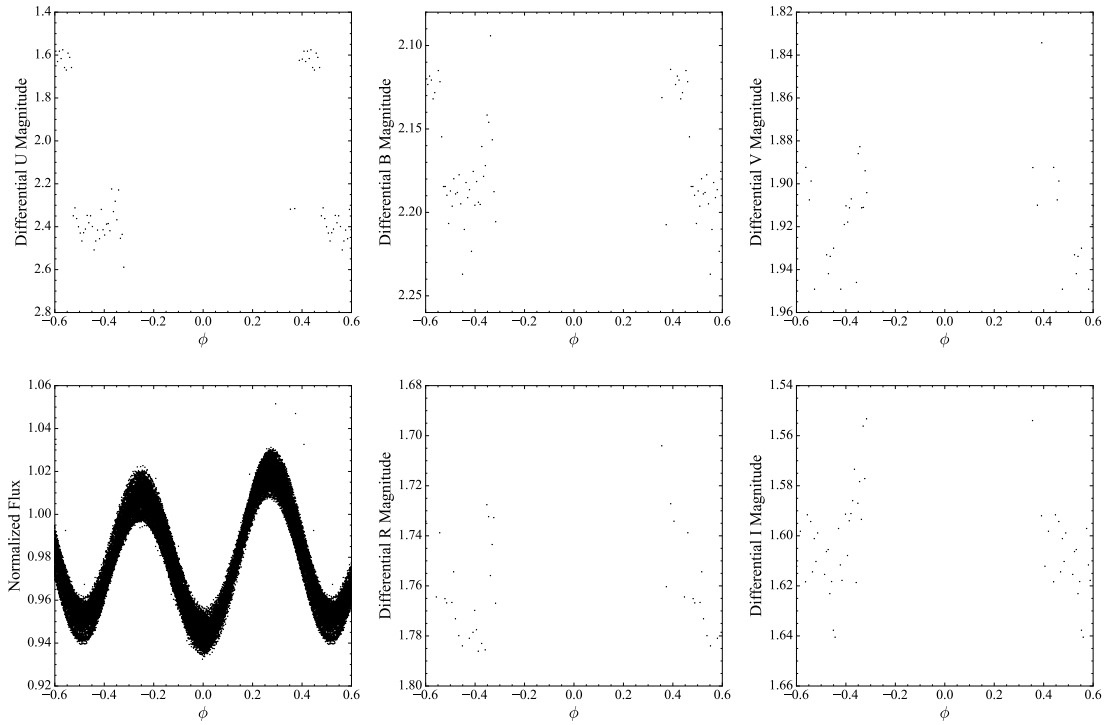


Figure 6.1: *Kepler* (bottom left), U (top left), B (top center), V (top right), $R_C$ (bottom center), and $I_C$ (bottom right) light curves for KIC 5195137.

KIC 5195137 is a full magnitude fainter than KIC 5282464 – a higher priority target – and so the signal-to-noise ratio is significantly lower for this system. High winds on the night of observation further reduced the data quality. Figure 6.1 shows the data I collected on this system. Figure 6.1 also shows KIC 5195137's *Kepler* data in the left panel, demonstrating that the system displays a moderate amount of temporal variation. KIC 5195137's SVP of 0.047 is just below the threshold for inclusion in the temporally varying system class.

I classified KIC 5195137 as a W Ursae Majoris-type system displaying partial eclipses. I could not determine an OES due to the lack of data at either maximum. The system's $B - V$ color index is $0.76 \pm 0.06$ during the secondary eclipse (neglecting interstellar reddening), corresponding to a primary component temper-

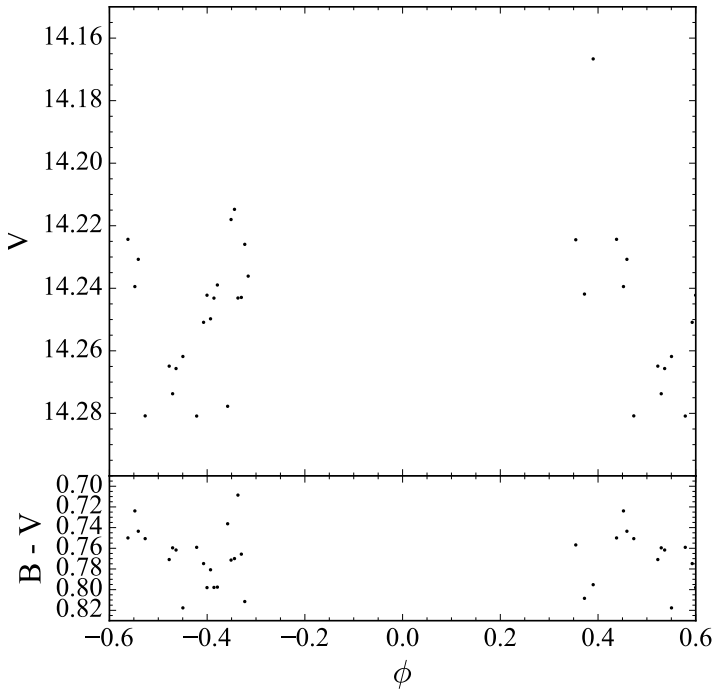


Figure 6.2: Apparent V magnitude and B − V color curve for KIC 5195137.

ature $T_1 \approx 5{,}386^{+155}_{-173}$ K by Flower (1996). Figure 6.2 shows KIC 5195137's B − V color curve using the magnitudes given in Table 6.2.

### 6.1.2 KIC 5282464

KIC 5282464 has an orbital period of 0.496 d and a $\mu$ of 0.73, suggesting a semi-detached or overcontact system. The KEBC has two flags for the system – FB and TM – indicating that the the eclipses are flat-bottomed (i.e. the system experiences total eclipses) and that the ETV indicates a third body. The system's *Kepler* data shows an OES of 0.038, a primary eclipse depth of 0.308, and an SVP of 0.006. *Gaia* reports a temperature of $7{,}478^{+72}_{-154}$ K for the system, while the *Gaia* color indicates a spectral type of F2 V. KIC 5282464's *Gaia* EDR3 data implies a luminosity of $6.14^{+0.16}_{-0.14}$ $L_\odot$. Hartman et al. (2004) and Pigulski et al. (2009) cite

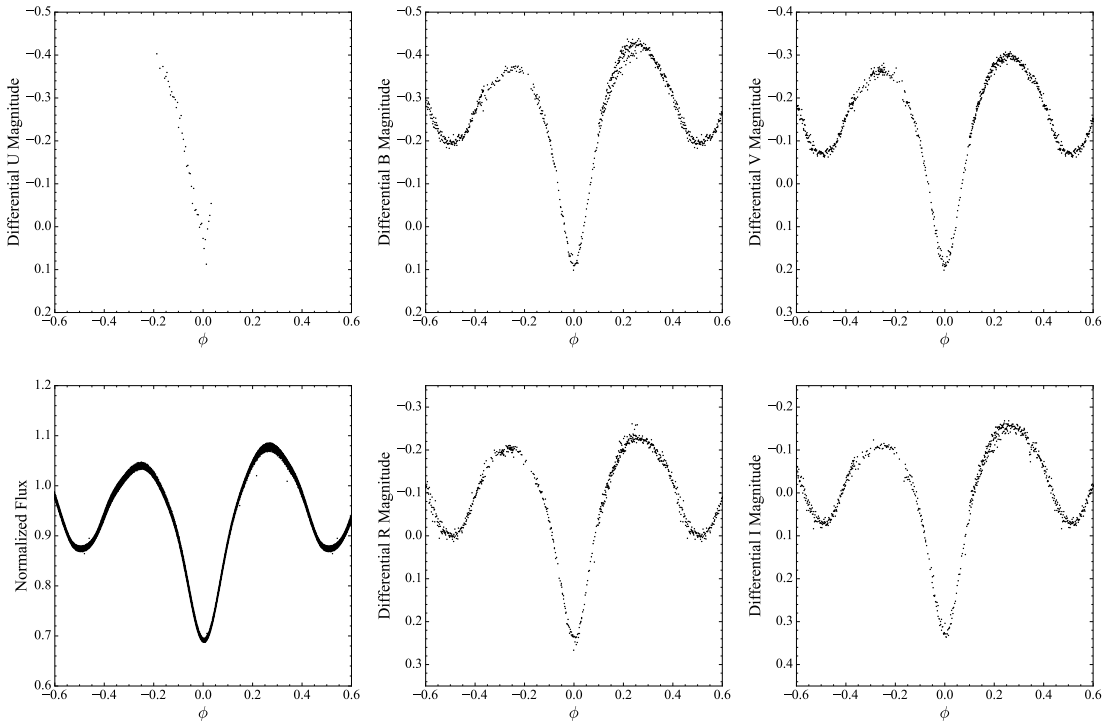


Figure 6.3: *Kepler* (bottom left), U (top left), B (top center), V (top right), $R_C$ (bottom center), and $I_C$ (bottom right) light curves for KIC 5282464.

the system as having been observed by HATnet and ASAS, respectively.

I observed KIC 5282464 on four nights with both SARA-KP and SARA-RM. Figure 6.3 shows the data I collected of KIC 5282464 along with the system's *Kepler* data. The night of July 1, 2019, marked the only occasion during my project in which I observed the same object with SARA-KP and SARA-RM simultaneously. These overlapping observations (which occurred at the dimmer maximum) allowed me to calibrate my data better. I conducted observations using the Johnson U filter on November 11, 2019, covering the primary eclipse. However, high winds at ORM that night reduced data quality in all filters, affecting the U observations the most due to the longer exposure time (~4 times the B exposure times). It proved problematic to combine data from all nights due to the bad pixel columns

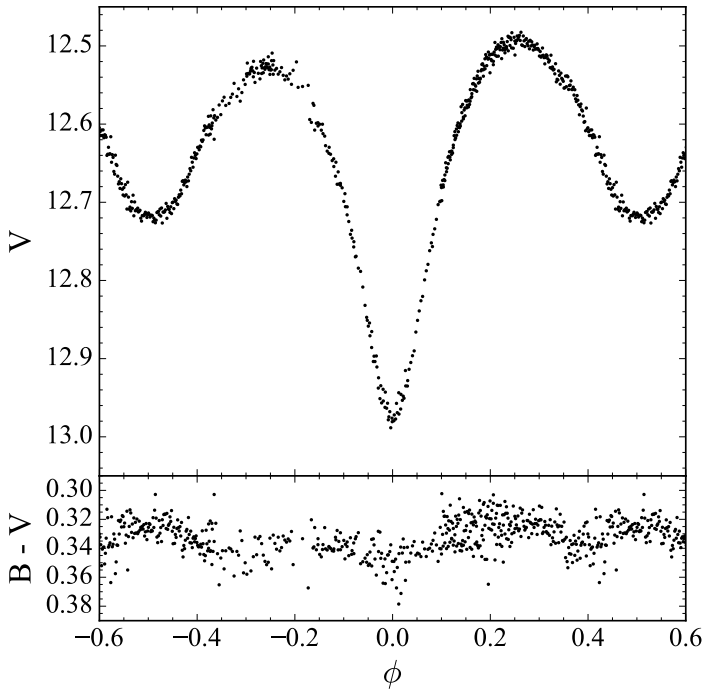


Figure 6.4: Apparent V magnitude and B − V color curve for KIC 5282464.

on SARA-RM's camera negatively affecting the photometry in combination with data scatter of unknown origin. I eventually succeeded in integrating the data from all nights by using the check star as the comparison star during periods that the comparison star proved unsuitable.

I classified KIC 5282464 as a $\beta$ Lyrae-type system displaying partial eclipses. I found the following OESs from my observations: $\mathrm{OES_B} = 0.053 \pm 0.003$, $\mathrm{OES_V} = 0.034 \pm 0.004$, $\mathrm{OES_{R_C}} = 0.027 \pm 0.002$, and $\mathrm{OES_{I_C}} = 0.044 \pm 0.003$, all in units of magnitude. The system's B − V color index is $0.33 \pm 0.02$ during the secondary eclipse, corresponding to $T_1 \approx 7{,}063^{+101}_{-99}$ K. Figure 6.4 shows KIC 5282464's B − V color curve using the magnitudes given in Table 6.2.

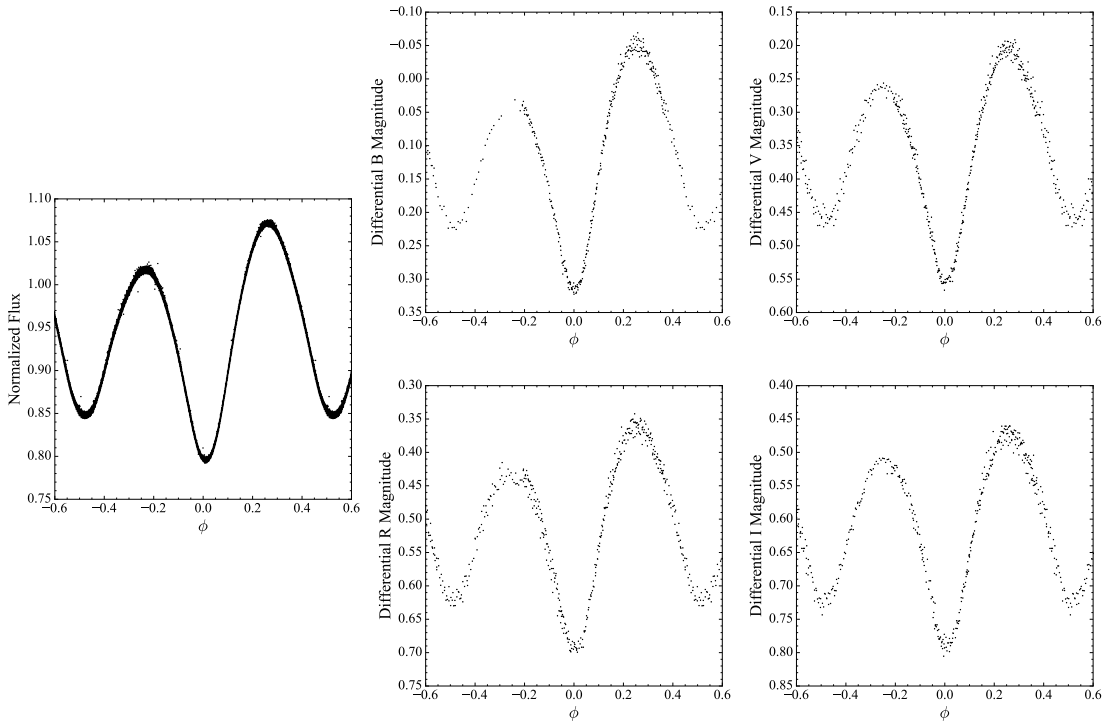


Figure 6.5: *Kepler* (left), B (top center), V (top right), $R_C$ (bottom center), and $I_C$ (bottom right) light curves for KIC 6223646.

### 6.1.3 KIC 6223646

KIC 6223646 has an orbital period of 0.365 d and a $\mu$ of 0.84, suggesting a strongly overcontact system or ellipsoidal variable. The KEBC has a single flag for the system (UNC), indicating that they are uncertain that the system is an eclipsing binary. The system's *Kepler* data shows an OES of 0.054, a primary eclipse depth of 0.204, and an SVP of 0.005. *Gaia* reports a temperature of $7{,}907^{+148}_{-461}$ K for the system, while the *Gaia* color indicates a spectral type of F2 V. Frasca et al. (2016) assigned a spectral type of B9 III to KIC 6223646 based on optical spectra, although the presence of a giant is inconsistent with the system's 0.365-day orbital period. KIC 6223646's *Gaia* EDR3 data implies a luminosity of $4.60 \pm 0.16$ $L_\odot$. Pigulski et al. (2009) cites the system as having been observed by ASAS.

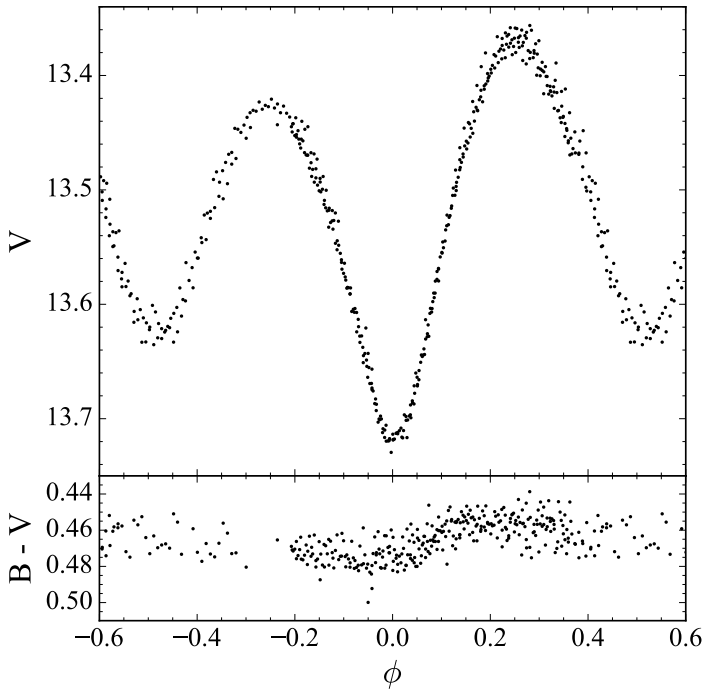


Figure 6.6: Apparent V magnitude and B − V color curve for KIC 6223646.

I observed KIC 6223646 on seven nights with both SARA-KP and SARA-RM, although observations at KPNO were hampered by errant moths in June 2019 and an offset of unknown origin affecting the B-filter data in October 2019. Ultimately, the only SARA-KP data on KIC 6223646 I used was the $VR_CI_C$ data from October 8, 2019, while I used all data collected with SARA-RM. Figure 6.5 shows the data I collected of KIC 6223646 along with the system's *Kepler* data.

I classified KIC 6223646 as a W Ursae Majoris-type system displaying partial eclipses. I found the following OESs from my observations: $OES_B = 0.087 \pm 0.003$, $OES_V = 0.057 \pm 0.004$, $OES_{R_C} = 0.071 \pm 0.004$, and $OES_{I_C} = 0.042 \pm 0.004$. The system's B − V color index is $0.46 \pm 0.05$ during the secondary eclipse, corresponding to $T_1 \approx 6{,}453^{+225}_{-213}$ K. Figure 6.6 shows KIC 6223646's B − V color curve.

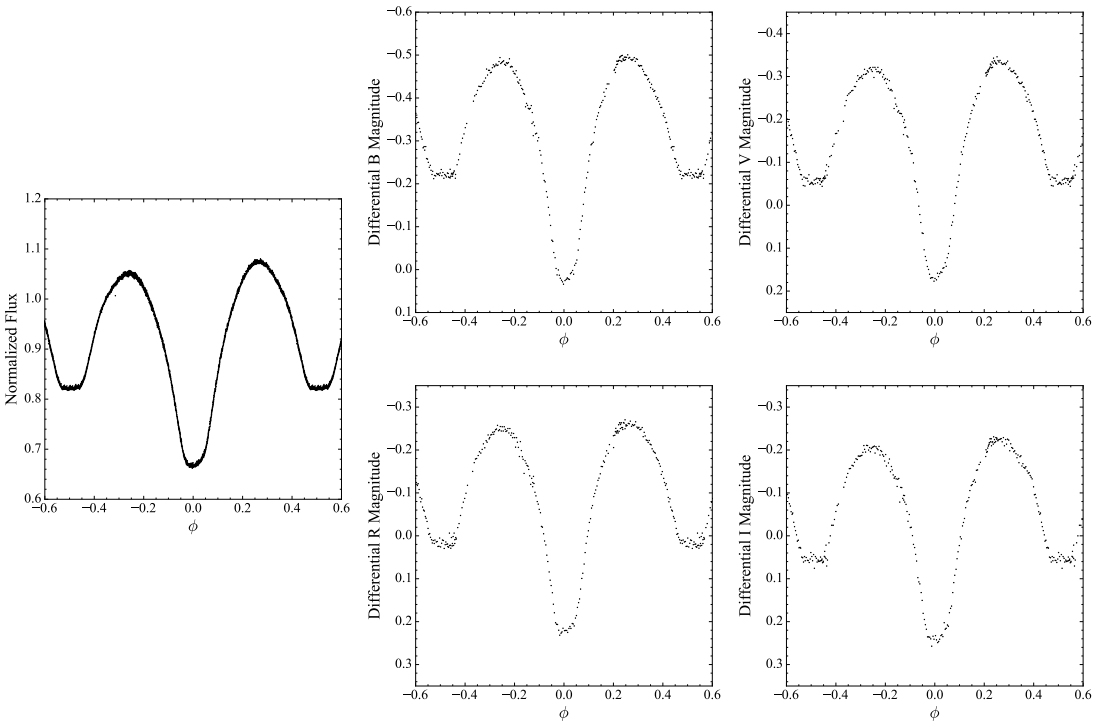


Figure 6.7: *Kepler* (left), B (top center), V (top right), $R_C$ (bottom center), and $I_C$ (bottom right) light curves for KIC 8696327.

### 6.1.4 KIC 8696327

KIC 8696327 has an orbital period of 0.879 d and a $\mu$ of 0.75, suggesting a semi-detached or overcontact system. The KEBC has a single flag for the system (FB), indicating that its light curve shows flat bottoms indicative of total eclipses. The system's *Kepler* data shows an OES of 0.023, a primary eclipse depth of 0.333, and an SVP of 0.017. *Gaia* reports a temperature of $6{,}144^{+499}_{-269}$ K for the system, while the *Gaia* color indicates a spectral type of F9 V. KIC 8696327's *Gaia* EDR3 data implies a luminosity of $9.41^{+1.05}_{-0.98}$ $L_\odot$. KIC 8696327's period of 0.8792289 d (75,965 s) is 43.028 times the *Kepler* long-cadence integration time of 1,765.5 s. This near-resonance introduced an oscillatory pattern into KIC 8696327's *Kepler* light curve similar to the one in KIC 8248967's (Figure 2.8).

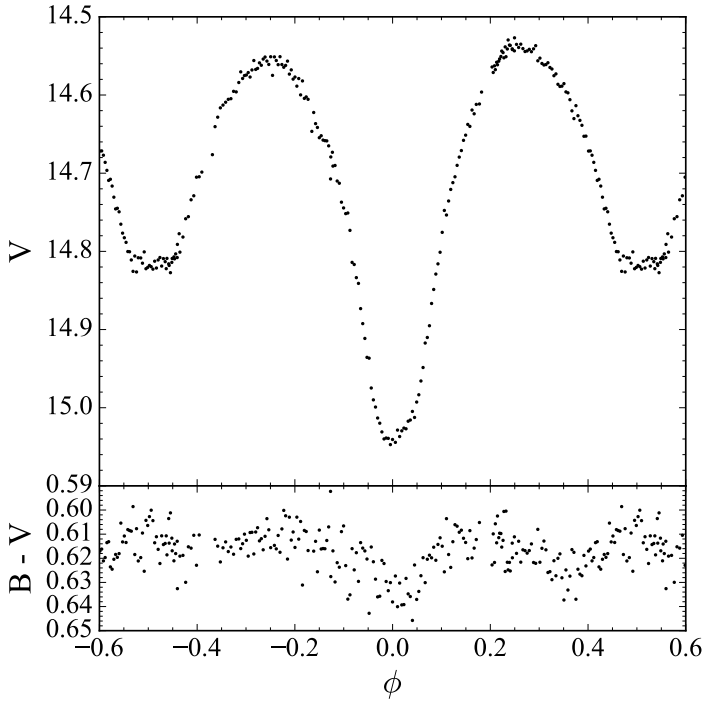


Figure 6.8: Apparent V magnitude and B − V color curve for KIC 8696327.

I observed KIC 8696327 on three nights with SARA-RM. These three nights occurred over only ten days, and I had no difficulty combining the data into a high-quality light curve. Figure 6.7 shows the data I collected of KIC 8696327 along with the system's *Kepler* data. My data does not show the oscillatory pattern seen in the *Kepler* data.

I classified KIC 8696327 as a $\beta$ Lyrae-type system displaying total eclipses with asymmetric minima. I found the following OESs from my observations: $\mathrm{OES_B} = 0.011 \pm 0.004$, $\mathrm{OES_V} = 0.018 \pm 0.004$, $\mathrm{OES_{R_C}} = 0.012 \pm 0.004$, and $\mathrm{OES_{I_C}} = 0.019 \pm 0.006$. The system's B − V color index is $0.61 \pm 0.04$ during the secondary eclipse, corresponding to $T_1 \approx 5{,}855^{+138}_{-147}$ K. Figure 6.8 shows KIC 8696327's B − V color curve.

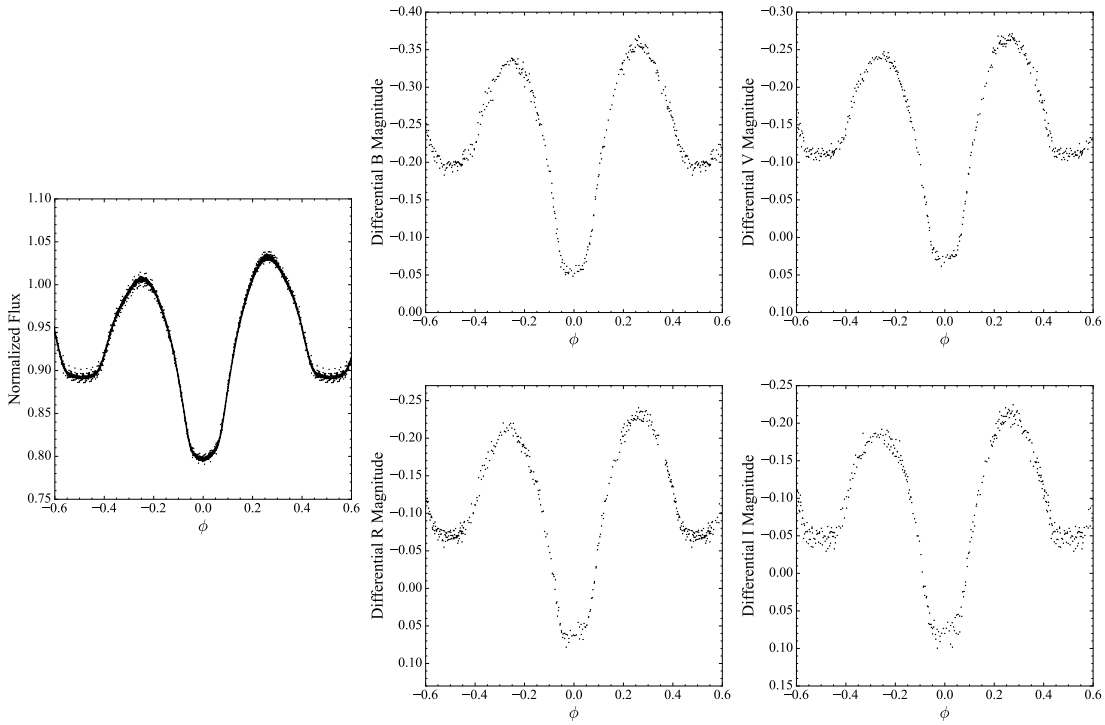


Figure 6.9: *Kepler* (left), B (top center), V (top right), $R_C$ (bottom center), and $I_C$ (bottom right) light curves for KIC 8822555.

### 6.1.5 KIC 8822555

KIC 8822555 has an orbital period of 0.853 d and a $\mu$ of 0.80, suggesting an overcontact system. The KEBC has a single flag for the system (FB), indicating that its light curve shows flat bottoms indicative of total eclipses. The system's *Kepler* data shows an OES of 0.025, a primary eclipse depth of 0.203, and an SVP of 0.006. *Gaia* reports a temperature of $6{,}185^{+293}_{-612}$ K for the system, while the *Gaia* color indicates a spectral type of F9 V. KIC 8822555's *Gaia* EDR3 data implies a luminosity of $9.03^{+0.71}_{-0.70}$ $L_\odot$.

I observed KIC 8822555 on six nights using SARA-RM. I was able to combine all of the data I collected on KIC 8822555 without issue. Figure 6.9 shows the data I collected of KIC 8822555 along with the system's *Kepler* data.

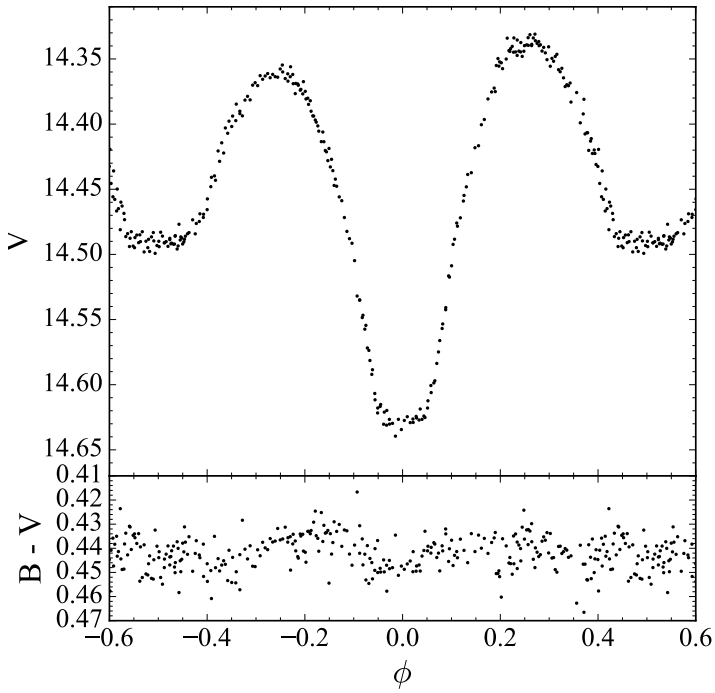


Figure 6.10: Apparent V magnitude and B − V color curve for KIC 8822555.

I classified KIC 8822555 as a $\beta$ Lyrae-type system displaying total eclipses with asymmetric minima. I found the following OESs from my observations: $\mathrm{OES_B} = 0.022 \pm 0.003$, $\mathrm{OES_V} = 0.023 \pm 0.003$, $\mathrm{OES_{R_C}} = 0.016 \pm 0.004$, and $\mathrm{OES_{I_C}} = 0.025 \pm 0.004$. The system's B − V color index is $0.45 \pm 0.02$ during the secondary eclipse, corresponding to $T_1 \approx 6{,}496^{+91}_{-87}$ K. Figure 6.10 shows KIC 8822555's B − V color curve.

### 6.1.6 KIC 9164694

KIC 9164694 has an orbital period of 1.111 d and a $\mu$ of 0.75, suggesting a semi-detached or overcontact system. The KEBC has a single flag for the system (FB), indicating that its light curve shows flat bottoms indicative of total eclipses. The system's *Kepler* data shows an OES of 0.050, a primary eclipse depth of 0.277, and

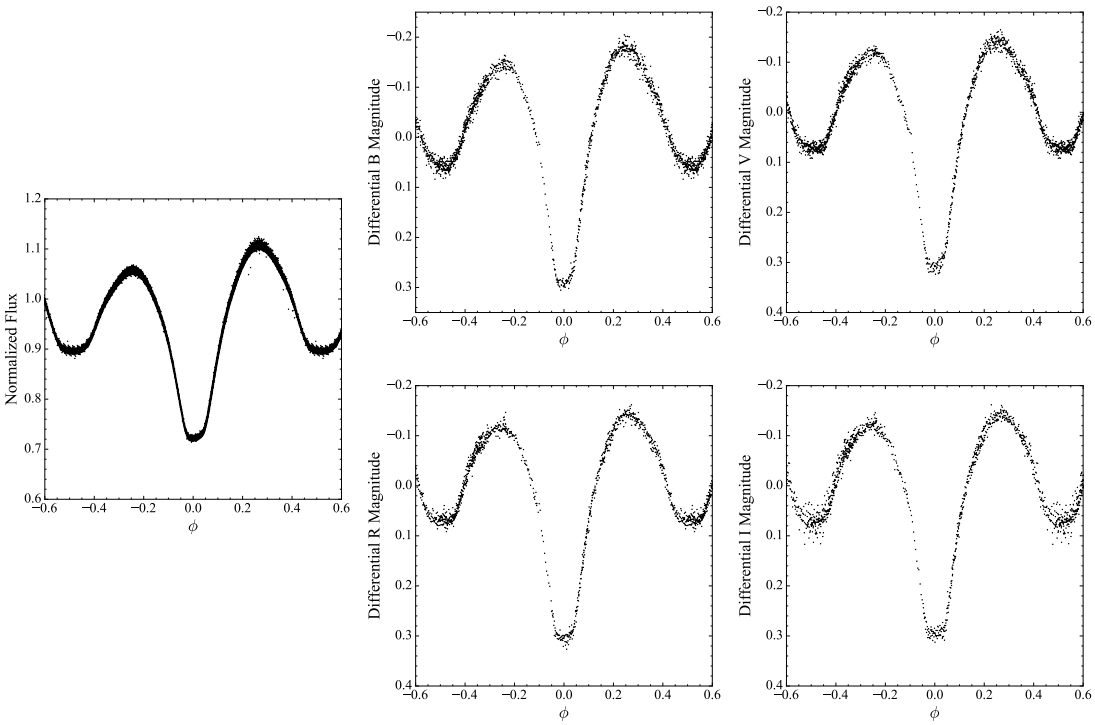


Figure 6.11: *Kepler* (left), B (top center), V (top right), $R_C$ (bottom center), and $I_C$ (bottom right) light curves for KIC 9164694.

an SVP of 0.007. *Gaia* reports a temperature of $6{,}113^{+210}_{-520}$ K for the system, while the *Gaia* color indicates a spectral type of F9.5 V. KIC 9164694's *Gaia* EDR3 data implies a luminosity of $12.12^{+0.88}_{-0.79}$ $L_\odot$. Davenport (2016) cites the system as having flares, while Kobulnicky et al. (2022) describes spectroscopic observations and binary modeling of KIC 9164694. I discuss Kobulnicky et al. (2022) and KIC 9164694's model in Section 3.6.

I observed KIC 9164694 on eleven nights using both SARA-KP and SARA-RM. I considered KIC 9164694 my highest priority target due to its strongly asymmetric minimum, and the long orbital period of 1.111 d required numerous nights to completely cover the light curve. I had no issues combining the data into a single light curve. Figure 6.11 shows the data I collected of KIC 9164694 along

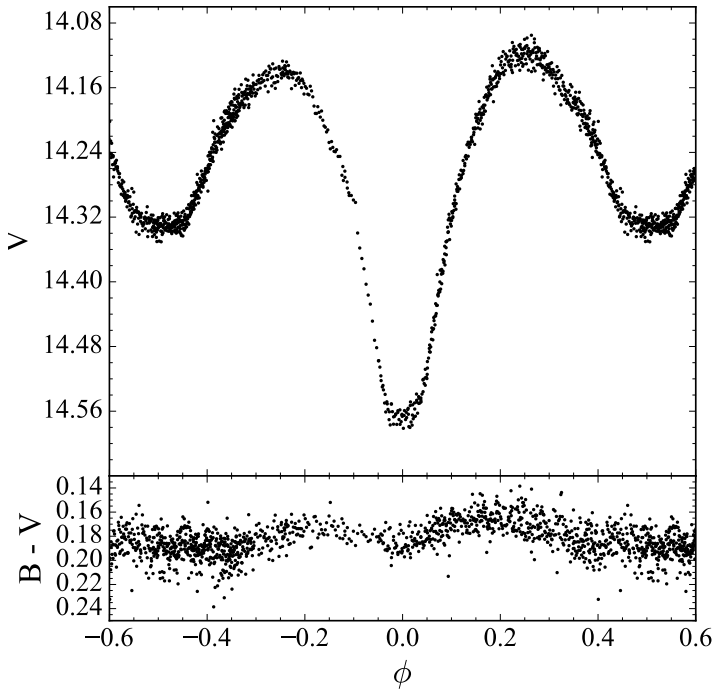


Figure 6.12: Apparent V magnitude and B − V color curve for KIC 9164694.

with the system's *Kepler* data.

I classified KIC 9164694 as a $\beta$ Lyrae-type system displaying total eclipses with asymmetric minima. I found the following OESs from my observations: $\mathrm{OES_B} = 0.032 \pm 0.003$, $\mathrm{OES_V} = 0.025 \pm 0.003$, $\mathrm{OES_{R_C}} = 0.027 \pm 0.003$, and $\mathrm{OES_{I_C}} = 0.021 \pm 0.003$. The system's B − V color index is $0.19 \pm 0.01$ during the secondary eclipse, corresponding to $T_1 \approx 7{,}825^{+61}_{-59}$ K. Figure 6.12 shows KIC 9164694's B − V color curve.

### 6.1.7 KIC 10861842

KIC 10861842 has an orbital period of 0.472 d and a $\mu$ of 0.76, suggesting a semi-detached or overcontact system. The KEBC has a single flag for the system (FB), indicating that its light curve shows flat bottoms indicative of total eclipses. The

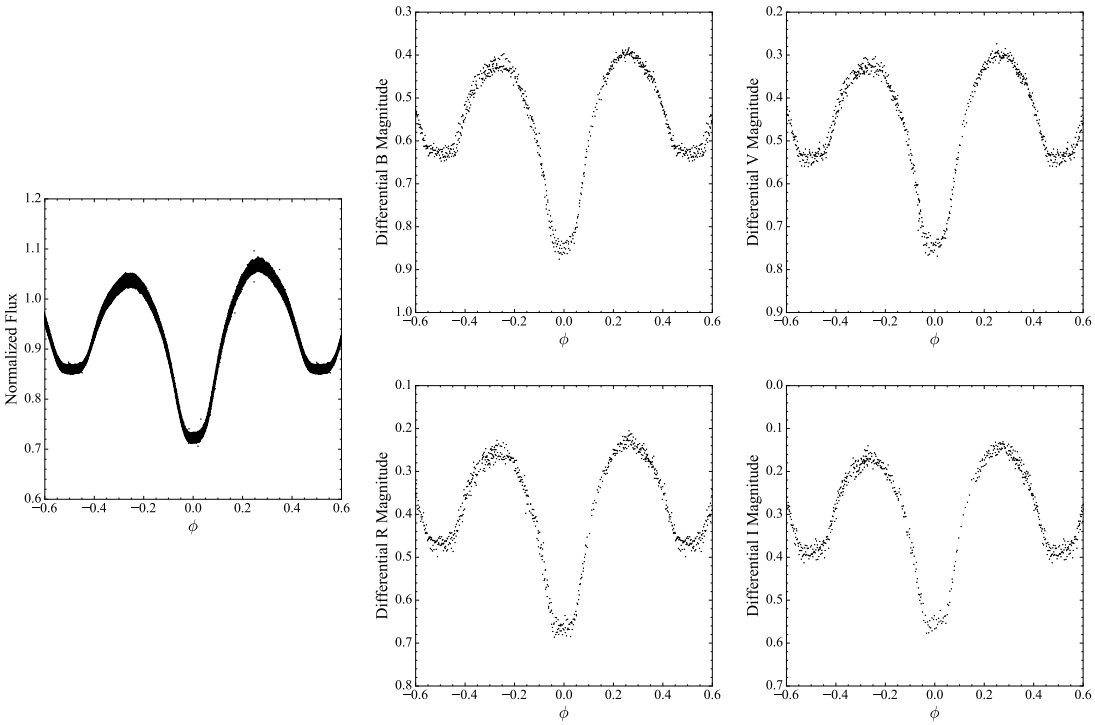


Figure 6.13: *Kepler* (left), B (top center), V (top right), $R_C$ (bottom center), and $I_C$ (bottom right) light curves for KIC 10861842.

system's *Kepler* data shows an OES of 0.029, a primary eclipse depth of 0.276, and an SVP of 0.013. *Gaia* reports a temperature of $6{,}955^{+246}_{-256}$ K for the system, while the *Gaia* color indicates a spectral type of F5 V. Frasca et al. (2016) assigned a spectral type of F1 V to KIC 10861842 based on optical spectra. KIC 10861842's *Gaia* EDR3 data implies a luminosity of $5.14^{+0.20}_{-0.21}$ $L_\odot$. Davenport (2016) cites the system as having flares, while Hartman et al. (2004) cites the system as having been observed by HATnet.

I observed KIC 10861842 on seven nights using both SARA-KP and SARA-RM. There was an offset in the data taken at SARA-KP on May 30, 2019, which I have attributed to partial occlusion by the dome based on the star shapes in the images. There were also some irregularities in the $I_C$ data taken at SARA-

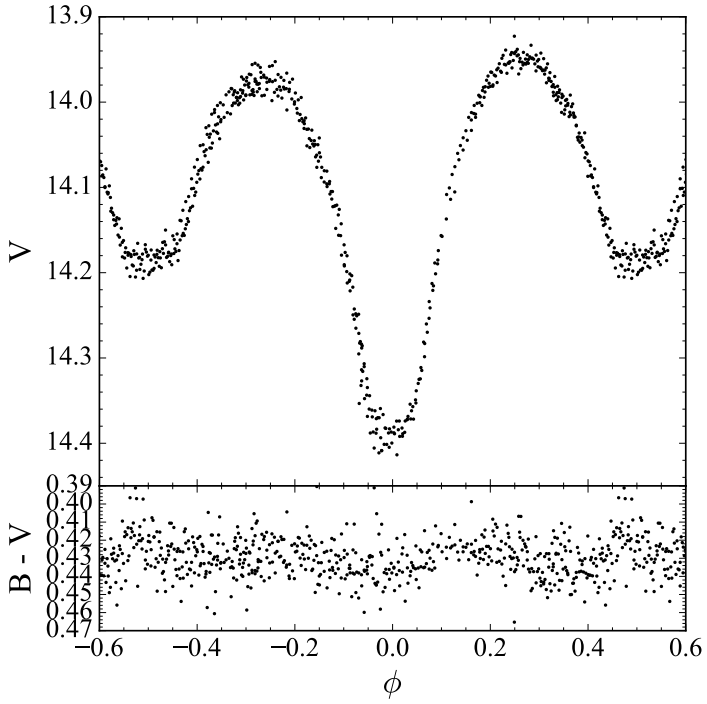


Figure 6.14: Apparent V magnitude and B − V color curve for KIC 10861842.

RM on May 5, 2019, so I discarded the $I_C$ data from that night. I integrated the remaining data into a single light curve. Figure 6.13 shows the data I collected of KIC 10861842 along with the system's *Kepler* data.

I classified KIC 10861842 as a $\beta$ Lyrae-type system displaying total eclipses with asymmetric minima. I found the following OESs from my observations: $OES_B = 0.024 \pm 0.003$, $OES_V = 0.024 \pm 0.003$, $OES_{R_C} = 0.020 \pm 0.003$, and $OES_{I_C} = 0.027 \pm 0.004$. The system's B − V color index is $0.42 \pm 0.16$ during the secondary eclipse, corresponding to $T_1 \approx 6{,}632^{+794}_{-668}$ K. Figure 6.14 shows KIC 10861842's B − V color curve.

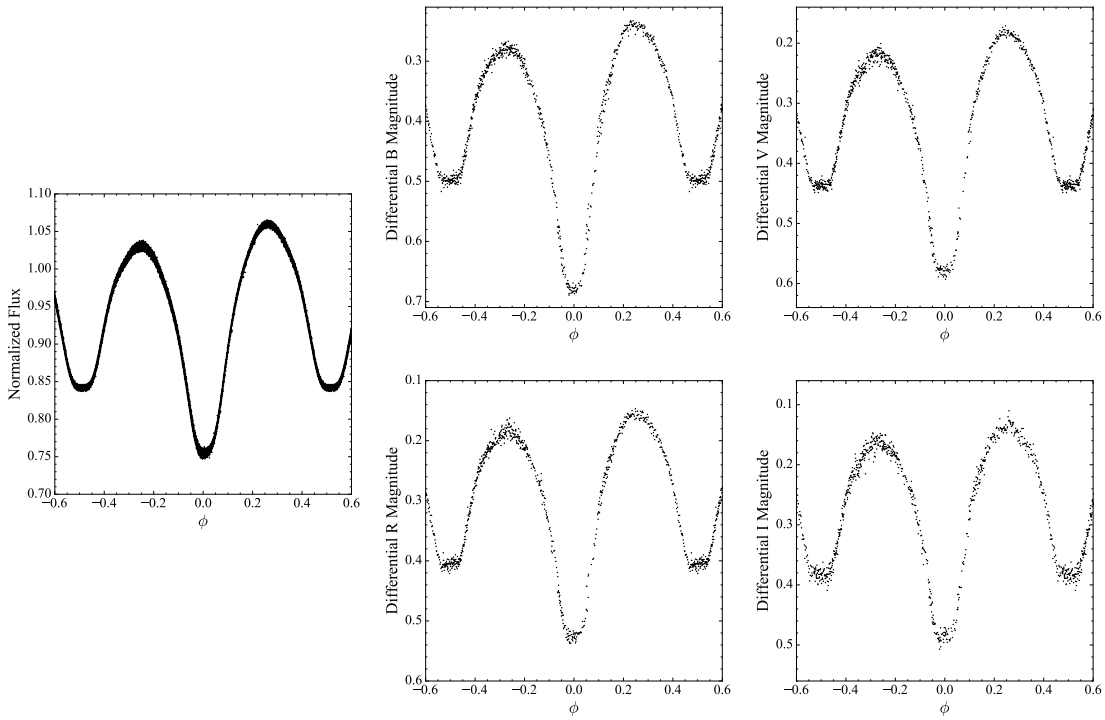


Figure 6.15: *Kepler* (left), B (top center), V (top right), $R_C$ (bottom center), and $I_C$ (bottom right) light curves for KIC 11924311.

### 6.1.8 KIC 11924311

KIC 11924311 has an orbital period of 0.455 d and a $\mu$ of 0.77, suggesting a semi-detached or overcontact system. The KEBC has two flags for the system (FB and TM), indicating that its light curve shows flat bottoms indicative of total eclipses and that its ETV indicates a third body. The system's *Kepler* data shows an OES of 0.030, a primary eclipse depth of 0.246, and an SVP of 0.007. *Gaia* reports a temperature of $6{,}515^{+414}_{-355}$ K for the system, while the *Gaia* color indicates a spectral type of F8 V. Frasca et al. (2016) assigned a spectral type of F3/F5 V to KIC 11924311 based on optical spectra. KIC 11924311's *Gaia* EDR3 data implies a luminosity of $5.24^{+5.99}_{-2.66}$ $L_\odot$, with the large error resulting from the system's poorly constrained distance. Kouzuma (2018), Devor et al. (2008), and Pigulski et al.

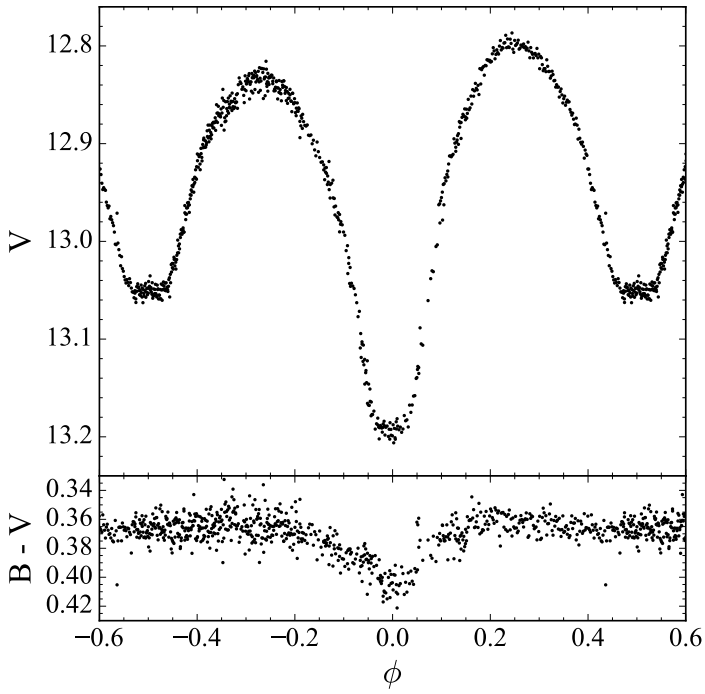


Figure 6.16: Apparent V magnitude and B − V color curve for KIC 11924311.

(2009) cite the system as undergoing mass transfer and having been observed by TrES and ASAS, respectively.

I observed KIC 11924311 on six nights using both SARA-KP and SARA-RM. Moth issues in late-May to early-June 2019 along with a phase offset of unknown origin on May 10, 2019, caused me to discard all SARA-KP data and use only the SARA-RM data. My suspicion is that the period of the system has changed since the *Kepler* observations due to the mass transfer Kouzuma (2018) described, causing the phase offset. Increasing the period by ~2.5 s over the *Kepler* period eliminates the offset, and such a period increase is within an order of magnitude of the $\dot{P} = 6.822 \times 10^{-7}$ $\frac{\mathrm{d}}{\mathrm{y}}$ value Kouzuma (2018) gave. Figure 6.15 shows the data I collected of KIC 11924311 along with the system's *Kepler* data.

I classified KIC 11924311 as a $\beta$ Lyrae-type system displaying total eclipses with

asymmetric minima. I found the following OESs from my observations: $OES_B = 0.039 \pm 0.003$, $OES_V = 0.035 \pm 0.002$, $OES_{R_C} = 0.027 \pm 0.002$, and $OES_{I_C} = 0.027 \pm 0.004$. The system's B − V color index is $0.37 \pm 0.08$ during the secondary eclipse, corresponding to $T_1 \approx 6{,}867^{+400}_{-371}$ K. Figure 6.16 shows KIC 11924311's B − V color curve.

## 6.2 The Varying: KICs 7433513 and 7885570

Unlike the other systems I observed, KICs 7433513 and 7885570 show rapid changes in their light curves. This required altering my observing approach, as I had to conduct my observations within a small time range to ensure the systems did not change appreciably.

### 6.2.1 KIC 7433513

KIC 7433513 has an orbital period of 0.701 d and a $\mu$ of 0.70, suggesting a semi-detached system. The KEBC has two flags for the system (FB and TI), indicating that its light curve shows flat bottoms indicative of total eclipses and that its ETV shows a pattern that is neither sinusoidal or parabolic. The system's *Kepler* data shows an OES of 0.035, a primary eclipse depth of 0.110, and an SVP of 0.080. *Gaia* reports a temperature of $5{,}337^{+99}_{-305}$ K for the system, while the *Gaia* color indicates a spectral type of K0 V. KIC 7433513's *Gaia* EDR3 data implies a luminosity of $1.32^{+0.12}_{-0.09}$ $L_\odot$.

I observed KIC 7433513 on four nights using both SARA-KP and SARA-RM. I conducted these observations over one week in September 2018. KIC 7433513's faint magnitude (V $\approx$ 15, the faintest system I observed) meant my data had

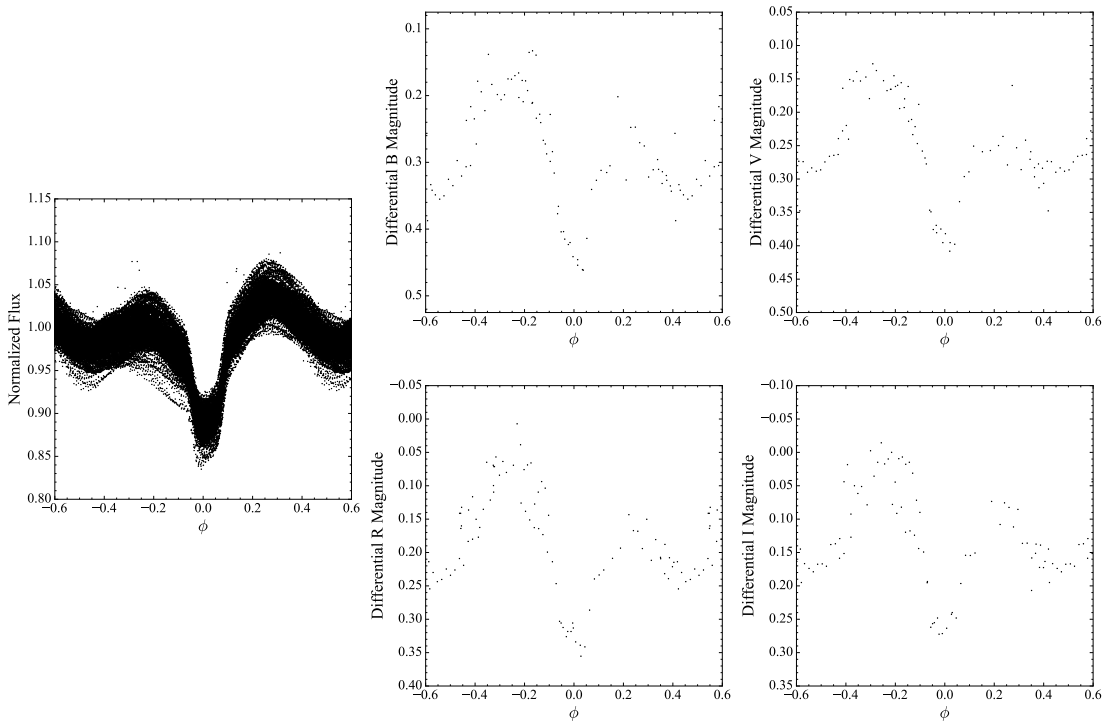


Figure 6.17: *Kepler* (left), B (top center), V (top right), $R_C$ (bottom center), and $I_C$ (bottom right) light curves for KIC 7433513.

poor time resolution and a significant amount of scatter. The core sample system KIC 7516345 lies only 3' from KIC 7433513, but it is considerably brighter ($V \approx 12$) and is saturated in my images as a result. I was able to combine all data I took into a single light curve. Figure 6.15 shows the data I collected of KIC 7433513 along with the system's *Kepler* data.

I classified KIC 7433513 as an Algol-type system displaying total eclipses with temporal variation and asymmetric minima. I found the following OESs from my observations: $OES_B = -0.072 \pm 0.011$, $OES_V = -0.077 \pm 0.010$, $OES_{R_C} = -0.100 \pm 0.013$, and $OES_{I_C} = -0.079 \pm 0.015$. The system's $B - V$ color index is $0.73 \pm 0.09$ during the secondary eclipse, corresponding to $T_1 \approx 5{,}470^{+281}_{-239}$ K. Figure 6.18 shows KIC 7433513's $B - V$ color curve, while Figure 5.2 shows ten-

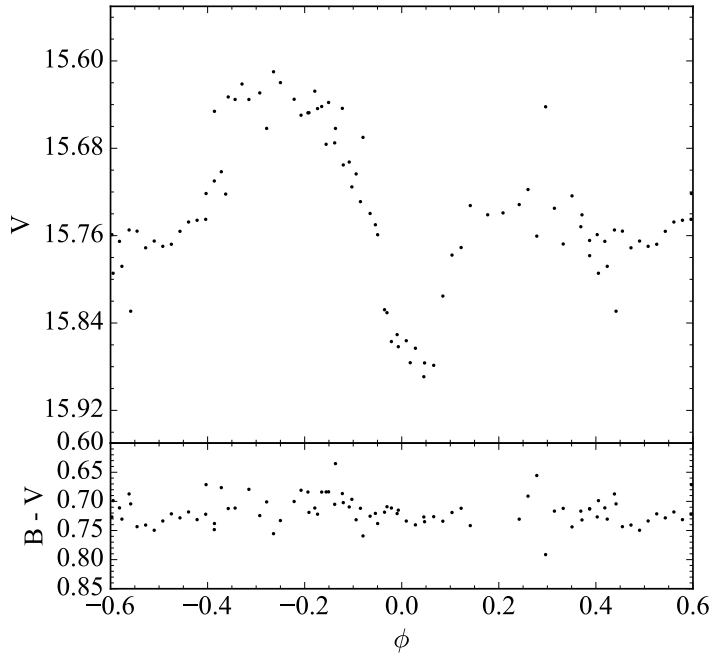


Figure 6.18: Apparent V magnitude and B − V color curve for KIC 7433513.

day intervals of KIC 7433513's *Kepler* data separated by several months. The apparent V magnitudes Figure 6.18 lists are over half a magnitude fainter than the APASS V magnitude Table 6.2 gives for KIC 7433513. I am unsure if the system has dimmed or if there was an error with my data processing.

### 6.2.2 KIC 7885570

KIC 7885570 has an orbital period of 1.729 d and a $\mu$ of 0.58, suggesting a detached or semi-detached system. The KEBC has two flags for the system (FB and TI), indicating that its light curve shows flat bottoms indicative of total eclipses and that its ETV shows a pattern that is neither sinusoidal or parabolic. The system's *Kepler* data shows an OES of 0.009, a primary eclipse depth of 0.147, and an SVP of 0.088. *Gaia* reports a temperature of $5{,}455^{+84}_{-57}$ K for the system, while

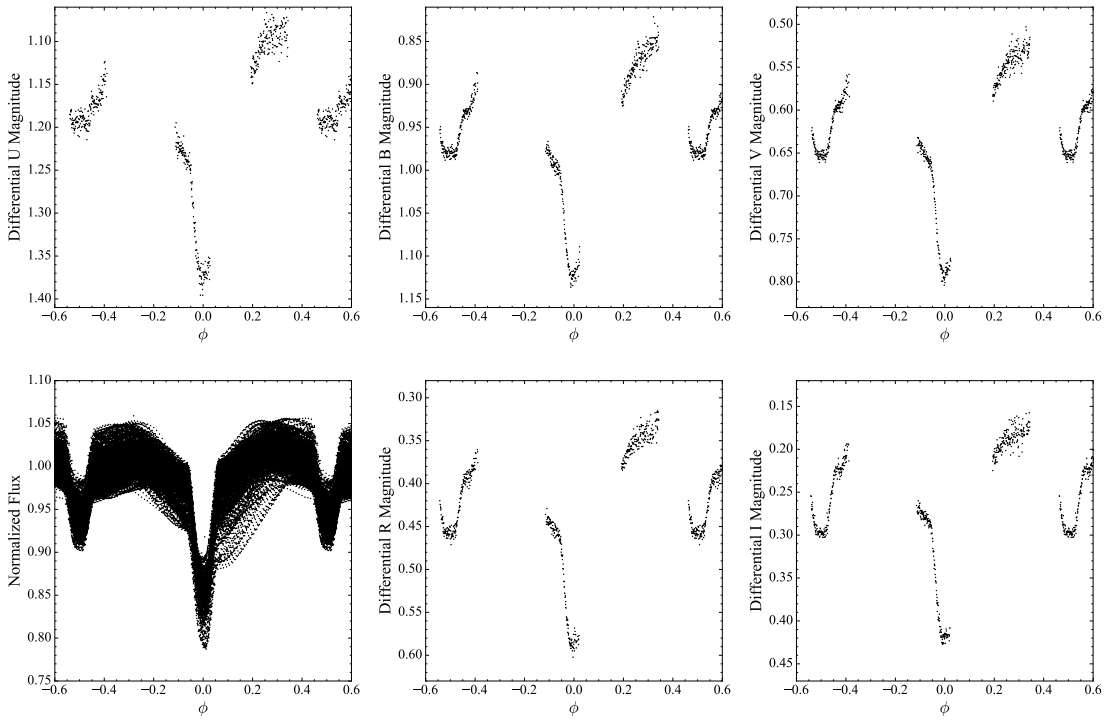


Figure 6.19: *Kepler* (bottom left), U (top left), B (top center), V (top right), $R_C$ (bottom center), and $I_C$ (bottom right) light curves for KIC 7885570.

the *Gaia* color indicates a spectral type of G7 V. Frasca et al. (2016) assigned a spectral type of G0 IV to KIC 11924311 based on optical spectra. KIC 7885570's *Gaia* EDR3 data implies a luminosity of $4.12^{+0.09}_{-0.08}$ $L_\odot$. Balona (2015) and Gao et al. (2016) cite the system as having flares, Davenport (2016) cites the system as having starspots, and Pigulski et al. (2009) cites the system as having been observed by ASAS. Finally, Kunt & Dal (2017) discusses KIC 7885570 in detail.

I observed KIC 7885570 on six nights using both SARA-KP and SARA-RM. Unfortunately, the weather at KPNO was uncooperative, with high humidity and intermittent clouds throughout each night. These environmental factors introduced a significant offset in the SARA-KP data. As a result, I discarded all SARA-KP data and used only the SARA-RM data. Figure 6.19 shows the data I collected of

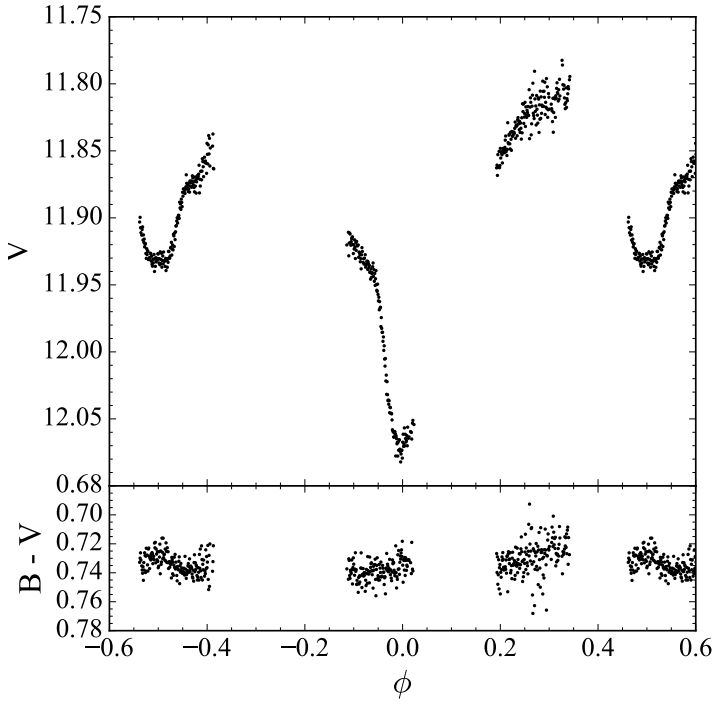


Figure 6.20: Apparent V magnitude and B − V color curve for KIC 7885570.

KIC 7885570 along with the system's *Kepler* data.

I classified KIC 7433513 as an Algol-type system displaying total eclipses with temporal variation and asymmetric minima. I could not determine an OES due to the lack of data at $\phi = -0.25$. The system's B − V color index is $0.73 \pm 0.01$ during the secondary eclipse, corresponding to $T_1 \approx 5{,}470^{+29}_{-28}$ K. Figure 6.20 shows KIC 7885570's B − V color curve, while Figure 6.21 shows ten-day intervals of KIC 7885570's *Kepler* data separated by several months.

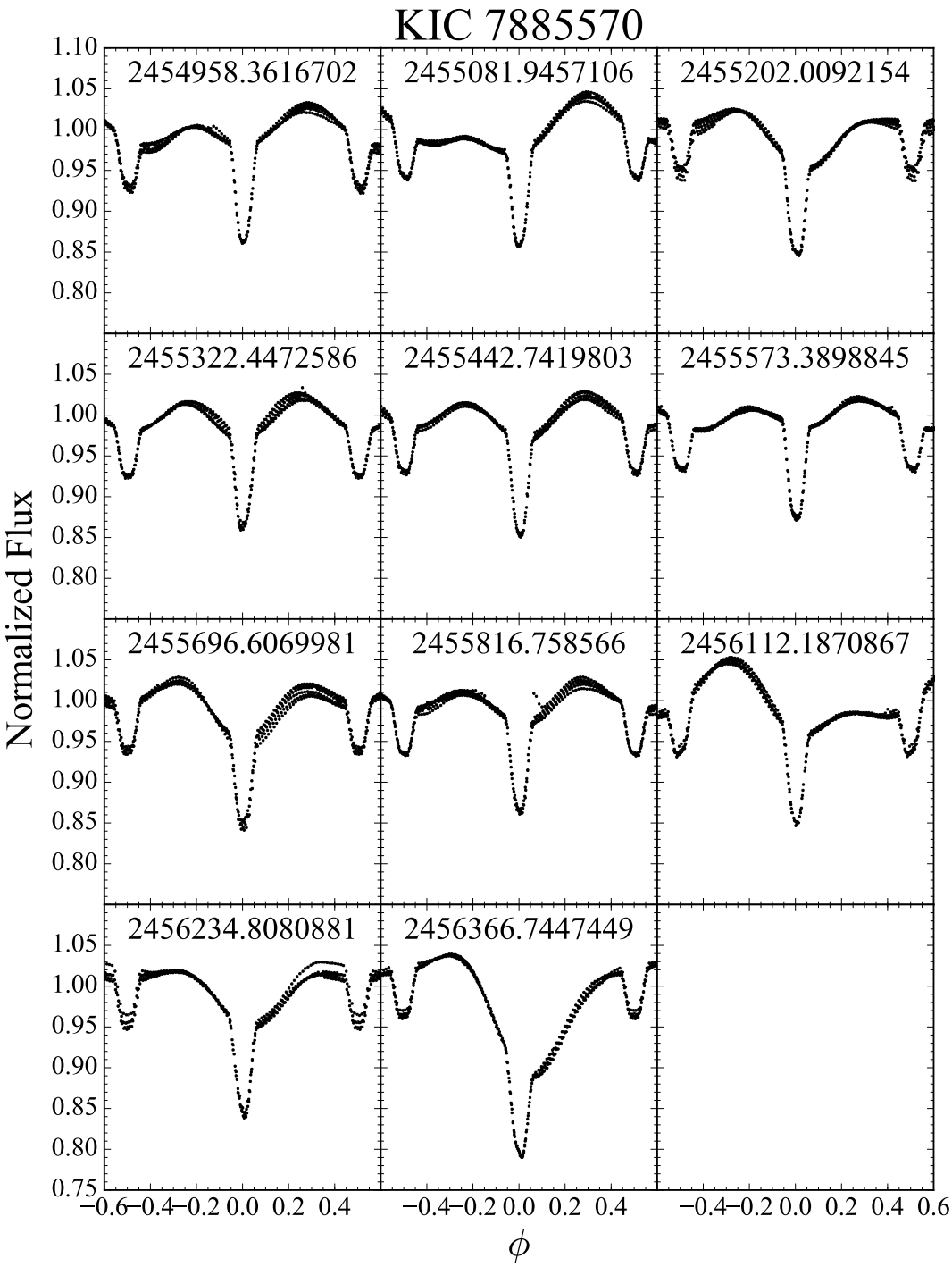


Figure 6.21: Ten-day slices of KIC 7885570's phased *Kepler* data separated by several months (similar to Figure 5.2 for KIC 7433513), showing the significant changes in the light curve over time. The number in each time slice's subplot is the BJD for the midpoint of that time slice.

# Chapter 7

# Conclusion

I have studied the characteristics of 258 O'Connell effect binaries observed by the *Kepler* space telescope to better understand the O'Connell effect's causes. My study supports several of the conclusions O51 and D84 made regarding the O'Connell effect, most prominently that the O'Connell effect likely has multiple causes and that the O'Connell effect is strongly correlated with interacting binaries. My study also found several new trends in O'Connell effect systems, including that there is a dearth of O'Connell effect systems with small eclipses, that O'Connell effect systems have the deeper eclipses than similar-looking systems, and that there are fundamental differences between positive and negative O'Connell effect systems. I will detail my study's main conclusions in this chapter. I will also discuss future opportunities to expand upon my study, including using *Gaia* DR3 data, the TESS Eclipsing Binary Catalog, and binary modeling of the systems I have observed.

## 7.1 Summary of Results

My results can be broadly broken into three categories: the distributions, the trends, and the correlations. In this section, I summarize the primary results of each category for the varied samples I studied. I will also discuss how these results intertwine with the larger picture of the O'Connell effect phenomenon. First, however, I present a summary of the core sample.

### 7.1.1 The Core Sample

I compiled the core sample by measuring the average OES of KEBC systems and selecting those with $|\mathrm{OES}| \geq 0.01$ (Section 2.1.3.2). This process found 212 systems, which is large enough to be a statistically significant sample. Additionally, given *Kepler*'s photometric precision and the span of continuous observation, the sample is complete, modulo *Kepler*'s selection function. The two previous major O'Connell effect studies, O51 and D84, each studied samples of approximately 50 systems, or about a quarter of my sample's size. Furthermore, they excluded W Ursae Majoris-type (O51) or overcontact (D84) systems and systems with varying OES (both studies), while my study includes these types of systems.

### 7.1.2 Distributions

Comparing the characteristic distribution between samples allows inferring the similarities and differences between the samples. This inference gives a clearer understanding of the processes that underlie what defines the samples in question. I discuss the distributions of my samples in this section, with an emphasis on how the samples differ.

I find that core sample systems are, on average, less luminous (Figure 4.3) and cooler (Figure 4.6) systems with much shorter periods (Figure 4.4) and larger $\mu$ (Figure 4.9) than eclipsing binaries at large. The first three facts, taken together, suggest that systems displaying a significant O'Connell effect may contain lower mass stars than other eclipsing binaries. It is well-known that stellar multiplicity increases with increasing mass (Duchêne & Kraus 2013), so it is not surprising that there are a significant number of higher-mass eclipsing binaries despite such stars being rarer. The period and $\mu$ (morphology parameter) distributions imply that O'Connell effect binaries have components close to one another, suggesting that the O'Connell effect may be a consequence of the components' proximity to each other.

The marginal sample's distributions are broadly similar to the core sample's. On average, marginal sample systems have shallower eclipses (Figure 4.23), longer periods (Figure 4.26), and are cooler (Figure 4.27) than core sample systems. The marginal sample's $\mu$ distribution also peaks at a smaller $\mu$ than the core sample's (compare Figures 4.9 and 4.30). Overall, these distribution differences imply that marginal sample systems contain components that are more separated and fill less of their critical lobes than core sample systems.

On average, temporally varying systems have much shallower eclipses (Figure 5.4), longer periods (Figure 5.5), and are cooler (Figure 5.6) than other complete sample systems. The temporally varying systems' $\mu$ distribution also peaks at a smaller $\mu$ than the complete sample's (compare Figures 4.9 and 5.7), much like the marginal sample's. In fact, the temporally varying systems' distributions are more similar to the marginal sample's than the core sample's, despite the fact that most temporally varying systems (61%) are in the core sample. Furthermore,

there are relatively few temporally varying systems around the peak of non-varying systems' $\mu$.

As a subset of the temporally varying systems, the temporally varying asymmetric minima systems have similar distributions to the former class (aside from being slightly more luminous). Therefore, I focus on the temporally stable asymmetric minima systems, which have deeper eclipses (Figure 5.15), are more luminous (Figure 5.16), and are hotter (Figure 5.18) than other complete sample systems, including the temporally varying asymmetric minima systems. These facts may indicate that temporally stable systems showing asymmetric minima have more massive components than other O'Connell effect binaries on average.

Finally, on average, the concave-up systems have significantly longer periods (Figure 5.29) and lower temperatures (Figure 5.30) than other complete sample systems. They also have much smaller values of $\mu$ (Figure 5.31). These facts imply that systems with a concave-up region are more detached systems, possibly with either lower-mass main-sequence or evolved components.

### 7.1.3 Trends

Trends in parameter space can reveal relationships between characteristics that the distributions and correlations cannot capture. For instance, trends can reveal subsets of otherwise uniform populations. I discuss the trends I have observed in the varied samples I studied in this section. I do not discuss the concave-up system class here due to its lack of apparent trends.

One trend the core sample shows is that, while both hot and cool systems present a positive O'Connell effect, hot systems displaying a negative O'Connell effect are nearly nonexistent (Figure 4.17's center panel). Another trend is that

$|\text{OES}| \to 0$ as the period increases (Figure 4.17's top-left panel). There is a dearth of systems with a small eclipse depth and a significant O'Connell effect that extends below the criterion cutoff (Figure 4.21's right panel) which cannot be explained as a selection bias. Core sample systems are clustered along an edge on the left side of Figure 4.18's center panel, indicating that core sample systems tend to have the deepest eclipses of systems with similar $\mu$. Finally, systems with $D \geq 0.6$ have $\mu \approx 0.6$ (Figure 4.18's center panel), while the four KEBC systems with $D \geq 0.8$ have a non-negligible O'Connell effect (Figure 4.21's left panel).

The marginal sample lacks many of the trends seen in the core sample. There is no asymmetry between the temperature with regards to having a positive or negative O'Connell effect. Additionally, |OES| does not go to zero as the period increases up to 10 d. Finally, Figure 4.32's center panel shows no clustering on the right side like Figure 4.18's center panel does. The fact that the marginal sample does not show these trends may suggest that those trends underlie the presence of a significant O'Connell effect. The marginal sample does not show any trends that do not appear in the core sample.

Like the marginal sample, the temporally varying systems do not share the same trends as the core sample, despite most temporally varying systems being from the core sample. However, there are several trends involving the SVP (scaled variance parameter) apparent in Figures 5.9 and 5.10. In particular, a large SVP (SVP $\geq$ 0.3) only occurs in ranges of OES ($-0.015 \leq \text{OES} \leq 0.005$), temperature ($4{,}500 \leq T \leq 5{,}100$ K), morphology parameter ($0.48 \leq \mu \leq 0.59$), and primary eclipse depth ($D \leq 0.01$) that are relatively narrow compared to the full distribution of those characteristics. As a result, it seems that the most extreme temporal variation only occurs under rather specific conditions.

Most of the asymmetric minima system class's trends involve the different parameter spaces occupied by the class's two subsets. Like the core sample, systems in the temporally stable subset lie exclusively on the edge in Figure 5.21's center panel, while systems in the temporally variable subset lie largely away from this edge. The differing areas of parameter space occupied by the subsets are particularly noticeable in Figure 5.20's lower-left panel (color versus period), Figure 5.21's center, center-left, and lower-center panels (eclipse depth versus morphology parameter and color and absolute magnitude versus morphology parameter, respectively), and Figure 5.22's upper-left, center-left, and lower-left panels (morphology parameter, eclipse depth, and absolute magnitude versus period, respectively). The different parameter spaces occupied by these subsets imply that these subsets are physically meaningful beyond their difference in SVP. Furthermore, these trends show that the temporally stable subset is more similar to the core sample than the temporally varying subset is.

### 7.1.4 Correlations

Determining how the varied characteristics are correlated in my samples provides insight into what the O'Connell effect's physical causes may be. The correlations also paint a better picture of the properties of O'Connell effect binaries that the distributions alone may not capture. In this section, I focus on the correlations that I believe are the most illuminating about the O'Connell effect and the systems showing it.

I found several correlations for the core sample, but one that proved most illuminating was the correlation between OES and temperature. Kouzuma's (2019) study of starspots implies a positive correlation between these two characteristics

for spotted systems. However, further investigation showed that this correlation was only significant for OES $< 0$, and it had the wrong sign in that regime. Two other illuminating correlations were those between eclipse depth and the morphology parameter (mirroring the trend I mentioned in Section 7.1.3) and between absolute magnitude and period. The latter correlation may suggest that O'Connell effect binaries tend to have the shortest possible periods for a given luminosity.

The marginal sample shows a strong correlation between |OES| and the morphology parameter that the core sample lacks. The marginal sample also has a strong correlation between |OES| and period, and its negative sign implies that |OES| $\rightarrow 0$ as the period increases. However, as I mentioned in Section 7.1.3, this trend does *not* appear in the data, and ODR's best fit for this correlation has a positive slope.

The most illuminating correlation I found in the temporally varying systems is the one between |OES| and eclipse depth. While this correlation appeared in both the core and marginal samples, they were weaker and visual inspection of the data did not support their existence (Figures 4.21 and 4.35). The same cannot be said for the temporally varying systems, and there appears to be a genuine correlation between the two in this subset (Figure 5.11).

I discuss the asymmetric minima systems' correlations by comparing the correlations between the class's two subsets. |OES| and temperature are correlated in both subsets, but while they are similarly strong, they have opposite signs. This fact could mean that the O'Connell effect's causes are not the same in both subsets. |OES| is correlated with period in the temporally stable subset but not the temporally varying subset, which is consistent with an O'Connell effect caused by mass transfer. Finally, while period is strongly correlated with luminosity in both

subsets, it is exceptionally correlated in the temporally stable subset. This very strong correlation may suggest that systems in the temporally stable subset have the shortest period for a system with their luminosity, much like the core sample.

The concave-up system class has a sample size of only seven systems, making it difficult to place much confidence in the correlations listed in Table 5.14. Nevertheless, the almost perfectly monotonic relationship between luminosity and period (Figure 5.34's lower-left panel) is striking. If it is a genuine representation of such systems, one interpretation is that such systems have the shortest period for a given luminosity. However, the fact that their trend lies at longer periods than the trend of the core sample and temporally stable asymmetric systems would imply that concave-up systems contain evolved stars.

## 7.2 The Larger Picture

I found a significant O'Connell effect in 212 eclipsing binaries, with the 211 KEBC systems representing 7.3% of the 2,907 KEBC systems. Considering that the KEBC appears to be complete (Kirk et al. 2016) and nearly unbiased for $K_p \leq 14$ (Wolniewicz et al. 2021) – modulo *Kepler*'s selection function – it is reasonable to conclude that this 7.3% prevalence of a significant O'Connell effect is representative of eclipsing binaries. In other words, a significant O'Connell effect is present in a non-negligible fraction of eclipsing binaries.

The core sample has a moderate correlation between OES and temperature. The dearth of hot ($T \geq 6{,}000$ K) systems displaying a negative O'Connell effect (discussed in Section 4.2.3.1) is a new result of my study. The ratio of positive to negative O'Connell effect systems was remarkably consistent, falling between 65%

and 70% for every sample I studied except the concave-up systems (57%). Since the KEBC should be completely unbiased regarding the OES's sign, this ratio signifies that the OES is positive about twice as often as it is negative. This fact, combined with the previously mentioned difference in temperature distribution, leads me to conclude that the physical causes of the O'Connell effect preferentially induce the maximum following the primary eclipse to be brighter than the one preceding it.

In addition to the correlation between OES and temperature, I found a weaker correlation between OES and period. Kouzuma's (2019) study on starspots predicted both correlations, and their presence could suggest that starspots predominately cause the O'Connell effect. However, looking at positive and negative O'Connell effects separately does not yield the same correlations. Instead, I found correlations that are inconsistent with Kouzuma's (2019) results. Therefore, I draw no conclusion as to whether starspots are a predominant cause of the O'Connell effect, but note that the idea is plausible. At the same time, it should not automatically be assumed that starspots cause the O'Connell effect in a given system without further supporting evidence for starspots.

The core sample distributions Section 4.1.2 gives show that a wide variety of systems exhibit a significant O'Connell effect. Additionally, the O'Connell effect manifests itself quite differently in different systems. Many systems have light curves that vary greatly with time, and the OES varies along with the light curve. Other systems have extremely stable light curves and a stable value of the OES. Chromospheric starspots naturally produce the hemispheric asymmetry necessary to produce the O'Connell effect, but such spots could migrate in the same manner as on the Sun. Balaji et al. (2015) found strong evidence that starspots in close

binary systems do migrate. Therefore, chromospheric spots could explain the O'Connell effect in systems with temporally varying light curves, but they seem ill-suited to explaining the O'Connell effect in systems with very stable light curves.

The four peculiar system classes Chapter 5 discussed are subsets of my sample worth discussing in their own right. Chromospheric activity in the form of flares and starspots can explain the features in most of these systems, including temporal variation, asymmetric minima (when occurring in totally eclipsing systems), KIC 10544976's activity, and concave-up regions. The exception is the temporally stable asymmetric minima systems, which may instead be a result of mass transfer in totally eclipsing systems.

The main, overarching trend I found during this study is that the O'Connell effect and binary interaction are strongly correlated. Binaries with widely separated components do not exhibit the O'Connell effect. Additionally, O'Connell effect binaries almost always show evidence of non-spherical stars, such as ellipticity effects in light curves. As Section 4.2.3.4 discussed, core sample systems appear to have the shortest possible period for a given luminosity, supporting the idea that binary interaction is related to the O'Connell effect. Furthermore, as Section 7.1.2 mentioned, marginal sample systems are farther apart than core sample systems, and they have a smaller OES by definition. Another piece of supporting evidence is W Crucis (Zoła 1996; Pavlovski et al. 2006), an O'Connell effect binary outside of my sample. W Crucis's light curve suggests an interacting, semi-detached system similar to many systems in my sample. However, W Crucis's orbital period is 198.5 d, the longest of an O'Connell effect binary that I am aware of and over 20 times longer than the longest period in my sample. Nevertheless, W Crucis's ellipsoidal variation shows that it has interacting components, reinforcing the

connection between binary interaction and the O'Connell effect.

## 7.3 Future Work

My study only begins to scratch the surface of the O'Connell effect, and significant future paths for expanding my study exist. Some of these paths involve modeling the systems I observed and examining the peculiar system classes in more detail. Other paths involve expanding my work in new directions, such as using tricolor *Gaia* DR3 photometric data of my sample, looking at the subset of systems with $K_p \leq 14$, investigating the *Kepler* eclipsing binaries Bienias et al. (2021) found, and performing my analysis on the recently released TESS Eclipsing Binary Catalog. I will discuss these future directions and more in this section.

### 7.3.1 Observations and System Modeling

The data I presented in Chapter 6 does not represent the final results of my observations. Due to time constraints, I could not correct my data for atmospheric reddening, nor could I transform my data to a standard magnitude system (e.g. Johnson–Cousins UBV$R_C I_C$) using standard field stars. Both processes are required before I can use the data for proper analysis. Therefore, Chapter 6 only provides a preliminary overview of my data.

I intend to use either PHOEBE Legacy or PHOEBE 2 to create models using my data, as described in Section 1.3. While PHOEBE 2 is currently in a less-polished state than PHOEBE Legacy, the former was designed to handle the high-volume *Kepler* data that the latter cannot. Therefore, I could fit my data and the *Kepler* data simultaneously with PHOEBE 2 to provide a better model

for each system. Additionally, Kobulnicky et al. (2022) provides radial velocity data for KIC 9164694 (Section 6.1.6), and other systems may also have available radial velocity data. Utilizing this data would provide absolute parameters for the systems, including mass, size, and luminosity. System modeling therefore has the potential to provide a significantly clearer picture of these systems than currently exists.

### 7.3.2 Peculiar System Classes

In future work, I will further discuss the peculiar system classes from Chapter 5. One aspect I will investigate is if there are any periodicities or quasi-periodicities in the temporally varying systems, which will further characterize the nature of their time variability. I will also try to correct the AP's issues I discussed in Section 2.2.7 by attempting further methods to find the conjunction phase. Finally, I will investigate if the degree of concavity is correlated with any of the characteristics I studied, although I suspect that the small sample size will impede drawing concrete conclusions.

### 7.3.3 *Gaia* Data Release 3

*Gaia* Data Release 3's (DR3; Gaia Collaboration et al. 2022) recent release provides the opportunity to greatly expand the parameter space available for study. According to Gaia Collaboration et al. (2022), *Gaia* DR3 provides photometry on nearly 2.2 million eclipsing binaries, which includes an unknown but presumably large fraction of my sample. *Gaia*'s tricolor photometry allows a temperature estimate for both stars based on the color during the eclipses, which is impossible to do with *Kepler*'s monochromatic data. The tricolor photometry also allows

the exploration of the O'Connell effect's temperature dependence, providing clues critical to understanding the O'Connell effect's physical causes.

### 7.3.4 TESS Eclipsing Binary Catalog

Like *Kepler*, the Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2015) was launched to observe exoplanet transits, and like *Kepler*, it observed many eclipsing binaries in the process. Recently, the team behind the KEBC published the TESS Eclipsing Binary Catalog (TEBC; Prša et al. 2022), which contains 4,584 eclipsing binaries that TESS observed. Applying to the TEBC the same criterion and analysis I applied to the KEBC is a natural next step. However, it is important to consider the differences between *Kepler* and TESS. TESS focuses on brighter systems ($I_C \lesssim 13$) than *Kepler*, which doubtlessly alters the sample's bias. TESS also observes most systems for only 27.4 d, although the results of my study imply that the shorter observing time should not affect O'Connell effect binary studies (aside from reducing sensitivity to temporal variation). On the other hand, TESS's long-cadence exposures are 2 min, which is only twice the length of *Kepler*'s *short*-cadence exposures. Additionally, TESS captures a full-frame image every 30 min (or every 10 min during its extended mission), instead of *Kepler*'s once a month. Therefore, TESS data has much better time resolution than *Kepler* data. Since TESS is still in operation, data will continue to be added to the TEBC in the coming years.

### 7.3.5 Unbiased Subset of the Sample

Wolniewicz et al. (2021) found that the *Kepler* selection function is essentially unbiased for systems with $K_p \leq 14$. Eighty-seven of the 258 complete sample

systems (33.7%) – 73 (34.4%) core sample and 14 (30.4%) marginal sample systems – meet this criterion. Studying an unbiased sample increases the likelihood that the study's results reflect the underlying population, making investigating this subset of my sample worthwhile.

### 7.3.6 Background *Kepler* Eclipsing Binaries

Bienias et al. (2021) found 547 short-period ($P \lesssim 0.5$ d) eclipsing binaries that the KEBC did not include. These systems were not *Kepler* targets themselves, but were instead background stars near observed targets. They are much fainter than KEBC systems, with an average magnitude of 18.2. Nevertheless, examining these systems is critical to ensure the completeness of my sample, particularly in light of the probability that the O'Connell effect is present in such short-period systems. Studying this sample will involve determining which systems meet the $|\mathrm{OES}| \geq 0.1$ criterion and applying the same analysis to them that I applied to the core sample. If the core sample is well-representative of O'Connell effect binaries, the expectation is that the results from this new set of systems would not strongly differ from the results I have presented in this dissertation.

### 7.3.7 Positive and Negative O'Connell Effect Differences

My study shows a clear difference between positive and negative O'Connell effect systems. The two most prominent differences are that a positive O'Connell effect is about twice as common as a negative O'Connell effect and that there are few, if any, negative O'Connell effect systems with $T \geq 6{,}000$ K. I intend to study the differences between these two subsets in further detail. My idea is to break my sample into these two subsets and analyze each subset similarly to the core

sample. By understanding how the distributions and correlations differ between the subsets, I should better understand what causes these systems to differ.

# Bibliography

Abdul-Masih, M., Prša, A., Conroy, K., et al. 2016, AJ, 151, 101, doi: 10.3847/0004-6256/151/4/101

Agresti, A. 2010, *Analysis of Ordinal Categorical Data* (John Wiley & Sons)

Akiba, T., Neugarten, A., Ortmann, C., & Gokhale, V. 2019, JAAVSO, 47, 186

Alard, C., & Lupton, R. H. 1998, ApJ, 503, 325, doi: 10.1086/305984

Alard, C. 2000, A&AS, 144, 363, doi: 10.1051/aas:2000214

Allard, F., Homeier, D., & Freytag, B. 2012, Philosophical Transactions of the Royal Society of London Series A, 370, 2765, doi: 10.1098/rsta.2011.0269

Almeida, L. A., de Almeida, L., Damineli, A., et al. 2019, AJ, 157, 150, doi: 10.3847/1538-3881/ab0963

Almenara, J. M., Alonso, R., Rabus, M., et al. 2012, MNRAS, 420, 3017, doi: 10.1111/j.1365-2966.2011.20175.x

Alonso, R., Brown, T. M., Torres, G., et al. 2004, ApJL, 613, L153, doi: 10.1086/425256

Altman, N. S. 1992, The American Statistician, 46, 175

Alton, K. B., & Wiley, E. O. 2021, JAAVSO, 49, 170

Andrae, R., Fouesneau, M., Creevey, O., et al. 2018, A&A, 616, A8, doi: 10.1051/0004-6361/201732516

Andronov, I. L., & Richter, G. A. 1987, Astronomische Nachrichten, 308, 235, doi: 10.1002/asna.2113080403

Applegate, J. H. 1992, ApJ, 385, 621, doi: 10.1086/170967

Armstrong, D., Pollacco, D., Watson, C. A., et al. 2012, A&A, 545, L4, doi: 10.1051/0004-6361/201219479

Avvakumova, E. A., Malkov, O. Y., & Kniazev, A. Y. 2013, Astronomische Nachrichten, 334, 860, doi: 10.1002/asna.201311942

Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, AJ, 156, 58, doi: 10.3847/1538-3881/aacb21

Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Demleitner, M., & Andrae, R. 2021, AJ, 161, 147, doi: 10.3847/1538-3881/abd806

Balaji, B., Croll, B., Levine, A. M., & Rappaport, S. 2015, MNRAS, 448, 429, doi: 10.1093/mnras/stv031

Balona, L. A. 2014, MNRAS, 437, 1476, doi: 10.1093/mnras/stt1981

Balona, L. A. 2015, MNRAS, 447, 2714, doi: 10.1093/mnras/stu2651

Basri, G., Walkowicz, L. M., Batalha, N., et al. 2010, ApJL, 713, L155, doi: 10.1088/2041-8205/713/2/L155

Batalha, N. M., Borucki, W. J., Koch, D. G., et al. 2010, ApJL, 713, L109, doi: 10.1088/2041-8205/713/2/L109

Bate, M. R. 2019, MNRAS, 484, 2341, doi: 10.1093/mnras/stz103

Bellet, A., Habrard, A., & Sebban, M. 2015, Synthesis Lectures on Artificial Intelligence and Machine Learning, 9, 1

Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393, doi: 10.1051/aas:1996164

Bessell, M. S. 1990, PASP, 102, 1181, doi: 10.1086/132749

Bevington, P. R., & Robinson, D. K. 2003, *Data Reduction and Error Analysis* (McGraw Hill)

Bienias, J., Bódi, A., Forró, A., Hajdu, T., & Szabó, R. 2021, ApJS, 256, 11, doi: 10.3847/1538-4365/ac10c0

Blättler, E., & Diethelm, R. 2000a, Information Bulletin on Variable Stars, 4982, 1

Blättler, E., & Diethelm, R. 2000b, Information Bulletin on Variable Stars, 4996, 1

Bódi, A., & Hajdu, T. 2021, ApJS, 255, 1, doi: 10.3847/1538-4365/ac082c

Boggs, P. T., Byrd, R. H., & Schnabel, R. B. 1987, SIAM Journal on Scientific and Statistical Computing, 8, 1052

Borkovits, T., Rappaport, S., Hajdu, T., & Sztakovics, J. 2015, MNRAS, 448, 946, doi: 10.1093/mnras/stv015

Borkovits, T., Hajdu, T., Sztakovics, J., et al. 2016, MNRAS, 455, 4136, doi: 10.1093/mnras/stv2530

Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977, doi: 10.1126/science.1185402

Bradstreet, D. H., & Steelman, D. P. 2002, in American Astronomical Society Meeting Abstracts, Vol. 201, American Astronomical Society Meeting Abstracts, 75.02

Breger, M. 2000, in Astronomical Society of the Pacific Conference Series, Vol. 210, Delta Scuti and Related Stars, ed. M. Breger & M. Montgomery, 3

Brown, T. M., Latham, D. W., Everett, M. E., & Esquerdo, G. A. 2011, AJ, 142, 112, doi: 10.1088/0004-6256/142/4/112

Brownlee, R. R. 1957, ApJ, 125, 372, doi: 10.1086/146314

Bryson, S., Coughlin, J., Batalha, N. M., et al. 2020, AJ, 159, 279, doi: 10.3847/1538-3881/ab8a30

Budding, E., Erdem, A., Çiçek, C., et al. 2004, A&A, 417, 263, doi: 10.1051/0004-6361:20034135

Burrows, D. N., Hill, J. E., Nousek, J. A., et al. 2005, SSRv, 120, 165, doi: 10.1007/s11214-005-5097-2

Butterworth, S. 1930, Wireless Engineer, 7, 536

Caldwell, D. A., Borucki, W. J., Jenkins, J. M., et al. 2004, in Planetary Systems in the Universe, ed. A. Penny, Vol. 202, 69

Campos-Cucarella, F., Gomez-Forrellad, J. M., & Garcia-Melendo, E. 1997, Information Bulletin on Variable Stars, 4426, 1

Chebyshev, P. L. 1867, Journal de Mathématiques Pures et Appliquées, 12, 177

Christopoulou, P. E., Parageorgiou, A., & Chrysopoulos, I. 2011, AJ, 142, 99, doi: 10.1088/0004-6256/142/4/99

Clark Cunningham, J. M., Rawls, M. L., Windemuth, D., et al. 2019, AJ, 158, 106, doi: 10.3847/1538-3881/ab2d2b

Cochran, W. T., Cooley, J. W., Favin, D. L., et al. 1967, Proceedings of the IEEE, 55, 1664

Collins, K. A., Kielkopf, J. F., Stassun, K. G., & Hessman, F. V. 2017, AJ, 153, 77, doi: 10.3847/1538-3881/153/2/77

Conroy, K. E., Prša, A., Stassun, K. G., et al. 2014, AJ, 147, 45, doi: 10.1088/0004-6256/147/2/45

Čotar, K., Zwitter, T., Traven, G., et al. 2019, MNRAS, 487, 2474, doi: 10.1093/mnras/stz1397

Cousins, A. W. J. 1976, MmRAS, 81, 25

Cox, A. N. 2000, *Allen's Astrophysical Quantities* (Springer)

Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, *2MASS All Sky Catalog of point sources.*

Davenport, J. R. A., Hawley, S. L., Hebb, L., et al. 2014, ApJ, 797, 122, doi: 10.1088/0004-637X/797/2/122

Davenport, J. R. A. 2016, ApJ, 829, 23, doi: 10.3847/0004-637X/829/1/23

Davidge, T. J., & Milone, E. F. 1984, ApJS, 55, 571, doi: 10.1086/190969

de Bernardi, C., & Scaltriti, F. 1979, A&AS, 35, 63

de Boer, K., & Seggewiss, W. 2008, *Stars and Stellar Evolution* (EDP Sciences)

Debski, B., Baran, A., & Zoła, S. 2014, Contributions of the Astronomical Observatory Skalnate Pleso, 43, 427

Debski, B., Zoła, S., & Baran, A. 2015, in Astronomical Society of the Pacific Conference Series, Vol. 496, Living Together: Planets, Host Stars and Binaries, ed. S. M. Rucinski, G. Torres, & M. Zejda, 293

Deeming, T. J. 1975, Ap&SS, 36, 137, doi: 10.1007/BF00681947

Devor, J. 2005, ApJ, 628, 411, doi: 10.1086/431170

Devor, J., & Charbonneau, D. 2006a, Ap&SS, 304, 351, doi: 10.1007/s10509-006-9155-3

Devor, J., & Charbonneau, D. 2006b, ApJ, 653, 647, doi: 10.1086/508609

Devor, J., Charbonneau, D., O'Donovan, F. T., Mandushev, G., & Torres, G. 2008, AJ, 135, 850, doi: 10.1088/0004-6256/135/3/850

Djurašević, G., Rovithis-Livaniou, H., & Rovithis, P. 2000, A&A, 364, 543

Djurašević, G., Vince, I., & Atanacković, O. 2008, AJ, 136, 767, doi: 10.1088/0004-6256/136/2/767

Dotter, A., Chaboyer, B., Jevremović, D., et al. 2008, ApJS, 178, 89, doi: 10.1086/589654

Drake, A. J., Graham, M. J., Djorgovski, S. G., et al. 2014, ApJS, 213, 9, doi: 10.1088/0067-0049/213/1/9

Dryomova, G. N. 2014, Astronomical and Astrophysical Transactions, 28, 191

Duchêne, G., & Kraus, A. 2013, ARA&A, 51, 269, doi: 10.1146/annurev-astro-081710-102602

Duerbeck, H. W. 1975, A&AS, 22, 19

Eggleton, P. P. 2012, Journal of Astronomy and Space Sciences, 29, 145, doi: 10.5140/JASS.2012.29.2.145

El-Sadek, M. A., Djurašević, G., Essam, A., et al. 2019, NewA, 69, 21, doi: 10.1016/j.newast.2018.11.006

Euler, L. 1767, Novi commentarii academiae scientiarum Petropolitanae, 144

Flower, P. J. 1996, ApJ, 469, 355, doi: 10.1086/177785

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

Fox-Machado, L., Echevarria, J., González-Buitrago, D., & Michel, R. 2016, IAU Focus Meeting, 29B, 536, doi: 10.1017/S1743921316006074

Frasca, A., Molenda-Żakowicz, J., De Cat, P., et al. 2016, A&A, 594, A39, doi: 10.1051/0004-6361/201628337

Friedman, J. H. 1984, *A variable span smoother*, Tech. rep., Stanford Univ CA lab for computational statistics

Furlan, E., Ciardi, D. R., Everett, M. E., et al. 2017, AJ, 153, 71, doi: 10.3847/1538-3881/153/2/71

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1, doi: 10.1051/0004-6361/201833051

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, A&A, 649, A1, doi: 10.1051/0004-6361/202039657

Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2022, arXiv e-prints, arXiv:2208.00211, doi: 10.1051/0004-6361/202243940

Gao, Q., Xin, Y., Liu, J.-F., Zhang, X.-B., & Gao, S. 2016, ApJS, 224, 37, doi: 10.3847/0067-0049/224/2/37

Gao, X.-Y., Li, K., Cai, Y.-W., et al. 2022, PASP, 134, 114202, doi: 10.1088/1538-3873/ac99fd

Gardner, T., Hahs, G., & Gokhale, V. 2015, JAAVSO, 43, 186

Gaulme, P., McKeever, J., Rawls, M. L., et al. 2013, ApJ, 767, 82, doi: 10.1088/0004-637X/767/1/82

Gautam, C., Tiwari, A., Suresh, S., & Ahuja, K. 2019, IEEE Transactions on Systems, Man, and Cybernetics: Systems

Gazeas, K. D., & Niarchos, P. G. 2006, MNRAS, 370, L29, doi: 10.1111/j.1745-3933.2006.00182.x

Gershberg, R. E. 1972, Ap&SS, 19, 75, doi: 10.1007/BF00643168

Gies, D. R., Williams, S. J., Matson, R. A., et al. 2012, AJ, 143, 137, doi: 10.1088/0004-6256/143/6/137

Gies, D. R., Matson, R. A., Guo, Z., et al. 2015, AJ, 150, 178, doi: 10.1088/0004-6256/150/6/178

Gilliland, R. L., Brown, T. M., Kjeldsen, H., et al. 1993, AJ, 106, 2441, doi: 10.1086/116814

Gilliland, R. L., Chaplin, W. J., Dunham, E. W., et al. 2011, ApJS, 197, 6, doi: 10.1088/0067-0049/197/1/6

Goodricke, J. 1783, Philosophical Transactions of the Royal Society of London Series I, 73, 474

Gray, D. F. 2008, *The Observation and Analysis of Stellar Photospheres* (Cambridge University Press)

Gray, J. D., Samec, R. G., & Carrigan, B. J. 1995, Information Bulletin on Variable Stars, 4177, 1

Gu, S.-h., Chen, P.-s., Choy, Y.-k., et al. 2004, A&A, 423, 607, doi: 10.1051/0004-6361:200400017

Guo, D.-F., Li, K., Liu, F., et al. 2022, MNRAS, 517, 1928, doi: 10.1093/mnras/stac2811

Hahs, G., Ortmann, C., & Gokhale, V. 2020, JAAVSO, 48, 57

Harmanec, P. 1988, Bulletin of the Astronomical Institutes of Czechoslovakia, 39, 329

Harmanec, P., Zasche, P., Brož, M., et al. 2022, arXiv e-prints, arXiv:2204.07045. https://arxiv.org/abs/2204.07045

Hartman, J. D., Bakos, G., Stanek, K. Z., & Noyes, R. W. 2004, AJ, 128, 1761, doi: 10.1086/423920

Hartung, J., Knapp, G., & Sinha, B. 2008, *Statistical Meta-Analysis with Applications* (Wiley). https://books.google.com/books?id=U1FelnouDrsC

He, H., Wang, H., Zhang, M., et al. 2018, ApJS, 236, 7, doi: 10.3847/1538-4365/aab779

Helfer, E., Smith, B., Haber, R., & Peter, A. 2015, *Statistical Analysis of Functional Data*, Tech. rep., Florida Institute of Technology

Henden, A. A., & Kaitchuck, R. H. 1990, *Astronomical Photometry: A Text and Handbook for the Advanced Amateur and Professional Astronomer* (Willmann-Bell)

Henden, A. A., Templeton, M., Terrell, D., et al. 2016, VizieR Online Data Catalog, II/336

Henden, A. A. 2019, JAAVSO, 47, 130

Hewitt, E., & Hewitt, R. E. 1979, Archive for History of Exact Sciences, 21, 129

Hilditch, R. W., Collier Cameron, A., Hill, G., Bell, S. A., & Harries, T. J. 1997, MNRAS, 291, 749, doi: 10.1093/mnras/291.4.749

Hoffman, D. I., Harrison, T. E., Coughlin, J. L., et al. 2008, AJ, 136, 1067, doi: 10.1088/0004-6256/136/3/1067

Hoffman, D. I., Harrison, T. E., & McNamara, B. J. 2009, AJ, 138, 466, doi: 10.1088/0004-6256/138/2/466

Horch, E. P., Howell, S. B., Everett, M. E., & Ciardi, D. R. 2014, ApJ, 795, 60, doi: 10.1088/0004-637X/795/1/60

Howell, S. B. 2006, *Handbook of CCD Astronomy* (Cambridge University Press)

Hrivnak, B. J. 1982, ApJ, 260, 744, doi: 10.1086/160294

Huber, D., Silva Aguirre, V., Matthews, J. M., et al. 2014, ApJS, 211, 2, doi: 10.1088/0067-0049/211/1/2

Iverson, K. E. 1962, in Proceedings of the May 1-3, 1962, spring joint computer conference, 345–351

Jansen, F., Lumb, D., Altieri, B., et al. 2001, A&A, 365, L1, doi: 10.1051/0004-6361:20000036

Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. 2010, ApJL, 713, L87, doi: 10.1088/2041-8205/713/2/L87

Jester, S., Schneider, D. P., Richards, G. T., et al. 2005, AJ, 130, 873, doi: 10.1086/432466

Johnson, H. L., & Morgan, W. W. 1953, ApJ, 117, 313, doi: 10.1086/145697

Johnston, K. B., Haber, R., Caballero-Nieves, S. M., et al. 2019, Computational Astrophysics and Cosmology, 6, 4, doi: 10.1186/s40668-019-0031-2

Johnston, K. B., & Haber, R. 2019, *OCD: O'Connell Effect Detector using push-pull learning.* http://ascl.net/1910.020

Kallrath, J., & Milone, E. F. 2009, *Eclipsing Binary Stars: Modeling and Analysis* (Springer), doi: 10.1007/978-1-4419-0699-1

Kałużny, J. 1990, AJ, 99, 1207, doi: 10.1086/115409

Kałużny, J. 1991, AcA, 41, 17

Kaplan, G. H. 2005, U.S. Naval Observatory Circulars, 179. https://arxiv.org/abs/astro-ph/0602086

Keel, W. C., Oswalt, T., Mack, P., et al. 2017, PASP, 129, 015002, doi: 10.1088/1538-3873/129/971/015002

Kendall, M. G. 1938, Biometrika, 30, 81

Kinemuchi, K., Smith, H. A., Woźniak, P. R., McKay, T. A., & ROTSE Collaboration. 2006, AJ, 132, 1202, doi: 10.1086/506198

Kirk, B., Conroy, K., Prša, A., et al. 2016, AJ, 151, 68, doi: 10.3847/0004-6256/151/3/68

Kjurkchieva, D., Vasileva, D., & Atanasova, T. 2017, AJ, 154, 105, doi: 10.3847/1538-3881/aa83b3

Knote, M. F., Kaitchuck, R. H., & Berrington, R. C. 2019, JAAVSO, 47, 194. https://arxiv.org/abs/1910.07627

Knote, M. F., Boudreaux, T. M., Achille, R. J., et al. in prep.

Knuth, D. E. 1997, *The Art of Computer Programming* (Pearson Education)

Kobulnicky, H. A., Molnar, L. A., Cook, E. M., & Henderson, L. E. 2022, ApJS, 262, 12, doi: 10.3847/1538-4365/ac75bd

Kochanek, C. S., Shappee, B. J., Stanek, K. Z., et al. 2017, PASP, 129, 104502, doi: 10.1088/1538-3873/aa80d9

Kolmogorov, A. 1933, Giornale dell'Istituto Italiano degli Attuari, 4, 83

Kopal, Z. 1955, Annales d'Astrophysique, 18, 379

Kouzuma, S. 2018, PASJ, 70, 90, doi: 10.1093/pasj/psy086

Kouzuma, S. 2019, PASJ, 71, 21, doi: 10.1093/pasj/psy140

Kovács, G., Zucker, S., & Mazeh, T. 2002, A&A, 391, 369, doi: 10.1051/0004-6361:20020802

Kovács, G., Hartman, J. D., & Bakos, G. Á. 2019, A&A, 631, A126, doi: 10.1051/0004-6361/201936207

Kozai, Y. 1962, AJ, 67, 591, doi: 10.1086/108790

Kruszewski, A. 1963, AcA, 13, 106

Kuiper, G. P. 1941, ApJ, 93, 133, doi: 10.1086/144252

Kunt, M., & Dal, H. A. 2017, AcA, 67, 345, doi: 10.32023/0001-5237/67.4.4

Kwee, K. K., & van Woerden, H. 1956, BAN, 12, 327

Lada, C. J. 2006, ApJL, 640, L63, doi: 10.1086/503158

Lagrange, J.-L. 1772, Prix de l'académie royale des Sciences de paris, 9, 292

Lanczos, C. 1956, *Applied Analysis* (Prentice-Hall)

Landolt, A. U. 1975, PASP, 87, 409, doi: 10.1086/129784

Lazaro, C., Arevalo, M. J., & Almenara, J. M. 2009, NewA, 14, 528, doi: 10.1016/j.newast.2009.01.010

Lenz, P., & Breger, M. 2005, Communications in Asteroseismology, 146, 53, doi: 10.1553/cia146s53

Leung, K. C., Zhai, D., & Zhang, Y. 1985, AJ, 90, 515, doi: 10.1086/113759

Leung, K. C. 1989, SSRv, 50, 279, doi: 10.1007/BF00215937

Li, K., Hu, S., Guo, D., et al. 2015, NewA, 41, 17, doi: 10.1016/j.newast.2015.04.010

Li, K., Xia, Q.-Q., Kim, C.-H., et al. 2021, ApJ, 922, 122, doi: 10.3847/1538-4357/ac242f

Liakos, A., & Niarchos, P. 2017, MNRAS, 465, 1181, doi: 10.1093/mnras/stw2756

Liakos, A. 2017, A&A, 607, A85, doi: 10.1051/0004-6361/201731247

Liao, W.-P., Li, L.-J., Zhou, X., & Wang, Q.-S. 2021, Research in Astronomy and Astrophysics, 21, 041, doi: 10.1088/1674-4527/21/2/41

Lin, J., Keogh, E., Lonardi, S., & Chiu, B. 2003, in Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery, 2

Lindegren, L. 2018. http://www.rssd.esa.int/doc_fetch.php?id=3757412

Liu, L., Qian, S. B., Liao, W. P., et al. 2011, AJ, 141, 44, doi: 10.1088/0004-6256/141/2/44

Liu, Q.-Y., & Yang, Y.-L. 2003, ChJA&A, 3, 142, doi: 10.1088/1009-9271/3/2/142

Lomb, N. R. 1976, Ap&SS, 39, 447, doi: 10.1007/BF00648343

Maceroni, C., & van't Veer, F. 1993, A&A, 277, 515

Maceroni, C., & van't Veer, F. 1996, A&A, 311, 523

Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94, doi: 10.3847/1538-3881/aa784d

Mancuso, S., Milano, L., & Russo, G. 1977, Ap&SS, 47, 277, doi: 10.1007/BF00642838

Massey, P. 1997, National Optical Astronomy Observatory

Matijevič, G., Prša, A., Orosz, J. A., et al. 2012, AJ, 143, 123, doi: 10.1088/0004-6256/143/5/123

Matson, R. A., Gies, D. R., Guo, Z., & Williams, S. J. 2017, AJ, 154, 216, doi: 10.3847/1538-3881/aa8fd6

McCartney, S. A. 1999, PhD thesis, The University of Oklahoma

McQuillan, A., Mazeh, T., & Aigrain, S. 2014, ApJS, 211, 24, doi: 10.1088/0067-0049/211/2/24

Mennickent, R. E., Garcés, J., Djurašević, G., et al. 2020, A&A, 641, A91, doi: 10.1051/0004-6361/202038110

Mennickent, R. E., & Djurašević, G. 2021, A&A, 653, A89, doi: 10.1051/0004-6361/202141098

Mergentaler, J. 1950, *On the Asymmetry of the Light-Curves of the Eclipsing Variables: Contributions from the Wrocław Astronomical Observatory* (Wrocławskie Towarzystwo Naukowe)

Michel, R., Echevarría, J., Cang, T. Q., Fox-Machado, L., & González-Buitrago, D. 2016, RMxAA, 52, 339. https://arxiv.org/abs/1606.03464

Milone, E. E. 1968, AJ, 73, 708, doi: 10.1086/110682

Miyakawa, K., Hirano, T., Sato, B., Fukui, A., & Narita, N. 2021, AJ, 161, 276, doi: 10.3847/1538-3881/abf500

Moe, M., Kratter, K. M., & Badenes, C. 2019, ApJ, 875, 61, doi: 10.3847/1538-4357/ab0d88

Molnar, L. A., Van Noord, D. M., Kinemuchi, K., et al. 2017, ApJ, 840, 1, doi: 10.3847/1538-4357/aa6ba7

Montet, B. T., Tovar, G., & Foreman-Mackey, D. 2017, ApJ, 851, 116, doi: 10.3847/1538-4357/aa9e00

Morbey, C. L., & Brosterhus, E. B. 1974, PASP, 86, 455, doi: 10.1086/129630

NegmEldin, M. A., Essam Elsayed, A., & Yousef, S. M. 2019, Research in Astronomy and Astrophysics, 19, 025, doi: 10.1088/1674-4527/19/2/25

Negu, S. H., & Tessema, S. B. 2018, Astronomische Nachrichten, 339, 478, doi: 10.1002/asna.201813466

O'Connell, D. J. K. 1951, Publications of the Riverview College Observatory, 2, 85

Oliphant, T. E. 2006, *A Guide to NumPy* (Trelgol Publishing USA)

Orosz, J. A., & Hauschildt, P. H. 2000, A&A, 364, 265. https://arxiv.org/abs/astro-ph/0010114

Papageorgiou, A., Kleftogiannis, G., & Christopoulou, P. E. 2014, Contributions of the Astronomical Observatory Skalnate Pleso, 43, 470

Papageorgiou, A., Catelan, M., Christopoulou, P.-E., Drake, A. J., & Djorgovski, S. G. 2018, ApJS, 238, 4, doi: 10.3847/1538-4365/aad8a9

Paunzen, E., & Vanmunster, T. 2016, Astronomische Nachrichten, 337, 239, doi: 10.1002/asna.201512254

Pavlenko, Y. V. 2003, Astronomy Reports, 47, 59, doi: 10.1134/1.1538496

Pavlenko, Y. V., Evans, A., Banerjee, D. P. K., et al. 2018, A&A, 615, A120, doi: 10.1051/0004-6361/201832717

Pavlovski, K., Burki, G., & Mimica, P. 2006, A&A, 454, 855, doi: 10.1051/0004-6361:20054733

Payne-Gaposchkin, C., & Gaposchkin, S. 1938, *Variable Stars* (Harvard Observatory)

Pearson, K. 1895, Proceedings of the Royal Society of London Series I, 58, 240

Pecaut, M. J., & Mamajek, E. E. 2013, ApJS, 208, 9, doi: 10.1088/0067-0049/208/1/9

Pigulski, A., Pojmański, G., Pilecki, B., & Szczygieł, D. M. 2009, AcA, 59, 33. https://arxiv.org/abs/0903.4921

Pizzocaro, D., Stelzer, B., Poretti, E., et al. 2019, A&A, 628, A41, doi: 10.1051/0004-6361/201731674

Plavec, M. 1958, in Liege International Astrophysical Colloquia, Vol. 8, Liege International Astrophysical Colloquia, 411–420

Pogson, N. 1856, MNRAS, 17, 12, doi: 10.1093/mnras/17.1.12

Pojmanski, G. 1997, AcA, 47, 467. https://arxiv.org/abs/astro-ph/9712146

Pollacco, D. L., Skillen, I., Collier Cameron, A., et al. 2006, PASP, 118, 1407, doi: 10.1086/508556

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2007, *Numerical Recipes 3rd Edition: The Art of Scientific Computing* (Cambridge University Press)

Pribulla, T., Vaňko, M., Chochol, D., Hambálek, Ľ., & Parimucha, Š. 2011, Astronomische Nachrichten, 332, 607, doi: 10.1002/asna.201111569

Prša, A., & Zwitter, T. 2005, ApJ, 628, 426, doi: 10.1086/430591

Prša, A. 2006, Univ. of Ljubljana. http://phoebe-project.org/static/legacy/docs/phoebe_science.pdf

Prša, A., Guinan, E. F., Devinney, E. J., et al. 2008, ApJ, 687, 542, doi: 10.1086/591783

Prša, A., Batalha, N., Slawson, R. W., et al. 2011, AJ, 141, 83, doi: 10.1088/0004-6256/141/3/83

Prša, A., Conroy, K. E., Horvat, M., et al. 2016, ApJS, 227, 29, doi: 10.3847/1538-4365/227/2/29

Prša, A., Kochoska, A., Conroy, K. E., et al. 2022, ApJS, 258, 16, doi: 10.3847/1538-4365/ac324a

Ramsay, G., Brooks, A., Hakala, P., et al. 2014, MNRAS, 437, 132, doi: 10.1093/mnras/stt1863

Rasband, W. S., et al. 1997, *ImageJ*, Bethesda, MD

Ribárik, G. 2002, Konkoly Observatory Occasional Technical Notes, 12, 1

Ribárik, G., Oláh, K., & Strassmeier, K. G. 2003, Astronomische Nachrichten, 324, 202, doi: 10.1002/asna.200310088

Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003, doi: 10.1117/1.JATIS.1.1.014003

Roberts, A. W. 1906, MNRAS, 66, 123, doi: 10.1093/mnras/66.3.123

Rodonò, M., Lanza, A. F., & Catalano, S. 1995, A&A, 301, 75

Rosen, S. R., Webb, N. A., Watson, M. G., et al. 2016, A&A, 590, A1, doi: 10.1051/0004-6361/201526416

Roweis, S. T., & Saul, L. K. 2000, Science, 290, 2323, doi: 10.1126/science.290.5500.2323

Rucinski, S. M. 1997, AJ, 113, 407, doi: 10.1086/118263

Russell, H. N., & Merrill, J. E. 1952, *The Determination of the Elements of Eclipsing Binaries* (Princeton University Observatory)

Samec, R. G., Charlesworth, S. D., & Dewitt, J. R. 1991, AJ, 102, 688, doi: 10.1086/115903

Samec, R. G., Melton, R. A., Figg, E. R., et al. 2010, AJ, 140, 1150, doi: 10.1088/0004-6256/140/5/1150

Savitzky, A., & Golay, M. J. E. 1964, Analytical chemistry, 36, 1627

Scargle, J. D. 1982, ApJ, 263, 835, doi: 10.1086/160554

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. 2012, Nature methods, 9, 671

Schwarz, G. 1978, Annals of Statistics, 6, 461

Schwarzenberg-Czerny, A. 1989, MNRAS, 241, 153, doi: 10.1093/mnras/241.2.153

Schwarzenberg-Czerny, A. 1996, ApJL, 460, L107, doi: 10.1086/309985

Sevilla-Lara, L., & Learned-Miller, E. 2012, in 2012 IEEE Conference on computer vision and pattern recognition, IEEE, 1910–1917

Shappee, B. J., Prieto, J. L., Grupe, D., et al. 2014, ApJ, 788, 48, doi: 10.1088/0004-637X/788/1/48

Shugarov, S. Y., Goranskij, V. P., Galkina, M. P., & Lipunova, N. A. 1990, Information Bulletin on Variable Stars, 3472, 1

Slawson, R. W., Prša, A., Welsh, W. F., et al. 2011, AJ, 142, 160, doi: 10.1088/0004-6256/142/5/160

Smirnov, N. 1948, The annals of mathematical statistics, 19, 279

Smith, K. L., Boyd, P. T., Mushotzky, R. F., et al. 2015, AJ, 150, 126, doi: 10.1088/0004-6256/150/4/126

Socia, Q. J., Welsh, W. F., Short, D. R., et al. 2018, ApJL, 864, L32, doi: 10.3847/2041-8213/aadc0d

Söderhjelm, S., & Dischler, J. 2005, A&A, 442, 1003, doi: 10.1051/0004-6361:20042541

Soszyński, I., Pawlak, M., Pietrukowicz, P., et al. 2016, AcA, 66, 405. https://arxiv.org/abs/1701.03105

Spearman, C. 1904, Am. J. Psychol, 15, 72

Stark, M. A., Hamlin, M. T., Anderson, J. R., et al. 2000, JAAVSO, 28, 25

Stellingwerf, R. F. 1978, ApJ, 224, 953, doi: 10.1086/156444

Struve, O. 1941, ApJ, 93, 104, doi: 10.1086/144249

Svechnikov, M. A., Istomin, L. F., & Grekhova, O. A. 1980, Peremennye Zvezdy, 21, 399

Swan, P. R. 1982, AJ, 87, 1608, doi: 10.1086/113252

Tamuz, O., Mazeh, T., & Zucker, S. 2005, MNRAS, 356, 1466, doi: 10.1111/j.1365-2966.2004.08585.x

Terrell, D., & Wilson, R. E. 2005, Ap&SS, 296, 221, doi: 10.1007/s10509-005-4449-4

Terrell, D., & Nelson, R. H. 2014, ApJ, 783, 35, doi: 10.1088/0004-637X/783/1/35

Thompson, S. E., Everett, M., Mullally, F., et al. 2012, ApJ, 753, 86, doi: 10.1088/0004-637X/753/1/86

Thompson, S. E., Fraquelli, D., Van Cleve, J. E., & Caldwell, D. A. 2016, *Kepler Archive Manual*, Kepler Science Document KDMC-10008-006

Tody, D. 1993, in Astronomical Society of the Pacific Conference Series, Vol. 52, Astronomical Data Analysis Software and Systems II, ed. R. J. Hanisch, R. J. V. Brissenden, & J. Barnes, 173

Tokunaga, A. T. 2000, in Allen's Astrophysical Quantities, ed. A. N. Cox (Springer), 143

Tran, K., Levine, A., Rappaport, S., et al. 2013, ApJ, 774, 81, doi: 10.1088/0004-637X/774/1/81

Tregloan-Reed, J., & Southworth, J. 2013, MNRAS, 431, 966, doi: 10.1093/mnras/stt227

Tylenda, R., Hajduk, M., Kamiński, T., et al. 2011, A&A, 528, A114, doi: 10.1051/0004-6361/201016221

Van Cleve, J. E., & Caldwell, D. A. 2016, *Kepler Instrument Handbook*, Kepler Science Document KSCI-19033-002

Van Cleve, J. E., Christiansen, J. L., Jenkins, J. M., et al. 2016, *Kepler Data Characteristics Handbook*, Kepler Science Document KSCI-19040-005

Van Der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science & Engineering, 13, 22

Van Rossum, G., & Drake Jr, F. L. 1995, *Python Reference Manual* (Centrum voor Wiskunde en Informatica Amsterdam)

van't Veer, F. 1973, A&A, 26, 357

Vaughan, A. H., & Preston, G. W. 1980, PASP, 92, 385, doi: 10.1086/130683

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: 10.1038/s41592-019-0686-2

Völschow, M., Schleicher, D. R. G., Banerjee, R., & Schmitt, J. H. M. M. 2018, A&A, 620, A42, doi: 10.1051/0004-6361/201833506

Willson, R. C., Gulkis, S., Janssen, M., Hudson, H. S., & Chapman, G. A. 1981, Science, 211, 700, doi: 10.1126/science.211.4483.700

Wilsey, N. J., & Beaky, M. M. 2009, Society for Astronomical Sciences Annual Symposium, 28, 107

Wilson, J. C., Hearty, F. R., Skrutskie, M. F., et al. 2019, PASP, 131, 055001, doi: 10.1088/1538-3873/ab0075

Wilson, R. E., & Devinney, E. J. 1971, ApJ, 166, 605, doi: 10.1086/150986

Wilson, R. E. 1979, ApJ, 234, 1054, doi: 10.1086/157588

Wilson, R. E. 2001, Information Bulletin on Variable Stars, 5076, 1

Wilson, R. E. 2006, in Astronomical Society of the Pacific Conference Series, Vol. 349, Astrophysics of Variable Stars, ed. C. Aerts & C. Sterken, 71

Wolniewicz, L. M., Berger, T. A., & Huber, D. 2021, AJ, 161, 231, doi: 10.3847/1538-3881/abee1d

Wood, D. B. 1972, *A Computer Program for Modeling Non-Spherical Eclipsing Binary Star Systems*. https://ntrs.nasa.gov/citations/19730008094

Woodward, E. J., & Wilson, R. E. 1977, Ap&SS, 52, 387, doi: 10.1007/BF01093875

Woźniak, P. R. 2000, AcA, 50, 421. https://arxiv.org/abs/astro-ph/0012143

Woźniak, P. R., Vestrand, W. T., Akerlof, C. W., et al. 2004, AJ, 127, 2436, doi: 10.1086/382719

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868, doi: 10.1088/0004-6256/140/6/1868

Yang, Y.-G., Qian, S.-B., Zhu, L.-Y., He, J.-J., & Yuan, J.-Z. 2005, PASJ, 57, 983, doi: 10.1093/pasj/57.6.983

Yoldaş, E. 2021, RMxAA, 57, 351, doi: 10.22201/ia.01851101p.2021.57.02.08

Zasche, P., Henzl, Z., & Mašek, M. 2022, A&A, 664, A96, doi: 10.1051/0004-6361/202243723

Zhang, X. B., Zhang, R. X., & Fang, M. J. 2002, A&A, 395, 587, doi: 10.1051/0004-6361:20021317

Zhou, D.-Q., & Leung, K.-C. 1990, ApJ, 355, 271, doi: 10.1086/168760

Zoła, S. 1996, A&A, 308, 785

Zoła, S., Baran, A., Debski, B., & Jableka, D. 2017, MNRAS, 466, 2488, doi: 10.1093/mnras/stw3138

# Appendix A

# Target List

Table A.1: The complete sample of 258 *Kepler* eclipsing binaries with an O'Connell effect.

| *Kepler* ID[a] | Period (days) | LC Class[b] | Morph. Param. | Color | BP Mag. | G Mag. | Lum. ($L_\odot$) | Spectral Type Lit.[c] | Spectral Type Color[d] | Ref.[ef] | Samp.[g] | Notable Type[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2159783 | 0.3739 | WU | 0.87 | 0.938 | 15.365 | 14.982 | 1.810 | F6 V[1] | G9 | 4, 5, 6, 22* | C | AM, S, M |
| 2161623 | 2.2835 | Al | 0.54 | 0.771 | 14.571 | 14.273 | 5.262 | — | F9.5 | — | C | — |
| 2302092 | 0.2947 | WU | 0.88 | 1.124 | 14.982 | 14.503 | 0.775 | — | K2 | 5, 7 | C | S |
| 2305372 | 1.4047 | Al | 0.58 | 0.929 | 14.214 | 13.846 | 3.498 | F8 Vs[1] | G9 | 2, 7, 13, 14*, 16, 19 | C | F |
| 2437038 | 0.2677 | WU | 0.87 | 1.120 | 16.501 | 16.149 | 2.004 | — | K2 | 3, 5, 6 | C | F, S, M |
| 2449084 | 0.7398 | BL | 0.70 | 0.842 | 15.319 | 14.983 | 5.926 | — | G4 | 3 | C | AM, F |
| 2569494 | 1.5233 | Al | 0.58 | 1.209 | 17.799 | 17.288 | 2.051 | — | K3 | 8, 13 | C | TV, AM |
| 2577756 | 0.8703 | Al | 0.63 | 0.911 | 15.860 | 15.491 | 2.262 | — | G8 | 2 | M | TV, AM, F |
| 2694741 | 0.3266 | WU | 0.93 | 0.936 | 12.639 | 12.256 | 1.304 | G4 V[1] | G9 | 3, 5 | C | F, S |
| 2695740 | 3.6159 | WU | 0.85 | 1.305 | 13.704 | 13.126 | 8.844 | — | K4 | 3, 5, 19 | C | TV, AM, F, S |
| 2854752 | 0.4704 | BL | 0.79 | 0.636 | 14.153 | 13.918 | 5.132 | — | F6 | 3, 6 | C | F, M |
| 2858322 | 0.4364 | BL | 0.77 | 0.984 | 15.197 | 14.789 | 4.118 | — | K0 | 5, 6 | C | S, M |
| 2972514 | 0.5035 | WU | 0.74 | 0.802 | 14.332 | 14.020 | 3.405 | — | G1 | — | C | — |
| 3124420 | 0.9490 | Al | 0.60 | 1.174 | 18.801 | 18.287 | 0.946 | — | K2.5 | — | C | TV, AM |
| 3218683 | 0.7717 | Al | 0.69 | 1.156 | 16.524 | 16.030 | 1.208 | — | K2.5 | 2, 5 | C | F, S |
| 3338660 | 1.8734 | Al | 0.60 | 0.944 | 15.087 | 14.707 | 12.053 | — | G9 | 3, 7, 8 | C | F |
| 3339563 | 0.8412 | BL | 0.72 | 1.030 | 14.624 | 14.193 | 2.240 | G0 IIIp[1] | K1 | 2, 3, 5 | C | TV, AM, F, S |
| 3342425 | 0.3934 | WU | 0.94 | 0.898 | 15.489 | 15.137 | 2.362 | — | G8 | 3, 5 | C | TV, AM, F, S |
| 3344427 | 0.6518 | Al | 0.57 | 1.448 | 16.423 | 15.776 | 0.274 | — | K5 | — | M | — |
| 3351945 | 1.0805 | Al | 0.50 | 1.288 | 15.359 | 14.794 | 0.210 | — | K3.5 | — | M | TV |
| 3557421 | 0.3937 | Al | 0.70 | 1.124 | 10.486 | 10.003 | 0.662 | — | K2 | 5, 16, 19 | M | S |
| 3656322 | 3.6637 | Al | 0.59 | 1.125 | 13.538 | 13.077 | 4.706 | — | K2 | 13 | M | TV |

Table A.1 *continued on next page*

Table A.1 *(continued)*

| *Kepler* ID[a] | Period (days) | LC Class[b] | Morph. Param. | Color | BP Mag. | G Mag. | Lum. ($L_\odot$) | Spectral Type Lit.[c] | Spectral Type Color[d] | Ref.[ef] | Samp.[g] | Notable Type[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3659940 | 0.8963 | Al | 0.55 | 1.047 | 15.619 | 15.187 | 0.899 | — | K1.5 | — | C | TV |
| 3662635 | 0.9394 | Al | 0.57 | 0.988 | 15.273 | 14.965 | 32.241 | — | K0 | 3 | C | TV, F |
| 3743834 | 0.2736 | WU | 0.76 | 1.218 | 14.775 | 14.241 | 0.402 | — | K3 | 3, 5 | C | F, S |
| 3832382 | 0.2726 | WU | 0.75 | 1.426 | 16.447 | 15.802 | 0.300 | — | K5 | 5 | C | S |
| 3839964 | 0.2561 | WU | 0.80 | 1.026 | 14.988 | 14.557 | 1.272 | — | K1 | 3, 7 | M | F |
| 3848919 | 1.0473 | Al | 0.57 | 1.059 | 14.409 | 13.984 | 1.269 | — | K1.5 | 21 | C | — |
| 3954227 | 0.5943 | Al | 0.56 | 1.434 | 16.689 | 16.097 | 0.359 | — | K5 | — | M | TV |
| 3965242 | 0.9967 | Al | 0.53 | 1.074 | 14.536 | 14.084 | 0.396 | — | K1.5 | — | C | TV |
| 3972629 | 0.2442 | BL | 0.85 | 1.140 | 15.448 | 14.957 | 0.403 | — | K2.5 | 3, 5 | C | F, S |
| 4037163 | 0.6354 | Al | 0.59 | 1.484 | 18.153 | 17.446 | 0.358 | — | K5.5 | 7, 8 | M | TV |
| 4049124 | 4.8045 | Al | 0.17 | 1.007 | 15.063 | 14.647 | 0.900 | — | K1 | 2, 13 | M | F |
| 4077442 | 0.6928 | Al | 0.59 | 1.325 | 14.213 | 13.629 | 0.785 | — | K4 | 21 | M | TV |
| 4168013 | 0.3109 | WU | 0.85 | 1.092 | 15.527 | 15.079 | 0.875 | — | K2 | — | M | — |
| 4241946 | 0.2844 | WU | 0.75 | 1.329 | 16.472 | 16.073 | 9.260 | — | K4 | 3 | C | F |
| 4249218 | 0.3162 | WU | 0.76 | 1.053 | 14.449 | 14.184 | 1.015 | G4 V[1] | K1.5 | 3, 5, 19 | C | F, S |
| 4349483 | 0.5091 | Al | 0.71 | 1.243 | 14.798 | 14.371 | 0.546 | — | K3.5 | — | C | TV |
| 4350454 | 0.9657 | BL | 0.80 | 0.829 | 13.934 | 13.608 | 9.305 | — | G3 | 3 | C | F |
| 4385109 | 0.2746 | WU | 0.90 | 1.181 | 15.580 | 15.071 | 0.535 | — | K2.5 | 3 | C | F |
| 4466691 | 0.4827 | BL | 0.78 | 0.688 | 14.576 | 14.308 | 5.792 | — | F8 | — | C | — |
| 4474193 | 0.3225 | WU | 0.89 | 1.067 | 15.628 | 15.173 | 1.022 | — | K1.5 | 3, 5 | C | TV, F, S |
| 4474637 | 3.8867 | Al | 0.58 | 0.911 | 18.111 | 17.780 | 3.450 | — | G8 | 13 | C | TV, AM |
| 4547308 | 0.5769 | WU | 0.88 | 0.552 | 12.617 | 12.457 | 21.215 | — | F4 | 3, 7 | C | F |
| 4660997 | 0.5626 | Al | 0.62 | 0.885 | 12.606 | 12.250 | 1.742 | G1 V[1] | G7 | 1, 3, 14, 16, 17, 19, 21 | M | F |

Table A.1 *continued on next page*

Table A.1 *(continued)*

| *Kepler* ID[a] | Period (days) | LC Class[b] | Morph. Param. | Color | BP Mag. | G Mag. | Lum. ($L_\odot$) | Spectral Type Lit.[c] | Spectral Type Color[d] | Ref.[ef] | Samp.[g] | Notable Type[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4859432 | 0.3855 | WU | 0.77 | 1.064 | 15.651 | 15.214 | 1.773 | — | K1.5 | 3, 7 | C | F |
| 4941060 | 0.3034 | WU | 0.75 | 1.160 | 15.892 | 15.403 | 0.628 | — | K2.5 | 5 | C | S |
| 4991959 | 0.3609 | WU | 0.76 | 0.943 | 15.442 | 15.053 | 1.300 | — | G9 | 3, 5, 6 | C | F, S, M |
| 5020034 | 2.1191 | Al | 0.57 | 0.884 | 16.110 | 15.761 | 7.522 | — | G7 | — | C | — |
| 5033682 | 0.3799 | WU | 0.94 | 0.892 | 13.640 | 13.278 | 2.000 | — | G8 | 3, 4, 5 | C | F, S |
| 5123176 | 0.7078 | WU | 0.82 | 0.622 | 11.713 | 11.481 | 11.153 | F5 IV[1] | F6 | 3, 5, 6, 16, 17, 19 | C | F, S, M |
| 5195137 | 0.3241 | WU | 0.92 | 1.007 | 14.426 | 14.010 | 1.099 | — | K1 | 3, 5 | C | F, S |
| 5215999 | 2.5312 | Al | 0.68 | 1.134 | 14.636 | 14.155 | 16.783 | — | K2 | 3, 5 | M | TV, F, S |
| 5282464 | 0.4964 | BL | 0.73 | 0.491 | 12.708 | 12.524 | 6.144 | — | F2 | 16, 19 | C | — |
| 5283839 | 0.3152 | WU | 0.93 | 0.892 | 15.614 | 15.253 | 1.486 | — | G8 | 3, 4, 5 | M | AM, F, S |
| 5294739 | 3.7360 | Al | 0.59 | 1.255 | 14.111 | 13.567 | 7.949 | — | K3.5 | 3, 13, 16, 19 | C | F |
| 5300878 | 1.2794 | Al | 0.50 | 1.387 | 15.583 | 14.963 | 0.303 | — | K4.5 | 5, 8 | C | CU, S |
| 5357682 | 0.7177 | BL | 0.83 | 0.741 | 14.801 | 14.511 | 5.242 | — | F9 | — | C | — |
| 5425950 | 0.3782 | WU | 0.76 | 0.979 | 15.534 | 15.136 | 1.323 | — | K0 | 5 | C | S |
| 5543482 | 0.2948 | WU | 0.80 | 1.154 | 16.171 | 15.680 | 0.667 | — | K2.5 | — | C | — |
| 5557368 | 0.2979 | WU | 0.93 | 1.086 | 13.220 | 12.756 | 0.765 | — | K2 | 3, 5, 16, 19 | C | F, S |
| 5563814 | 0.3082 | WU | 0.91 | 1.135 | 15.634 | 15.184 | 0.997 | — | K2 | 3, 5 | C | F, S |
| 5615528 | 0.3392 | Al | 0.70 | 0.852 | 14.912 | 14.672 | 1.714 | — | G5 | — | M | TV |
| 5774375 | 1.5500 | Al | 0.60 | 1.092 | 14.375 | 13.915 | 3.604 | — | K2 | 2, 3 | C | TV, F |
| 5820209 | 0.6561 | BL | 0.81 | 1.310 | 14.095 | 13.519 | 50.251 | — | K4 | 6, 11 | C | M |
| 5956588 | 0.2729 | WU | 0.83 | 1.083 | 16.217 | 15.755 | 0.625 | — | K2 | — | C | — |
| 5984000 | 0.3231 | WU | 0.82 | 0.998 | 14.845 | 14.435 | 1.074 | — | K0.5 | 3, 5 | C | F, S |
| 6034812 | 0.3673 | WU | 0.93 | 0.765 | 13.143 | 12.844 | 2.633 | F5[1] | F9.5 | 3 | M | TV, F |

Table A.1 *continued on next page*

Table A.1 *(continued)*

| *Kepler* ID[a] | Period (days) | LC Class[b] | Morph. Param. | Color | BP Mag. | G Mag. | Lum. ($L_\odot$) | Spectral Type Lit.[c] | Spectral Type Color[d] | Ref.[ef] | Samp.[g] | Notable Type[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 6044064 | 5.0632 | Al | 0.54 | 1.070 | 15.527 | 15.081 | 6.145 | — | K1.5 | 2, 3, 13, 23* | C | CU, F |
| 6050116 | 0.2399 | WU | 0.79 | 1.280 | 14.948 | 14.392 | 0.262 | — | K3.5 | 3, 5, 19†, 24* | C | F, S |
| 6103049 | 0.6432 | Al | 0.59 | 0.931 | 15.462 | 15.083 | 3.839 | — | G9 | 2, 7 | M | F |
| 6106771 | 0.2638 | WU | 0.77 | 1.032 | 16.041 | 15.598 | 0.554 | — | K1.5 | 3 | C | F |
| 6128248 | 0.2917 | WU | 0.76 | 1.238 | 15.808 | 15.267 | 0.421 | — | K3 | 5 | C | S |
| 6191574 | 1.0345 | Al | 0.50 | 1.560 | 15.013 | 14.310 | 0.103 | — | K6 | 2 | M | TV, F |
| 6197038 | 9.7517 | Al | 0.46 | 1.224 | 13.919 | 13.381 | 7.136 | — | K3 | 2 | C | TV, CU, F |
| 6205460 | 3.7228 | Al | 0.58 | 1.065 | 12.957 | 12.558 | 4.282 | — | K1.5 | 1, 2, 3, 13, 16, 19 | C | AM, F |
| 6223646 | 0.3649 | WU | 0.84 | 0.474 | 13.655 | 13.482 | 4.600 | B9 III[1] | F2 | 19 | C | — |
| 6264091 | 0.3250 | WU | 0.97 | 0.936 | 15.524 | 15.144 | 1.685 | — | G9 | 3, 5 | C | TV, F, S |
| 6267702 | 0.2832 | BL | 0.90 | 1.128 | 16.021 | 15.783 | 0.469 | — | K2 | 3 | M | TV, F |
| 6283224 | 0.7665 | Al | 0.57 | 1.128 | 16.021 | 15.541 | 0.825 | — | K2 | 2 | M | F |
| 6421483 | 2.1584 | Al | 0.58 | 1.119 | 13.595 | 13.120 | 2.415 | — | K2 | 3 | C | TV, AM, F |
| 6424124 | 0.3855 | WU | 0.91 | 0.872 | 14.275 | 13.920 | 2.102 | F7 III[1] | G6 | 3, 5, 6, 19 | C | F, S, M |
| 6431545 | 0.3168 | WU | 0.91 | 0.963 | 14.146 | 13.751 | 1.230 | G2 V[1] | G9 | 3, 5*, 7*‡ | M | TV, F, S |
| 6436038 | 0.3763 | WU | 0.96 | 0.917 | 15.588 | 15.219 | 1.949 | — | G8 | 3 | C | TV, F |
| 6471048 | 0.2972 | WU | 0.77 | 1.159 | 13.287 | 12.784 | 0.549 | K0 V[1] | K2.5 | 5, 19 | C | S |
| 6615041 | 0.3401 | WU | 0.76 | 0.877 | 14.056 | 13.713 | 1.304 | Flat[1] | G7 | 5, 7, 19 | C | S |
| 6697716 | 1.4432 | Al | 0.44 | 1.172 | 14.918 | 14.417 | 0.828 | — | K2.5 | — | M | CU |
| 6791604 | 0.5288 | BL | 0.70 | 0.555 | 14.098 | 13.896 | 5.373 | — | F4 | 6 | C | M |
| 6836140 | 0.4877 | WU | 0.73 | 1.139 | 15.215 | 14.737 | 0.950 | — | K2.5 | 1, 3, 5 | C | TV, F, S |
| 6962018 | 1.2699 | Al | 0.45 | 1.657 | 15.577 | 14.785 | 0.132 | — | K6.5 | — | C | — |
| 7035139 | 0.3097 | WU | 0.80 | 1.011 | 15.950 | 15.543 | 0.851 | — | K1 | 5 | M | S |

Table A.1 *continued on next page*

Table A.1 *(continued)*

| *Kepler* ID[a] | Period (days) | LC Class[b] | Morph. Param. | Color | BP Mag. | G Mag. | Lum. ($L_\odot$) | Spectral Type Lit.[c] | Spectral Type Color[d] | Ref.[ef] | Samp.[g] | Notable Type[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7037319 | 0.5781 | BL | 0.84 | 0.592 | 13.808 | 13.587 | 9.375 | — | F5 | 3 | C | F |
| 7119757 | 0.7429 | Al | 0.64 | 1.134 | 16.118 | 15.636 | 1.839 | — | K2 | 2, 3, 5, 7 | M | TV, F, S |
| 7173910 | 0.4022 | WU | 0.82 | 0.878 | 14.719 | 14.360 | 2.263 | — | G7 | 5 | M | S |
| 7199183 | 0.3241 | WU | 0.97 | 1.003 | 16.086 | 15.667 | 1.012 | — | K0.5 | 3 | C | TV, F |
| 7199353 | 0.2820 | WU | 0.78 | 1.054 | 14.953 | 14.530 | 0.623 | — | K1.5 | 3, 5 | C | F, S |
| 7259917 | 0.3847 | BL | 0.70 | 0.711 | 12.132 | 11.867 | 2.267 | F3 V[1] | F9 | 5, 17, 19 | C | S |
| 7272739 | 0.2812 | WU | 0.77 | 1.060 | 13.445 | 12.982 | 0.588 | — | K1.5 | 3, 5, 7, 17, 19[†] | C | F, S |
| 7284688 | 0.6460 | Al | 0.64 | 0.942 | 11.918 | 11.930 | — | G2 V[1] | G9 | 2, 5, 9 | M | TV, F, S |
| 7367833 | 0.2862 | Al | 0.77 | 0.940 | 15.793 | 15.412 | 1.864 | — | G9 | — | M | — |
| 7433513 | 0.7011 | Al | 0.70 | 0.966 | 15.872 | 15.482 | 1.319 | — | K0 | — | C | TV, AM |
| 7458285 | 0.6606 | BL | 0.71 | 0.813 | 13.848 | 13.527 | 6.189 | F0 IV[1] | G1 | 6 | C | M |
| 7516345 | 0.4919 | BL | 0.75 | 0.506 | 12.499 | 12.311 | 6.761 | — | F3 | 5, 19[†] | C | S |
| 7542091 | 0.3905 | WU | 0.83 | 0.860 | 12.711 | 12.369 | 2.021 | G1 V[1] | G6 | 3, 4, 5, 17, 19 | C | F, S |
| 7546429 | 0.3521 | WU | 0.92 | 0.998 | 15.608 | 15.196 | 1.553 | — | K0.5 | — | C | — |
| 7584739 | 0.9116 | BL | 0.72 | 0.651 | 13.691 | 13.452 | 10.721 | F3 V[1] | F6 | 3, 6, 17[†], 19[†] | C | F, M |
| 7584826 | 0.6223 | Al | 0.62 | 0.992 | 15.654 | 15.255 | 1.053 | — | K0.5 | 5 | C | TV, S |
| 7610486 | 0.4577 | WU | 0.80 | 0.940 | 16.107 | 15.788 | 2.076 | — | G9 | 5 | C | S |
| 7627059 | 0.3237 | WU | 0.81 | 0.957 | 15.362 | 14.976 | 1.308 | — | G9 | 3 | C | F |
| 7657914 | 0.4745 | BL | 0.72 | 0.581 | 13.178 | 12.979 | 4.336 | F5 Vaw[1] | F5 | 17, 19 | C | — |
| 7667885 | 0.3148 | Al | — | 1.413 | 17.450 | 16.827 | 0.262 | G[2] | K4.5 | 3, 20[*] | C | F |
| 7671594 | 1.4103 | Al | 0.34 | 2.369 | 16.980 | 15.719 | 0.017 | — | M2.5 | 2, 13 | C | TV, CU, F |
| 7685689 | 0.3252 | WU | 0.78 | 0.955 | 15.955 | 15.579 | 1.101 | — | G9 | 7 | C | — |
| 7696778 | 0.3312 | WU | 0.79 | 1.209 | 15.931 | 15.409 | 0.906 | — | K3 | — | C | — |

Table A.1 *continued on next page*

Table A.1 *(continued)*

| *Kepler* ID[a] | Period (days) | LC Class[b] | Morph. Param. | Color | BP Mag. | G Mag. | Lum. ($L_\odot$) | Spectral Type Lit.[c] | Spectral Type Color[d] | Ref.[ef] | Samp.[g] | Notable Type[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7740566 | 0.2971 | WU | 0.81 | 0.951 | 16.030 | 15.696 | 1.019 | — | G9 | 3 | C | F |
| 7773380 | 0.3076 | WU | 0.94 | 1.041 | 14.863 | 14.434 | 0.927 | — | K1.5 | 3, 4, 5 | C | F, S |
| 7871200 | 0.2429 | WU | 0.75 | 1.335 | 13.614 | 13.030 | 0.315 | — | K4 | 3, 6, 17, 19, 25* | C | F, M |
| 7877062 | 0.3037 | WU | 0.82 | 0.943 | 14.285 | 13.898 | 0.956 | — | G9 | 3, 5, 7, 19 | C | F, S |
| 7879399 | 0.3927 | WU | 0.74 | 0.900 | 12.980 | 12.599 | 1.744 | — | G8 | 19† | C | — |
| 7885570 | 1.7293 | Al | 0.58 | 0.890 | 12.046 | 11.686 | 4.115 | G0 IV[1] | G7 | 1, 2, 3, 19, 26* | M | TV, AM:SC, F |
| 7950962 | 0.8270 | WU | 0.94 | 0.524 | 10.794 | 10.621 | 17.477 | — | F3 | 3, 19 | C | F |
| 7959612 | 0.2432 | WU | 0.81 | 1.297 | 17.513 | 16.947 | 0.258 | — | K4 | — | C | — |
| 8029708 | 0.3136 | BL | 0.75 | 1.165 | 17.215 | 16.731 | 0.499 | — | K2.5 | — | C | — |
| 8112324 | 0.5759 | Al | 0.59 | 1.086 | 15.126 | 14.675 | 0.850 | — | K2 | — | C | TV |
| 8190491 | 0.7779 | BL | 0.95 | 0.826 | 14.554 | 14.226 | 11.116 | — | G2 | 5, 7 | C | S |
| 8211618 | 0.3374 | WU | 0.74 | 0.969 | 16.068 | 15.667 | 1.121 | — | K0 | — | C | — |
| 8211824 | 0.8411 | Al | 0.59 | 1.215 | 16.600 | 16.075 | 0.479 | — | K3 | 2 | C | TV, F |
| 8230815 | 0.3452 | WU | 0.93 | 0.854 | 13.640 | 13.298 | 1.666 | G0.5 V[1] | G5 | 3, 5 | C | F, S |
| 8248967 | 0.5925 | WU | 0.83 | 0.799 | 15.439 | 15.112 | 6.960 | — | G1 | — | C | AM |
| 8285349 | 0.6671 | BL | 0.90 | 0.712 | 12.437 | 12.166 | 3.293 | F3/F5 V[1] | F9 | 2, 3, 5, 18 | M | TV, F, S |
| 8294484 | 1.0127 | Al | 0.58 | 0.832 | 15.152 | 14.827 | 2.269 | — | G3 | — | C | TV, AM |
| 8298344 | 0.3029 | WU | 0.93 | 1.120 | 15.780 | 15.305 | 0.681 | — | K2 | 3, 5 | C | TV, F, S |
| 8386048 | 0.8174 | BL | 0.71 | 1.031 | 17.059 | 16.631 | 4.314 | — | K1.5 | 3 | C | F |
| 8431389 | 0.3511 | WU | 0.84 | 0.938 | 15.058 | 14.689 | 1.774 | — | G9 | 6 | C | M |
| 8479107 | 0.7676 | Al | 0.53 | 1.087 | 15.464 | 15.010 | 0.582 | — | K2 | 2, 5 | M | TV, F, S |
| 8608490 | 1.0828 | Al | 0.51 | 1.125 | 15.427 | 14.955 | 0.563 | — | K2 | 1, 3, 5 | C | F, S |
| 8620561 | 0.7820 | Al | 0.53 | 1.408 | 17.439 | 16.808 | 0.223 | — | K4.5 | — | C | — |

Table A.1 *continued on next page*

Table A.1 *(continued)*

| *Kepler* ID[a] | Period (days) | LC Class[b] | Morph. Param. | Color | BP Mag. | G Mag. | Lum. ($L_\odot$) | Spectral Type Lit.[c] | Spectral Type Color[d] | Ref.[ef] | Samp.[g] | Notable Type[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8690104 | 0.4088 | WU | 0.77 | 0.872 | 15.261 | 14.913 | 1.994 | — | G6 | 5, 6, 7 | C | S, M |
| 8694926 | 0.3715 | WU | 0.80 | 0.995 | 16.050 | 15.693 | 7.570 | — | K0.5 | — | M | — |
| 8696327 | 0.8792 | BL | 0.75 | 0.708 | 14.850 | 14.575 | 9.412 | — | F9 | — | C | AM |
| 8703528 | 0.3999 | WU | 0.75 | 0.990 | 13.411 | 12.992 | 1.462 | — | K0 | 3, 6, 17, 19† | C | F, M |
| 8715667 | 0.4057 | WU | 0.85 | 1.034 | 13.873 | 13.516 | 3.695 | — | K1.5 | 3, 5, 19† | C | F, S |
| 8822555 | 0.8528 | BL | 0.80 | 0.742 | 14.570 | 14.282 | 9.029 | — | F9 | — | C | AM |
| 8842170 | 0.3494 | WU | 0.90 | 1.025 | 15.200 | 14.771 | 1.598 | — | K1 | 10 | C | AM |
| 8846978 | 1.3791 | Al | 0.77 | 1.098 | 13.716 | 13.254 | 2.929 | — | K2 | 3, 5, 11 | C | TV, F, S |
| 8904448 | 0.8660 | BL | 0.74 | 0.516 | 14.008 | 13.814 | 16.603 | A2 V[1] | F3 | 3, 7, 19 | C | F |
| 8934111 | 0.3645 | WU | 0.92 | 0.943 | 15.553 | 15.169 | 1.699 | — | G9 | 3, 5 | C | F, S |
| 8949316 | 0.6044 | Al | 0.51 | 2.247 | 17.317 | 16.148 | 0.046 | — | M2 | 8 | M | — |
| 8977390 | 0.6533 | WU | 0.86 | 0.781 | 13.427 | 13.125 | 8.199 | F0 IV[1] | G0 | 3, 19 | C | F |
| 9071104 | 0.3852 | WU | 0.82 | 0.820 | 13.958 | 13.627 | 2.249 | — | G2 | 4, 5, 18, 19 | C | S |
| 9071373 | 0.4218 | BL | 0.81 | 0.836 | 13.313 | 12.974 | 2.733 | G0 V[1] | G3 | 3, 5, 18 | C | F, S |
| 9091810 | 0.4797 | Al | 0.70 | 0.970 | 13.006 | 12.601 | 1.418 | — | K0 | 1, 3, 5, 7, 11 | C | F, S |
| 9101279 | 1.8115 | Al | 0.58 | 0.535 | 13.916 | 13.721 | 10.832 | — | F4 | 7*, 13 | C | — |
| 9119652 | 2.2162 | Al | 0.43 | 1.156 | 14.821 | 14.327 | 0.864 | — | K2.5 | 13 | C | CU |
| 9137992 | 1.4700 | Al | 0.64 | 0.936 | 13.388 | 13.004 | 3.479 | G5.5 IIb[1] | G9 | 3, 18 | C | TV, F |
| 9145707 | 0.3208 | WU | 0.75 | 0.993 | 15.152 | 14.740 | 0.927 | — | K0.5 | 3, 5, 6, 18 | C | F, S, M |
| 9145846 | 0.2945 | WU | 0.96 | 1.246 | 15.818 | 15.276 | 0.181 | — | K3.5 | 3, 5 | C | TV, F, S |
| 9161428 | 0.3026 | WU | 0.92 | 0.967 | 15.515 | 15.123 | 1.049 | — | K0 | 3 | C | F |
| 9164694 | 1.1114 | BL | 0.75 | 0.744 | 14.386 | 14.107 | 12.125 | — | F9.5 | 3, 15* | C | AM, F |
| 9274472 | 0.3021 | WU | 0.90 | 1.103 | 14.420 | 13.948 | 0.697 | — | K2 | 3, 5, 18 | C | F, S |

Table A.1 *continued on next page*

Table A.1 *(continued)*

| *Kepler* ID[a] | Period (days) | LC Class[b] | Morph. Param. | Color | BP Mag. | G Mag. | Lum. ($L_\odot$) | Spectral Type Lit.[c] | Spectral Type Color[d] | Ref.[ef] | Samp.[g] | Notable Type[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9283826 | 0.3565 | WU | 0.84 | 0.854 | 13.460 | 13.124 | 1.749 | — | G5 | 3, 5, 6, 7, 19 | C | AM, F, S, M |
| 9290838 | 0.3630 | WU | 0.92 | 1.011 | 15.890 | 15.489 | 1.402 | — | K1 | 3, 5 | C | F, S |
| 9307423 | 0.5586 | WU | 0.73 | 1.402 | 16.762 | 16.150 | 0.965 | — | K4.5 | — | C | — |
| 9328852 | 0.6459 | Al | 0.54 | 1.402 | 15.919 | 15.291 | 0.282 | — | K4.5 | 1, 2, 3, 18† | C | TV, AM:SC, F |
| 9345163 | 0.2451 | WU | 0.77 | 1.242 | 16.018 | 15.478 | 0.305 | — | K3.5 | 3 | C | F |
| 9345838 | 1.0459 | BL | 0.75 | 0.636 | 12.392 | 12.159 | 16.703 | F6 III[1] | F6 | 3, 5, 15*, 17, 19 | C | F, S |
| 9388303 | 0.2667 | WU | 0.78 | 1.095 | 16.075 | 15.652 | 0.432 | — | K2 | 3 | C | F |
| 9419603 | 1.0524 | BL | 0.77 | 0.944 | 15.635 | 15.255 | 5.589 | — | G9 | 3, 5 | M | TV, AM, F, S |
| 9450883 | 0.3637 | WU | 0.94 | 0.897 | 15.543 | 15.188 | 1.689 | — | G8 | 3 | C | TV, F |
| 9466316 | 0.3525 | BL | 0.84 | 1.078 | 16.155 | 15.697 | 1.136 | — | K2 | 5 | M | TV, S |
| 9489411 | 6.6889 | Al | 0.66 | 1.666 | 14.559 | 13.748 | 12.113 | — | K6.5 | 3, 11, 19 | C | F |
| 9527167 | 0.2809 | WU | 0.97 | 0.987 | 15.858 | 15.453 | 0.939 | — | K0 | 3 | C | TV, F |
| 9532591 | 0.3584 | WU | 0.95 | 1.071 | 15.920 | 15.481 | 1.192 | — | K1.5 | 3, 5, 10 | C | F, S |
| 9533706 | 0.2841 | WU | 0.78 | 0.971 | 15.797 | 15.407 | 0.734 | — | K0 | — | C | — |
| 9551200 | 0.7119 | Al | 0.56 | 1.185 | 15.807 | 15.299 | 0.245 | — | K2.5 | 2 | M | TV, F |
| 9654476 | 0.3963 | Al | 0.71 | 0.966 | 15.208 | 14.812 | 1.060 | — | K0 | 3, 5 | C | F, S |
| 9662581 | 0.2580 | WU | 0.79 | 1.221 | 16.213 | 15.681 | 0.367 | — | K3 | 3 | C | F |
| 9664382 | 0.2908 | WU | 0.75 | 1.344 | 17.145 | 16.572 | 2.643 | — | K4 | — | C | — |
| 9717924 | 0.3956 | BL | 0.72 | 0.721 | 17.730 | 17.450 | 1.141 | — | F9 | — | C | AM |
| 9767437 | 0.7127 | Al | 0.67 | 1.131 | 13.296 | 12.814 | 0.705 | — | K2 | 2 | C | TV, F |
| 9772642 | 0.2754 | WU | 0.80 | 1.175 | 14.101 | 13.590 | 0.470 | K2[1] | K2.5 | 3, 5, 19 | C | F, S |
| 9777984 | 0.2585 | WU | 0.76 | 1.592 | 18.471 | 17.733 | 0.177 | — | K6 | 7 | C | — |
| 9786165 | 0.3326 | WU | 0.73 | 1.086 | 17.576 | 17.164 | 1.064 | G[2] | K2 | 3, 20* | C | AM:SC, F |

Table A.1 *continued on next page*

Table A.1 *(continued)*

| *Kepler* ID[a] | Period (days) | LC Class[b] | Morph. Param. | Color | BP Mag. | G Mag. | Lum. ($L_\odot$) | Spectral Type Lit.[c] | Spectral Type Color[d] | Ref.[ef] | Samp.[g] | Notable Type[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9832227 | 0.4579 | WU | 0.95 | 0.794 | 12.599 | 12.284 | 3.562 | — | G1 | 3, 4, 5, 19, 27* | C | F, S |
| 9838047 | 0.4362 | WU | 0.85 | 0.905 | 13.828 | 13.462 | 5.130 | — | G8 | 5, 7 | C | S |
| 9882280 | 0.2891 | WU | 0.77 | 1.010 | 14.988 | 14.560 | 0.656 | — | K1 | 3 | C | F |
| 9935245 | 0.7849 | Al | 0.63 | 1.065 | 16.632 | 16.190 | 1.303 | — | K1.5 | — | C | TV, AM |
| 9935311 | 0.2738 | BL | 0.82 | 1.355 | 15.048 | 14.454 | 0.291 | — | K4.5 | 18 | C | TV |
| 9954225 | 1.3407 | Al | 0.56 | 0.824 | 14.786 | 14.460 | 2.905 | — | G2 | — | C | — |
| 9957351 | 0.4956 | Al | 0.60 | 1.352 | 17.586 | 16.981 | 0.801 | — | K4.5 | — | C | — |
| 9991887 | 0.3224 | WU | 0.88 | 1.077 | 15.633 | 15.178 | 0.949 | — | K2 | 3, 5 | C | F, S |
| 9994475 | 0.3184 | WU | 0.77 | 0.969 | 14.772 | 14.365 | 1.081 | — | K0 | 3, 7, 18 | C | F |
| 10020247 | 0.3456 | WU | 0.83 | 1.035 | 16.409 | 15.973 | 0.870 | — | K1.5 | — | C | — |
| 10123627 | 0.2949 | WU | 0.96 | 1.347 | 16.059 | 15.510 | 0.406 | — | K4 | 3, 5 | C | F, S |
| 10128961 | 0.3477 | WU | 0.86 | 1.007 | 15.945 | 15.553 | 2.238 | — | K1 | 5 | C | S |
| 10189523 | 1.0139 | Al | 0.48 | 1.149 | 16.304 | 15.814 | 0.466 | — | K2.5 | 2 | M | TV, F |
| 10226388 | 0.6607 | WU | 0.77 | 0.465 | 10.861 | 10.691 | 18.761 | F0 IV[1] | F2 | 7, 19 | C | — |
| 10259530 | 0.7072 | WU | 0.86 | 0.423 | 12.210 | 12.067 | 14.057 | — | F1 | 6, 17, 18 | C | M |
| 10322582 | 0.2913 | WU | 0.88 | 0.912 | 15.092 | 14.730 | 0.989 | — | G8 | 5, 6 | C | TV, S, M |
| 10346522 | 3.9891 | Al | 0.59 | 1.021 | 14.794 | 14.386 | 17.732 | — | K1 | — | C | — |
| 10350225 | 0.2821 | BL | 0.80 | 0.995 | 15.373 | 15.014 | 0.771 | — | K0.5 | 3, 5 | C | TV, AM, F, S |
| 10351767 | 0.6351 | Al | 0.58 | 0.801 | 14.833 | 14.649 | 0.794 | — | G1 | 2, 3, 5 | M | TV, F, S |
| 10389982 | 0.4440 | BL | 0.74 | 0.539 | 11.934 | 11.765 | 4.815 | — | F4 | 18 | C | — |
| 10395202 | 0.3059 | WU | 0.81 | 0.991 | 14.974 | 14.573 | 0.978 | — | K0.5 | 3, 18 | C | F |
| 10528299 | 0.3998 | WU | 0.81 | 0.775 | 14.146 | 13.837 | 2.672 | G0 V[1] | G0 | 3, 5, 18 | C | AM, F, S |
| 10544976 | 0.3505 | Al | 0.55 | 1.302 | 19.292 | 18.845 | 0.006 | DA+M4 V[3] | K4 | 3, 28* | M | WD, F |

Table A.1 *continued on next page*

Table A.1 *(continued)*

| *Kepler* ID[a] | Period (days) | LC Class[b] | Morph. Param. | Color | BP Mag. | G Mag. | Lum. ($L_\odot$) | Spectral Type Lit.[c] | Spectral Type Color[d] | Ref.[ef] | Samp.[g] | Notable Type[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10676927 | 1.0311 | Al | 0.56 | 1.016 | 17.298 | 16.868 | 0.919 | — | K1 | 2, 8 | C | TV, AM, F |
| 10680475 | 0.3498 | WU | 0.78 | 0.788 | 14.275 | 13.970 | 1.898 | — | G0 | 3, 5, 19 | C | F, S |
| 10711646 | 0.7377 | Al | 0.51 | 1.530 | 16.496 | 15.798 | 0.148 | — | K5.5 | — | C | TV |
| 10723143 | 0.3076 | WU | 0.79 | 1.108 | 15.790 | 15.330 | 0.662 | — | K2 | 5 | C | S |
| 10727655 | 0.3534 | WU | 0.74 | 0.872 | 13.760 | 13.412 | 1.684 | — | G6 | 3, 7, 19, 25* | C | F |
| 10737574 | 0.3656 | WU | 0.90 | 0.927 | 14.213 | 13.834 | 1.476 | G5[1] | G9 | 3 | C | F |
| 10743600 | 0.8170 | Al | 0.58 | 0.886 | 18.562 | 18.194 | 0.986 | — | G7 | — | C | — |
| 10797126 | 0.4324 | BL | 0.77 | 0.711 | 14.148 | 13.871 | 3.535 | — | F9 | 5 | C | S |
| 10802917 | 0.2725 | WU | 0.74 | 1.235 | 16.555 | 15.998 | 0.353 | — | K3 | — | C | — |
| 10815379 | 0.8905 | BL | 0.79 | 0.700 | 11.465 | 11.199 | 10.340 | — | F8 | 3, 19 | C | F |
| 10854621 | 0.3759 | WU | 0.81 | 0.861 | 15.615 | 15.271 | 1.888 | — | G6 | 5 | C | S |
| 10857342 | 2.4159 | Al | 0.57 | 0.262 | 14.454 | 14.359 | 21.036 | — | A7 | — | C | — |
| 10861842 | 0.4719 | BL | 0.76 | 0.578 | 14.298 | 14.078 | 5.138 | F1 V[1] | F5 | 3, 19 | C | AM, F |
| 10877703 | 0.4369 | WU | 0.80 | 0.909 | 14.759 | 14.392 | 2.393 | — | G8 | 5, 19 | C | S |
| 10924462 | 0.3754 | Al | 0.67 | 1.201 | 14.121 | 13.603 | 0.443 | — | K3 | 3, 5 | C | F, S |
| 10979669 | 0.9258 | BL | 0.83 | 0.609 | 12.478 | 12.248 | 17.801 | F7 IV[1] | F5 | 3 | C | F |
| 11017077 | 0.7086 | BL | 0.68 | 0.810 | 13.410 | 13.087 | 4.421 | F5 V[1] | G1 | 2, 5, 18, 19 | C | F, S |
| 11036301 | 0.2852 | WU | 0.95 | 1.213 | 15.543 | 15.018 | 0.597 | — | K3 | 3, 5 | C | F, S |
| 11076176 | 0.7000 | Al | 0.58 | 1.123 | 16.154 | 15.671 | 0.623 | — | K2 | 2 | M | TV, AM, F |
| 11084647 | 0.3117 | WU | 0.83 | 1.019 | 15.746 | 15.332 | 1.000 | — | K1 | 5 | C | S |
| 11127048 | 0.2993 | WU | 0.94 | 0.985 | 15.636 | 15.227 | 0.911 | — | K0 | 3, 5 | C | TV, F, S |
| 11154110 | 0.5297 | BL | 0.74 | 0.585 | 13.647 | 13.426 | 6.240 | — | F5 | 18†, 19† | C | — |
| 11175495 | 2.1910 | BL | 0.66 | 0.291 | 12.998 | 12.896 | 13.763 | A4 V[1] | A7 | 12, 29* | M | — |

Table A.1 *continued on next page*

Table A.1 *(continued)*

| *Kepler* ID[a] | Period (days) | LC Class[b] | Morph. Param. | Color | BP Mag. | G Mag. | Lum. ($L_\odot$) | Spectral Type Lit.[c] | Spectral Type Color[d] | Ref.[ef] | Samp.[g] | Notable Type[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 11193447 | 0.6259 | Al | 0.64 | 0.768 | 13.439 | 13.618 | — | — | F9.5 | 2, 5 | M | TV, F, S |
| 11198068 | 0.4002 | Al | 0.71 | 1.168 | 16.019 | 15.513 | 0.850 | — | K2.5 | 5 | C | S |
| 11244501 | 0.2967 | BL | 0.89 | 0.982 | 10.251 | 9.840 | 0.954 | G2 V[1] | K0 | 3, 5, 19 | M | TV, F, S |
| 11246163 | 0.2792 | WU | 0.79 | 1.024 | 15.032 | 14.603 | 0.623 | — | K1 | 7 | C | — |
| 11247386 | 0.3943 | BL | 0.76 | 0.471 | 14.442 | 14.269 | 9.022 | — | F2 | 7*‡ | C | — |
| 11295026 | 0.2909 | WU | 0.90 | 1.053 | 13.188 | 12.749 | 0.780 | — | K1.5 | 3, 5 | C | F, S |
| 11303416 | 0.3208 | WU | 0.94 | 1.213 | 15.606 | 15.080 | 0.720 | — | K3 | 5 | C | S |
| 11336707 | 0.2513 | WU | 0.91 | 1.327 | 15.801 | 15.476 | 0.302 | — | K4 | 3, 5 | C | F, S |
| 11341950 | 0.2884 | WU | 0.74 | 1.032 | 15.419 | 14.994 | 0.664 | — | K1.5 | 3, 18† | C | F |
| 11347875 | 3.4551 | Al | 0.86 | 1.491 | 14.806 | 14.124 | 1.306 | — | K5.5 | 2*‡, 3, 19 | C | TV, AM, CU, F |
| 11395645 | 0.4969 | BL | 0.75 | 0.674 | 15.235 | 14.968 | 5.903 | — | F7 | 18 | C | AM |
| 11404698 | 0.5902 | Al | 0.58 | 1.569 | 15.569 | 14.863 | 0.167 | — | K6 | 2 | C | TV, F |
| 11494583 | 0.2483 | WU | 0.94 | 1.341 | 16.047 | 15.452 | 0.308 | — | K4 | 3, 5, 6 | C | F, S, M |
| 11498689 | 0.3063 | WU | 0.77 | 1.017 | 15.382 | 14.964 | 0.850 | — | K1 | 5, 18 | C | S |
| 11566174 | 0.2769 | WU | 0.77 | 1.005 | 15.425 | 15.022 | 0.718 | — | K0.5 | 3, 5 | C | F, S |
| 11671660 | 8.7103 | WU | 0.76 | 1.237 | 13.836 | 13.294 | 10.755 | — | K3 | 11 | C | TV |
| 11704155 | 0.2739 | WU | 0.94 | 0.995 | 13.048 | 12.639 | 1.041 | G8 V[1] | K0.5 | 3, 5, 19 | C | TV, F, S |
| 11717798 | 0.3747 | WU | 0.84 | 0.974 | 14.447 | 14.034 | 1.496 | — | K0 | 3, 5, 6 | C | F, S, M |
| 11819135 | 1.9016 | Al | 0.57 | 0.693 | 15.253 | 14.988 | 5.553 | — | F8 | 12‡ | C | — |
| 11920266 | 0.4674 | WU | 0.91 | 0.698 | 11.789 | 12.083 | 2.619 | F7 IV[1] | F8 | 3, 5 | M | TV, F, S |
| 11924311 | 0.4451 | BL | 0.77 | 0.701 | 13.091 | 12.982 | 5.239 | F3/F5 V[1] | F8 | 6, 18, 19 | C | AM, M |
| 12055517 | 0.6098 | BL | 0.77 | 1.250 | 16.409 | 15.866 | 1.105 | — | K3.5 | 2 | C | TV, F |
| 12104285 | 0.2515 | WU | 0.87 | 1.323 | 16.440 | 15.845 | 0.297 | — | K4 | 3, 5 | C | F, S |

Table A.1 *continued on next page*

Table A.1 *(continued)*

| *Kepler* ID[a] | Period (days) | LC Class[b] | Morph. Param. | Color | BP Mag. | G Mag. | Lum. ($L_\odot$) | Spectral Type Lit.[c] | Spectral Type Color[d] | Ref.[ef] | Samp.[g] | Notable Type[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 12109575 | 0.5317 | Al | 0.64 | 1.433 | 14.002 | 13.352 | 0.231 | K5 Ve[1] | K5 | 2, 5 | C | TV, F, S |
| 12164751 | 2.6301 | Al | 0.60 | 0.828 | 14.467 | 14.136 | 12.426 | — | G3 | 3, 13 | C | F |
| 12218858 | 0.4267 | Al | 0.65 | 0.948 | 13.800 | 13.426 | 0.814 | — | G9 | 18 | M | TV |
| 12400729 | 0.9317 | Al | 0.52 | 1.211 | 15.883 | 15.355 | 0.301 | — | K3 | — | M | TV |
| 12458797 | 0.2413 | WU | 0.94 | 1.268 | 14.434 | 13.876 | 0.317 | — | K3.5 | 3 | C | TV, F |
| 12602985 | 0.2509 | BL | 0.88 | 1.232 | 15.795 | 15.259 | 0.533 | — | K3 | 3 | M | TV, F |

[a] Link leads to system's KEBC page

[b] Al – Algol; BL – $\beta$ Lyrae; WU – W Ursae Majoris

[c] 1 – Frasca et al. (2016); 2 – Ramsay et al. (2014); 3 – Almenara et al. (2012)

[d] Calculated using Dr. Mamajek's table; a luminosity class of V is implied for each system

[e] 1 – Balona (2015); 2 – Gao et al. (2016); 3 – Davenport (2016); 4 – Tran et al. (2013); 5 – Balaji et al. (2015); 6 – Kouzuma (2018); 7 – Borkovits et al. (2016); 8 – Furlan et al. (2017); 9 – Smith et al. (2015); 10 – Pizzocaro et al. (2019); 11 – Gaulme et al. (2013); 12 – Liakos & Niarchos (2017); 13 – Kjurkchieva et al. (2017); 14 – Matson et al. (2017); 15 – Kobulnicky et al. (2022); 16 – Hartman et al. (2004); 17 – Hoffman et al. (2008) or Hoffman et al. (2009); 18 – Devor et al. (2008); 19 – Pigulski et al. (2009); 20 – Ramsay et al. (2014); 21 – Clark Cunningham et al. (2019); 22 – Debski et al. (2014); 23 – Yoldaş (2021); 24 – NegmEldin et al. (2019); 25 – Blättler & Diethelm (2000a) or Blättler & Diethelm (2000b); 26 – Kunt & Dal (2017); 27 – Molnar et al. (2017), Socia et al. (2018), Pavlenko et al. (2018), and Kovács et al. (2019); 28 – Almenara et al. (2012) and Almeida et al. (2019); 29 – Liakos (2017)

[f] * – paper discusses this system; † – signal produced by this system is included in the paper's sample but is conflated with another source; ‡ – system is mentioned in the paper but is not included in the paper's sample

[g] C – Core Sample; M – Marginal Sample

[h] TV – Temporal Variations; AM – Asymmetric Minima; WD – White Dwarf; CU – Concave-Up; :SC indicates the system shows the indicated feature only with its short cadence data; F – Flares; S – Spots; M – Mass Transfer