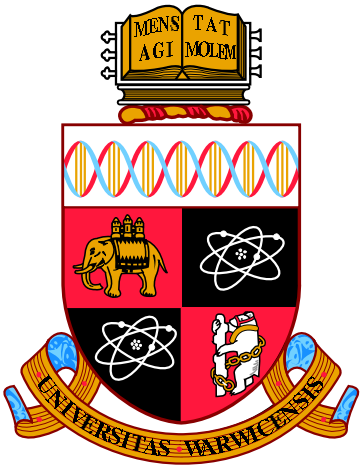

On The Abundance Of Circumbinary Exoplanets

by

David John Armstrong

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Physics

March 2015

THE UNIVERSITY OF WARWICK

Contents

List of Tables

List of Figures

Acknowledgments

I would like to thank firstly my supervisor, Don Pollacco, who has kept on providing advice throughout four years, two universities, and more unscheduled office visits than I can count. Without him this thesis would never have been submitted. He has managed to give me both a great degree of freedom while at the same time making it clear that most projects needed doing by last week, a difficult thing to pull off! Also in need of thanks are both of my second supervisors, Chris Watson and Pete Wheatley, for all the help they have given.

I would also like to thank all of the people at Warwick and QUB, in the Astronomy departments, the rest of Physics, Mountains (in particular for swig points) and Karate. All of them have made this PhD much more fun and interesting than it could have been without them.

Most importantly I would like to thank my family, who have supported me through more years of university than they probably expected without complaint. They have always encouraged me, and without them I would never have been able to take my interests this far. Finally this acknowledgement would never be complete without Nessa, who has been incredibly supportive and kind for the last two years, and who provided a great example of how to finish a PhD under any circumstances.

Declarations

Several chapters of this thesis has been published, in the papers Armstrong et al. (2013b, Chapter 3), Armstrong et al. (2012, Chapter 5), Armstrong et al. (2013a, Chapter 6), and Armstrong et al. (2014, Chapter 7). Chapter 3 contains some collaborative work, specifically in the N-body simulation code used to test the derived transit timing variation limits. This code was built and run by David Martin at the Geneva Observatory.

The majority of this thesis was written in August 2014, and so developments in the field between that date and the date of submission may not be recorded.

Abstract

Circumbinary planets are bodies that orbit both components in a binary star system. This thesis focuses on transits of these planets, which with the aid of the Kepler space telescope have recently led to the discovery of several such objects. First, transit timing variations - departures from strict periodicity in the transit times - are studied. These arise from both the motion of the host stars and relatively rapid precession of the planet's orbit. Limits on the maximum possible transit timing variation are derived, and tested against N-body simulations of simulated circumbinary systems.

These limits are then utilised to form a search algorithm designed to find these planets in light curves, focusing on data from the WASP and Kepler observatories. This search algorithm uses an individual transit search to identify potential transit signatures, then forms periodograms allowing for the possible timing variations. It is used to identify several new candidate planets, as well as confirm detections of previously known circumbinary systems. In addition a number of interesting multiple stellar systems are identified including the as yet unexplained KIC2856960, which display multiple eclipses, significant tidal heating or rapid orbital evolution on the timescale of the 4 year Kepler observations.

In 2013 unbiased stellar radii for the eclipsing binaries of the Kepler dataset were not available. A catalogue is produced, derived from spectral energy distribution fits to data from the KIS, HES and 2MASS surveys of the Kepler field, which gives temperatures for these stars accurate to ~300K. These are then used to find calibrated stellar masses and radii. These parameters, in combination with the search algorithm, are used to study circumbinary planet rates of occurrence in the Kepler dataset. The known sample of eclipsing binaries is tested for detectability, and a Monte Carlo population synthesis used to find probability density functions for these rates. These are a function of the as yet unknown circumbinary planetary inclination distribution, and show that the rate of occurrence of circumbinary planets is consistent with that of single stars if these planets are in the majority coplanar with their host binaries. However, if they are more misaligned, to a degree greater than that implied by a 5° Gaussian distribution, their rate of occurrence becomes significantly higher. Furthermore, it is confirmed that planets of Jupiter size and greater occur less often in circumbinary configurations, and that circumbinary planets are preferentially found around binaries with periods longer than ~7 days.

Chapter 1

Introduction

1.1 History

The first extrasolar planet discovered was found in a binary star system. γ Cephei Ab orbits the primary star of the γ Cephei system, and was first tentatively claimed as a planet by Campbell et al. (1988) (though not generally accepted until 2003, Hatzes et al., 2003). Despite the study of planets in binaries having seen much recent development their history starts with that of exoplanets in general.

Binaries make up a large fraction of stars in the Galaxy; in all more than half of nearby F7-M4 type dwarfs are in multiple star systems (Raghavan et al., 2010; Delfosse et al., 2004; Halbwachs et al., 2003b; Fischer & Marcy, 1992; Duquennoy & Mayor, 1991). If they harbour planets with the same occurrence rates as single stars (e.g. Howard et al., 2010), then their planets represent a vast and in comparison far less explored population. For the first 10-15 years of planet discoveries, the focus for planets in dual star systems was on those planets which orbited one member of the system (aside from rare departures, Sigurdsson, 2003) and particularly giant planets in these systems due to observational constraints. If the binary separation was wide enough, these could often be treated as planets around single stars, with some observational adjustments. Indeed, a common method of searching for these systems is to target the hosts of already known planets and look for stellar companions (e.g. Mugrauer et al., 2014; Baines et al., 2008). These circumstellar planets presented interesting theoretical challenges, and the formation processes leading to them were not (and in many ways still aren't) understood. It was thought that for close stellar companions (in this context $<$ ~100 AU) planet formation could be inhibited through for example disc truncation (Artymowicz & Lubow, 1994; Zsom et al., 2011; Alexander, 2012), and affected by processes such as heating or stirring caused by

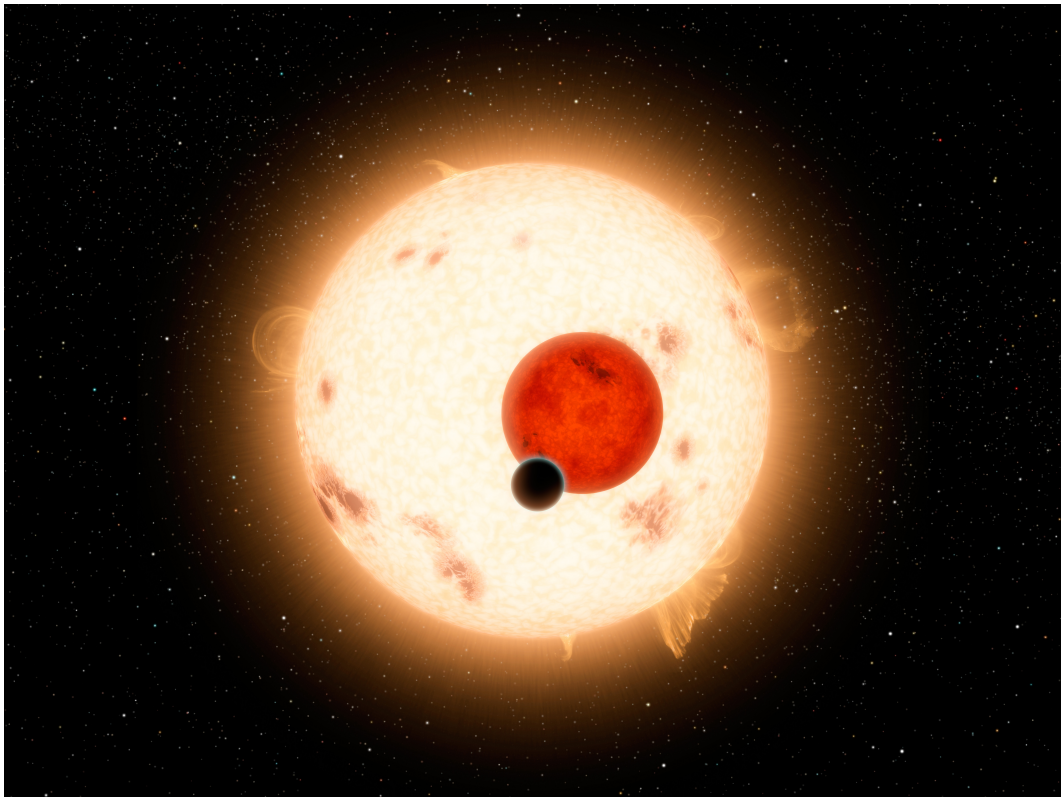

Figure 1.1: Artist's impression of the circumbinary planet Kepler-16b (Credit NASA)

the torque of the companion (Thébault et al., 2006; Boss, 2006; Mayer et al., 2005). Moreover, simulations showed that the two main theories of planet formation (core accretion and disc instability, see Section 1.2.2) should produce planets at different rates for these close binaries (Duquennoy & Mayor, 2005), raising the intriguing possibility of testing these formation theories using the relevant planetary occurrence rates. The effect of stellar multiplicity on planet occurrence rate has been heavily studied in the context of these circumstellar planets, and it has been shown that the rate is comparable to or lower than the single star case in multiple stellar systems (e.g. Bergfors et al., 2012; Lillo-Box et al., 2012; Roell et al., 2012; Mugrauer & Neuhäuser, 2009; Adams et al., 2013). Recently Wang et al. (2014) have shown that planet formation is likely inhibited by a close stellar companion within 20 AU.

Within the population of planets in binary star systems is the even smaller cadre of circumbinary (CB) planets. These planets orbit both stars of their host binary, and present an very different area of study to their circumstellar cousins. Table 1.1 shows the current state of this population, discovered through a wide range of detection methods. Each method has its strengths (see Section 1.2.4) but it is only recently with the Kepler space telescope that the transit method has come into play, due to the long periods and diluted transits typical of CB planets. Kepler has produced the first systems with sub-jupiter mass planets and led to a significant increase in the number of planets available to work with.

The circumbinary planets present several questions. High among them, as with the circumstellar sample, is their formation. How does planet formation proceed in a circumbinary disc? A number of theoretical studies have been performed, and are discussed in Section 1.2.2. They could also shed light on circumbinary discs (e.g. Alexander, 2012), which represent the first stage of formation. Observationally to date there are few constraints. This has now begun to change with the recent Kepler sample, from which initial constraints may be derived which will be important indicators of the history of CB systems. Their distributions may show the imprint of formation pathways as well as subsequent dynamics, and the effect of scattering or migration. In all CB planets represent a rich area of investigation, which shows a great deal of promise.

1.2 Planets in Binary Star Systems

1.2.1 Classification

Planets in binary systems are classified into two main categories, S and P. S-type planets orbit one component of a binary system, with the other stellar companion

Table 1.1: Currently known Circumbinary planets. Unknown quantities are left blank. Note that several planets discovered through eclipse timing are not included, as doubt has been expressed over the stability of the planets in these systems (e.g. Wittenmyer et al., 2013a). See also Zorotovic & Schreiber (2013) for an interesting analysis of planets orbiting post-common-envelope binaries.

Planet	Method	P_{planet} (Sep)	P_{planet}/P_{binary}	R_{planet}	M_{planet}	e_{planet}	Reference
Kepler-16b	Transit	228.8d	5.57	0.754 R_{Jup}	0.333 M_{Jup}	0.0069	Doyle et al. (2011)
Kepler-34b	Transit	288.8d	10.4	0.764 R_{Jup}	0.220 M_{Jup}	0.18	Welsh et al. (2012)
Kepler-35b	Transit	131.5d	6.34	0.728R_{Jup}	0.127M_{Jup}	0.042	Welsh et al. (2012)
Kepler-38b	Transit	105.6d	5.62	4.35 R_{Earth}	0.384M_{Jup}		Orosz et al. (2012a)
Kepler-47b	Transit	49.5d	6.65	2.98 R_{Earth}		< 0.035	Orosz et al. (2012b)
Kepler-47c	Transit	303.2d	40.7	4.61 R_{Earth}		< 0.411	Orosz et al. (2012b)
PH1/ Kepler-64b	Transit	138.5d	6.9	0.52–0.56 R_{Jup}	< 0.531 M_{Jup}	0.054–0.1	Schwamb et al. (2013)
							Kostov et al. (2013)
Kepler-413b	Transit	66.3d	6.55	4.3 R_{Earth}	67 M_{Earth}	0.12	Kostov et al. (2014)
NN Ser c	Ecl Timing	7.9yr	~22000		2.3 M_{Jup}		Marsh et al. (2013)
							Beuermann et al. (2010)
NN Ser d	Ecl Timing	15.3yr	~43000		7.3 M_{Jup}		Marsh et al. (2013)
							Beuermann et al. (2010)
PSR B1620-26	Psr Timing	99.5yr	~190		2.5 M_{Jup}		Sigurdsson (2003)
ROXs 42B	Imaging				6–14 M_{Jup}		Bowler et al. (2014)
							Kraus et al. (2013)
FW Tauri	Imaging				6–14 M_{Jup}		Bowler et al. (2014)
							Kraus et al. (2013)
ROSS 458 C	Imaging				6–12 M_{Jup}		Bowler et al. (2014)
SR12 C	Imaging				12–15 M_{Jup}		Bowler et al. (2014)
2MASS J01-55(AB)b	Imaging	(84 AU,proj)			12–14 M_{Jup}		Delorme et al. (2013)

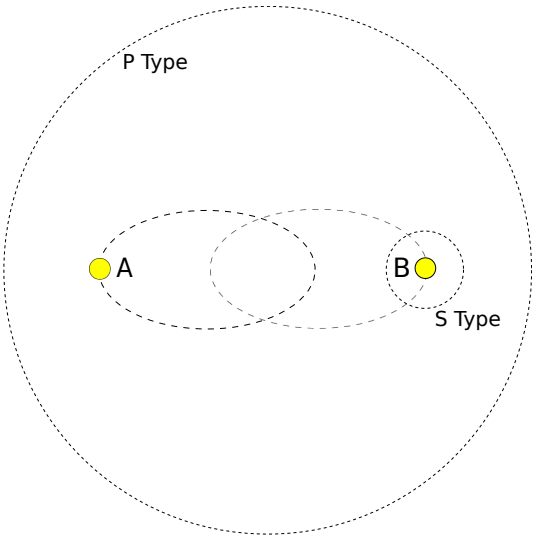

Figure 1.2: S and P type planetary orbits.

at generally much wider separation. P-type systems orbit both components of the binary, and are known as circumbinary planets. The majority of this work focuses on the latter P-type systems.

1.2.2 Planet Formation in Binary Systems

Giant planet formation is thought to occur through one of two main routes. In core accretion, kilometre size Planetessimals (formed from coagulation of dust and grains) must accumulate collisionally, and once a certain mass is reached begin to accumulate gas from the surrounding disc, until stopped by exhaustion of the gas supply. In the alternative method, disc instability, locally bound regions of dust and gas in the disc collapse to form the resulting planets (see Seager et al. (2010) for a review of both). Each of these methods is strongly affected by the relative velocity of the disc material, $\triangle v$. This represents the main area of potential impact of a binary, for both circumstellar and circumbinary discs. Torque from the binary companion can stir up the disc, potentially raising relative velocities. This parameter is crucial - if $\triangle v$ is too high, impacts between Planetessimals will tend to fragment rather than accrete, preventing planet formation and undoing past accretion (Zsom et al., 2011). Moreover, the effective gravitational cross section of a capturing Planetessimal depends strongly on $\triangle v$. This is the distance within which two Planetessimals with defined velocities must pass to become gravitationally bound. For larger relative velocities this cross section rapidly falls, decreasing potential accretion further and reducing the chance of significant areas of the disc becoming bound. The rel-

ative effect of changes in $\triangle v$ is different for each of the main proposed accretion mechanisms, such that if the change in $\triangle v$ caused by a companion star is known, the effect on the planet formation rate becomes an indicator of which mechanism dominates. This formation rate can be tested through the planetary occurrence rate, although this is also dependent on protoplanetary disc occurrences as well as subsequent evolution (such as planet ejection or capture).

In the circumbinary disc case, work by Meschiari (2012) shows that planet formation could be hindered through raised $\triangle v$ over a range out to several AU from the system centre of mass (they used the Kepler-16 system as the host binary, which has a period of 41d). This has implications for the history of the known (transiting) planets, whose orbits lie well within this zone, implying that they must have formed further out and later migrated to their present positions (Kley & Haghighipour, 2014a; Lines et al., 2014; Pelupessy & Portegies Zwart, 2013). However, circumbinary planet formation is far from fully understood, with the specific extent of the formation suppressing region in doubt (Rafikov, 2013). Furthermore, it may be that certain regions in the disc - 'dead zones' - of lower $\triangle v$, in fact aid planet formation relative to the single star case (Martin et al., 2013; Rafikov, 2012). The outcome of these competing effects is an area of ongoing study.

Another possible formation route for circumbinary planets is the so-called second generation option. This is where planets form from material ejected during the later stages of a star's life, such as strong stellar winds or common envelope ejecta. It has been invoked to possibly explain the detection of planets orbiting evolved binaries such as NN Ser (Beuermann et al., 2010). At present this method seems feasible (see e.g. Schleicher & Dreizler, 2014) and can help to explain cases where the survival of planets through stellar evolution is unlikely.

1.2.3 Orbits and Stability

The orbit of a planet in a binary system can be described by the keplerian orbital elements (a, e, f, i, ω, Ω) much as for a single planet. The last three of these elements are defined in Figure 1.3. They are the argument of periapsis, ω, longitude of ascending node, Ω and inclination i and describe the orientation of the orbital plane of the body under consideration, with respect to a defined reference plane and direction. The semimajor axis a, eccentricity e and true anomaly f describe the position and motion of the body within that orbital plane.

For typical binary systems, the reference plane used is the plane of the sky. This remains the case here, with the reference direction generally set so as to make Ω_{binary} zero. However, simultaneously defining the planetary orbit from this refer-

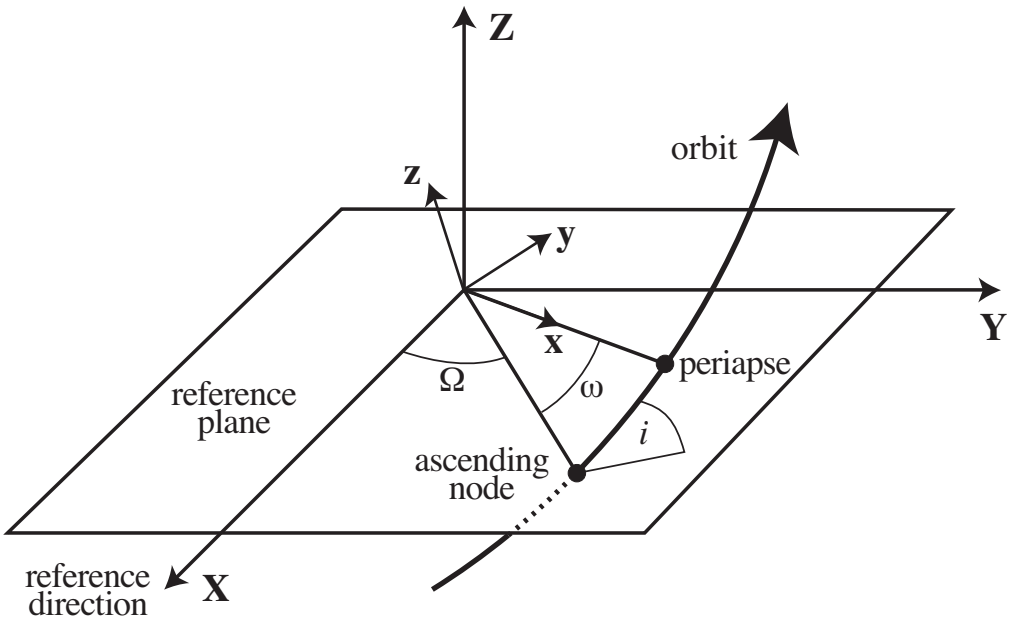

Figure 1.3: Keplerian orbital elements (Credit Seager & Winn, 2010a)

ence can lead to problems. It seems intuitive that the orbit of a circumbinary planet is a function of its past formation history and evolution, both strong functions of the binary (and its presumed past circumbinary disc). As such, the planetary orbital elements are likely to bear some strong relation to the binary's, although the form of this (in terms of planet inclination, eccentricity and period distributions) is still a subject of much debate (see for example Chapter 7). Given this, it makes sense to define these planetary orbital parameters with respect to the binary plane itself. This then becomes the reference plane, and the reference direction the line of nodes (positive to the ascending node) of the binary orbit. The complete three body system is then defined by 11 parameters, 5 (as Ω_{binary} is zero by definition) for the binary and 6 for the planet.

In a Keplerian two body system, all of these parameters excepting the true anomalies are conserved. However, due to the torque of the binary in this case the planetary orbital elements become functions of time. This manifests as a nodal libration, where a torque from the host binary produces a precession in the longitude of the ascending node of the planet's orbit along with long term oscillation in the planetary inclination (see e.g. Doolin & Blundell, 2011; Farago & Laskar, 2010a). Oscillations can also be seen in the eccentricity and semimajor axis (see for example Leung & Lee, 2013). This may have serious consequences for observability, with the

planet Kepler-16b expected to stop transiting in 2018 (and resume in 2042) due to precession of its orbit (Doyle et al., 2011).

There is a long history of analysis of stability in the three-body problem. Recent work includes Jaime et al. (2012), who found stable orbital regions for ~160 binaries in the solar neighbourhood, and Doolin & Blundell (2011), who studied the effect of binary mass fraction and planetary inclination on suites of simulated test particles (see Figure 1.4 for a summary of their results). Prior to these, Dvorak et al. (1989); Holman & Wiegert (1999); Pilat-Lohinger et al. (2002); Musielak et al. (2008) have all studied the issue, with varying constraints. Holman provides particularly useful analytic limits for orbital stability in the zero eccentricity planet case which are utilised later in this work, taking into account varying binary mass ratio and binary orbital eccentricity. The later studies explore the complex parameter space involved, finding among more subtle effects that both high binary eccentricity and lower binary mass fraction serve to increase the minimum stable orbital radius of a circumbinary planet. To give a guiding level, this minimum radius typically requires $P_{planet}/P_{binary} > \sim$4-5. Intriguingly, there are indications of a potential 'pileup' of planets near this minimum radius (Welsh et al., 2013) - almost all of the Kepler transiting circumbinaries orbit close to the minimum. It has been shown that this could be due to migration in the protoplanetary disc, which is then halted near the disc's inner edge where the competing torques balance (Pierens & Nelson, 2007).

1.2.4 Detection Methods

Many of the usual methods of planet detection have been applied to the circumbinary case. As seen in Table 1.1, there has been success using transits (where the planet obscures a proportion of the stellar flux), eclipse timing (where the light travel time effect on the binary eclipses or other periodic signal, due to motion of the host binary induced by the planet is measured, see Section 2.7), and direct imaging (where the planet itself is resolved, see for example the SPOTS project, Thalmann et al., 2013). Radial velocity detection requires removing the signal of the binary orbit, and while no discoveries have been made this way as yet it will likely soon produce further circumbinary detections (see for example the TATTOINE project, Konacki et al., 2009). Each of these methods can present difficulties specific to the circumbinary case, the most relevant of which are discussed in the next section.

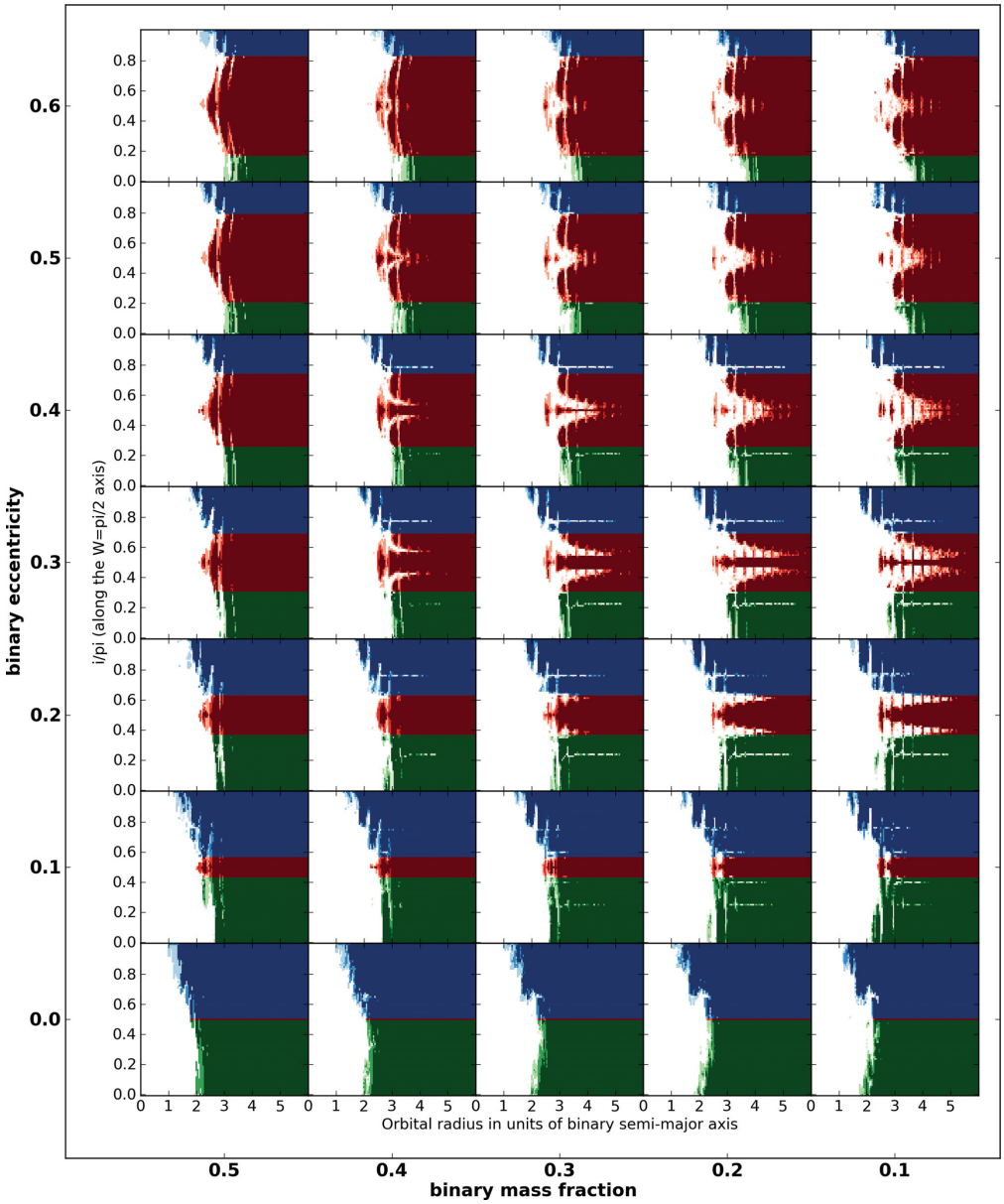

Figure 1.4: Orbital stability plotted as a function of orbital radius a/ab and inclination i on the $\Omega = \pi/2$ axis, across binary eccentricity and mass fraction parameter space. Shaded areas represent stable orbits. Colours: Green: prograde ($i < \pi/2$). Blue: retrograde ($i > \pi/2$). Red: island of libration centred at ($i = \pi/2$, $\Omega = \pi/2$). W represents the longitude of the ascending node, Ω, here. (Credit Doolin & Blundell, 2011)

1.2.5 Detection Issues

There are various binary specific issues associated with the above detection methods. In this work I focus on transits and to a lesser degree eclipse timing, so these are presented with the most detail here.

(a) **Transit Timing Variations**

Planetary transits in the single star case occur with fixed periodicity, duration and depth, in the absence of exceptional factors such as evaporating planets or rapidly evolving orbits (see e.g. Mandel & Agol, 2002, for the basic transit equations). Transit timing variations (TTVs) are possible due to perturbations by other planets in the system (see e.g. Nesvorny et al., 2012; Hadden & Lithwick, 2014) but are typically on the order of seconds or minutes in magnitude. Adding another star, as might be expected, induces changes to this system. TTVs can become not only present but strong and unavoidable. They have two main sources, the geometric timing variation caused by the changing position of the binary stars as they orbit their centre of mass, and a timing variation caused by the sometimes relatively rapid precession of the planet's orbit. Transit durations are also affected due to the changing relative velocity between star and planet, which depends on the binary star's orbital phase. These TTVs are generally on the order of days, far larger than the transit durations, which makes a complete change in search algorithm methodology necessary. See Chapter 3 for more detail on TTVs. Search algorithms have been developed to accommodate such TTVs (Ofir, 2008; Carter & Agol, 2013, also Chapter 4) and along with simple by-eye searching have produced the transiting planets of Table 1.1.

(b) **Azimuthal Period**

This is the time which on average the planet takes to traverse 2π radians in a fixed reference frame - i.e. the time interval between successive conjunctions. It is offset from the Keplerian period which can be derived from the planet's semi-major axis and the binary mass. In Leung & Lee (2013) it is shown that the azimuthal period is shorter than the Keplerian orbital period for circumbinary planets. The effect of this can be seen in many of the published transiting circumbinary planets so far. Taking the observed times of transit of these planets and estimating a period from the mean transit interval (which is equivalent to the azimuthal period), the estimated period is generally found to be a few days under the published Keplerian period. This is not an error, but a mark of the difference between the azimuthal period and Keplerian period. The effect is clear for Kepler-16b: The maximum observed TTV (i.e. transit spread in time

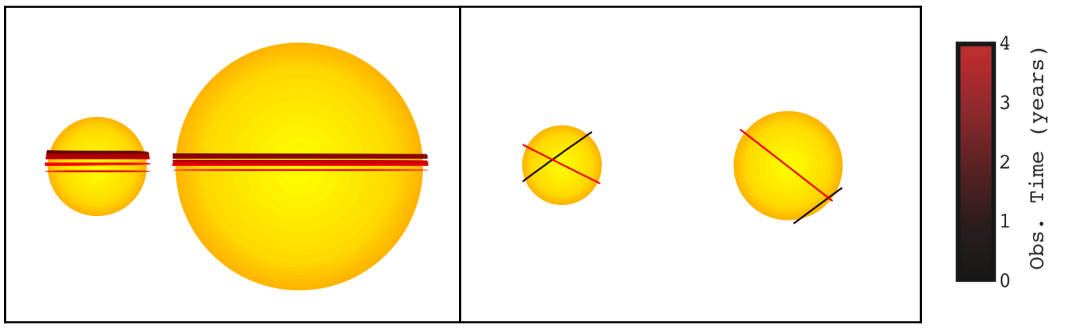

Figure 1.5: Chords of transits observed in an N-body simulation of a planet orbiting a binary star. The 'consecutive' transits (left) occur for a planet coplanar with the binary, and have many more transits than the non-coplanar, 'sparse' case (right). (Credit Martin & Triaud, 2014)

when phase folded) at the published Keplerian period (228.78d) is ~13d, but at the azimuthal period (225.72d), it is ~4.5d, significantly lower. This azimuthal period is the important quantity when considering circumbinary planets from an observational perspective.

(c) **Non-Coplanarity**

If a circumbinary planet is not close to coplanarity with its host binary (such that it is within a few degrees of the binary orbital plane), then due to the motion of the binary stars it will often 'miss' them while crossing the observer's line of sight, exhibiting transits only on some if any orbits and making detection much less likely. This has become known as a 'sparse' transiting system. The constraint is relaxed for binary stars where the mass of one star is much greater than that of its companion (such that the more massive star's orbit is smaller than its radius) or for contact binaries. Furthermore, for systems that are not exactly coplanar, the precession of the planetary orbit (see Chapter 3) will take it in and out of a transiting configuration such as in the case of Kepler-16b. On the other hand, coplanar planets produce what are known as 'consecutive' transits, where at least one transit is seen on every crossing of the binary by the planet. The chords of transits found in an N-body simulation run of test circumbinary planets are shown in Figure 1.5.

(d) **Eccentricity**

As part of the source of TTVs of circumbinary exoplanets is due to precession of the planet's orbit, highly eccentric planets will show more variations. While this does not reduce their detection chances as much as the above points, it

increases the difficulty caused by these variations, further 'blurring' the planet's transit signal. The 'blurring' effect of eccentricity is then scaled by the period of the precession of the planet's orbit. Planets that precess faster will experience more transit timing variations over a given timescale.

(e) **Eclipse Timing**
While this method has no extra disadvantage against the single star case (being not possible there, unless the star has a regular strong periodic signal which can be measured accurately enough to show light travel time) some indications of difficulty have become apparent. Firstly a number of circumbinary planets discovered in this way have later been refuted on stability constraints - the proposed orbits would result in planet ejection or collapse into the host star within timescales on the order of as little as 1000 years in one case (Hinse et al., 2014; Wittenmyer et al., 2013b; Horner et al., 2013). The origin of the original detected signals remains unclear. Should these planets have all been real, it has been shown that the consequences for second generation planet formation are extreme, requiring exceptionally high formation rates (Zorotovic & Schreiber, 2013). Eclipse timing variations can arise through light curve noise or stellar activity, particularly systematic variations such as starspots. Other forms of period variation due to for example mass transfer or the Applegate effect (Applegate, 1992) can lead to non-planetary signals. The eclipse timing method has however been used to successfully place upper limits on planetary masses (Orosz et al., 2012b).

1.2.6 Proposed Trends

Despite the present low numbers of transiting circumbinary planets, some trends have begun to become evident. The transiting planets explore a very different area of parameter space to other methods. At present, and largely due to the Kepler mission, transits are the only method sensitive to the detection of sub-Jupiter mass planets in relatively short (up to ~a few hundred day) orbits. Among these planets Welsh et al. (2013) have recently pointed out that there is a tendency for the planets to be found near the inner stability limit for their host binary (generally P_{planet}/P_{binary}~4-5, Holman & Wiegert, 1999). See Table 1.1 for the actual values. Another interesting trend is the lack of giant planets - although the transit method has a bias towards larger planet radii as they are easier to detect, the majority of the known transiting circumbinary planets have sub-Jupiter radii and masses. The final trend to point out here is with the host binaries themselves. No transiting planet has

been found so far orbiting a binary with period less than 7 days, despite the Kepler sample of binaries being heavily biased towards shorter periods. The investigation of these trends, and of the circumbinary planet population in general, is the central aim of this thesis, and further analysis of them can be found in Chapter 7.

1.3 Eclipsing Binary Stars

1.3.1 Planets Orbiting Eclipsing Binary Stars

When searching for planets via transits, we require relatively close binary stars as targets. This is due to the minimum orbital stability (Section 1.2.3) - even a 5d binary can only host planets with periods greater than ~25d, well above the limit of current ground based transit surveys. In the Kepler or WASP datasets (see Section 1.4), given that a large scale radial velocity survey of the entire dataset is implausible, the main way to find such binaries is via their eclipses. As such, large catalogues of eclipsing binaries exist in both surveys. A further advantage can be found in studying eclipsing binaries over non-eclipsing - there are indications that circumbinary planets should be preferentially coplanar with their hosts (Foucart & Lai, 2013). If true this implies that a circumbinary planet orbiting an eclipsing binary is a priori more likely to transit. Moreover, if it does transit and is coplanar, then so called 'consecutive' transits will occur (Section 1.2.5) making detection much easier. As such, eclipsing binaries represent an ideal sample for initial surveys for these planets.

1.3.2 Information from the Eclipses

A number of parameters can be extracted from the light curve of an eclipsing binary, many of which can be fed back to refine searches for potential planets. The first and most obvious of these parameters is the **binary orbital period**, which can be found via a variety of period searching methods (see Section 2.6).

The **eccentricity** parameters are also derivable. The locations of the primary and secondary eclipses, along with their widths are functions of the eccentricity and argument of periapsis of the binary orbit. Note that primary eclipse is when the primary star (that with the highest surface brightness) is eclipsed by the secondary. The primary star at time $t = 0$ has a true anomaly of zero, whereas the secondary has a true anomaly of π (an alternative way of looking at this is to use different arguments of periapsis, ω, offset by π for each star). A circular orbit produces primary and secondary eclipses of equal width, located at a separation in phase of 0.5.

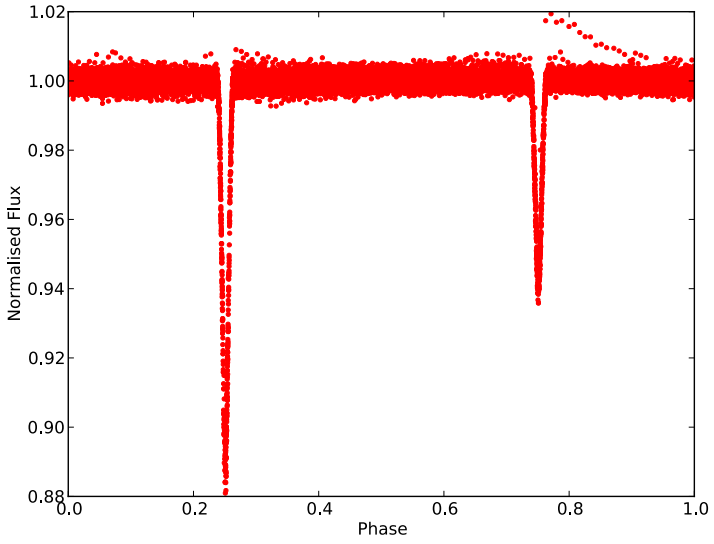

Figure 1.6: The light curve of the eclipsing binary KIC2019056, phase folded at the orbital period of 7.13 days.

The eclipse durations constrain $e \sin \omega$ (Kopal, 1959), through

$$e \sin \omega = \frac{d_{\text{sec}} - d_{\text{pri}}}{d_{\text{sec}} + d_{\text{pri}}} \tag{1.1}$$

with e and ω defined in Section 1.2.3, and d_{pri} and d_{sec} the durations (in either phase or time space) of the two eclipses. $e \cos \omega$ can then be constrained through the eclipse locations (Kallrath & Milone, 2009), via

$$e \cos \omega = \frac{\pi}{2} \left(\phi_{sec} - \phi_{pri} - \frac{1}{2} \right) \tag{1.2}$$

where ϕ represents the phase of eclipse minima. These equations are independent, allowing e and ω to be determined absolutely. They do however involve an $i \approx \frac{\pi}{2}$ approximation, which is good for detached eclipsing binaries (where the stellar separation is much greater than the stellar radii) but less so for contact or near contact binaries where a large range of inclinations can still lead to eclipses. When ellipsoidal variations (the change in flux due to the changing projected surface area of a distorted non-spherical star) are significant, they can be seen over a very wide range of inclination as eclipses are not necessary.

The depths of the eclipses give information on the **temperature ratio** of the two stars. The precise eclipse depth is a function of the chord followed by one star across the other, itself a function of the system inclination, as well as the limb

darkening (see e.g. Claret, 2004) of the star being eclipsed. For a total eclipse (where one star is completely within the projected radius of the other) it is possible to use the ratio of eclipse depths to give the ratio of mean surface brightnesses, which can then be converted into a temperature ratio. A version of this method, which also takes account of partial eclipses, is described in Chapter 6.

It is also possible to derive inclinations and scaled radii from the light curves, through the depths and widths of eclipses relative to the total flux or phase. Inclination in particular is correlated with third light (from another star, whether bound to the binary or not). See Prsa et al. (2011) for a discussion. For contact binaries (where each star overflows its roche lobe) it is also possible to derive mass ratios and 'fill out factors', a measurement of the extent to which this lobe is extended.

1.4 Data Sources

1.4.1 WASP

WASP (the Wide Angle Search for Planets, Pollacco et al., 2006) is a long running mission which has found over 135 giant planets orbiting bright stars. The project has two sites, on La Palma in the northern hemisphere and SAAO (the South African Astronomical Observatory) in the south. The data used in this project come from the northern observatory, which will be focused on here. Each telescope consists of 8 Canon 200mm, f/1.8 telephoto lenses, each with attached science grade e2v CCD detector (Ikon-L devices manufactures by Andor). This results in an extremely large field of view, of ~480 square degrees. With this large field objects covering ~3/4 of the sky have been observed, for up to ~8 years in some cases.

Part of the pipeline for planet detection within WASP involves the use of the BLS algorithm (Kovacs et al., 2002, Section 2.6.1). This algorithm strongly detects eclipsing binaries as well as planets, and while these binaries represent unwelcome false positives for planet detection they are flagged for use elsewhere. This has led to a catalogue of ~7000 eclipsing binaries (as of early 2014), being available within the WASP data. These binaries have periods mostly less than 10 days, and in the large majority of cases less than 5. Above this WASP's sensitivity decreases due to the observing window of its targets. These periods are output from the detection pipeline. However, each object appears in multiple fields within the catalogue (measured using different combinations of CCDs and detrending methods), and each has the BLS algorithm run on it separately. As such, sometimes different periods arise in different fields (generally harmonics of the true period) and this must be dealt with in using the data (see Section 4.2.1).

1.4.2 Kepler

The NASA Kepler satellite is a mission producing extremely high precision, near continuous light curves of ~155,000 stars on the level of 30 ppm for 12th magnitude stars over 6.5 hour timescales (Koch et al., 2010; Batalha et al., 2010a; Haas et al., 2010). This mission began science operations on 13 May 2009, and to date ~1500d of data are publicly available on the NASA Data Archive[1]. After this period the spacecraft ceased operations due to reaction wheel failure, but may begin the new redesigned K2 mission soon (Howell et al., 2014). Due to technical constraints during the primary mission, the satellite must reorient itself each quarter year, and so data is provided as a separate file per object and per quarter, each quarter consisting of three months of observations. 17 quarters were taken, all of which are now publicly accessible. The majority of data are taken in long cadence mode (29.4 minute sampling, formed of summed shorter exposures), while short cadence observations (1 minute sampling) on some selected targets are available. See the Data Characteristics Handbook on the NASA Data Archive[2] for precise observation periods. In raw format, the data show a variety of noise signals, the details of which are explained in the Kepler Data Processing Handbook[2]. Software tools for accessing and manipulating the data files have been made available by the Guest Observer office (Still & Barclay, 2012).

In preparation for the Kepler mission Brown et al. (2011) produced the Kepler Input Catalog (KIC), providing broadband observations and spectral energy distribution fit parameters for stars in the Kepler field of view. The parameters presented by the KIC are used in target selection for both the primary Kepler purpose of planet hunting and other guest observer programs, as well as to provide estimated radii for candidate planets (Batalha et al., 2013). They have been subject to later testing, through for example population synthesis (Farmer et al., 2013).

The satellite has a number of known issues, including a particular problem with Quarter 2 data. In this quarter the guidance system was influenced by two variable stars, leading to unnecessary telescope attitude adjustments and hence systematics in target light curves. From Quarter 3 the guiding algorithm was updated and this ceased to be an issue. Other significant issues occurring over the telescope lifetime include the loss of one CCD module (causing follow on temperature drops across the focal plane), as well as two reaction wheel failures. The first was in Q14 followed by a safe mode event while the second stopped operations. Several less severe instrumental systematics are present, including differential velocity aberra-

[1] http://archive.stsci.edu/kepler/

[2] http://archive.stsci.edu/kepler/documents.html

tion, focus changes correlated with temperature changes, attitude error when the telescope crosses the axis of a reaction wheel such that that wheel has no bearing on the telescope pointing for a short period, feature depth changes between quarters due to varying apertures and contamination levels, and cosmic ray impacts. Many of these effects are removed during detrending. It is worth noting that due to the high precision of the flux data, even if instrumental systematics could all be removed significant astrophysical noise (at least in the context of planet searching) remains in terms of stellar variability. Kepler has provided a large array of hitherto unknown precisely measured variable stars (Debosscher et al., 2011).

Detrending of the data is performed by the Kepler science team using the Presearch Data Conditioning pipeline. This was updated in 2012 and uses a Bayesian maximum likelihood estimation method to remove systematics while in theory preserving astrophysical signals. It works by producing a series of cotrending basis vectors, which are generated from the raw light curves of the 50% most correlated target light curves on a given CCD channel. The strongest components are then used to characterise the elements of noise for that channel. This involves the implicit assumption that the dominant noise signatures are highly correlated based on target proximity (specifically within a channel). These basis vectors are then fit to each light curve, with allowances made for noise correlation with target magnitude and spatial proximity. See Stumpe et al. (2012) for an overview and examples, and Smith et al. (2012) for a full description of the Bayesian fitting. The cotrending basis vectors produced during this process are available, and can be utilised to perform corrections on individual or groups of targets if motivated by a specific science case.

1.4.3 Kepler Eclipsing Binary Catalogue

Many of the targets observed by the Kepler satellite have proven to be eclipsing binaries. These are catalogued in the Kepler Eclipsing Binary Catalogue (KEBC) (Prsa et al., 2011; Slawson et al., 2011; Matijevič et al., 2012), and number well over 2000. With this catalogue as a guide, many interesting results have been found (e.g. Carter et al., 2011; Rappaport et al., 2012; Bloemen et al., 2012; Armstrong et al., 2012; Lee et al., 2013) not least those of the circumbinary planets already described. The catalogue has been hosted online, and the latest version is available at http://keplerebs.villanova.edu. This at the time of writing represents version 3, for the which the accompanying publication is under preparation. After the formal release of version 2 some papers have been released detailing measurements of binary eclipse times made by the Kepler Eclipsing Binary Working Group (Conroy et al., 2014, Orosz et. al. in prep).

Binaries in the catalogue are discovered through the same pipeline which searches for planets (Jenkins et al., 2010). Outputs from this search which are not good planet candidates (due to for example the presence of secondary eclipses, or too deep primaries) are catalogued as eclipsing binaries. A full description of the methods for finding these systems in the Kepler data can be found in Prsa et al. (2011) and the subsequent catalogue papers. Parameters were initially estimated using the EBAI code (Prsa et al., 2008), a trained neural network based system. With EBAI model light curves are used to 'train' the network, allowing it to quickly correlate features of the phased light curve with a defined set of output parameters. This is a promising technique for eclipsing binary investigation, which could allow for rapid human-free preliminary analysis of the millions of binaries which may be discovered by future missions such as Gaia (Perryman et al., 2001).

In versions 1 and 2, temperature ratio, mass ratio, eccentricity, inclination, scaled radius $(R_1 + R_2)/a$ and fill out factor parameters were presented (see Prsa et al. (2011) for full descriptions). These have been removed in the current online version (due to possible errors, Prsa private communication) and remaining are the period and morphology parameter. The Morphology parameter represents the shape of the light curve - it is low for detached binaries, rising to near unity for over contact types. A typical cutoff to separate detached binaries from contact types is between 0.5–0.7. See Matijevič et al. (2012) for a full discussion. Added to these are the so-called polyfit parameters. These result from the fit of a chain of nth order polynomials, connected at 'knots', to the phased binary light curve. The knots are allowed to move. This was found to be an effective fitting method for binary light curves, as demonstrated in Matijevič et al. (2012); Prsa et al. (2008). They provide the depth, width, and separation of the primary and secondary eclipses, allowing temperature ratios to be estimated and eccentricity parameters to be derived as described in Sections 6.2.4 and 1.3.2 respectively.

1.5 Layout

This thesis is laid out as follows. In Chapter 2, several methods used later are discussed. Chapter 3 derives new constraints on the transit timing variations of circumbinary planets, and tests them against an N-body integrator. Chapter 4 describes how these constraints were used to create a search algorithm for the WASP and Kepler datasets, and the data preparation necessary. It also describes some results from the algorithm. Chapter 5 discusses the results in more detail, focusing on the multiple stellar systems which were also discovered. Chapter 6 uses the Kepler

eclipsing binaries along with other photometric surveys to produce spectral energy distribution fits to each binary, and so create a catalogue of stellar temperatures for the Kepler systems. In Chapter 7, each of these sections is combined to allow rates of occurrence of circumbinary planets to be observationally constrained for the first time. Finally Chapter 8 concludes, and suggests possible avenues of future work.

Chapter 2

Methods

In this chapter a number of methods utilised later in the thesis are explained. These cover a wide range of areas, ranging from simulation methods to periodogram and data vetting. Some are well known standards, MCMC for example, whereas the circumbinary population synthesis is a Monte Carlo process developed for specific use here.

2.1 Markov Chain Monte Carlo

Monte Carlo methods are those where random sample generation is used to solve a deterministic problem. In a general Markov chain, the current sample value determines the distribution used to generate the next sample value. In Markov Chain Monte Carlo (MCMC) these principles are combined, such that a distribution of random samples are generated one by one, each sample controlling the generation of the next. These samples can be seen as statistical tests, performed usually on a simulation designed to represent a real problem to be investigated. A sample is equivalent to an experiment, but performed on the simulation rather than a physical setup. As such, MCMCs produce distributions of experimental samples, which can be used to study a wide range of situations.

Practically, MCMCs are used to study the joint posterior distribution of a multivariate problem. The chain of samples produced by the Markov chain can be shown to converge to the desired posterior distribution (Ibe, 2009). A suitable burn in phase is defined to allow this convergence, after which the sample chain is drawn from the posterior distribution, and so can be used to investigate maximum-

likelihood values and associated errors.

Here MCMC procedures were implemented using the python module PyMC (Patil et al., 2010), and the method described follows the underlying code of this module. The most typical method for generating the chain is the Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1969). At each sample stage, parameters are iterated using a jump function. Each variable X is iterated via:

$$X_{i+1} = X_i + D(X)R \tag{2.1}$$

where D(X) is the distribution set for variable X (often gaussian, but not always), from which possible new values are drawn, and R the jump size scale for all parameters. It is possible to make use of known priors for a variable (through defining a particular distribution it should be drawn from). A commonly used prior is for example truncating the normal distribution at some limiting values. R is tuned at defined intervals through the MCMC chain such that ~25% of new sets of parameters are accepted. Whether new sets of parameters are accepted (becoming the next chain value) or not depends on their comparison to the data. For each set, χ^2 is calculated, via

$$\chi^2 = \sum_k \frac{(\text{model}_k - \text{data}_k)^2}{\sigma_k} \tag{2.2}$$

where k represents each data point, data_k and σ_k its value and error, and model_k the current parameter set's prediction for its value. If χ^2 for the new parameter set is less than for the previous set the new set is automatically accepted. If it is higher, the new set is only accepted with probability $\exp(-\triangle\chi^2/2)$. The chain is often then thinned (a certain proportion of samples discarded) to avoid correlations between adjacent sample calls.

This method produces (if a suitable burn in phase was discarded) the joint posterior probability distribution for the set of variables being studied. The χ^2 surface for this parameter space can sometimes resemble a mountainous ice step in the Himalaya, full of unpredictable gullies and local minima. Barring these dangers however, the values of each variable can be used to investigate possible correlations between the variables, and the distribution of each. This distribution can be used to produce confidence intervals as described in Section 2.3.1, although these do not generally take account of systematic noise in the data. In this thesis this is investigated by applying the MCMC to separate datasets (see Chapter 6).

2.2 Circumbinary Population Synthesis

The eventual aim of the project is to produce rates of occurrence for circumbinary planets, which requires some synthesis of circumbinary planet populations. Using these populations, detected numbers of planets can be used to get probability density functions for the underlying occurrence rate. This allows us to find what occurrence rates with what probability are consistent with the sample of binaries and known number of planets seen. Creating this sample of binaries requires a debiasing process - a testing of which binaries we are able to detect planets within for a given planet parameter range - the method for which will be developed and implemented in later chapters. This removes binaries from the sample which are providing no information on the presence or absence of planets, due to for example excessive light curve noise.

The usual method for carrying out a study of this kind involves generating completeness fractions - recovery rates for possible planets within the binary sample as a function of the parameters involved. In this case there are several arguments against this route. Firstly, each binary has a very different 'window' of planet parameters within which a planet is observable. Using completeness fractions treats the whole sample as a uniform block, whereas using the method below we can take account of each binary individually. Furthermore, to test such completeness fractions planet signal recovery would need to be investigated using injected simulated planet signals covering a wide range of parameter space, in planet inclination, radius and period, which is computationally expensive. Under this method such injections need only take place during the debiasing process, and over a much reduced number of trial parameters. Finally the method automatically generates distributions for the occurrence rates, allowing good error determination (which takes account of the specific binaries of the sample), whereas in the normal method Poisson errors must be approximated. As such, I proceeded using a Monte Carlo process whereby large numbers of plausible circumbinary systems are simulated around the binaries of the sample and their observational properties explored. In this work the analysis is restricted to consecutive transits (where the planet transits on every orbit), as it is complex to quantify the detectability of 'sparse' transiting planets which only transit occasionally. The necessary workflow is summarised in Figure 2.1, and laid out below.

1. A sample of binaries within which a planet could have been detected if it was present is found (see Chapter 7 for detail). A trial occurrence rate is selected.

2. Using this occurrence rate, a proportion of those binaries are randomly assigned a planet.

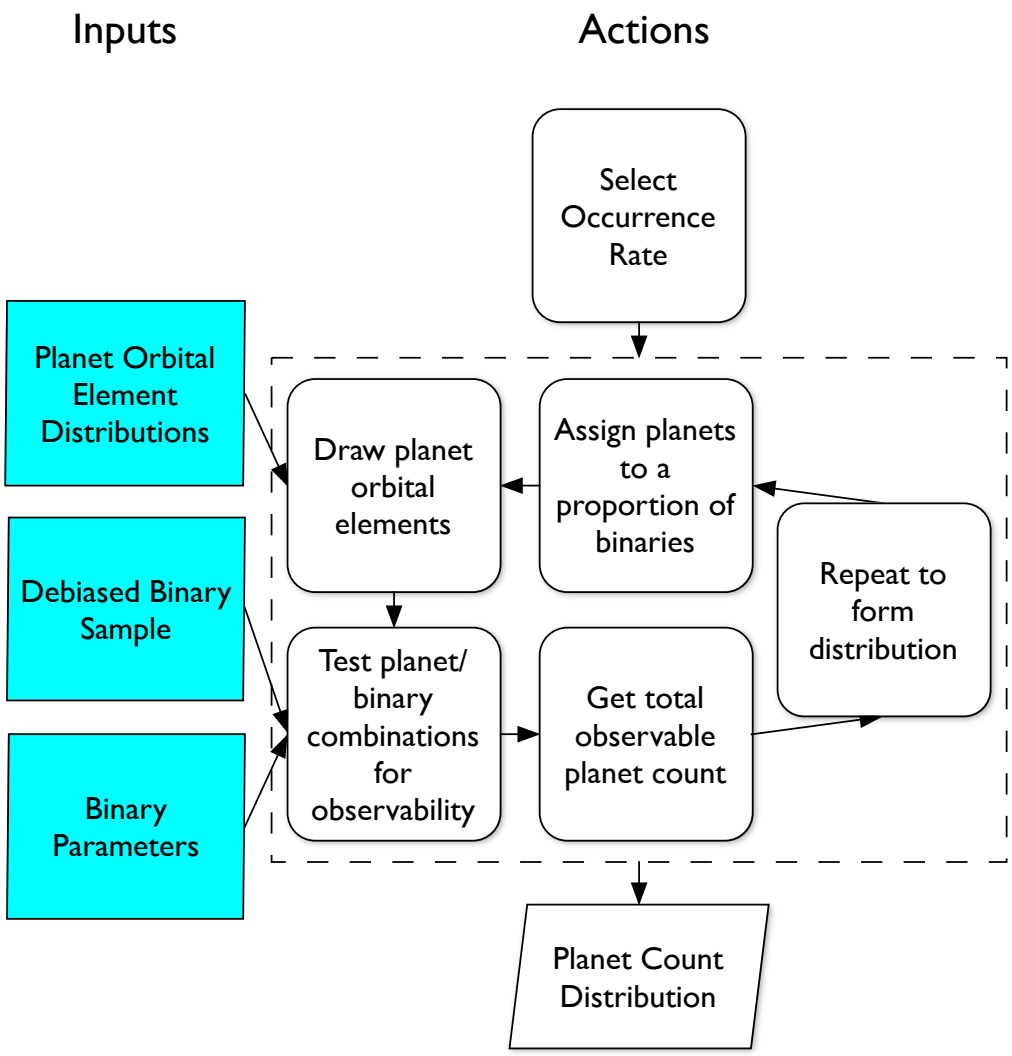

Figure 2.1: Flowchart describing the process followed in investigating circumbinary planet occurrence rates, as used in Chapter 7.

3. These planets are given orbital parameters, formed from various input distributions (see Chapter 7 for the specific sources)

4. Each planet and binary combination is tested, to check whether it would transit consecutively. Consecutive transits imply that the planet transits at least one of the host stars on every planetary orbit, and requires a high degree of coplanarity between the planet and the binary. Those which do are counted to give a total observable planet count for this iteration.

5. This process is repeated to form a distribution of observable planet counts for the trialled occurrence rate.

6. The process is then repeated over a range of occurrence rates.

At the end of the above chain, we obtain distributions of total planet count for each of a range of occurrence rates. Using a specific known number of detected planets (within the defined binary sample), these can be inverted - for each occurrence rate, the number of trials where this known number of planets was recovered can be found. This then gives a distribution over the range of occurrence rates, and when normalised, represents a probability density function for the occurrence rate, given the observed number of planets.

Formally speaking, we are looking to obtain $P(R_{occ}|N_{obs})$, where R_{occ} represents the occurrence rate and N_{obs} the observed planet count (generally within a defined period and radius range). Using Bayes' Theorem, this can be written as

$$P(R_{occ}|N_{obs}) = \frac{P(N_{obs}|R_{occ})P(R_{occ})}{P(N_{obs})} \tag{2.3}$$

The above method measures $P(N_{obs}|R_{occ})$. Strictly it measures $P(N_{obs}|(R_{occ}, \text{Model}))$ - some testing is performed (see Chapter 7) to find the effects of the model (i.e. the distributions) used. $P(R_{occ})$ then represents the prior on the occurrence rate. Given the current lack of knowledge regarding circumbinary occurrence rates, this was set to be uniform. $P(N_{obs})$ is equal to the sum over all possible occurrence rates of $P(N_{obs}|R_{occ})P(R_{occ})$ and is accounted for when normalising the resulting distribution.

We are left with $P(R_{occ}|N_{obs})$, for as many occurrence rates as trials are run for. Normalised, and under the assumption that the tested occurrence rates map the variability of the underlying distribution successfully, this forms a probability density function for R_{occ}. Different binary samples (and associated observed planet counts) can be made for different ranges of parameter space - the sample of binaries

within which we can detect planet with $R_p < 6R_\oplus$ is different from that for planets with $R_p < 10R\oplus$ for example. Repeating the above process for such samples allows occurrence rates to be derived for any chosen range of parameter space.

2.3 Probability Density Functions

As described in Section 2.2, while investigating circumbinary planets we can obtain probability density functions (PDFs) for their occurrence rate. These can be used to form various conclusions regarding the rates, and the planets themselves. The application of these methods is left to Chapter 7, but the general background is described here.

A probability density is defined by the differential of the cumulative distribution function F, i.e. for PDF f,

$$f_X(x) = \frac{dF_X(x)}{dx} \tag{2.4}$$

for a random variable X. For any set A, the PDF can give the probability that the value of X lies within A, as

$$P(X \text{ in } A) = \int_A f_X(x)dx \tag{2.5}$$

In the context of the rates of occurrence studied here, the probability that the occurrence rate lies between two values can be extracted from the derived PDF in this way.

2.3.1 Values and Errors

Given a PDF for a variable, there are several ways of extracting the probable values for that variable. For typical Gaussian distributions the mean and standard deviation are used, but these do not generally make sense for unusual (and asymmetric) functions. Two of the more commonly used options are the expected value or the mode. The mode is clear, it is the value corresponding to the maximum of the PDF, and the value which would be obtained most often if a large census of values could be pulled from the PDF. The expectation value E represents an average of the possible values the variable studied can take, weighted by the probability of each of those values, i.e.

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)dx \tag{2.6}$$

and is the same as the population mean. A useful form for presenting errors on variables represented by non-Gaussian PDFs is a confidence interval. This is the region of the PDF within which a certain percentage of the total probability is contained. As such we can present for example 50 or 95% confidence intervals. For general intervals, the confidence interval is defined by

$$\alpha = \int_{x0}^{x1} f_X(x)dx \tag{2.7}$$

with α between 0 and 1, and representing the confidence interval desired (for example 0.95 for a 95% interval). The limits x0 and x1 represent the limiting values for the variable corresponding to this interval.

2.3.2 Differences between two populations

It is often useful to compare two PDFs formed from different underlying variables. There are two tests which are used later in this work: what is the probability that two variables are the same, and in a similar sense what is the probability that one is less than another. For testing if two variables are the same,

$$P(X = Y) = \int_{-\infty}^{\infty} f_X(x) f_Y(x)dx \tag{2.8}$$

and for one less than another

$$P(X < Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{x} f_X(t)dt f_Y(x)dx \tag{2.9}$$

when written directly in terms of PDFs. Probabilities derived in this way can then be used to exclude these events at various levels of confidence.

2.4 N-body Simulations

Due to the gravitational torque of the host binary, a circumbinary planet is subject to dynamical perturbations which are not subject to Kepler's Third Law. To fully describe these, N-body integration codes must be used, where body positions are calculated individually for sequential timesteps. An N-body (here N=3) code is used for testing various approximations to circumbinary planetary motion, as well as producing realistic transit times and durations for simulations of planets. This code was provided by D. Martin, and is described fully in Martin & Triaud (2014). In short, a fourth-order Runge-Kutta algorithm is used to integrate the N-body equations of motion. Since this integrator does not inherently conserve energy,

the total system energy was calculated over time to ensure that it was conserved such that the energy loss fraction remained below approximately 10^{-7}. Orbital elements of planets and stars are tracked (at least several times per orbit), and any becoming unstable flagged, such that an inherent stability check is provided by the process. Using these derived planet and star positions, along with stellar radii, transit durations and times can be extracted.

2.5 Centroid Analysis

A key component of the data validation procedure used by the Kepler science team is analysis of the photocentres of interesting flux sources (here termed the 'centroid'. This can be measured from the distribution of flux across the aperture for a given source. This method is used principally as a vetting procedure, to determine if the source of a detected transit like event is located at the same position as the primary flux source (Torres et al., 2010; Batalha et al., 2010b). This can give precisions of a few milli arc seconds for bright sources, a significant improvement on the ~4 arcsec Kepler pixels.

There are two main methods for finding the centroid of a target. The first uses the weighted flux across the aperture, and produces time dependent row and column centroid values from these equations.

$$r = \frac{\sum_n i_n p_{s,n}}{\sum_n p_{s,n}} \tag{2.10}$$

$$c = \frac{\sum_n j_n p_{s,n}}{\sum_n p_{s,n}} \tag{2.11}$$

where i_n and j_n are the row and column coordinates of pixel n, and $p_{s,n}$ is the background subtracted flux of that pixel. r and c give the row and column centroids respectively. In the case of an offset background source causing a detected flux variation, the row and column centroid shift would be expected to be correlated with the event. There are other sources of centroid variation, including pointing drift, instrumental systematics, and differential velocity aberration, but these would not be expected to correlate with a potential planetary signal. Further detail can be found in Fraquelli & Thompson (2012).

Another method of centroid determination uses the pixel response functions (PRFs, Bryson et al., 2010a) of the Kepler spacecraft. These are the expected point spread functions for a point source as measured at various positions on the Kepler CCD. They can be fit to a given cadence (or set of binned cadences) to provide a

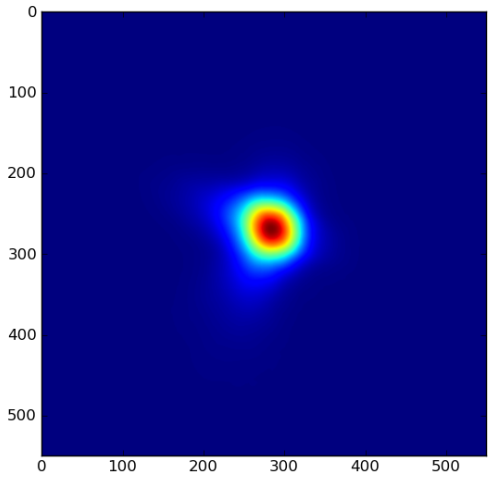

Figure 2.2: A typical Pixel Response Function for a Kepler target. Axes show hundredths of a pixel.

source location. This method becomes particularly useful for investigating individual events, such as potential transits. Points in and out of transit can be binned, to produce an increased significance pixel map in each case. The sum of these images (known as the 'Direct' image) represents the main flux source, whereas the difference between them ('Difference' image) shows the source of the flux variation. As such, a significant difference in the PRF location between the direct and difference images represents a likely background source for the event.

For illustration, the flux weighting method is applied to a known false positive - a Kepler object showing potential eclipses but which has been marked by the Kepler science team as showing high centroid variations. The time series of flux weighted centroid values can be seen in Figure 2.3. The row values show significant correlation with eclipses in the light curve, although the column values do not, implying that the true source of the eclipses is a background source located in the same column as the target but offset from it. These techniques will be used for candidate vetting in later chapters.

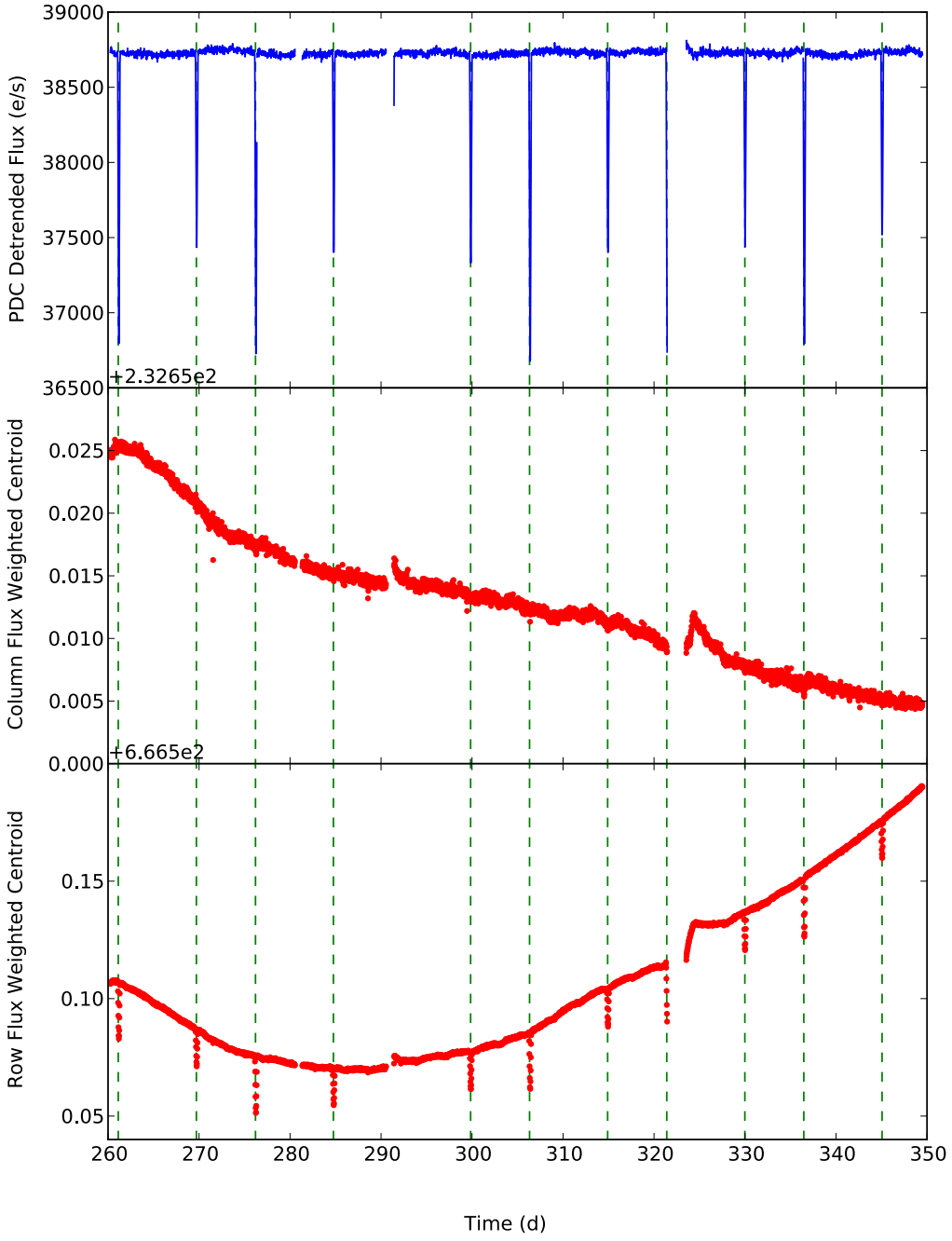

Figure 2.3: The light curve and flux weighted centroids of the quarter 3 data for KIC3655332. Candidate eclipses can be seen in the flux, which correlate with significant offsets in the centroid row position.

2.6 Period Detection

2.6.1 BLS

The Box-Least-Square method was first proposed by Kovacs et al. (2002), and has since become one of the most popular search algorithms for exoplanet transit hunting (see for example Hellier et al., 2014; Collier Cameron et al., 2006). The circumbinary search algorithm developed in this thesis is similar in form to the BLS method.

Under this process, the target light curve is phase folded at a series of trial periods. At each of these periods a two-level box shape is fit to the phase curve. The algorithm is very useful in its simplicity - only the two discrete levels, the position of the box, and its fractional width are fit. For fractional widths significantly less than unity (relative to the period searched, with 0.01-0.05 being typical values) this shape reasonably approximates an exoplanet transit. The periodogram output by BLS is characterised by the signal residue, given by

$$SR = MAX((\frac{s^2}{r(1-r)})^{\frac{1}{2}}) \tag{2.12}$$

$$s = \sum_{i=i_1}^{i_2} w_i x_i \tag{2.13}$$

$$r = \sum_{i=i_1}^{i_2} w_i \tag{2.14}$$

where i_1 and i_2 represent the data point indices at start and end of the fit box, x the data fluxes, and w their weights (taken from for example the flux errors). The maximum is found for all possible values of i_1 and i_2, subject to limits generally set for each search. It must then be repeated for all trial periods. The spacing between trial periods should be such that the phase of each data point changes by less than the expected transit duration, such that no signals are missed between trial periods.

It has been shown by Kovacs et al. (2002) that for planetary signals BLS tends to perform better than the other methods available. For more specific cases (such as the circumbinary planetary transits studied here) this is not the case, but BLS retains some power nonetheless.

2.6.2 AOV

The analysis of variance method (AOV, Schwarzenberg-Czerny, 1989) is used in Chapter 4 as a potential method for detecting binary periods. While for sharp planet-like signals BLS is an improved method, for the more variable (and in the

case of contact systems, continuously varying) binary signals AOV has significant advantages. These include no preference for sinusoidal signals (as seen in Fourier methods) as well as a very well characterised probability distribution. It is commonly used for period detection in astrophysics (as in for example Nataf et al., 2010; Prsa et al., 2011; Devor, 2005).

AOV, as with BLS, phase folds data at a series of trial periods and bins it into r bins. The AOV statistic, Θ_{AOV}, is then calculated for the phased data. Θ_{AOV} is given by

$$\Theta_{AOV} = \frac{s_1^1}{s_2^2} \tag{2.15}$$

where

$$s_1^2 = \frac{\sum_{i=1}^{r} n_i(\bar{x}_i - \bar{x})^2}{r-1} \tag{2.16}$$

$$s_2^2 = \frac{\sum_{i=1}^{r}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2}{n-r} \tag{2.17}$$

with n the total number of observations, $\bar{x}$ their average, subscript i represents a data bin and ij data point j within bin i. For a full analysis of the power of this statistic see Schwarzenberg-Czerny (1989).

2.6.3 PDM

Phase Dispersion Minimisation (PDM, Schwarzenberg-Czerny, 1997; Stellingwerf, 1978) is another period analysis technique often used in astrophysical contexts. It also has the advantage of being useable in situations with missing or unevenly sampled data where Fourier techniques become awkward, and again of not preferring sinusoidal variations. The technique again phase folds data at a series of trial periods, then bins it into a suitable number (typically 10-100) of bins. In PDM, the variance of each bin is measured, and compared to the variance of the total phase curve. Through minimising this ratio, the 'smoothest' functional form of the phase curve can be found. While at a disadvantage for sharp and localised periodic variations, for more continuously varying functions PDM can work well. The output statistic is given as

$$\Theta_{PDM} = \frac{s^2}{\sigma^2} \tag{2.18}$$

where σ is the whole phase curve variance and s^2 represents the individual

bin variances s_j. It is formed from these via

$$s^2 = \frac{\sum(n_j - 1)s_j^2}{\sum n_j - M} \tag{2.19}$$

with n_j the number of points in bin j and M the total number of bins. The PDM and AOV techniques are very similar in their nature and associated advantages. They are both explained here as each was experimented with while developing period searching algorithms later.

2.7 Eclipse Timing Analysis

A popular method of searching for companions to binaries has been the analysis of eclipse timings. This relies on the light travel time effect (LTTE), the variation in arrival time of light as the binary orbits the centre of mass of the whole system. Strictly this can be applied to any periodic signal, as long as it can be measured accurately enough, and so there is also hope for seeing this effect in pulsating stars (Shibahashi & Kurtz, 2012; Murphy et al., 2014). The eclipse timing variation caused by the light travel time effect is given by (following the form of Rappaport et al., 2013, hereafter R13):

$$ETV_{LTTE} = A_{LTTE}((1 - e_3^2)^{\frac{1}{2}} \sin E_3(t) \cos\omega_3 + (\cos E_3(t) - e_3) \sin\omega_3) \tag{2.20}$$

$$A_{LTTE} = \frac{G^{\frac{1}{3}}}{c(2\pi)^{\frac{2}{3}}} \left(\frac{m_3}{m_T^{\frac{2}{3}}} \sin i_3 \right) P_3^{\frac{2}{3}} \tag{2.21}$$

where the subscript 3 represents the third body, T the total system, and P, i, e, m and ω are the period, inclination, eccentricity, mass and argument of periapsis of the relevant body and orbit, with E and M representing the eccentric and mean anomalies. An example O-C curve (observed - calculated eclipse times) is shown in Figure 2.4. A third star in the system causes eclipse time variations through another effect, the physical delay. This is where the presence of a third body directly lengthens the binary orbital period, the effect becoming stronger for closer third stars, and larger binary periods. Rather than the light travel time effect, this is an actual change in the physical binary period. For a circular coplanar third star, the effect is constant and so will not be seen. However, for eccentric third bodies the distance from the tertiary star to the binary centre of mass changes

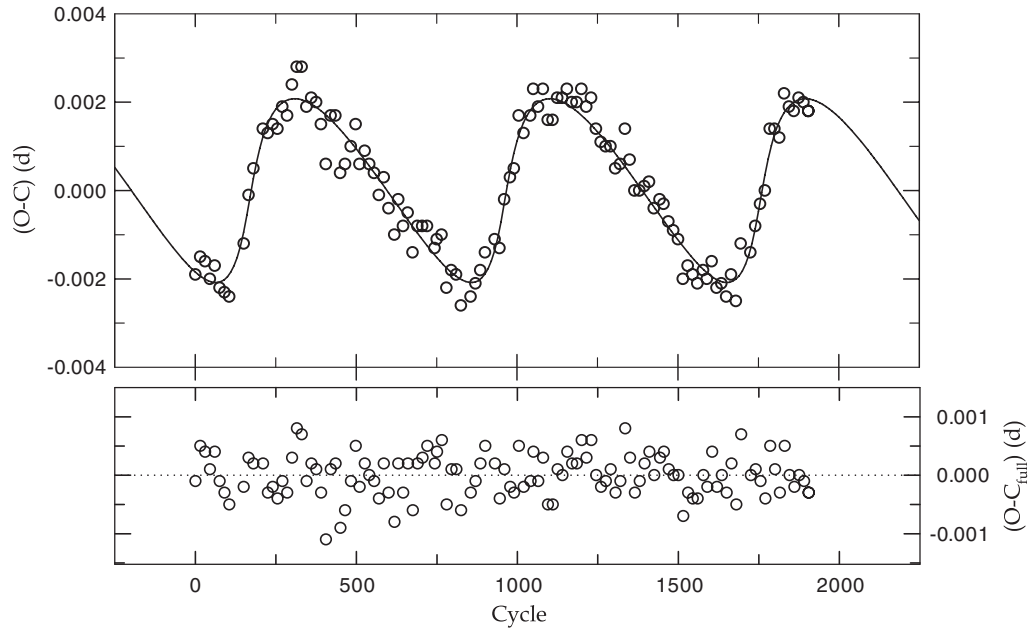

Figure 2.4: An O-C (Observed -Calculated) plot for KIC2856960. Cycles represent binary periods. Credit Lee et al. (2013)

with the third body period, and so variations in eclipse time on this period will be seen. Further, if the third body is inclined to the binary orbital plane, although the distance to the binary centre of mass may remain constant, the tidal interactions caused by the third body do vary, leading to an observable effect. Further detail, and analysis of the comparative strengths of these two effects (light travel time and physical) can be found in R13.

There are various other processes which can cause eclipse timing variations in binary systems. Any form of period evolution, from for example mass transfer between the binary components or the Applegate effect (spin-orbit angular momentum transfer due to gravitational quadrupole coupling, see Applegate, 1992) can alter the sequence of eclipse times. Changes in eclipse times can also be induced through apsidal motion (in eccentric binary orbits) or light curve noise (both instrumental or stellar activity related).

Such measured eclipse timing variations, if they are likely from a third body (from the form of the variation for example) can be used to measure the projected mass of the companion. This allows triple stellar systems to be analysed without the need for radial velocity measurements, and even allows measurement of planet masses if the eclipse times can be measured accurately enough (see Beuermann

et al. (2010) for example, for some albeit high mass planets). With the Kepler data, however, this form of planet mass measurement can generally only be used to place upper limits on the mass. Generally, accuracies of order a minute have been achieved with Kepler long cadence data, in for example Conroy et al. (2014), which is still a significant improvement on the ~29min cadence.

A variety of similar methods have been used by various groups to study eclipse timings in the Kepler binary sample. Gies et al. (2012) made templates of the binary phase curve, then fit the lowest 20% of points in these sequentially to the light curve, producing 14 possible (period > 700d) candidates out of a sample of 41. R13 fit parabolas to 3 points near every local minima of the light curve, generating 39 candidates. More recently Conroy et al. (2014) used a polynomial chain fit (see Section 1.4.3) to map the binary phase curve, then also used sequential fitting, generating 236 candidates among the shorter period Kepler binaries. Following on from R13, Tran et al. (2013) found several hundred short period binaries exhibiting eclipse timing variations (often random walks or quasi-periodicities), and attributed these to the presence of slowly evolving starspots. All the mentioned candidates are triple star systems rather than planets; Sybilski et al. (2010) showed that Kepler should only be sensitive down to ~10 M_{Jup}, and that for accuracies in eclipse time measurements of ~3 s. Using the more typical ~a minute accuracy, the minimum detectable mass is closer to $100M_{Jup}$.

Here a version of the sequential fitting methods is used, where a model is created of the eclipsing binary variation. This is done through fitting periodic splines to the phased binary light curve. For studying individual objects the spline fit can be tuned in each case. This means that we are mapping the shape of the variation, with no link to an underlying system model. The resulting spline is then fit to each binary eclipse sequentially. In cases where there is a lot of noise, a small amplitude timing variation signal, or very few points per binary eclipse (as is the case for very short period binaries) then multiple eclipses can be binned to increase the accuracy, at the loss of some time resolution. The resulting chain of eclipse times can then be used to constrain possible third bodies in the system. The method can be influenced by any system where the binary phase curve changes significantly over the Kepler baseline (such as for example rapid binary orbital evolution, as seen in Chapter 5).

2.8 Human Input

The WASP project uses human eyeballing as a critical part of the pipeline of planet discoveries. Similarly it is used in later chapters when searching through the results

of the search algorithms applied to various datasets. There is one key advantage (and one key disadvantage) to the significant use of human input.

Firstly, the advantage: it works. While a wide range of statistics, analytical methods and algorithms provide invaluable aid in discovering planets or other signals, they are all vulnerable to unusual (and yet potentially extremely interesting) signals which were not part of their design. Similarly even for expected signals, there may be issues with sensitivity. In these areas simply looking at every light curve (alongside more objective algorithm outputs) can reveal systems or events worthy of study, which may not have been seen if relying purely on output statistics. While time consuming, it is rarely more so than developing such an output statistic, for example.

However, for statistical purposes there is an obvious weakness: subjectivity. While odd, weak and interesting signals may be seen, they will not always be, and quantifying the relevant proportions in order to study underlying populations statistically is non trivial. As such, for that kind of work formal output statistics are invaluable aids in providing well understood samples which can be used.

Chapter 3

The Transit Timing Variations of Circumbinary Planets

An efficient analytical method to predict the maximum transit timing variations of a circumbinary exoplanet is derived and presented. This gives limits on the potential location of transits for coplanar planets orbiting eclipsing binaries, which are tested against numerical N-body simulations of a distribution of binaries and planets. The model is applied to the planets Kepler-16b, -34b and -35b. It is shown to be fast, efficient and accurate to approximately 1% in predicting limits on possible times of transit over a three year observing campaign. The model can easily be used to, for example, place constraints on transit timing while performing circumbinary planet searches on large datasets. It is adaptable to use in situations where some or many of the planet and binary parameters are unknown.

3.1 Introduction

The purpose of this chapter is to derive constraints on the observational characteristics of a transiting circumbinary exoplanet. This is done through our knowledge of the host binary system, using a fast method which requires no complex modelling. The derivation is limited to TTVs in coplanar circumbinary systems, placing general limits on the magnitude of such variations, through constraining the location of possible transits. These constraints are of use to surveys for such planets, where we

can place limits on and aid the design of new automated searches, such as the QATS algorithm (Carter & Agol, 2013). While it is possible with numerical simulations to predict exact times of transit for circumbinary systems, this analytical model allows (under some approximations) constraints to be placed on systems where some or many orbital parameters are not yet known, including the majority of eclipsing binaries in the Kepler Eclipsing Binary Catalogue (Prsa et al., 2011).

TTVs on the transits of circumbinary planets have two main sources. The first is a geometrical timing variation (referred to as Effect I) resulting from the changing positions of the host binary stars. This leads to a range in time in which transits can occur, similar to more 'usual' TTVs, and is derived in Section 3.2.1. The second is a precessional variation (referred to as Effect II), a long term oscillation in time around a constant periodicity of the potential location of transits, caused by precession of the planet's orbit (which is itself caused by torques arising from the non point mass nature of the binary). It is treated in Section 3.2.1.

We make use of several unusual terms in this chapter, and define them here for clarity. First, a 'crossing', or 'crossing region'. This is the region of a circumbinary exoplanet's orbit where the planet crosses the binary star orbit, from the observer's perspective. It may only transit the stars within this crossing region, but will generally spend most of its time in the region out of transit. Second the 'azimuthal' period of a circumbinary planet is used extensively. There are several periods which may be relevant to a circumbinary planet, and we make use of two here - the azimuthal period and the Keplerian period. The azimuthal period is the period which on average the planet takes between successive alignments with the observer, i.e. to traverse 2π radians relative to a fixed reference vector and plane. The Keplerian period is an osculating period taken at a particular epoch, derivable from Kepler's third law via the binary mass and planet semi-major axis. These two periods are not equivalent, and were discussed in more detail in Section 1.2.5.

3.2 Models

3.2.1 Analytic Approximation

The derivation proceeds using Keplerian orbital equations for both the stars and planets of a circumbinary system, and hence is an approximation only. It does not consider three-body effects that perturb the orbits of the binary and planet (although precession of the planet's argument of periapsis is included). TTVs of transits of only one star at a time are considered, through this chapter star 1. To consider transits of star 2, swap the indices 1 and 2 in Equation 3.3.

Geometrical Timing Variations - Effect I

These variations arise from the movement of the binary stars within their orbit. As such the limits of this orbit are used, coupled with the time the planet takes to cross said orbit. An equation for the duration of a transit in a single star/planet system (Equation 3.1, from Seager & Winn (2010b), their Equation 14) is used. A crossing (defined in Section 3.1) of a circumbinary planet is analogous to the transit of a single star by a planet passing in front of it; conceptually, we just replace the single star with a 'metastar' of diameter equal to the maximum extent of the binary's orbit, giving

$$T_{GTV} = \frac{P_\mathrm{p}}{\pi} \arcsin\left(\frac{R_\mathrm{metastar}}{a_\mathrm{p}}\right) \frac{\sqrt{1-e_\mathrm{p}^2}}{1+e_\mathrm{p}\sin(\omega_\mathrm{p})}, \tag{3.1}$$

where T_{GTV} represents the duration of the crossing, subscript p represents the planet, P the azimuthal period, a the semi-major axis, e the eccentricity and ω the argument of periapsis. The approximations were made that the impact parameter $b_\mathrm{p} \ll R_\mathrm{metastar}$, the inclination of the planet $i_\mathrm{p} = \pi/2$ and $R_\mathrm{p} \ll R_\mathrm{metastar}$. To find R_metastar we must derive the extent of the binary's orbit, projected onto the sky.

Consider the eclipsing binary orbit to be in the x-z plane, with the z axis being along the line of sight of the observer. By doing this we take the binary orbit to have inclination $\pi/2$, a reasonable approximation for detached eclipsing binaries and for this purpose. Take the motion of star 1 in the x plane, projected onto the sky. From Seager et al. (2010) (their Equation 53, with $\Omega = 0$), this is given by

$$X = \beta(f) a_\mathrm{b}, \tag{3.2}$$

where

$$\beta(f_\mathrm{b}) = \frac{M_2}{M_1+M_2} \frac{(1-e_\mathrm{b}^2)}{1+e_\mathrm{b}\cos(f_\mathrm{b})} \cos(\omega_\mathrm{b}+f_\mathrm{b}), \tag{3.3}$$

and a_b represents the semi-major axis of the binary, $M_{1,2}$ the mass of stars 1 and 2 respectively, e_b the binary eccentricity, f_b the true anomaly of the binary and ω_b its argument of periapsis. Taking the zero points of the derivative with respect to f_b of Equation 3.2 gives us the minimum and maximum values of X - the extent of the star's motion projected onto the sky. The values of the true anomaly of the binary at these points are given by

$$f_0, f_1 = \arcsin[-e_\mathrm{b}\sin(\omega_\mathrm{b})] - \omega_\mathrm{b}. \tag{3.4}$$

Equation 3.4 has two solutions within the range $0, 2\pi$. Inserting both into Equation 3.2 gives the maximum and minimum values for X. We term these X_1 and X_0. Which of X_0 and X_1 is the minimum and which the maximum depends on ω_b, but is unimportant here.

The radius of the 'metastar' is given by

$$R_\mathrm{metastar} = \frac{|X_1| + |X_0|}{2}, \tag{3.5}$$

and a scaled radius by

$$R_\mathrm{m,scaled} = \frac{R_\mathrm{metastar}}{a_\mathrm{b}} = \frac{|\beta(f_1)| + |\beta(f_0)|}{2}. \tag{3.6}$$

Substituting Equation 3.5 into Equation 3.1 leads to

$$T_{GTV} = \frac{P_\mathrm{p}}{\pi} \arcsin\left[R_\mathrm{m,scaled}\left(\frac{P_\mathrm{b}}{P_\mathrm{p}}\right)^{\frac{2}{3}}\right] \frac{\sqrt{1-e_\mathrm{p}^2}}{1 + e_\mathrm{p}\sin(\omega_\mathrm{p})}, \tag{3.7}$$

where the ratio of semi-major axes has been substituted to the equivalent ratio of periods using Kepler's third law, allowing the use of the azimuthal period outlined in Section 3.1. In the presented form T_{GTV} represents the duration of a crossing, and as such a range of time within which transits can occur. The argument of periapsis, ω_p, is a function of time due to precession of the planetary orbit; assuming a constant precession rate it can be estimated analytically using Equation 5 of Doolin & Blundell (2011), hereafter DB, which is derived from that of Farago & Laskar (2010b).

Lacking knowledge of the present system alignment, it is possible to take a 'safe' approximation by using the value of ω_p which gives the maximum T_{GTV}, i.e. $\omega_\mathrm{p} = 3\pi/2$. This corresponds to when the planet transits near its apoapse, and hence is travelling relatively slowly so that the range of transit times is extended. Using this constant value of T_{GTV} is often more practical. For systems with low planetary eccentricity the variation caused by varying ω_p is small (on the order of a few percent in T_{GTV}).

Precessional Timing Variation - Effect II

This variation is caused by the precession of the planet's orbit. For an eccentric planetary orbit, this precession will result in shifts in the time of potential transits away from the 'expected' time for a constant periodic signal. The magnitude of these shifts at a given time depends on the instantaneous value of ω_p.

We assume a constant precession rate for the planetary orbit, such that

$$\frac{d\omega_{\rm p}}{dt} = \frac{2\pi}{P_\omega}, \tag{3.8}$$

where P_ω represents the period of precession of the planet's periapsis (specifically its longitude of the ascending node, but the effect is the same for coplanar planets), and can be estimated analytically through the equation of DB.

For a planet precessing in the prograde direction, this change in $\omega_{\rm p}$ represents time 'gained', a portion of its orbit which it does not have to cover before aligning with the observer once more. The differential amount of time saved (i.e. period shifted) in this way is given by

$$\frac{dP_{\rm p}}{d\omega_{\rm p}} = \frac{dt}{df_{\rm p}}, \tag{3.9}$$

where dP represents an apparent change in the period of the planet, and $f_{\rm p}$ is the true anomaly of the planet, with $dt/df_{\rm p}$ evaluated at $f_{\rm p} = \pi/2 - \omega_{\rm p}$, the value of $f_{\rm p}$ at transit conjunction.

There are two contributions here, a constant term from the precession and a varying oscillation induced by the effect of the eccentricity of the planet's orbit. The constant term can be found simply, by realising that the planet 'loses' one full orbit of time in one precessional period. For a constant precession rate, this gives a constant rate of time loss of $P_{\rm p}/P_\omega$, which must be subtracted from Equation 3.9 to find the oscillation term. When using the azimuthal period of the planet (as defined in Section 3.1), or searching observationally for transits this constant term is automatically accounted for, which is why it must be removed here.

Continuing the derivation, the standard Keplerian orbital equation for $df_{\rm p}/dt$ (Seager et al., 2010, their Equation 32) is taken, evaluated at $f_{\rm p} = \pi/2 - \omega_{\rm p}$,

$$\frac{df_{\rm p}}{dt} = \frac{2\pi}{P_{\rm p}} \frac{[1 + e_{\rm p}\sin(\omega_{\rm p})]^2}{(1 - e_{\rm p}^2)^{\frac{3}{2}}}, \tag{3.10}$$

where it has been approximated that $P_{\rm p} \simeq P_{\rm p}(1 + P_{\rm p}/P_\omega)$, for this equation only. Combining Equations 3.8 and 3.9 gives us the oscillation term

$$T_{PTV} = \int_{t_0}^{t} dP_{\rm p} = \int_{t_0}^{t} \frac{dt}{df_{\rm p}} - \frac{P_{\rm p}}{P_\omega} d\omega_{\rm p}, \tag{3.11}$$

which, after inserting Equation 3.10, becomes

$$T_{PTV} = -\frac{P_{\rm b}}{P_\omega}\int_{t_0}^{t}\left(\frac{(1-e_{\rm p}^2)^{\frac{3}{2}}}{(1+e_{\rm p}\sin[\omega_{\rm p}(t)])^2}-1\right)dt, \tag{3.12}$$

where the negative sign accounts that this is time gained or equivalently an apparent shortening of the planetary period, and applies for prograde precession. T_{PTV} represents an oscillation of the location of possible transits with time. An example of its effect is given through application to a demonstration simulated system in Section 3.2.2.

Combined TTV Limits - Practical Use

Equations 3.7 and 3.12 can be combined to provide limits on the TTVs of transiting coplanar circumbinary planets. At a given epoch, T_{PTV} represents the offset around some zero point that the range of possible transit times would be centred around, whereas T_{GTV} represents the extent of the range around this offset. Constraints are presented here for practical use, in the situation where one transit has been detected and limits need placing on the times of future transits. In this case, as we do not know where in the possible transit range the first transit fell, we must use double the range to cover all possible times, giving the following limits:

$$t_{\rm min}(i) = t_0 + iP_{\rm p} + T_{PTV}(t_0+iP_{\rm p}) - T_{GTV}(t_0+iP_{\rm p}) \tag{3.13}$$

and

$$t_{\rm max}(i) = t_0 + iP_{\rm p} + T_{PTV}(t_0+iP_{\rm p}) + T_{GTV}(t_0+iP_{\rm p}), \tag{3.14}$$

where t_0 represents the time of first transit, and i an index for the orbit under consideration (each orbit may contain more than one transit, though in practice this is unusual). $t_{\rm min}$ and $t_{\rm max}$ represent the minimum and maximum times between which possible transits must fall within, on each orbit.

Over short ($\ll P_\omega$) timescales T_{GTV} is the dominant contribution (in some systems, such as those with low eccentricity planets, it is always so), and T_{PTV} may be neglected. Using the maximum possible value of T_{GTV} (by setting $\omega_{\rm p} = 3\pi/2$ in Equation 3.7) provides a 'safe' (in that the result will always be an overestimate) way of neglecting the time and ω_p dependence of T_{GTV}. Similarly, if little is known about a proposed circumbinary system, parameters in the above equations can be easily approximated with only small and quantifiable errors introduced.

The effects of T_{GTV} and T_{PTV} are shown for a demonstration circumbinary

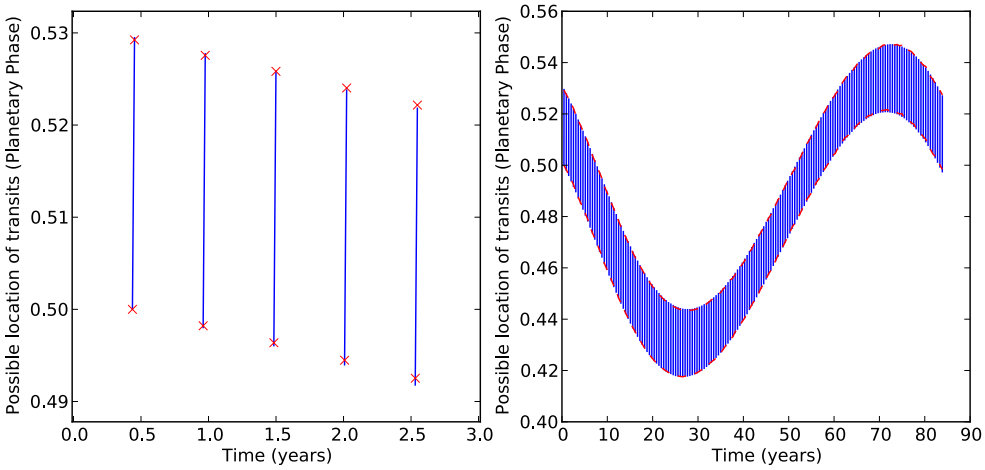

Figure 3.1: **Left** Three years and five crossings of a simulated planet. Transits must occur on the lines. The crosses represent the predicted maximum and minimum time for each crossing region derived from our analytical equations. The length in phase of each line represents T_{GTV}, while the shifting of the lines in phase represents T_{PTV}. The phase is calculated through phase folding over the planetary azimuthal period (191.5d). The starting epoch t=0 is arbitrary, as are absolute values of the phase. **Right** As left for a full planetary precessional period. The dashed line shows the analytical equation prediction, realigned with the numerical model every three years (chosen as a representative length for an observing campaign.) Realignment is justified as this is how the equations would be used in practice, with a single detected event representing a zero point to which the equations would be aligned.

planet in Section 3.2.2.

3.2.2 Numerical Model

Approach

A numerical model was used to test the above analytical framework. This is described in Section 2.4. To calculate the azimuthal period numerically the time intervals between the planet passing each of the two boundaries of the projected star orbit were averaged. The azimuthal period is the mean of these two averages. An alternative method is to average the interval between system centre of mass crossing times, which will converge to the same value but more slowly because it is only based on one crossing point, not two. Over time, the average interval between consecutive transits will converge to the azimuthal period.

Demonstration

Here the numerical and analytical models were applied to a simulated system (chosen from the simulations of Section 3.3 as a system with a typical error) to demonstrate the effects of the derived timing variations. This system has a binary star with period $P_{\rm b} = 14.10$ d, eccentricity $e = 0.13$, stellar masses of $M_1 = 1.22$ and $M_2 = 1.07$ $M_\odot$ and argument of periapsis $\omega = 282.3^\circ$. The planet has azimuthal period 191.5d and eccentricity 0.16, leading to a precessional period for the planet of 84.2 years from the numerical model. Figure 3.1 shows the potential locations of planetary transits derived from the numerical model, using times of potential transit phase-folded at the above azimuthal period. Potential transits must occur on each solid line. The variations seen are discussed below.

Geometrical Timing Variations - Effect I

The Effect I geometrical timing variations derived in Section 3.2.1 arise from the significant motion of the host binary stars. Given that it can take several days for the planet to traverse the full extent of the binary orbit, there can be large TTVs, as evidenced by Equation 3.7. This contribution to the TTVs corresponds to the length of the lines in Figure 3.1. The magnitude of the Effect I term itself oscillates with the precession period of the planet, due to the changing speed of the planet at crossing, as different regions of its eccentric orbit line up with the observer.

Precessional Timing Variations - Effect II

The other variation, an oscillation in phase or equivalently oscillation in apparent period, is due to the precession of the planet's orbit causing transits to correspond to different phases, as seen in Figure 3.1(right). The oscillation is in particular caused by the changing instantaneous effect of the precession on a planet in an eccentric orbit. This is different to the contribution of precession in the Effect I geometrical case, which varies T_{GTV} due to the changing planetary velocity. The Effect II precessional variation becomes significant over timescales approaching the planetary precessional period, typically decades. The amplitude of this variation is strongly dependent upon the planetary precessional period and eccentricity.

3.3 Results

3.3.1 Setup

The accuracy of the model of Section 3.2.1 was tested using the numerical model (Section 3.2.2) applied to a simulated distribution of 1000 single planet circumbinary systems, 799 of which were stable over 1200 years (longer than the maximum planetary precession period found, and significantly longer than the majority). A more thorough stability analysis was not necessary for the purposes of testing the equations in this paper. The binary star periods and eccentricities were taken from Halbwachs et al. (2003a), which presented an unbiased distribution taken from radial velocity surveys, expanding upon the work of Duquennoy & Mayor (1991). The primary star masses were taken from the Kepler catalog of all stars monitored, using an empirical calibration from Torres et al. (2010) to calculate the mass based on the metallicity, temperature and surface gravity. The secondary star mass was determined using the mass ratio distributions found in Halbwachs et al. (2003a), for binaries with periods less than and greater than 50 days. The radii of the stars were unimportant for this test.

For the planets, since no circumbinary planet distribution is known as yet, the period and eccentricity distributions were taken from data for planets orbiting single stars. Only radial velocity data were used to avoid the bias towards small periods seen in transit surveys. The planet was taken as a massless test particle, as its mass has a minimal effect on the dynamics. The planet radius was also unimportant for this simulation, as it has no effect on the dynamics. For each circumbinary system the minimum planet period was four times that of the binary, as a rough stability constraint (Holman & Wiegert, 1999), although many systems still proved to be

unstable (particularly those with high eccentricities). The maximum planet period was set at 500 days, long enough that TTVs in such systems are unlikely to be of interest in the near future. All systems were exactly coplanar. Each of these systems was integrated numerically over its expected precession period (calculated from the equation of DB) with a time step of 30 minutes. The system's azimuthal period was then calculated from the time it took the planet to orbit the system centre of mass on average.

To test the analytical model Equations 3.7 and 3.12 were used to predict the limits on possible transit time of the simulated planets. The precession period was split into three-year baselines (chosen as the length of a representative observing campaign). For each of the three-year baselines for each system, the predicted and numerical limits were initially aligned (as would be the case when detecting the first transit of a candidate planet) and then the system and predicted limits were allowed to evolve. At each crossing, the deviations between the upper analytical and numerical limits and lower analytical and numerical limits were averaged, and the same averaged for all crossings within each of the three-year baselines.

3.3.2 Test Results

Results are represented as a percentage of the numerically integrated crossing time found at each planetary crossing. As such, an error of 100% represents analytically predicted transit limits which are misaligned by one crossing time on average. Figure 3.2 shows the histogram of percentage errors found for the 799 stable systems. The peak shows an error of 0.4%. The median error is 0.84%. 43 systems are not shown in Figure 3.2 for clarity. These represent badly predicted single systems, with percentage errors higher than 20% (4 of them have errors over 100%). These larger error systems are discussed in Section 3.4.

3.3.3 Application to Kepler-16b, -34b and -35b

The numerical model was applied to the known systems Kepler-16b, -34b and -35b, and times of possible transit were extracted. Azimuthal periods of 227.06d, 283.13d and 127.30d were found for -16b, -34b and -35b respectively. They are slightly offset from those found in an analytic study by Leung & Lee (2013). These are compared to Keplerian periods from the respective discovery papers of 228.78d, 288.82d and 131.46d (Doyle et al., 2011; Welsh et al., 2012). It should be noted that care must be taken regarding the different reference frames parameters for these planets can be published under. This is exacerbated by the instantaneous and highly variable

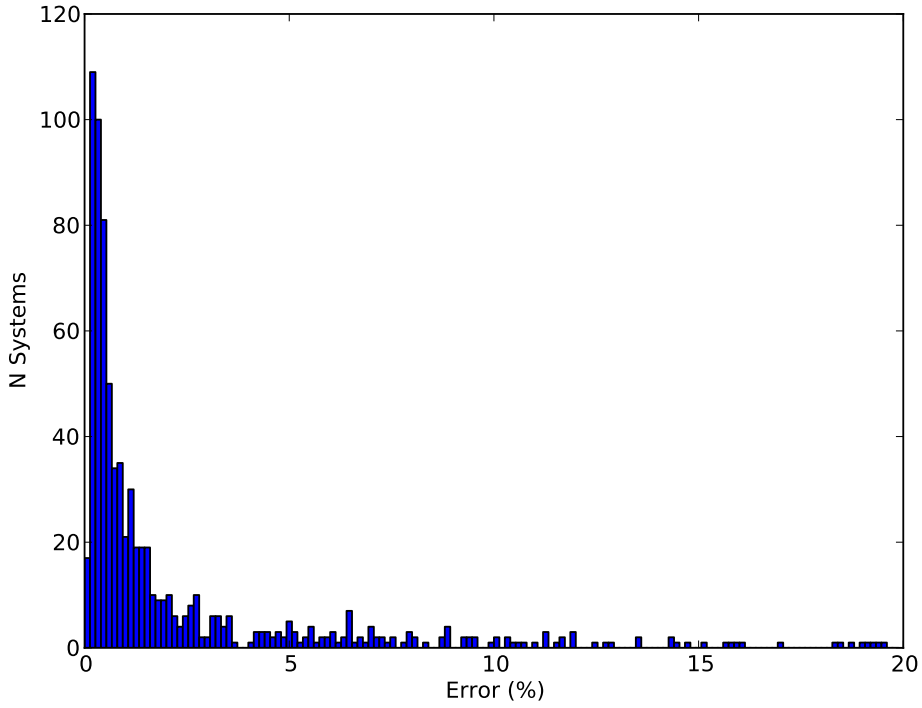

Figure 3.2: The error in the analytical model of Equations 3.7 and 3.12 from comparing them to simulated numerical limits on the possible transit locations of 756 systems. The difference between the analytical and numerical models is expressed as a percentage of the numerical planet crossing duration at each crossing. 43 additional systems with greater than 20% error are not shown for clarity. 4 of these systems have errors over 100%.

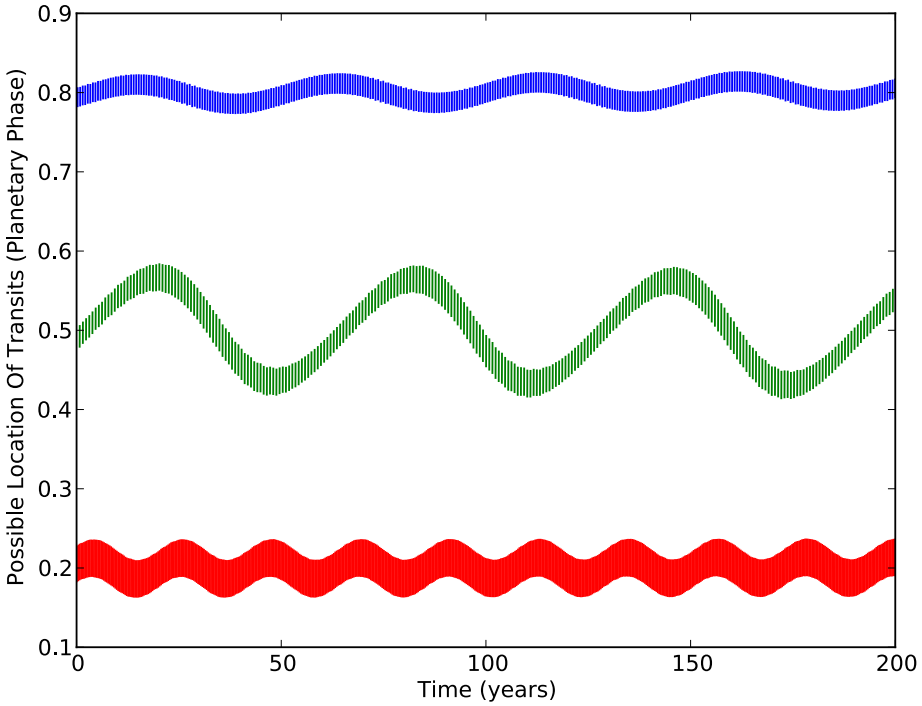

Figure 3.3: The variation in planetary phase of potential transit times, derived from the numerical model. Transits must occur within the thick bands. The thickness of these bands is defined by the geometrical timing variation. Their oscillation is defined by the precessional timing variation. From top to bottom, the lines show Kepler-16b, -34b and -35b. The phase is calculated using the planetary azimuthal period in each case. Absolute values of the phase are arbitrary. The starting epoch t=0 is also arbitrary.

nature of many of the usual planetary parameters. Figure 3.3 shows the potential locations of planetary transits derived from the numerical model, using times of potential transit phase-folded at the above azimuthal periods. Potential transits must occur within the thick band for each planet. The thickness of each band represents the Effect I, geometrical timing variation, and the oscillation in phase of the band represents the Effect II, precessional variation. The amplitude of this Effect II variation is strongly dependent upon the planetary precessional period and eccentricity. The period of the Effect II oscillations is equivalent to the planet's precessional period, ~48yr, ~63yr and ~21yr for Kepler-16b, -34b and -35b respectively.

A typical three year region is shown for each planet in Figures 3.4, 3.5 and

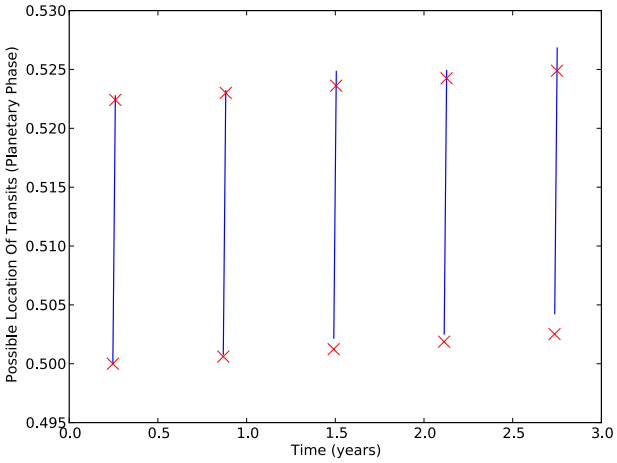

Figure 3.4: A typical 3 year region of the Kepler-16b curve of Figure 3.3. Transits must occur on the lines. The crosses represent the predicted maximum and minimum time for each crossing region derived from our analytical equations. The length in phase of each line represents T_{GTV}, while the shifting of the lines in phase represents T_{PTV}. Absolute values of the phase are arbitrary. The starting epoch t=0 is also arbitrary.

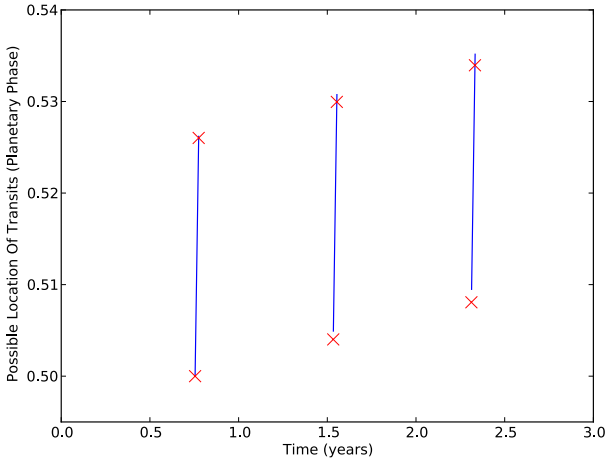

Figure 3.5: As Figure 3.4 for Kepler-34b.

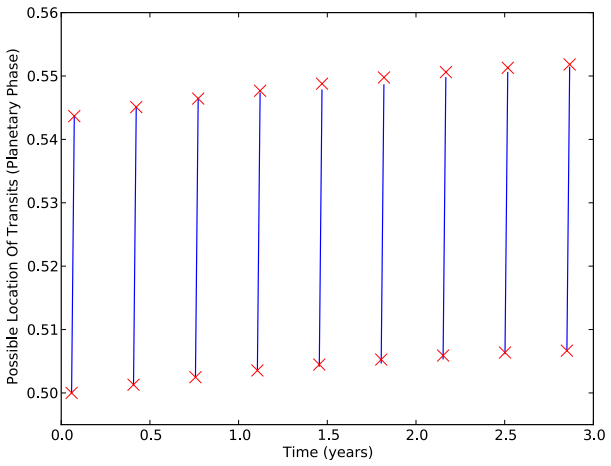

Figure 3.6: As Figure 3.4 for Kepler-35b.

3.6, with the analytical model prediction for each crossing. Note the slight secondary oscillation in Figure 3.4, which arises from the numerical model. This is an additional dynamical effect likely due to a non-Keplerian effect of the host binary, and is stronger for Kepler-16b than for -34b or -35b. This effect is not studied further here.

3.4 Discussion

3.4.1 Accuracy

Figure 3.2 shows the accuracy of Equations 3.7 and 3.12 in predicting the possible times of transit of coplanar circumbinary planets - a median percentage error of 0.84% of the planet crossing time across the test set of 799 stable systems, over three years of observations. This can be used as an error when using Equations 3.13 and 3.14 to predict possible times of transit, where the percentage error should be applied to both $t_{\rm max}$ and $t_{\rm min}$. It should be noted that our stated errors depend on the time baseline covered - they will be reduced for baselines lower than three years, and increased for those higher. The stated errors should, however, be indicative for a general observing campaign. Limitations on the accuracy arise primarily from non-Keplerian effects (beyond simple constant precession of the planet's orbit, which is accounted for). This is demonstrated by the 43 systems with errors greater than 20%, including 4 with errors greater than 100%. These, and the scattered systems found at over 5% in Figure 3.2, are systems which appear to be stable but which

show strong dynamical effects that are not accounted for, such as shorter period additional oscillations of ω_p or other effects we do not investigate here. The underlying dynamics behind these are beyond the scope of this work. Encouragingly, it seems that such effects are strong only in a small minority of cases - the analytical model missed the possible transit range entirely in only 0.5% of the tested stable systems.

3.4.2 Applications

Equations 3.7 and 3.12 (and in practice Equations 3.13 and 3.14) will prove useful particularly for current and future searches for circumbinary planets. These equations are also useful in reverse, for making first estimates of planet parameters using the observed transit variations of a newly discovered planet candidate. They provide a link between our theoretical knowledge of a circumbinary planetary system and the observational transit signatures which may arise from it, without requiring complex modelling or N-body integrations. This can be used to place limits on the potential transit times of candidate planets around a binary star, for the purpose of constraining searches for the transits of unknown planets, or similarly placing limits on attempts to find the transits of circumbinary planets discovered through another method. Importantly, this analytic framework can be used on systems where detailed knowledge of the component stars and orbital parameters is lacking, something impossible for N-body models. Full use of Equations 3.13 and 3.14 requires knowledge of the binary system, specifically the individual stellar masses, binary orbital eccentricity, argument of periapsis and binary period, as well as the argument of periapsis and eccentricity of the planet (while the planetary period is involved, for general searches for unknown planets a series of trial periods could be used). Lacking some or all of these details, it is possible to make useful conclusions through using simplifying assumptions - taking $M_2 \ll M_1$ for example removes the need for knowledge of the stellar masses while only overestimating the Effect I timing variation limit by at most a factor of 2 (i.e. placing loose but still useful limits on transit timing).

3.5 Summary

1. There are two key contributions to the timing variations affecting transits of circumbinary planets. These are geometrical, Effect I, from the motion of the binary stars, and precessional, Effect II, from the precession of the planet's orbit. Other contributions, from for example other planets in the system, are

generally on the order of minutes or less in amplitude and negligible compared to these.

2. An analytic framework to quickly estimate each of these terms has been derived and validated, for a planet coplanar with its host binary.

3. This can be used to place limits on the location of possible transits. In particular, the equations can be approximated using minimal knowledge of the system (in contrast to a more detailed numerical integrator), making them useful for searching datasets for transits of such planets. Specifically, full use of the equations require the individual stellar masses, binary eccentricity, argument of periapsis and binary period, as well as the period, argument of periapsis and eccentricity of the planet. It is simple to approximate the parameters or use trial values where necessary, as described at various points above.

Chapter 4

Search Algorithm Development and Results

This chapter describes a search algorithm used to look for circumbinary planetary transits in the WASP and Kepler datasets. The data preparation processes used are discussed, along with search algorithm development. Results are presented for previously published planets and new candidates found with the algorithm.

4.1 Starting Point

At the beginning of this project no algorithms had yet been developed to search for the transits of circumbinary planets (except Ofir (2008), whose algorithm required an unusable number of parameters to enact). The first such discovered system was in fact found through intense human eyeballing of the Kepler curves (Doyle et al., 2011). In order to allow their systematic study it became necessary to create a more automated approach, which was initially targeted at the WASP dataset. The starting point for this approach was the Box-Least Squares algorithm, which was described in Section 2.6.1. This trials a series of possible planetary periods, phase folding the available data at each and searching for a transit like box. As seen below, the final algorithm follows a similar pattern.

4.2 Data Preparation

4.2.1 Initial Attempt on WASP

The first target of the project was the WASP dataset, particularly the subset of low mass eclipsing binary stars (EBLMs) which had been flagged by researchers while looking for planets. There were ~750 of these EBLMs at the time of searching. While the majority of this chapter deals with the Kepler dataset, the process followed on the WASP data was generally similar. In this case detrending was performed by the WASP pipeline (Pollacco et al., 2006). This incurs some risk, as the pipeline was not designed to search for the more unusual transits of circumbinary planets. These transits can in principle have very different durations to normal planetary transits, for example lasting up to half the orbital period of the binary. The shape of these transits can also change, as the star moves at the same time as the planet. This could allow the impact parameter and so limb darkening to change during transit, as well as causing additional ingress or egress regions. While there is no particular reason that these variations would be removed by the WASP pipeline, the risk stems from the high potential for unusual and potentially unexpected transits.

Beyond the WASP pipeline, it was necessary to remove the previously flagged EBLM signal from each light curve. The specific method is identical to the Kepler data, and described in Section 4.2.3. This requires a knowledge of the period of the binary however, and although given the WASP detection this was known, it was found that the accuracy of the period detection was not enough by itself. In some cases the multiple fields presented by WASP (see Section 1.4.1) gave different, often harmonic, periods, in others they were simply slightly off from the true period. As such several methods of independent period searching were trialled. These were analysis of variance (AOV, see Section 2.6.2) and phase dispersion minimisation (PDM, see Section 2.6.3). The results from each were similar, but largely due to ease of use analysis of variance was settled on for producing the used binary periods.

Eventually it was realised that several intrinsic problems with the WASP data were precluding the detection of any significant signals. These are summarised here:

1. As WASP is ground based, despite the 8 year baseline of the data the window function is such that gaps exist that tend to preclude the detection of any signals with period $>$ ~10 days, as the chance that all transits will fall in light curve gaps becomes too high. It has become clear (see Chapter 7) that circumbinary planets are preferentially found around longer period binaries, with host periods $>$ ~7 days. Using the stability limit of Holman & Wiegert

(1999), this leads to minimum planet periods of ~30–40d, well above the WASP range.

2. Again as it is ground based, the WASP data is not continuous. Due to the large transit timing variations of circumbinary planets, this means that even a short period planet could have transits which fall entirely in the gaps in the data. More importantly, while the phase curve for single stellar hosted planets can still be fully mapped up to a certain period, this is not possible for circumbinary planets - it is possible to fill a phase curve, and still miss all transits, making it hard to use the WASP data for statistics of detections or non-detections.

3. Most importantly, as circumbinary transits will not be strictly periodic, we cannot use several transits to build up detection significance. Each transit needs to be clear enough to allow reasonable confidence in it, without other periodic transits to support it. While support can be gained by searching quasi-periodically (see below), to allow this at least one or two transits must give fairly strong confidence. While for single stellar hosts the WASP project has done extremely well through using multiple transits to beat down the noise, this method is less practical here.

4.2.2 Covariance Basis Vectors

Moving to the Kepler data, the challenges in dealing with the data changed. In Kepler, there is significantly less white noise, at a typical level of 20ppm. However, systematics caused by the instrument become increasingly important, to the level of dominating the light curve error in the majority of cases. The sources of these systematics are discussed in Section 1.4.2. The Kepler team produces their own detrended data, known as the presearch data conditioning (PDC) light curves. At the beginning of this project, these were known to have issues, removing astrophysical signals and introducing data artefacts in some cases. The Kepler team updated this process to use a Bayesian technique (called PDC-MAP) which is described in Smith et al. (2012). This is a detrending process deliberately targeted at finding the single stellar hosted planets which make up Kepler's primary objective. As such it has been warned that non-transit like events with durations over ~12 hours, such as might be found in the circumbinary case, can be adversely affected by PDC-MAP. During the PDC-MAP process covariance basis vectors are generated, of which the strongest 16 are extracted and can be accessed. These represent the dominant systematic trends found in each channel and quarter of the Kepler data, and are formed

from the normalised light curves of every targeted star in the appropriate channel.

It was decided to detrend the light curves from instrumental and systematic effects using these covariance basis vectors. In addition to the concern over PDC-MAP's robustness to long duration events, CBVs allow a more manual testing of the light curves, as it possible to run the process oneself rather than having preprocessed data made available by the Kepler team. While it would be ideal to individually tune the detrending of all light curves (through for example varying the number of basis vectors to use in each case), the sample size made this impractical. Typically there are 2–3 stronger vectors followed by a chain of weaker contributors. Here 5 vectors were used for all targets. This ensures the strongest components are accounted for, while avoiding using a large number, which can result in removing astrophysical signals. The PDC-MAP process typically used 8, before continuing with further detrending. As such we are likely to be more robust against removing real signals, at the risk of leaving some systematics in place. Detrending was enacted using the PyKE code (Still & Barclay, 2012). A quality flag (SAP_QUALITY) is given in the data, and marks a number of reasons for poor quality data points, including attitude tweaks, reaction wheel crossings, safe and pointing modes, and cosmic ray detections. A full list can be found in the Kepler Archive Manual (Fraquelli & Thompson, 2012). All points without SAP_QUALITY$= 0$ were cut.

Once detrended in this way, light curve quarter data was stitched together through dividing by the median flux value of each quarter, forming single near-continuous light curves for each binary. Figure 4.1 shows the effect of detrending for a typical Kepler light curve, going from raw data to a single detrended, stitched light curve.

4.2.3 EB signal removal

At this stage the signal of the known eclipsing binary must be removed, without affecting any potentially planetary signals. This is done using a modified whitening procedure, whereby the light curve is phase-folded at the binary period provided by the Kepler catalogue (or the AOV periods in the WASP case). The phased curve was then binned using 200 equal width bins, and the median of each bin determined. As any points exhibiting tertiary transit events will be distributed across the light curve when phase folded on the binary period, taking the median will exclude them from the whitening process. This principle could be affected by resonant orbits between the planet and binary, but as long as there are more out of eclipse points than in at a given phase location the planetary transits should be unaffected. Bins with standard deviation greater than 50% larger than the median

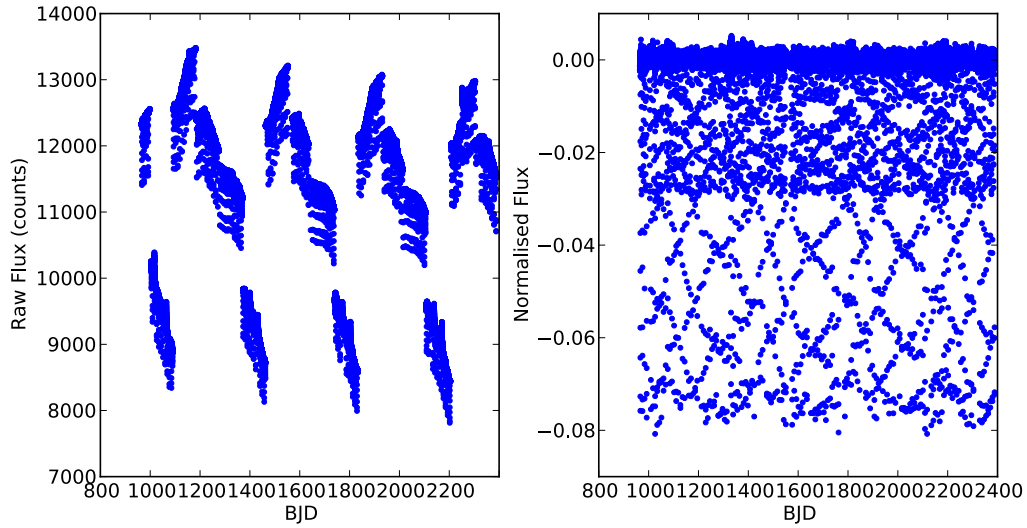

Figure 4.1: The light curve of KIC001026032, before (left) and after (right) detrending.

standard deviation across the dataset are then further subdivided into 5 higher resolution bins, mitigating the effect of sharp variations such as eclipse ingresses and egresses. The median of each bin is then subtracted from each point in the bin, resulting in a series of data points distributed around a median adjusted flux of zero. This method has the advantage of requiring no knowledge of the signal being removed, and as such can be applied to many variable objects. However it does have the disadvantage of occasionally leaving residual binary signals around regions of sharp variation, particularly ingress or egress points for detached binary eclipses. To lessen this effect, a scan was incorporated into the code to mark significant binary remnants. Bins were flagged as before ($\sigma > 50\%$ larger than the median standard deviation across the dataset) but this time incorporating the higher resolution bins now in place. Bins failing this test are marked as red in Figure 4.2. Further to this, each curve was checked by eye for strictly periodic binary remnants, and any found manually removed. Examples of potential pitfalls as the result of this process are shown in Figure 4.2. In each of these the marking of problematic points, combined with the manual option to remove them, resulted in a useable light curve.

Although the catalogue values were used for the binary periods when phase folding to remove binary eclipses, for the initial stages of the project these values were far less valuable. The Kepler EB catalogue is currently on version 3; earlier versions used reduced portions of the data to calculate periods. As the longer the timeline of the data, the more accurate a period must be to maintain a good phase

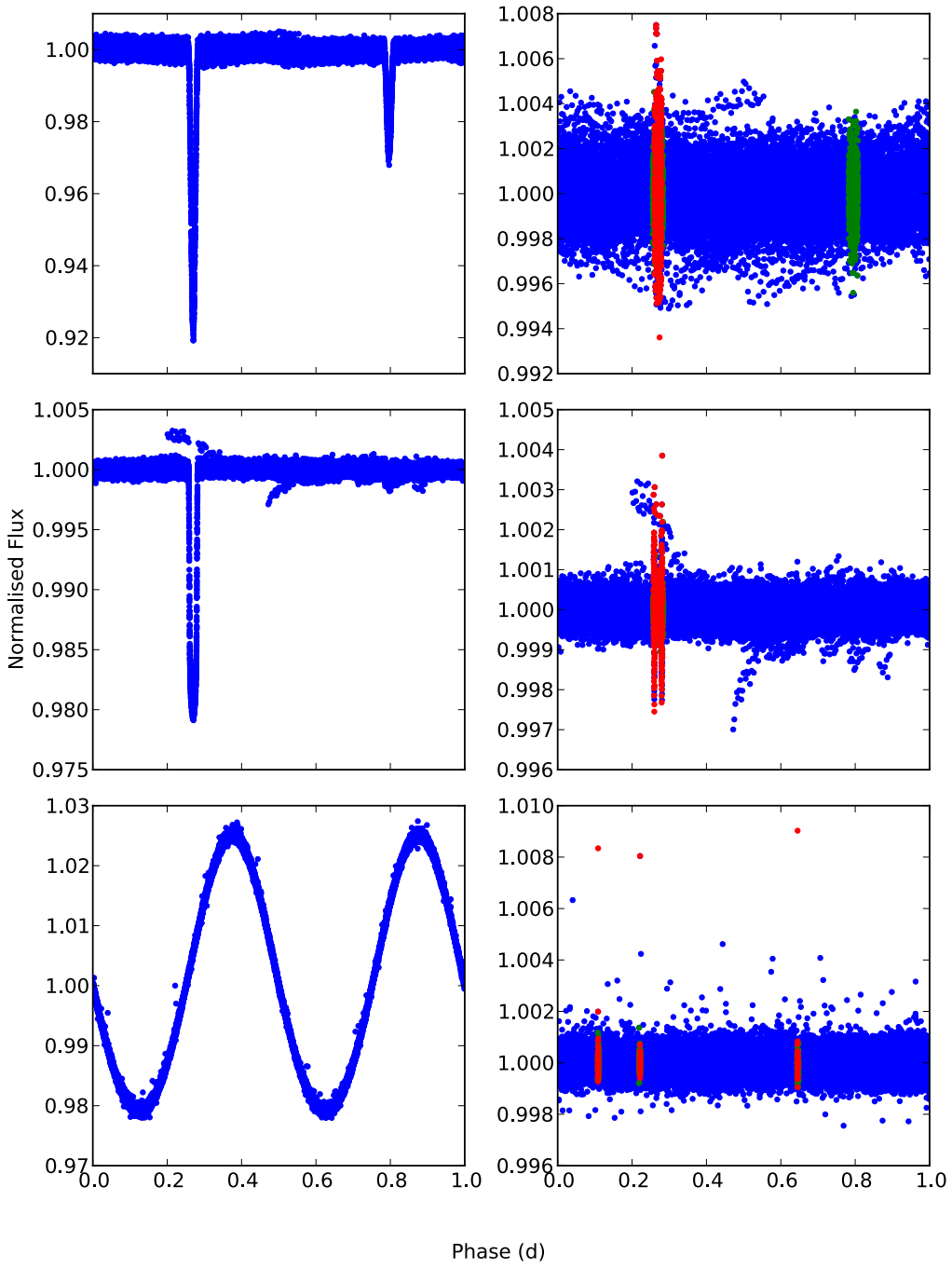

Figure 4.2: Examples of pre (left) and post (right) whitening light curves, with potential issues. Green points show points marked as being in eclipse, but successfully whitened, red points show points which retain some deviation after the whitening process. From top to bottom we see: general eclipse remnants, remnants of the sharp ingress and egress, and false point marking due to 3 discrepant points.

curve, this tended to result in periods insufficiently accurate for the data given. As such, initially Kepler object periods were refined using the AOV technique.

4.3 Individual Event Search

After these detrending and whitening procedures some noise remains in the light curve, the majority due to stellar activity or remaining instrumental effects. Several forms of search were tried, whereby the light curve was directly phase folded and processed as described in Section 4.4, or had further detrending steps applied first. The most successful results came from an algorithm designed to scan for individual transit events, applied before any periodicity tests were carried out. This also had the advantage of not introducing extra steps into the process - as the eventual goal of the search was to test statistically the properties of circumbinary planets, the fewer potentially biasing steps which are involved the better. Transits in a continuous light curve (such as those from Kepler) are strong candidates for such an individual search. Although they may not be deeper than the overall light curve variation (and may be well hidden to the naked eye without careful and close expansion of the light curve) the extremely fast change which occurs as the planet passes in front of the stellar surface is a very useful signal. This may be less clear in circumbinary systems - the ingress and egress may be extended due to the motion of the host star - but in general the clear two-level nature of the transit persists. As such, as long as enough transits exist to increase confidence that we are not seeing some form of background blend, this is a promising avenue of exploration.

The specific procedure used was as follows. The non-stellar component of the remaining noise often occurs surrounding gaps in the data; as such, 0.5 d regions on either side of gaps (defined by a greater than 0.5 d space between two adjacent datapoints) were ignored. A spacing of 0.5 d corresponds to ~20 data points in the Kepler long cadence. Typically these gaps are between Kepler quarters, or at regions where previously flagged 'bad' data points have been removed. Points falling at known binary eclipse times were also screened (only for well detached binaries with morphology <0.2, such that the eclipses could not take up a large proportion of the light curve). These binary eclipse times were taken from the data given by the Kepler EB catalogue.

A box was then passed across the light curve at a 0.1 d resolution (i.e. 4-5 data points). This box width was chosen to maximise resolution while minimising the time it took the algorithm to run - smaller box widths lead to excessively long computational times. At each step, a 3 d window centred on the current box was

taken. Three days was chosen to give a significantly long baseline, while still being short enough to track variability. Periodic noise or stellar variability with timescale less than the baseline fitting region will obscure planetary signals. A 3D polynomial was fit to this region, excluding the central 0.1 d box. Gaps were not fit across, due to discontinuities in the data often marked by a significant gap. This fit was repeated for 20 iterations, with points $> 5\sigma$ from the best fit excluded each time so as to avoid including transits (or other sharp discontinuous events) which were not in the central box but within the 3 d window. The effect of these iterations can be seen in Figure 4.3, for a transit of Kepler-16b. The offset of the central box from the resulting best fit baseline, relative to the standard deviation of the 3 d region around the best fit baseline, was then taken and stored. After the whole light curve is tested, any times with offset significances $> 3\sigma$ of the whole set of significances are passed on to the periodicity test. Transits of longer duration than 0.1 d are then represented as consecutive event detections, while those shorter become difficult to resolve in the long cadence Kepler data anyway.

4.4 Quasi-periodic events

Due to the large transit timing variations (TTVs) on the order of several days in circumbinary planet signals, as well as the possibility for multiple transits to appear on any given orbit, events cannot be held to be strictly periodic and the usual methods for forming periodograms cannot be used. Periodicity of the detected single event times was tested by phase-folding these times at a series of trial periods. In the case of a real planetary signal, at periods close to the true (specifically, azimuthal, see Chapter 1) planetary period the event times will group in phase. There is a defined maximum width to this grouping, which was derived in Chapter 3. As such a box with the width of this maximum, is passed over the phase folded array of times at each trial period (see Figure 4.4 for an illustration). To obtain a useable box width, several approximations needed to be made to the formulae. Only the geometric contribution (equation 3.7) was used (i.e. from the motion of the stars, and ignoring planetary precession). The necessary parameters were taken from the Kepler EB Catalogue, as well as eccentricity parameters as described in Section 1.3.2 for each binary. Stellar mass ratios were set to give equal mass stars - while not representing the true mass ratios of the systems studied, this lets the equation give the largest, and therefore most robust, upper limit on the TTVs. The tighter this limit is made the more efficient is the search algorithm, but using this more robust limit ensures that transits are not missed.

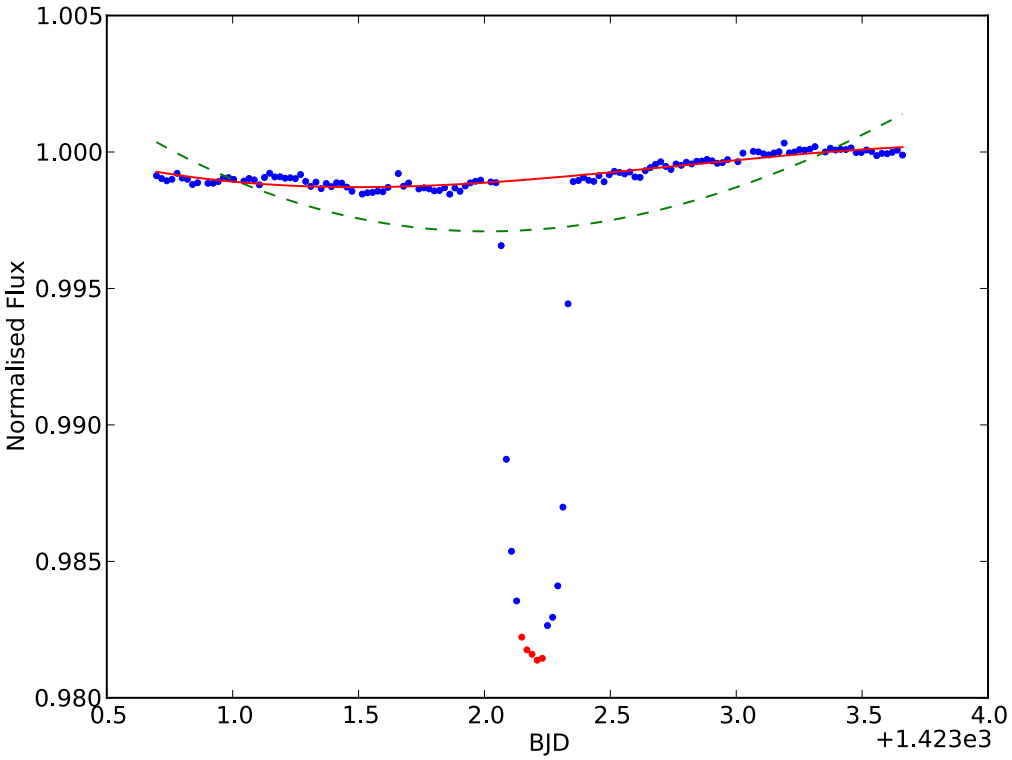

Figure 4.3: A transit of Kepler-16b, with the 3 day region used for finding the baseline plotted. Red points show the central box being tested. The green dashed line is the initial polynomial fit to the baseline, whereas the red line is the final fit used after 20 progressive clipping iterations.

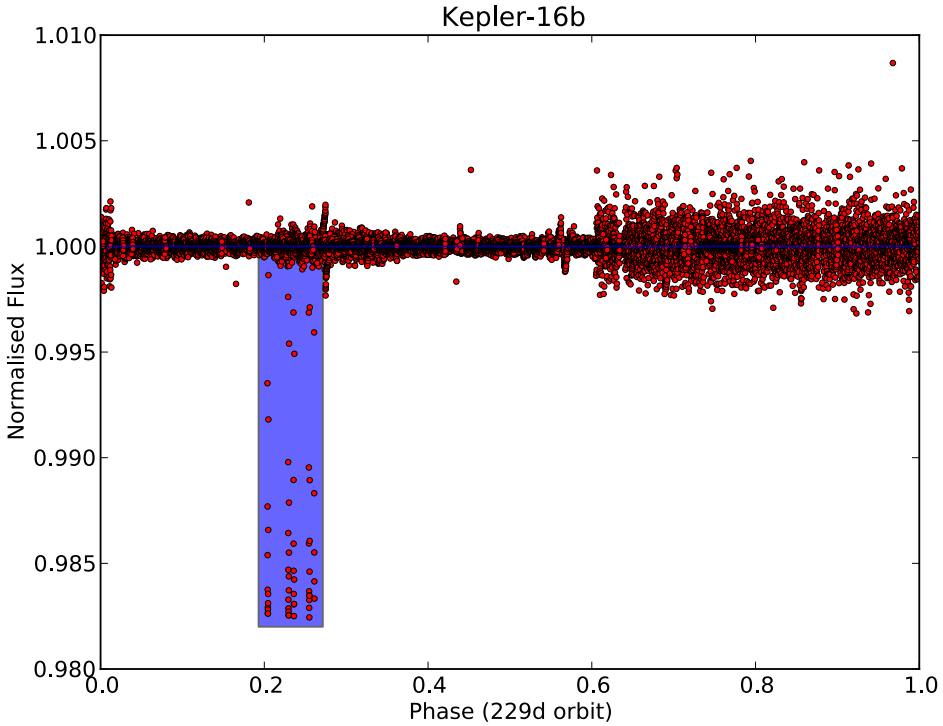

Figure 4.4: The light curve of Kepler-16b, phase folded at the planetary period. The blue box illustrates the search box used, and lies over the region showing planetary transits. The excess noise at phase¿0.6 is due to a small number of noisier quarters occurring later in the Kepler mission lifetime.

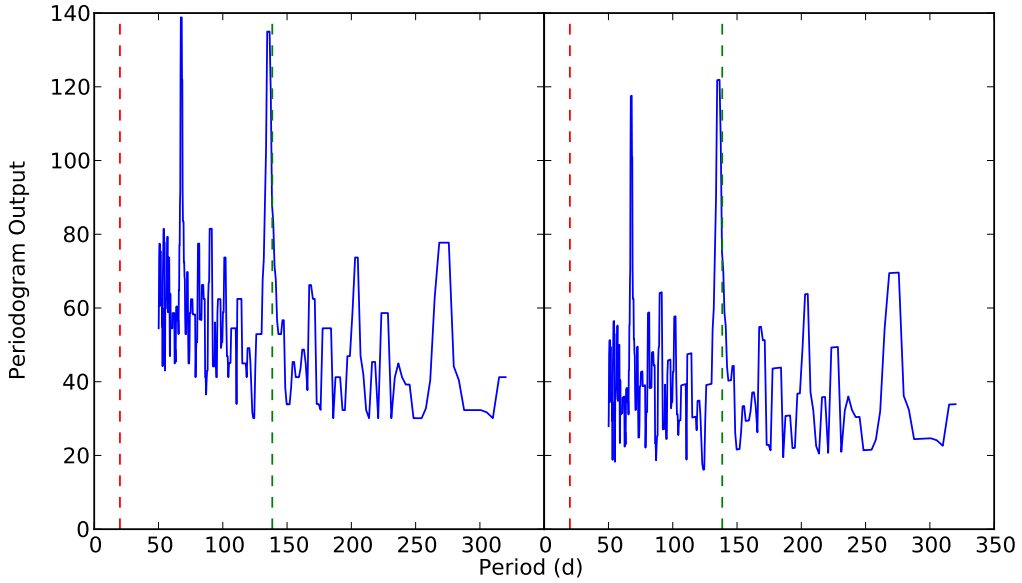

Figure 4.5: The periodogram of PH1, before correction (left) and after (right). Red line shows host binary period. Green line shows published planetary period.

For Kepler our test periods range from 320 d down to either 2 d or 2.5 $P_{\rm bin}$, whichever is longer. A uniform frequency resolution was used, with $\triangle f = 5 \times 10^{-5}$. The 320 d limit is used so as to avoid a hard periodogram limit at the same point as our limit for statistical purposes, which was 300 d. It is well below the full duration (~4 years) of the Kepler data, which allows for at least 4 orbits of a 300 d period planet. Two days is where individual transits may become hard to detect in the long cadence (30 min resolution) Kepler data, and 2.5 $P_{\rm bin}$ is set such as to be well within the inner stability limit given by Holman & Wiegert (1999), which is typically 4-5 $P_{\rm bin}$ for circular binaries. While these parameters were used for the statistical work of Chapter 7, they were extended when simply searching for new systems. Shorter periods were tried, and even (somewhat optimistically) periods below that of the host binary. However these did not lead to any major findings.

The total significance of all the event times within the box was then saved at each trial period, forming a periodogram over the whole tested range. As the maximum box width and so number of data points contained within the box grows for smaller planet to binary period ratios, a preference for shorter periods is introduced. This can be seen in Figure 4.5, which shows a periodogram (of the planet PH1). This was removed from the periodograms by applying a weighting of the inverse of the box width to each given period.

This periodicity search has some inherent weaknesses. In particular, harmonics of the true planetary period are detected strongly. This is due to the width of the search box used - for a period at half the true period, the planet to binary period ratio is halved, which significantly increases the box width. This allows more phase space to be counted as significant, while still picking up the same detected transits. Any significant other signal, be it noise or stellar activity, can then contribute to the significance of that period. It should also be noted that this periodicity test returns detections at or close to the mean transit interval or its harmonics. This is different, and generally a few days less than, the instantaneous Keplerian period, which is what has been published. See Section 1.2.5 for a fuller discussion.

The periodicity checking part of the search algorithm went through several stages of development before reaching the process described above. In particular a variety of search runs were trialled where the full light curve was phase folded, then scanned using the same transit timing variation defined box described (i.e. without individual event searching), measuring the reduced chi square of points within the box. This resulted in some success (a few of the candidates presented here and in Chapter 5 were detected first in this way) but preferentially found non-planetary periodic events, primarily pulsations. A number of refinements were enacted; an 'inverse box' operating on increases rather than decreases in flux and used to measure light curve noise, smoothing of the signal by minimising the chi square outside the box (as opposed to maximising that inside), and smoothing functions used to balance uneven densities of points at certain periods, such as harmonics of Kepler's quarters. However, the method described combines both success in planet finding with a degree of simplicity compared to these other attempts, making it the preferred option.

4.5 Output Statistic

While it is possible to detect planetary candidate signals from the periodograms by eye, a quantifiable output statistic is much more preferable. Due to the large TTVs, and the possibility of multiple transits appearing on a single planetary orbit, this must be formed using a more unusual method. In particular these issues combined with the search algorithm lead to a tendency to detecting harmonics - here the maximum peak of a typical planet detection periodogram is often a harmonic of the true period. This is due to the maximum TTV region tested by the algorithm at each period. If any planetary transits do not completely fill this region, it is possible for harmonics (which have a different maximum TTV, but one which may

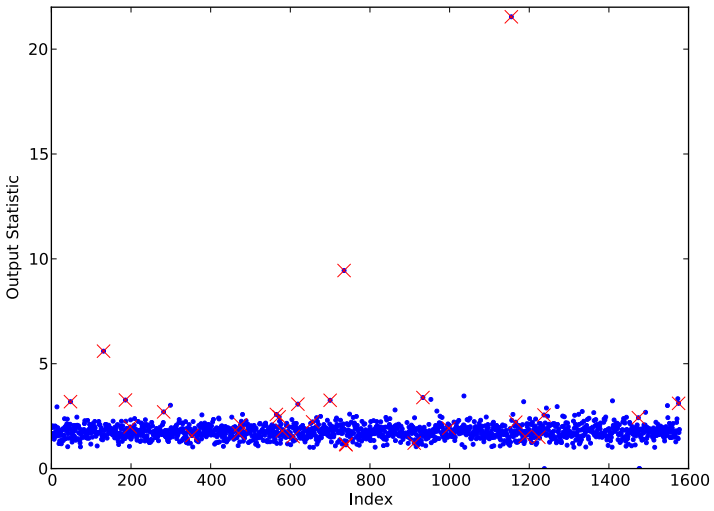

Figure 4.6: Θ_s for the non-overcontact binaries of the sample. Known planets, candidates and multiple systems (from Tables 4.1, 4.2 and 5.1) are marked with red crosses.

still cover all the known transits) to give strong periodogram peaks. This effect is not so prevalent for noise, and so the presence of strong harmonics can be used to advantage – when finding an output statistic those at $P_p/2$, $P_p/3$, $2P_p$ and $3P_p$ were used, where P_p is the tested planet period. Using additional harmonics to form the average was found to not significantly improve the results. The mean value of the maximum peak and these harmonic peaks was taken, each divided by the median value of the periodogram to take account of the different levels of noise between objects. This led to a statistic given by

$$\Theta_s = \frac{\sum \theta_p(P_{\text{resonance}})}{N_{\text{resonances}}\tilde{\theta_p}} \tag{4.1}$$

where θ_p represents the periodogram statistic (i.e. the sum of event significances found in the best fitting box, see Section 4.4), $\tilde{\theta_p}$ its median, $P_{\text{resonance}}$ the period of one of the above mentioned resonances, and the sum is over these resonances. The distribution of Θ_s is shown in Figure 4.6, with known planets and triple stars marked.

4.6 Human Eyeballing

While the statistic of Section 4.5 allows for a degree of objectivity, which is used particularly in Chapter 7, to ensure no interesting objects were excluded the entire

sample was also individually checked by two researchers. This led to the discovery of several candidates (both planetary and not) which due to several missed transits (through either light curve gaps or misaligned planets) or particularly high light curve noise would not otherwise have been found. In these cases good periodograms can be produced by the model (in that they find a correct period), but with peaks at too low significance to stand out from the noise. Kepler-413b, which does not transit for ~3/4 of the light curve, was found in this way, as was the similar candidate KIC9632895.

4.7 Results

4.7.1 Known Planets

An ideal test sample of previously published circumbinary planets is available in the Kepler data. These were listed in Table 1.1, and are Keplers -16b, 34b, 35b, 38b, 47b and c, 64b (PH1), and 413b. There is in addition an as yet unpublished but hinted at third planet in the Kepler-47 system. If functioning as it should, we would expect the search algorithm to detect these planets. Fortunately it does with high significance, in all cases but the Kepler-47 and 413 systems. The detected significances are shown in Table 4.1. Furthermore, each of the known systems except Kepler-47 was highlighted as among the strongest candidates by both researchers who eyeballed the data in a blind test. The periodograms of 6 of the 7 systems are shown in Figure 4.7. Kepler-64b (PH1) is discussed below. Kepler-47 has a particularly disadvantageous light curve, showing strong and rapid stellar activity induced variations. This, along with the masking of several of the ~300d planet (Kepler-47c), deeper transits, resulted in only a small detection. For the ~50d planet (Kepler-47b), the transits were particularly shallow, and while visible to the eye, were masked by the systematic noise in the algorithm caused by the stellar activity. Kepler-413b, while the correct period was detected, has reduced significance due to its misaligned nature - it only transits on some of its orbits within the Kepler data. As such, it is good that the search algorithm picked up the period (and led to the system being flagged as a strong candidate), and is not surprising the the formal detection significance was low.

The overall dataset (of non-overcontact binaries) has a median Θ_s of 1.73, with 95% of the values below 2.31. As such, the detection significances for the planets shown in Table 4.1 are high, as would be hoped for a well functioning algorithm. The two notable exceptions are Keplers 47b and 413b, which were discussed above.

At this stage of running the algorithm, PH1 had not yet been discovered. It

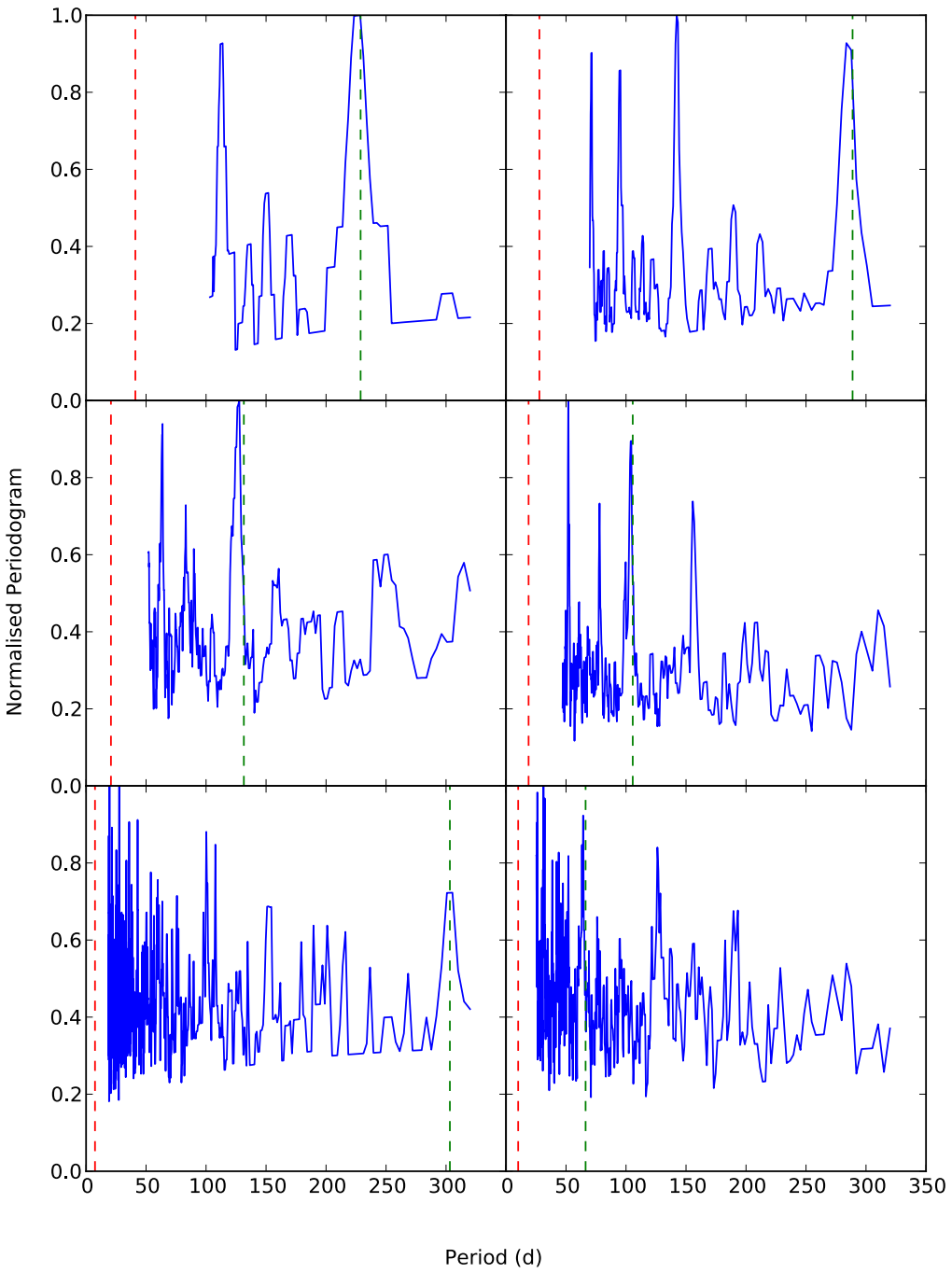

Figure 4.7: Normalised periodograms of (from top left): Kepler-16,34,35,38,47,413. Red dashed lines show the host binary periods. Green dashed lines show the known planetary period. Note that the published period is generally the Keplerian period and is a few days larger than the azimuthal period, which is what the periodograms will detect.

Table 4.1: Known Planet Significances. Note Detected Periods are azimuthal, and expected to be up to a few days below published periods (or harmonics).

Planet	$P_{published}$ (d)	$P_{detected}$ (d)	Θ_s	Comment
Kepler-16b	228.8	228.4	3.11	
Kepler-34b	288.8	142.2	3.39	$(P/2)$
Kepler-35b	131.5	127.7	2.21	
Kepler-38b	105.6	52.0	3.07	$(P/2)$
Kepler-47b	303.2	27.6	1.50	Non-detection
Kepler-64b (PH1b)	138.5	136.4	2.70	
Kepler-413b	66.3	31.2	1.52	$(P/2)$

Table 4.2: Candidate Planets

Kepler ID	P_{bin} (d)	e_{bin}	$P_{candidate}$ (d)
5473556	11.26	0.15	550 or 1110
6504534	28.16	0.094	~170
9632895	27.32	0.093	~240

presented a strong signal, the periodogram of which is shown in Figure 4.8. As a strong new candidate, it became the subject of dedicated follow up (such as transit modelling - Section 4.8, and binary eclipse timing - Section 2.7). A directors discretionary time proposal was also in place at the Nordic Optical Telescope to obtain radial velocities of the host stars. While this follow up work was in progress, the planet was published by two other research groups, officially becoming PH1 or Kepler-64b. As such we changed strategy to investigate overall occurrence rates.

4.7.2 New Candidates

Beyond PH1 three strong candidate planetary systems were detected, two of which lay within our period limit of 300 d. While these have not been followed up beyond initial vetting, they are discussed here. Table 4.2 gives the input and derived parameters for these candidate systems.

The shortest period object, KIC6504534, shows three clear transits, with several gaps where others would be expected to fall. These transits imply a ~170.3d planet, showing transit timing variations of at least 0.1 d, as well as transit duration variations of a similar magnitude. While this planet is not confirmed, the presence of clear transits with strong timing and duration variations supports the hypothesis that it represents a real signal.

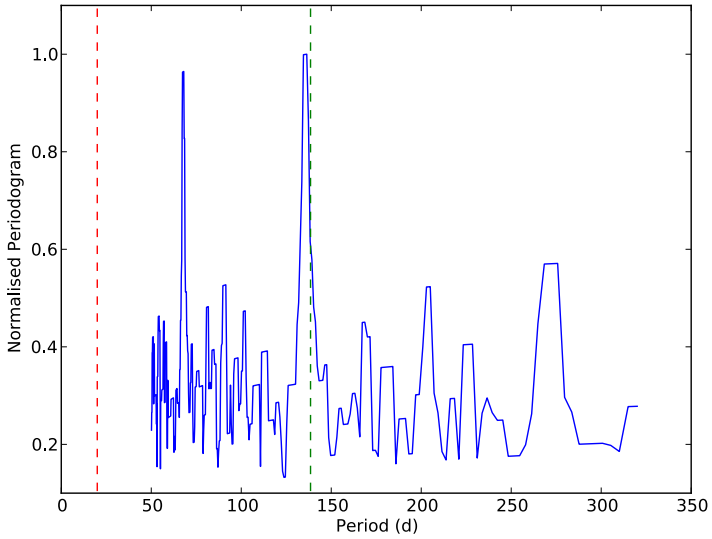

Figure 4.8: Normalised periodogram of Kepler-64 (PH1). Red dashed line shows the host binary period. Green dashed line shows the known planetary period. Note that the published period is a few days larger than the azimuthal period, which is what the periodogram detects.

KIC 5473556 was mentioned in Welsh et al. (2012) as showing a single transit. There are now two, implying a period of 550 or 1100 d (due to a gap in the light curve where a transit could have been missed). This candidate does not have enough transits to show TTVs, leaving the possibility of a background blend open. The remaining candidate, KIC9632895, shows three transits, implying an ~240 d period with TTVs of magnitude over 1d. There are however light curve regions where consecutive transits should lie, implying that this candidate is on a slightly misaligned orbit.

Beyond these 3 strong candidates, a number of weaker candidate objects were found. One class of these are the single-transit-events, light curves which show one single event which is clearly a potential transit but have no other events to back it up. These transit features could arise from systematic noise, but could also represent either long period or misaligned planets. The list of Kepler IDs in this category is shown in Table 4.3. A proposal is in place to investigate the host stars of these systems. Several other objects were found which exhibit multiple transit-like signatures. These all have problems preventing them becoming strong candidates, such as transits which are weak or unclear, minimal TTVs, or transits suspiciously resonant with the binary eclipses. They are also listed in Table 4.3.

The search algorithm also detected several eclipses too deep to be planets.

Table 4.3: Single Transit Events and Weaker Planet Candidates

Weaker Planet Candidates Kepler ID	Single Transit Events
2452440	2442084
3547874	2719873
3757778	4144236
4076952	5039442
4847832	6105491
6187893	6889235
7031714	9007918
7684873	9596187
7747425	9957668
8454250	10068795
8700506	10809677
8702921	12367310
10092506	
12306808	

Many of these are already known multiple star systems, and these and some other interesting systems are shown in Chapter 5.

4.8 Keplerian Transit Model

In a further attempt to test the validity of these candidates, and to provide some clues as to their underlying parameters, a three body model was developed. This was designed as a first order method to test transit times, and functioned using Keplerian orbital motion, treating the binary as a point mass from the perspective of the planet. Using a test set of planet parameters the three bodies could be allowed to orbit, and times of transit of the planet on either host star extracted. Unfortunately the combined three body effects (e.g. precession, orbital perturbation), while being weak enough to allow the approximations of Chapter 3 to function with only small errors, nevertheless preclude accurate mapping of transit times themselves. An attempt was made to fit the observed transits of PH1, while they were as yet unpublished, and hence derive some parameters of the system. However, without accurate parameters for the host stars (found through radial velocity data), the parameter space showed both multiple minima and even in these, no good fit to the transits.

4.9 Summary

This Chapter presented a new search algorithm derived from the equations of Chapter 3. This algorithm was then applied to the WASP and Kepler datasets, and a number of new candidate planetary systems presented. The results of the search algorithm on previously known planetary systems were also discussed. These results will be used in Chapter 7 to study the statistical properties of circumbinary planets.

Chapter 5

Detected Non-Planetary Systems

In the course of running and developing the search algorithm described in Chapter 4, several other results were obtained of a non-planetary nature. These are generally interesting binaries and triple stars, and are discussed in this chapter. In particular there is a focus on KIC002856960. At the time of discovery it was unclear whether this was a hierarchical triple star or a circumbinary planet, and it is presented as if both scenarios are possible. At that stage, I published the discovery in Armstrong et al. (2012). A significant part of this Chapter is drawn from that publication. After this work, Lee et al. (2013) published evidence for the hierarchical triple scenario, leading to the understanding that KIC002856960 consists of a K star orbited by a close pair of M dwarfs. The M dwarfs eclipse each other on a 6.2 hour period, and produce further unusual transits of the K star every 204 days. At the time of the work only 6 public quarters of Kepler data were available, displaying three of these 204 day period events. Since then, many more quarters have been made public, and display several more. Furthermore, new work undertaken on the system in collaboration with T. Marsh has shown that even the K star/double M dwarf interpretation is in doubt, as described at the end of this Chapter. The system remains an intriguing mystery.

Table 5.1: Multiple Systems

Kepler ID				
2835289	2856960	3832716	4150611	4247791
5007640	5255552	5897826	5952403	6144827
6462863	6543674	6678383	6964043	7289157
7456992	7622486	7668648	7670485	7871200
8478994	8948424	9471755	9790965	9843451
9909735	10030943	10091110	11601584	

Table 5.2: Other Signals

Kepler ID	Comment
5880661	Strong pulsation signal on 10 P_{binary}
6606653	Cepheid variable like signature in resonance with P_{binary}
8113154	Broad 5d long faint regions on 40d period
8164262	Heartbeat
9835416	Excited pulsations, possibly as result of heartbeat type binary
10223618	For several consecutive quarters binary secondary eclipses gain an additional 1% dip just after eclipse

5.1 Multiple Star Systems

A wide variety of triple and higher order stellar systems have become apparent during the course of implementing the search algorithm. Although these systems were not investigated in great detail, they represent the 'easy to find' end of the signals we could hope to discover, and moreover are systems of some interest. Such systems, especially when as close as can be found in the Kepler data (all orbital periods within a few hundred days) can exhibit eclipse timing variations, rapid orbital evolution, and are particularly useful clues for stellar evolution, given the coeval nature of the stars involved. Also, as is the case for planets in binary systems in some cases, it is possible to determine their absolute masses and radii. Table 5.1 gives a list of such systems we found in the Kepler data. Many of these systems have been found by others, and there are likely more systems in the Kepler data. Each show signals consistent with additional stellar companions. Some of these systems are liable to represent blended binaries, but many are true triple or multiple component stars. Table 5.2 shows other interesting signals which arose during the search. Some particularly interesting multiple star examples are discussed below.

5.1.1 7871200

KIC7871200 shows clear, ~2% depth eclipses on a 38 d period. These eclipses occur in pairs, separated by ~1.5 d. Initially it was though that these eclipses could be transits of a large planet, and the 1.5 d range the geometric range allowed for by the motion of the host stars. However, 1.5 d is too large, given the short, 0.24 d period of the host binary. Under Equation 3.7, this would lead to a maximum 0.4 d crossing for circular planet and binary. Adding eccentricity to the planet does not have a large enough effect to account for the discrepancy. This leaves open the possibility of the tertiary events being eclipses (primary and secondary) of a third star, on an extremely eccentric orbit. To produce the observed eclipse locations and widths this star would have to have orbital eccentricity of 0.72. It may also be a contaminant, which would imply a separate, blended, high eccentricity binary rather than triple star.

5.1.2 5952403

KIC5952403 is a 7th magnitude star which is also known as HD181068. It was discovered and partially analysed using the eclipse timing code (see Section 2.7), producing the O-C curve shown in Figure 5.1. The system, discussed in Derekas et al. (2011), contains a red giant star in a 45 d orbit with a pair of red dwarfs, themselves orbiting each other every 0.9 d. During eclipses of the giant star by the smaller pair, the short period eclipses change to become less deep, and for one of the dwarf stars even small brightenings, providing a direct window onto the comparative stellar surface brightnesses.

5.2 Orbital Perturbations

A number of systems show the signature of strong and rapid orbital variation, in the form of eclipse depths which change dramatically (occasionally even appearing or disappearing) over the ~4 year course of the Kepler data. An example is shown in Figure 5.2. These stars are undergoing rapid inclination changes, possibly due to the presence of an unseen stellar companion. They are flagged in the current edition of the Kepler EB catalogue, a represent interesting avenues into exploring observationally dynamical interactions on short timescales.

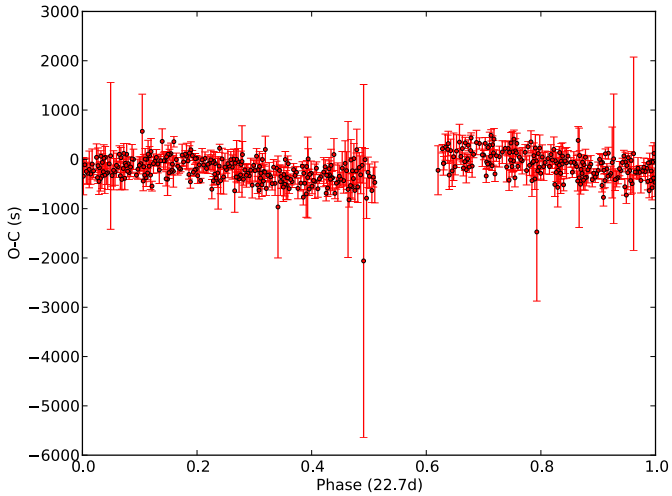

Figure 5.1: Observed - Calculated eclipse times for the longer period orbit of HD181068. Half the orbital period is used, as it was unclear at the time what the true period was. This was subsequently made clearer through radial velocity (Derekas et al., 2011).

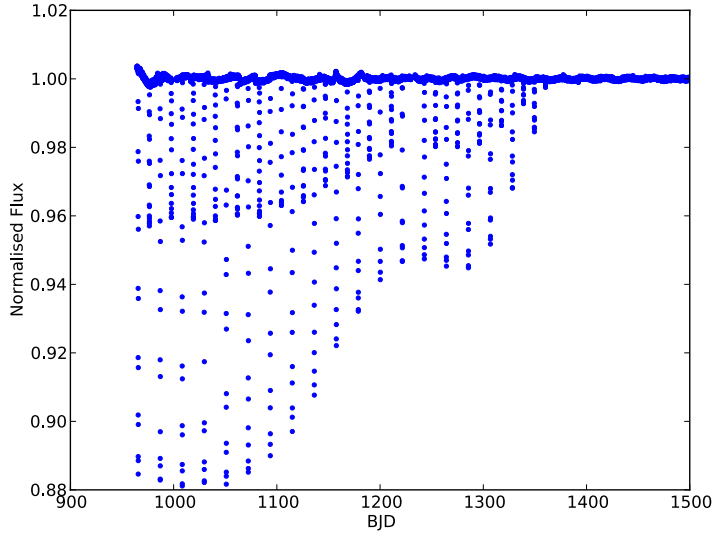

Figure 5.2: A section of the light curve of KIC010319590. The binary eclipses can be seen rapidly reducing in depth over only a few hundred days.

5.3 Heartbeat Stars

One of the first tertiary signals discovered arose from KIC8164262, which shows eclipses on an ~87 d period. These have an unusual brightening immediately after eclipse. Prior to the Kepler survey, this class of binary stars was unknown. They represent highly eccentric binaries, whereby at periastron the stellar separation reduces to the point where the stars become tidally distorted, and pulsations can be induced. A typical light curve is shown in Figure 5.3. This distortion and excitation results in a short lived brightening, in a similar but more concentrated fashion to the more common ellipsoidal variations. The shape and duration of the effect all depend on the orbit involved, primarily the eccentricity, argument of periapsis, and inclination. The first analysis of this type of star system, and the potential applications, can be found in Thompson et al. (2012), who also note that the amplitude of the brightening appears to increase with increasing stellar temperature. This possibly arises due to the dependence of the amplitude of the tidal distortions on both stellar surface gravity as well as the size of the tidal force from the secondary (see Pfahl et al., 2008), implying an increased amplitude for later type stars which have lower surface gravity. Where pulsations are excited, there is the possibility of performing asteroseismology using the frequency spectrum of the light curve (Chaplin et al., 2013) providing a new insight into the stellar properties.

5.4 KIC002856960

5.4.1 Overview

KIC002856960 is an object in the Kepler EB catalogue, displaying previously known short period variations as shown in Fig. 5.4. There are three extra new events present in the public data, in quarters 1, 4 and 6, and these are separated and presented in Fig. 5.5. They display a unique appearance, indicative of the presence of at least three bodies in the system.

The short period variations have not been amenable to modelling by the catalogue team. As such, while the assumption is made that the source of the short period variation is an eclipsing binary system (of a perhaps complex nature), no binary parameters beyond the ephemeris of the short-period variation are provided. An updated binary orbital period for this signal was derived using all of the publicly available quarters (1-6 at the time) via χ^2 minimisation of a binned form of the data using the phase dispersion minimisation technique (see Sect. 2.6.3) and is presented along with the catalogue value for T_0 in Table 5.3. The original binary signal is

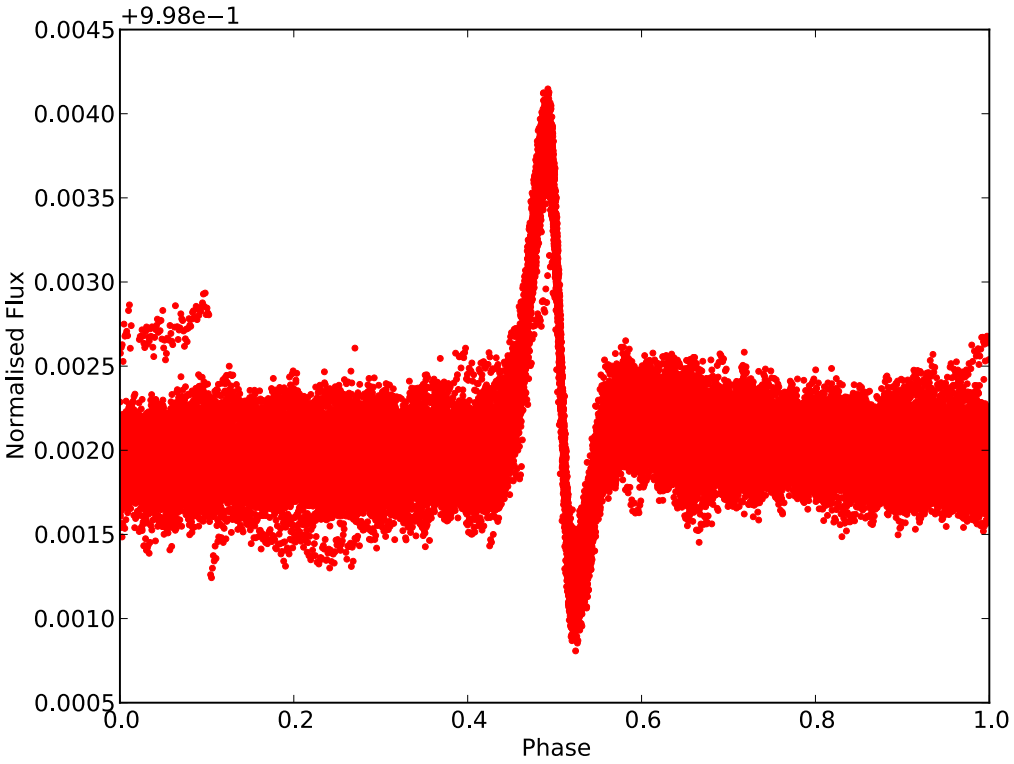

Figure 5.3: KIC003547874, phase folded at the binary period of 19.69 d. The light curve shows a classic heartbeat star shape.

approximately 1% in depth, notably shallower than the newly detected signal.

From here the bodies in this system will be referred to as Components AB (for the two members of the binary system), and either b or C depending on whether the third body is considered as a planet or star respectively. The analysis is limited to considerations of three bodies only, for simplicity. Each primary new event (hereafter 'crossing', as each event consists of a crossing with the binary orbit by a third body) contains multiple transits of A or B before C, or b/C before A or B (the distinction depends on the system architecture, see Sect. 5.4.4).

5.4.2 Tertiary Events

Ephemerides for the three observed crossings are given, along with durations and depths, calculated from the deepest point observed in each crossing. The implications of these results for the interpretation of the nature of the system are discussed in Sect. 5.4.4. The crossing duration is given for comparison with the binary period - the first crossing in the quarter 1 light curve is cut off by the end of the quarter before completion of the crossing, and hence a duration is not provided.

If secondary eclipse events in the long period signal (of component b/C by AB, or AB by C) are present, they are constrained by the variability of the data to have an upper limit to their depth of ~0.4%, and no evidence for any is observed in the phased light curve.

Time based parameter errors are dominated by the long cadence, in the absence of a full model. Crossings are taken as beginning midway between the data point immediately before flux variations are seen and its successor (the choice of these points is clear on viewing the data), and the absolute errors set such that the beginning of ingress could happen anywhere between these points. End times are derived similarly. All parameter errors derive from formal propagation from these values or the raw flux errors as appropriate.

5.4.3 Vetting

It has been noted that the depths for each transit are markedly different; this is expected but quarter to quarter variations in the depth, along with the large size of the *Kepler* pixels of ~4 arc sec (Koch et al., 2010) can cause suspicion as to whether the object involved is blending with a background source. The photometric aperture used changes between quarters to optimise the photometric precision, meaning that a significant background source may contribute a different amount of flux each quarter. To mitigate this concern, KIC002856960 has been subjected to a pixel

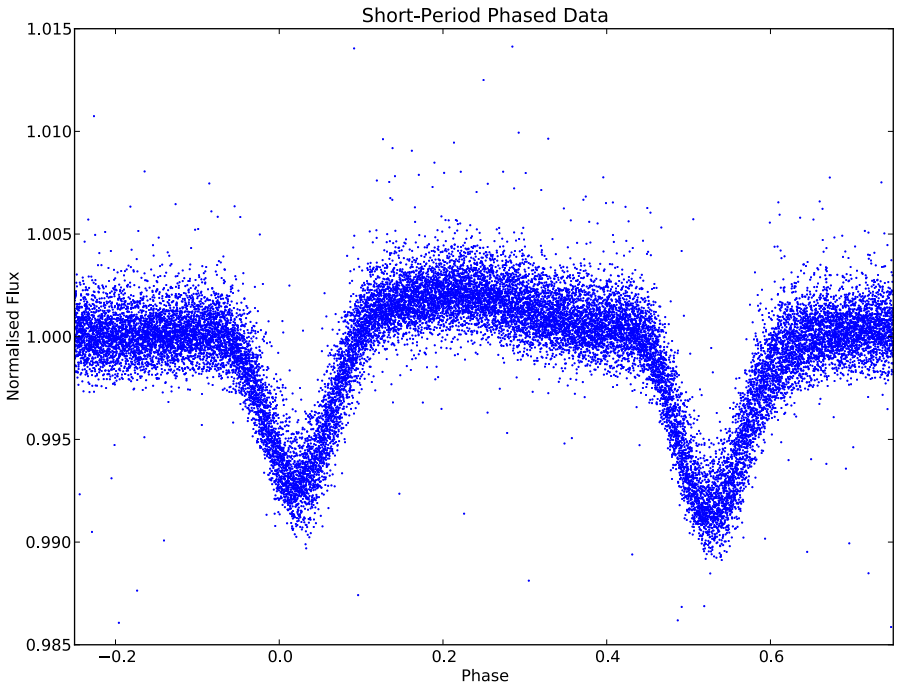

Figure 5.4: The binary variation of KIC2856960 (eclipses of A and B by each other), shown as phase folded data from all 6 quarters (scale shown is relative offset to the median of each quarter). The tertiary event has not been removed, and as such some points are omitted below the figure window. These are displayed more clearly in Fig. 5.5. Error bars are not plotted, as the spread of points gives a much clearer indication of the true error.

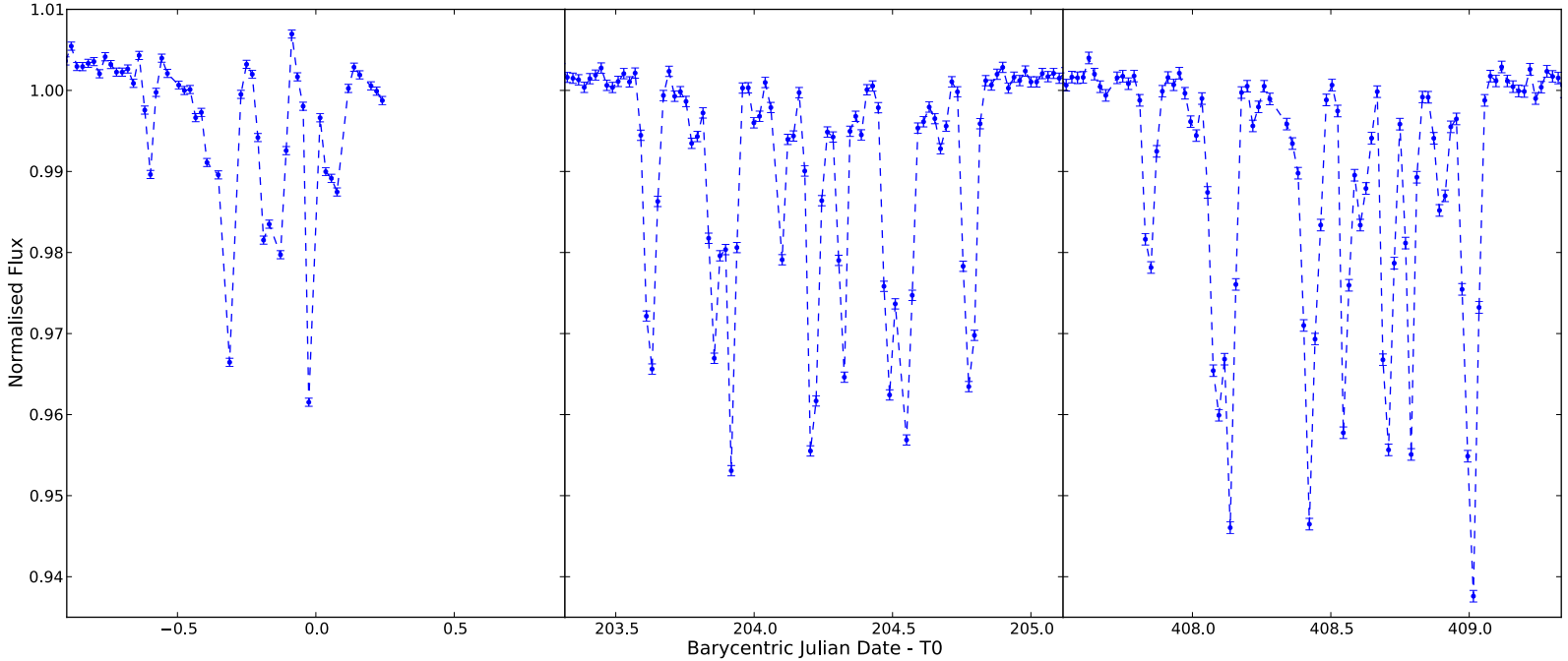

Figure 5.5: The three tertiary crossings available in the public data, in quarters 1, 4 and 6. Crossings are referred to as Crossing 1, 2 and 3, from left to right respectively. Light curves are displayed after removal of the binary signal. Dashed lines are simple connections of data points, for clarity, and do not stem from a model.

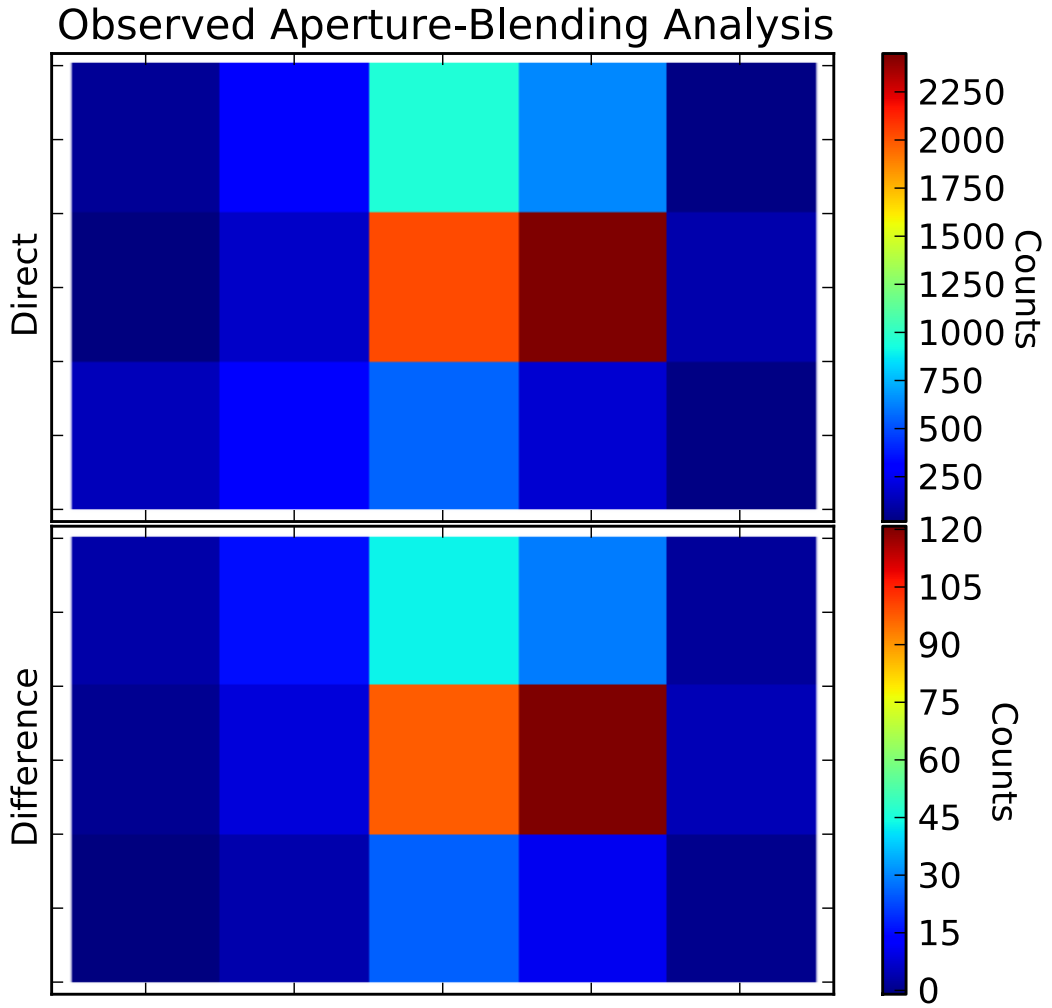

Figure 5.6: Images formed from the observed Q4 aperture for KIC002856960, using the method as utilised by Torres et al. (2010). The direct image shows the location of the primary source, whereas the difference image shows the location of the flux variation. No discernible difference is observed.

Table 5.3: Derived parameters for the KIC002856960 system. Parameters are labelled AB or b/C, standing for signals associated with the binary stars (AB) or companion (b/C). BJD stands for Barycentric Julian Date. T_0 given as T_0 - 2400000. A duration for crossing 1 is not provided as this crossing is cut off by the end of a quarter before completion.

Parameter	Units	Result	Error (absolute)
Period (AB)	d	0.2585082	...
T_0 (AB)	BJD	54964.652115	...
Period (b/C)	d	204.2163	0.030
T_0 (b/C)	BJD	54997.743313	0.024667
Duration (b/C)	d		
Crossing 2		1.2464	0.0204
Crossing 3		1.2669	0.0204
Depth (b/C)	%		
Crossing 1		4.1171	0.0588
Crossing 2		4.8179	0.0669
Crossing 3		6.3207	0.0749

response function-based centroid analysis as utilised by Torres et al. (2010) and described in Section 2.5.

Fig. 5.6 shows the difference image produced, and the comparison direct image (as a reminder, the direct image shows the location of the primary flux source, and the difference image the location of the flux variation). Blends originating from an offset of order 1 pixel are easily ruled out by inspection. The fit of a pixel response function (a modified point spread function - Bryson et al. (2010b)) limited the region where a blending source could be located to within a radius of order 0.3 pixels from the target, but was complicated by the presence of the binary signal. Despite the difficulties this does rule out a large region of the photometric aperture as the location of a background varying source.

The nature of the tertiary signal requires that three bodies are present in the system; given the constraints on background sources given above a blended background object would have to lie very close to our line of sight to mimic the signals seen, and be a triple star itself. The depth of the tertiary signal also indicates that, if it was caused by a background source, the contaminating source would need to be of very significant brightness compared to the target. No evidence for such a source is seen. As such, it is very likely that KIC002856960 is indeed the source of the signals observed. Estimating the probability of this is however non-trivial given the prior selection of this object as interesting.

5.4.4 Discussion

Summary

Our primary interest lies in the tertiary signal detected. The public data contains three crossings (Fig. 5.5) appearing in quarters 1, 4 and 6. These crossings occur on a period much longer than that of the binary AB, and are indicative of the presence of a companion third body b/C either crossing or being crossed by the binary AB. The short period binary completes ~5 orbits during the course of the crossing, meaning that during a crossing multiple transits will be produced as the components A or B move in and out of alignment with the slower moving third body b/C. The shape of the transits seen within a crossing will be dependent on the velocity and relative positions of all components, as well as the components' shape and surface brightness. In the absence of a full model or complete mapping of the flux variations (due to the long cadence data only just resolving the inner transits within each crossing) we have presented the simplest observable parameters available: ephemeris, depth and overall crossing duration. Since the ingress/egress times are expected to depend on the location of all three bodies at the start/end of each tertiary event, an accurate period cannot be determined with only three crossing events. This situation will be largely improved once the currently private data becomes accessible.

The structure and nature of this system remain open. Here we discuss the two most likely system geometries: an object in orbit around a close binary pair (system (AB)b or (AB)C), or a close binary pair in orbit around a larger star (system C(AB), where C is the more massive star in this case). We consider these two options in turn, acknowledging that there may be other more unusual configurations possible. The equal depths of the AB binary eclipses (Fig. 5.4) imply that, if the binary components AB were on the main sequence, they should be of roughly equal mass. Given their short period, we find it plausible that the binary components may have evolved away from the main sequence, although without further data this cannot be constrained. In either case, the equal depth of the AB binary eclipses shows that the components A and B must have the same surface brightness.

Circumbinary Object

In this scenario the tertiary body b/C is of lower mass than the combined binary pair AB. A secondary eclipse of b/C by A or B that is not observable implies that the majority of the flux must be emitted from the components AB, or that the orbit of b/C be particularly eccentric. Here we use the KIC colour derived parameters to

explore the nature of this possible system.

The KIC temperature (4733K) and stellar radius imply a star of luminosity $0.26L_{\odot}$. If both central stars AB were of this type, the 0.4% limit on secondary eclipses limits the luminosity of b/C to be $\sim 0.002L_{\odot}$ in this scenario. This is consistent with an object of spectral type M5 or cooler (Reid & Hawley, 2005). Presenting a specific companion radius is beyond our capability at this stage, but below we present example companion radii leading this configuration to support a circumbinary dwarf star, brown dwarf or perhaps transiting planet around a close binary pair, in the configuration (AB)b/C. The extremely shallow AB eclipses evident in the binary lightcurve could possibly be explained as the result of a grazing geometry, although this is difficult give the width in phase of the close binary eclipses.

The deepest observed transit (6.32%), on a single star of KIC stellar radius $0.757R_{\odot}$ (Brown et al., 2011), implies a companion radius of $1.85R_{J}$, or $0.19R_{\odot}$. Assuming the AB components are of equal radii (and equal surface brightness as shown by the depth of their mutual eclipses), dilution would increase the companion radius to $0.27R_{\odot}$. Given the very short period of AB it is possible that the colour derived stellar radius will be an overestimate, allowing for the companion radius to decrease. There remains the possibility of the observed eclipses of A or B by component b/C being grazing, in which case the 'true' transit depth would be greater and the companion radius also.

Circumstellar Binary

This scenario assumes that the tertiary body is of greater mass than the binary pair. As such it would be the source of the majority of the flux, and can be taken to be the object described by the KIC (consistent with a K star). This would then be orbited by the lower mass stellar binary pair AB, in the architecture C(AB). Taking the KIC stellar radius of the primary K star component C, we again derive a radius for a single component of the binary system A or B of $0.19R_{\odot}$. The system would then consist of a central K star orbited by a close pair of remarkably small dwarf stars. This has the important advantage of explaining the shallow AB binary eclipse depths observed through dilution by the K star. The lack of secondary eclipses (eclipses of A or B by component C) is of interest. To produce secondary eclipses of order 0.4% or less, such that they would not be visible in the data as shown, requires that each component of the binary be contributing less than 0.4% of the system flux (for total eclipses). This is not possible given the ~1% binary eclipse depths observed (where the components would each need to be contributing at least 1% of the system flux). As such this configuration requires that the secondary eclipses be grazing or

missed, which requires either an eccentric orbit or similarly grazing primary eclipses. Without further modelling we cannot take either configuration further.

5.4.5 KIC2856960 Conclusion

At this stage we are unable to conclusively distinguish between the two system architectures (AB)b/C or C(AB) as outlined above. We also acknowledge that there may well be other stable system configurations unexplored here. Spectral follow up and the currently private *Kepler* data, in combination with the presented lightcurves, will allow for a much greater understanding of this intriguing system. When its nature is fully understood the absolute determination of many of the system parameters will be possible. This includes the companion mass in the (AB)b/C case, if its dynamical effect on the binary stars proves significant as in Doyle et al. (2011) and Welsh et al. (2012) prove significant. We hope that this will prompt follow up studies, and will continue to develop our understanding of the object as more data becomes available. If the (AB)b case proves true, this system represents the first discovered transiting planet around such a relatively short period eclipsing binary. If the C(AB) or (AB)C cases prove true this system will contain stars with some of the lowest masses accurately determined. These will be low enough to imply fully convective interiors, an area where directly measured physical properties are few (e.g. Carter et al., 2011; Morales et al., 2009; Irwin et al., 2011).

5.4.6 Eclipse Timing Variations

After the above work was published in Armstrong et al. (2012), it became clear that significant eclipse timing variations were present in the short, 6.2 hour binary signal. A procedure was developed (see Section 2.7) to extract these, through sequential fitting of the binary eclipses. The results for KIC2856960 can be seen in Figure 5.7, and show a clear periodic feature at the same period as the long period light curve signature. The points where the long period tertiary events occur can be seen in the Figure, at the regions of high error.

The signal seen is of relatively large amplitude, of ~250 s, and strongly favours the triple star interpretations of this system. An MCMC fit to these eclipse times (performed by T. Marsh) gave orbital parameters as shown in Table 5.4. These values use a mass for the close binary of $0.46 M_{\odot}$, as used by Lee et al. (2013, hereafter L13) to allow easier comparison. L13 published an analysis of the system using the same eclipse timing variations. Their derived values are shown in the same Table for comparison, and are in agreement with this work. Our values are

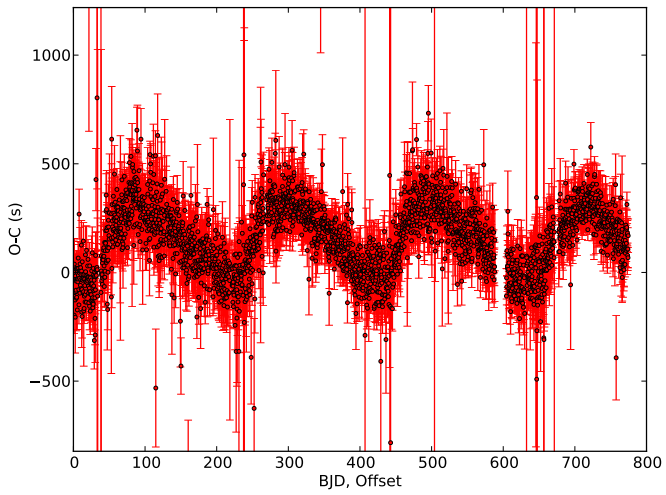

Figure 5.7: Observed - Calculated times for the 0.26 d binary eclipses. A strong trend can be seen at the ~200 d period of the tertiary light curve events. The 4 higher error regions correspond to where these tertiary events disturb the fit.

Table 5.4: Orbital parameters of the dwarf pair around the K star

Parameter	Units	O-C Fit	L13 value
Period	d	205 ± 4	205 ± 2
e		0.56 ± 0.03	0.61 ± 0.08
ω	rad	2.9 ± 0.1	3.02 ± 0.07
asini	AU	0.709 ± 0.006	0.73 ± 0.04
msini	$M_{\odot}$	0.67 ± 0.03	0.76 ± 0.05

more accurate in some cases due to the method used - I was able to measure eclipse times for individual close binary orbits, whereas L13 required binning of several eclipses to make each measurement. From these results it could be concluded that KIC002856960 is a hierarchical triply eclipsing triple star system, consisting of a K star orbited by a pair of eclipsing M dwarfs. However, as described in the next Section, this interpretation is incompatible with the details of the light curve.

5.4.7 Light Curve Fitting

Recently, an attempt to provide a fit to the full light curve has been attempted. Such a fit proved exceptionally hard to obtain. Furthermore, certain features of the light curve lead to constraints (dynamical and directly through Kepler's law) which are inconsistent with a triple star scenario - the previously proposed solution is

provably impossible (Marsh et al in prep). Confusingly, it is however possible under a particular combination of orbital parameters to obtain a fit which is reasonable, if far from precise. A four star attempt, while making an improvement (as would be expected given the additional free parameters), has yet to find a full fit. It seems that this system remains a mystery, which may be improved by the obtaining of long term radial velocities.

Chapter 6

Kepler Binary Star Temperatures

To allow testing of the search algorithm and to enable the population synthesis required to investigate circumbinary planet occurrence rates stellar radii and to a lesser degree masses were needed. Obtaining these individually (via, for example, spectra) for the over 2000 systems in the catalogue was not possible, leaving the option of using main sequence calibrations from known stellar temperatures. This chapter deals with how those temperatures were obtained. The calibrations themselves are explained in Chapter 7.

The aim of this chapter is to produce a catalogue similar to the Kepler Input Catalogue (KIC) for the Kepler eclipsing binary systems, taking into account the new knowledge of their binary nature and the information presented by the various photometric surveys of the Kepler field. In this way we can improve upon the KIC for these binary systems, through extended wavelength coverage (particularly inclusion of the U band), and consideration of both stars. Although the primary star often dominates the observed flux, not including the secondary (as in the KIC) can lead to biases. This work was published in Armstrong et al. (2013a), and the majority of this Chapter is drawn from this paper.

6.1 Data

6.1.1 KEBC

Data are taken from the Kepler Eclipsing Binary Catalogue (KEBC). The targeted objects' KIC Identification Numbers are listed in the Catalogue, along with the primary and secondary eclipse depth ratio, which were used to produce an estimate of the temperature ratio T_2/T_1 of each binary as described in Sect. 6.2.4. A version of the KEBC as presented online on 18-09-2013 was used to give KIC IDs and eclipse depth ratios, yielding 2610 systems. Thirteen of these systems contained multiple entries in the catalogue - in each of these cases only the first catalogue entry was used. Specific KIC IDs used are presented with our results in Section 6.4.

6.1.2 HES

Photometry from the Howell-Everett Survey (HES)(Everett et al., 2012) was used. The survey consists of the three optical filters Johnson U B and V, in the Vega system (Morgan et al., 1953), and contains data on 2424 objects of the 2610 from the KEBC. Errors were taken as presented in the HES catalogue.

6.1.3 KIS

In parallel to the Howell-Everett Survey, data from the Kepler INT Survey (KIS) (Greiss et al., 2012a,b), Data Release 2 was used. This allows models to be fit to two independent sets of photometry separately, increasing reliability and allowing bad data to be more easily flagged. The KIS provides data in the RGO U, Sloan g, r and i bands, in the Vega system. Errors provided in the catalogue are photometric only, systematic errors were added to the photometric errors in quadrature. Systematic errors for each band were taken from Table 3 of Greiss et al. (2012a), which gives the systematic offset used in calibrating each band to the KIC (a 0.05 mag systematic error was used for the U band).

Some objects in the KIS are observed more than once - these are available as separate sets of data, duplicates or triplicates (no object had more than 3 sets of data), for the same object. There are 2439 of the KEBC systems present in the KIS, of which 764 are duplicates and 111 triplicates. These multiple dataset systems were treated as independent objects during the subsequent analysis. Data for individual bands were filtered using the KIS class flag. Only bands of data with class -1 (stellar) or -2 (probably stellar) were used.

6.1.4 2MASS

To each of the HES and KIS surveys was added data from 2MASS (Skrutskie et al., 2006). This consisted of the Johnson J H and 2MASS specific Ks bands, in the Vega system. The combined total photometric uncertainties were used as presented in the 2MASS catalogue. Data were accepted if the SNR in that band was ≥ 5 and the object was not flagged as blended or contaminated. 2590 objects were found in the 2MASS catalogue.

6.1.5 Combination

For each object the above survey data was combined to produce two partially independent datasets, HES + 2MASS, hereafter UBVJHK, and KIS + 2MASS, hereafter UgriJHK. Each dataset is used separately in what follows, allowing comparison between results derived from the HES and KIS surveys and hence increasing reliability.

6.2 Model

6.2.1 Setup

A Markov Chain Monte Carlo code utilising the Metropolis-Hastings algorithm was used, enacted using the python module PyMC (Patil et al., 2010). For further details on the underlying method, see Chapter 2. We assumed each system to be composed of two stars, and fit the combined contribution from these two stars to the observed colour data. This intrinsically assumes that the data were taken when the stars were out of eclipse. This is reasonable for many eclipsing binaries, for those in overcontact systems (where the two stars are permanently in contact) it is less so but hard to avoid. In these cases it is only the apparent radii of the eclipsed star which will change; this is in general fit poorly anyway (see Section 6.3), and the fit temperatures should be unaffected.

6.2.2 Model Atmospheres

Model stellar atmospheres were taken from Castelli & Kurucz (2004) (CK). These cover a grid of $[3500K < \mathrm{T} < 50000K]$, $[-2.5 < [M/H] < 0.5]$, and $[0.0 < \mathrm{Log\ g} < 5.0]$, of which for computing efficiency purposes we used temperatures up to 13000K (some systems were run with higher temperature limits, see Sect 6.4.1). The grid spacing was 250K in Temperature (1000K for atmospheres above 13000K), 0.5 in $[M/H]$ (with an additional point at $[M/H] = 0.2$) and 0.5 in Log g. Model values were interpolated linearly between the two closest grid points of each parameter.

Although the CK atmospheres depend on surface gravity and metallicity as well as temperature, we found that these parameters were retrieved extremely poorly (See Sect 6.3). As such we did not include them in our model, using CK atmospheres with the KIC surface gravity and solar metallicity for each system.

To make use of the CK atmospheres they must be integrated over a response function for the relevant filter to produce band-integrated flux densities. Filter transmission curves were used as detailed in the respective papers of the HES and KIS. For the 2MASS data, relative spectral response functions from Cohen et al. (2003) were used. These provide an absolute flux calibration using the calibrated spectrum of Vega, matching with the Vega system magnitudes of the HES and KIS. All integrations were performed over F_λ, flux measured per unit wavelength, with an additional factor of λ included to account for photon counting.

6.2.3 Interstellar Extinction

The extinction relations of Cardelli et al. (1989) were used, with a constant R_V of 3.1, resulting in two analytical relations relevant for optical wavelengths (U to i) and IR wavelengths (JHK). This allowed extinction in each band to be calculated as a function of that of the V-band. The specific conversion factor for each photometric band depends on the spectrum of the star under question (due to the distribution of stellar flux within the band), but it was found that making the simplifying assumption of an extinction factor for each band calculated at the central wavelength of the relevant band had negligible effect for the stars considered here. V band extinctions were taken from the A_V values of the KIC, with the mean value of the KIC EBs of 0.4 mag taken for systems where no KIC values were available. This applied to 244 systems of the 2610; these systems are flagged in the presented catalogue. The KIC A_V values arise from their fit to the spectral energy distributions. While the KIC values for A_V are by no means perfect (see Brown et al. (2011) for a full discussion) it was found that fitting them ourselves did not constrain them, and resulted in an additional free parameter in what is already a large parameter space. As such the KIC values were used both to constrict the parameter space and to allow easier comparison between our results and the $T_{\rm eff}$ of the KIC.

6.2.4 Generation of T_2/T_1

As direct values for T_2/T_1 were not available at the time of submission, they were estimated from the ratio of secondary to primary eclipse depth (as $T_2/T_1 \simeq (\mathrm{depth_{sec}}/\mathrm{depth_{pri}})^{0.25}$). For circular binaries, this represents a good proxy for the

temperature ratio. For increasingly eccentric orbits, due to the possibility of different surface areas being occulted in primary and secondary eclipse, the eclipse depth ratio becomes an increasingly less accurate estimator of T_2/T_1. We formed a distribution for T_2/T_1 in these systems as described below. Using this, T_2/T_1 can be drawn from this distribution as part of the MCMC code.

Known Parameters

The known parameters used were binary period, eccentricity and argument of periapsis, sourced from the KEBC as described in Chapter 1. Initial values of T_2/T_1 were produced directly from the ratio of eclipse depths. In some cases no eclipse depth ratio or phase measurements were available, usually due to non-detection of the secondary eclipse. In these cases eccentricity was set to zero, and T_2/T_1 given an uninformative prior (see below).

Measurement Scatter

This was taken to be a gaussian error of 0.025 and 0.05 for over contact and non-overcontact binaries respectively, estimated from the test Figures 8 and 10 of Prsa et al. (2011). Overcontact systems were defined as having morphology parameter greater than 0.7 (see Matijevič et al., 2012). While these figures come from older versions of the Kepler EB catalogue, they represent a reasonable estimate of the uncertainty in recovering T_2/T_1 from eclipsing binary light curves.

Eccentricity Correction

This arises from the possibility of different surface areas being occulted in primary and secondary eclipse. The ratio of areas acts as a correction to the temperature ratio, in the form

$$\frac{T_2}{T_1} = \frac{A_1}{A_2}\left(\frac{D_2}{D_1}\right)^{\frac{1}{4}} \tag{6.1}$$

where D represents an eclipse depth and A the stellar surface area occulted in eclipse. The area eclipsed can be calculated from Equation 1 of Mandel & Agol (2002), and is a function of the projected orbital separation at eclipse, plus the stellar radii. The projected orbital separation at eclipse is then itself a function of period, eccentricity, argument of periapsis, system mass and inclination. This leaves the radii, mass and inclination unknown. We estimated the possible effect of this correction by drawing 10000 random sample binaries for each KEBC binary, with primary radii drawn

uniformly between 0.2 and 1.56 $R_{\odot}$. $1.56R_{\odot}$ is the Torres et al. (2010) radius of a T=12000K star with surface gravity 4.5 and 0.02 metallicity, chosen as an upper limit as the majority of our catalogue lie below this temperature. Secondary radii were drawn uniformly between 0.2 and the primary radius. Inclinations were constrained such that both stars eclipse. Given these radii, masses were found from the zero-age main sequence, solar metallicity models of Girardi et al. (2000). The area correction was calculated for each sample, and applied to Equation 6.1, giving a distribution of T_2/T_1. This was found to be of the form

$$\begin{aligned} P(\frac{T_2}{T_1}) \propto\ & \gamma\delta(\mu) \\ & + (1-\gamma)\left(\alpha_1 e^{\beta_1|\frac{T_2}{T_1}-\mu|} + \alpha_2 e^{\beta_2|\frac{T_2}{T_1}-\mu|}\right) \end{aligned} \tag{6.2}$$

where μ represents the initial value of T_2/T_1, produced from Equation 6.1 when A_1/A_2 is set to unity. This consists of two components, a delta function at μ (from the samples where both eclipses were total, leading to zero correction), and a continuous distribution which was found to be best fit by a sum of exponentials. The remaining parameters of Equation 6.2 were found for each KEBC binary via a fit to the observed distribution. This first order treatment of the eccentricity effect ignores limb darkening and starspots, which are both poorly constrained here.

Combined T_2/T_1 Distribution

The above inputs were combined into a final distribution of T_2/T_1, suitable for input as a prior into the MCMC. This distribution is found by convolving the measurement and eccentricity correction terms. For simplicity, the measurement gaussian was convolved with the delta function component of the eccentricity correction only. This is equivalent to approximating the exponential terms to dominate outside of approximately one gaussian standard deviation. This led to a final T_2/T_1 distribution of the form

$$\begin{aligned} P(\frac{T_2}{T_1}) = \frac{1}{N}\Bigg(& \gamma\mathcal{N}(\mu,\sigma) \\ & + (1-\gamma)\left(\alpha_1 e^{\beta_1|\frac{T_2}{T_1}-\mu|} + \alpha_2 e^{\beta_2|\frac{T_2}{T_1}-\mu|}\right)\Bigg) \end{aligned} \tag{6.3}$$

where N is a normalisation constant and σ is set as described in Sect 6.2.4. In cases

where no eclipse depth ratio or duration information was available, the prior on T_2/T_1 was set to be a gaussian with mean 0.7 and standard deviation 0.5, representing a very weak constraint.

6.2.5 Fit Parameters

Four fit parameters were used. These comprised the primary star temperature T_1, secondary star temperature T_2 (constrained through the temperature ratio as measured from the lightcurves), radius ratio of the stars R_2/R_1, and the primary radius to system distance ratio R_1/D. Note that stellar radii as used provide no allowance for non-sphericity of stars, and as such represent an 'effective radius', particularly in the case of overcontact eclipsing binary systems. No constraining relations were used, each parameter was allowed to vary according to its prior (see Section 6.2.6). Observables were treated as having normally distributed errors, and comprised each available colour band.

6.2.6 Input Data and Priors

The fits were performed to the combined UBVJHK dataset (6 colour bands), and separately to the UgriJHK dataset (7 colour bands). Each covered the wavelength range 0.36 to 2.16 Å. Missing (not available from the relevant photometric survey catalogue) or bad as defined in Sections 6.1.2, 6.1.3, and 6.1.4) data was given an error of 10^5 magnitudes to ensure it did not affect the fit. The KEBC temperature ratio was treated as an observable with distribution as described in Section 6.2.4, and used to constrain the temperature of the secondary star from that of the primary. We assumed priors on the model as detailed in Table 6.1. Where no KIC $T_{\rm eff}$ was available for an object, 5000K was used as the prior mean for T_1. We tested the effect of the prior on T_1 by running 1000 of the binary systems with firstly a prior of 5000K with standard deviation 2000K, and secondly a prior of the KIC $T_{\rm eff}$ with the same standard deviation, in each case with no extinction. The offset in the means of the T_1 distributions was 14K, so no significant systematic effect is caused by this prior. The standard deviation of the difference between the two cases, excluding non-converged systems, was ~200K, well within the errors quoted in Section 6.4.2. As such the choice of this prior has no significant effect on the retrieved values. The 'primary' star of each system was chosen using the KEBC temperature ratios - these values were taken for the purposes of fitting, even when they were greater than unity. In the final catalogue the primary star values have been set as the star with the dominant flux contribution, as calculated from the temperature and radius

ratios of the model output.

6.3 Testing

The model was tested on a simulated distribution of 1000 binary systems, for both the UBVJHK and UgriJHK datasets. These systems were generated with separate distributions for each physical stellar parameter, as no complete unbiased distribution could be found for these parameters in binary stars. The distributions used are laid out in Table 6.2 (with all values constrained to be above zero), which were designed to cover the expected parameter space for the Kepler mission EBs. The distribution of temperature ratio T_2/T_1 used to generate the sample was taken as a Gaussian approximation to the distribution of the KEBC. Note that this form of test involves generating fake colours using the very model atmospheres and filter transmissions used to fit them. Also while it involves realistic parameter values, these do not combine to represent 'real' stars. Hence this is purely a test of information content in the used colour bands. No significant difference was seen between each dataset, as expected from their similar colour bands and wavelength ranges.

Simulated colour bands were generated via integrating over the CK model atmospheres as detailed above. The MCMC was run for 50000 iterations with a burn in period of 20000 iterations. No significant extinction was included in this test (it was set to 0.02 magnitudes). The retrieved values of T_1, T_2, R_1/D and R_2/R_1 as compared to their input values for each dataset are shown in Figures 6.1, 6.2, 6.3 and 6.4. The agreement in all parameters except R_2/R_1 shows that 20000 is a sufficient number of iterations to allow convergence for the majority of systems. To remove as many as possible of those few remaining unconverged, a higher number of iterations is specified for the real data. The retrieval of surface gravity and metallicity was extremely poor. These parameters have very little impact on the observed colours within their error. As such these parameters were not included in the model.

For two stars with well separated temperatures it should be possible to fit each stellar atmosphere and hence obtain information on both stellar temperatures and also the ratio of the radii of the stars (to each other and to the system distance). For the systems in the KEBC however, the two stellar temperatures in general proved to be too close to allow this, as the peak emission of each star is often located too close in wavelength to that of the other star. This led to the well-retrieved information being in general the primary star temperature T_1 (and through the temperature ratio the secondary temperature T_2) along with a combination of parameters which can be termed the binary solid angle, equal to $(R_1^2 + R_2^2)/D^2$.

Table 6.1: Model Priors

Parameter	Distribution	Parameters			
		Mean	Standard Deviation	Lower Limit	Upper Limit
T_1	Normal	KIC T_{eff}	2000K	3500K	13000K (see Sect 6.4.1)
T_2/T_1	See Sect 6.2.4				
R_1/D	Normal	$0.003 R_\odot \text{pc}^{-1}$	$0.05 R_\odot \text{pc}^{-1}$	$10^{-5} R_\odot \text{pc}^{-1}$	$0.2 R_\odot \text{pc}^{-1}$
R_2/R_1	Normal	0.8	0.3	0.01	3.0

Table 6.2: Distribution of Test Parameters

Parameter	Distribution	Parameters Mean, Lower Limit,	 Standard Deviation Upper Limit
T_1	Uniform	3500K	10000K
T_2/T_1	Normal	0.9123	0.1668
R_1	Normal	$0.8R_\odot$	$0.2R_\odot$
D	Uniform	50 pc	1500 pc
R_2/R_1	Normal	1.0	0.4
Log g_1	Normal	4.5 cgs	0.2 cgs
Log g_2	Normal	4.5 cgs	0.2 cgs
$[M/H]$	Uniform	-2.5	0.5
A_V	Normal	0.05 mag	0.02 mag

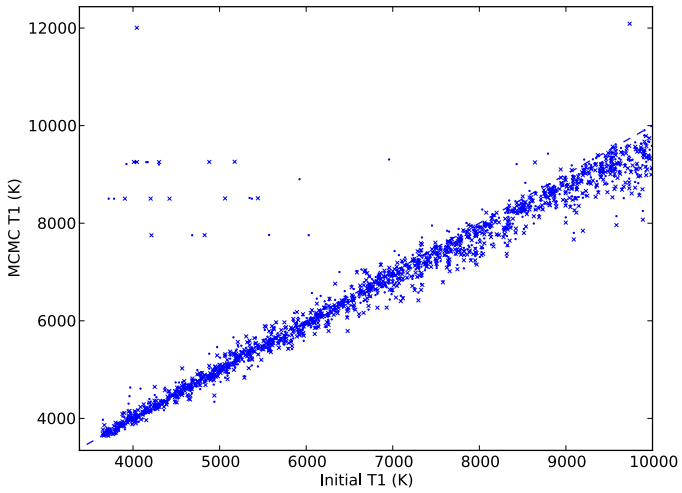

Figure 6.1: Input and MCMC fit primary star temperatures T1 for 1000 simulated sets of stelar parameters. The dashed line shows a perfect match. Dots represent the UBVJHK dataset and crosses the UgriJHK. The small number of systems highly deviant from the dashed line have not converged.

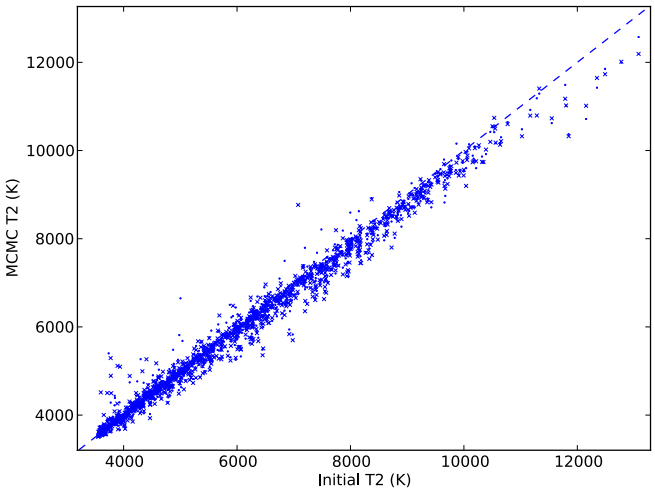

Figure 6.2: As Figure 6.1 for T2

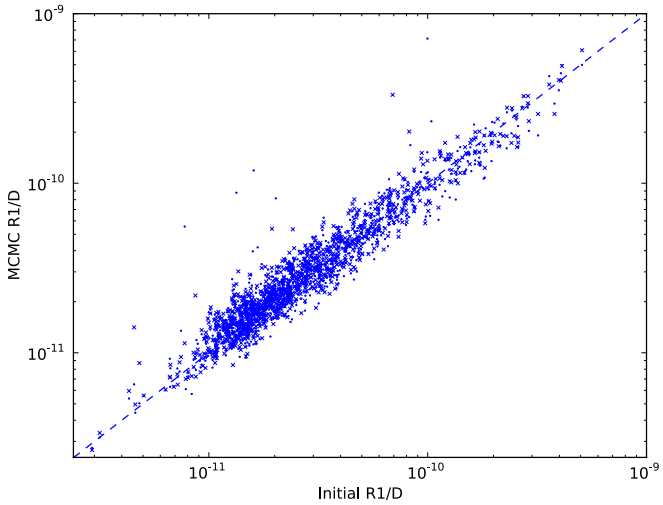

Figure 6.3: As Figure 6.1 for R_1/D

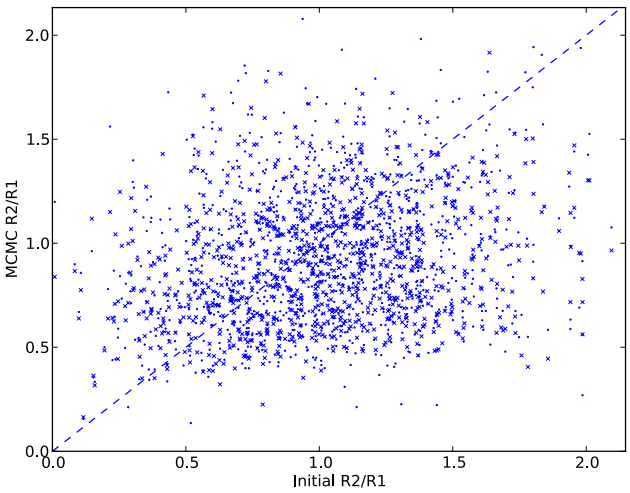

Figure 6.4: As Figure 6.1 for R_2/R_1

The relevance of the binary solid angle as opposed to R_1/D and R_2/R_1 depends on the temperature ratio. For ratios significantly different from unity the individual components R_1/D and R_2/R_1 become well retrieved. The difference between input and MCMC fit values for R_2/R_1 using the UBVJHK colours is shown as a function of T_2/T_1 in Figure 6.5 (with an additional 1000 systems with lower temperature ratios added to illustrate the correlation). Note the systematic offset of about -0.2 even at lower values of T_2/T_1. In what follows both R_1/D and R_2/R_1 are given, as each are in some cases accurate, but users should note the above in choosing whether to use these values individually or combined into the binary solid angle mentioned. The relevance of R_2/R_1 should be determined from Figure 6.5 in line with the needs of the user. It should be noted that when T_2/T_1 approaches unity for main sequence stars R_2/R_1 should also be close to unity.

6.4 Results

6.4.1 Fitting Parameters

Each object was run through the MCMC for 100000 iterations, including a burn in period of 30000 iterations (significantly more than was used in the test of Section 6.3). At 250 iteration intervals through the burn in phase, the model was tuned. Objects with KIC $T_{\rm eff}$ <= 9000K were run using a high temperature limit of 13000K for efficiency. Other objects (including those with no KIC value for $T_{\rm eff}$) were run

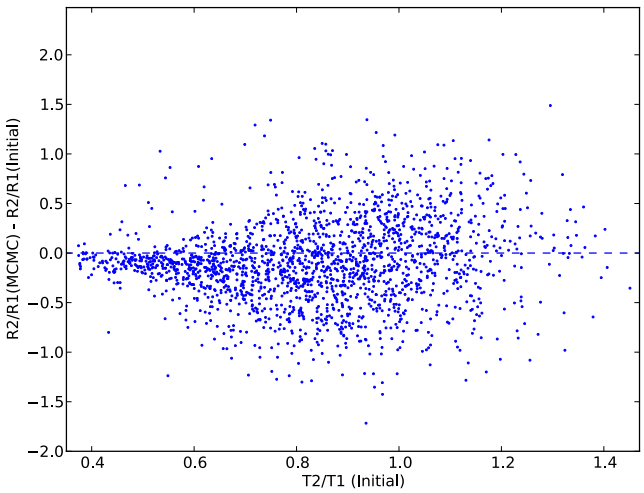

Figure 6.5: The dependence of R_2/R_1 fit quality on T_2/T_1. The 1000 simulated sets of parameters used in Figure 6.1 plus 1000 additional sets with lower temperature ratios are shown.

with a limit of 50000K, and are flagged in the catalogue.

In forming final values for each object, output parameters were checked for consistency between both datasets, and also between duplicate or triplicate UgriJHK datasets where available. T_1 was used as the parameter for checks (with T_1 defined as the temperature of the star dominating the system flux) , as this was the strongest recovered parameter while testing. First, all fits with 3 or less colour bands were excluded. Then, in the cases where more than one fit was available, each fit was checked against each of the others to see if it was within 3σ (using the maximum σ of the two fits being checked). If the fits were thus consistent, they were both included in the weighted average (using the inverse of their MCMC derived errors as weights) to calculate the final set of values. This process was repeated for all possible fit combinations (the maximum number of fits for a system is 4, one UBVJHK and 3 UgriJHK). In this way highly deviant fits (for example due to bad photometric data) can be excluded from the final values where possible. For systems where no good fits were present, final values were formed using the weighted average of all available fits, and errors by the standard deviation of those fit values. These systems are flagged in the catalogue to highlight the systematic difference between their fits and/or the lack of photometry available.

6.4.2 Errors

The MCMC derived errors were used to represent the gaussian noise associated with our model fitting. These have median values of 120K for T_1, 310K for T_2, $8.2 \times 10^{-5} R_{\odot}/pc$ for R_1/D and 0.17 for R_2/R_1. The effect of extinction on the presented values was also estimated, as the KIC extinctions are known to have particularly high error. Runs of the model on the UBVJHK dataset with the KIC extinction values (as in the final run) and with no extinction were compared. The error on the KIC extinction values (~0.3) is of the order of the values themselves (mean ~0.4), so the difference in model fits generated by removing extinction represents a reasonable estimate of the error they could cause. There was a median difference between fits with and without extinction of 350K and 540K for T_1 and T_2, $7.3 \times 10^{-5} R_{\odot}/pc$ for R_1/D, and for R_2/R_1 0.15. As such, extinction has a significant effect. These median extinction effects are treated as a 1σ additional Gaussian error on the presented values, and give the combined error in the catalogue. While this is an estimate, neither the KIC extinction values nor their errors are characterised enough for a more detailed approach to be meaningful.

It should be noted that this error estimation does not take account of systematic offsets caused by contamination ('third light'), bad data, or bad extinction values, and objects affected will have larger errors. While bad or contaminated photometric data was removed where it was labelled as such, and errors were given to the input temperature ratios to reduce the impact of bad values, these issues will remain for some objects.

The adopted 1σ errors are then formed from a combination of the effect of the extinction systematic and fitting noise values. Errors are presented individually for each system, but to give a guideline the median catalogue error is 370K in T_1, 620K in T_2, 2.5×10^{-12} in R_1/D, and 0.23 in R_2/R_1. For high temperatures ($> 9000K$) the errors on the temperature are larger, as seen in Figure 6.1) and discussed in Section 6.5.1. The errors on the radius parameters will vary with T_2/T_1, as described in Sections 6.3 and 6.5.1. The derived temperature errors are consistent with the temperatures of various systems in the KEBC which have been studied in detail, see Section 6.5 for detail.

6.4.3 Catalogue

The format of our results is presented in Table 6.4 (Armstrong et al., 2013a, the full catalogue available online). For each object the results are given for each fitted dataset, along with 'final' values. These are formed from the average of the available

Table 6.3: Catalogue Flags

Column	Flag Description
1	1 if object run with 50000K temperature limit
2	1 if no KIC T_{eff}, A_V or Log g values available
3	1 if candidate or proven three-body system
4	1 if no KEBC eclipse information available
5	1 if no matching fits available with 4+ colour bands

'good' fits - fits can be excluded as explained in Section 6.4.1. Entries are flagged for various reasons; a summary of the used flags, in order in which they appear in the catalogue, is given in Table 6.3.

6.4.4 Distributions

Distributions of our output parameters were formed using the 'final' values as detailed in the results catalogue. Systems for which no good final value could be formed were discarded (this left 2457 of the original 2610 systems, in at least one dataset), While the R_1/D distribution is generally uninformative due to the unknown distance, the T_1 and T_2 distributions, and in combination the T_2/T_1 distribution, is worth noting. The R_2/R_1 distribution is in general poorly fit so is again uninformative. The results for T_1 and T_2 are shown in Figures 6.6 and 6.7. In presenting the temperature ratio, the total distribution (Figure 6.8) is shown, and also the distribution split by stellar spectral type. The results for 'cool' ($T_1 < 5200K$, roughly K and M stars), solar type ($5200 \leq T_1 < 7500K$, roughly F and G stars) and 'hot' ($T_1 >= 7500K$, roughly A stars) are given. These are presented normalised to their sample sizes in Figures 6.9, 6.10, and 6.11, and are discussed in Section 6.5.

It is also worth comparing these results to the KIC itself. Assuming a single star as the KIC does will tend to focus on the primary star as the dominant source of flux. As such the KIC T_{eff} is compared to our T_1 in Figure 6.12. A general trend of an increase in our temperatures over the KIC's can be seen - this is expected, as in each of these systems an extra contribution from a usually cooler star has been included. The effect of our increased temperature limit over the KIC is also notable.

Table 6.4: Catalogue Format. UBV = UBVJHK, KIS=UgriJHK

Column Header	Description	Column Header	Description
KIC ID	Kepler Input Catalogue Identifier	T_2 KIS_1, Error	-1 if None
T_1 Final, Error	Final Primary Temperature and error (K)	R_1/D KIS_1, Error	-1 if None
T_2 Final, Error	Final Secondary Temperature and error(K)	R_2/R_1 KIS_1, Error	-1 if None
R_1/D Final, Error	Primary Radius / System Distance and error	KIS_1 Bands Used	1 if used, order – – – UgriJHK
R_2/R_1 Final, Error	Component Radii Ratio and error	T_1 KIS_2, Error	-1 if None
T_2/T_1 Input	Input prior temperature ratio	T_2 KIS_2, Error	-1 if None
Flags	4 digit flag of fit quality, see Table 6.3	R_1/D KIS_2, Error	-1 if None
Output Source	Final values source, UBV, KIS_1,KIS_2,KIS_3, 1 if used	R_2/R_1 KIS_2, Error	-1 if None
T_1 UBV, Error	-1 if None	KIS_2 Bands Used	1 if used, order – – – UgriJHK
T_2 UBV, Error	-1 if None	T_1 KIS_3, Error	-1 if None
R_1/D UBV, Error	-1 if None	T_2 KIS_3, Error	-1 if None
R_2/R_1 UBV, Error	-1 if None	R_1/D KIS_3, Error	-1 if None
UBV Bands Used	1 if used, order UBV– – – – JHK	R_2/R_1 KIS_3, Error	-1 if None
T_1 KIS_1, Error	-1 if None	KIS_3 Bands Used	1 if used, order – – – UgriJHK

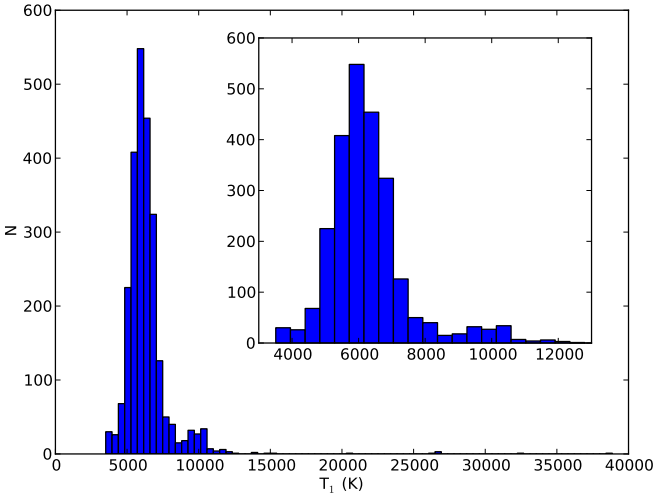

Figure 6.6: The distribution of primary star temperature T_1, drawn from the 'final' catalogue values. Systems with no consistent fits or lacking 4+ photometric bands are excluded. The cooler temperatures where the majority of our sample lies are shown in the inset.

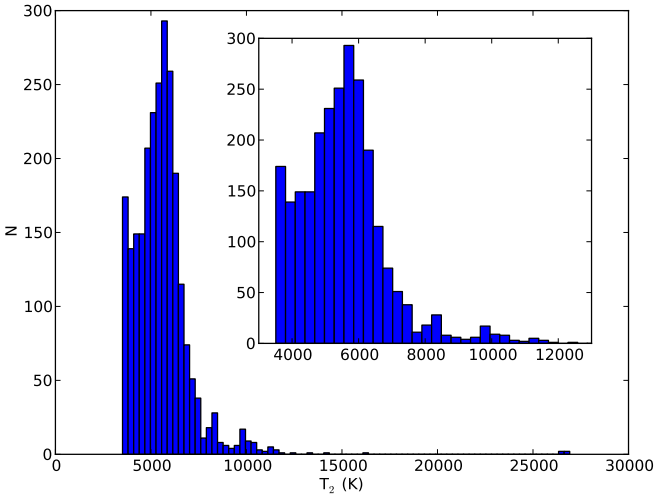

Figure 6.7: As Figure 6.6 for secondary star temperature T_2.

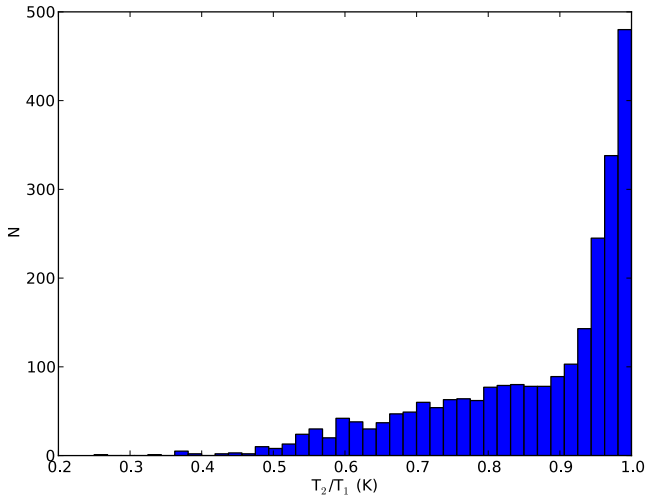

Figure 6.8: The total distribution of temperature ratio T_2/T_1, drawn from the 'final' catalogue values. Systems with T_2/T_1 are included as their inverse.

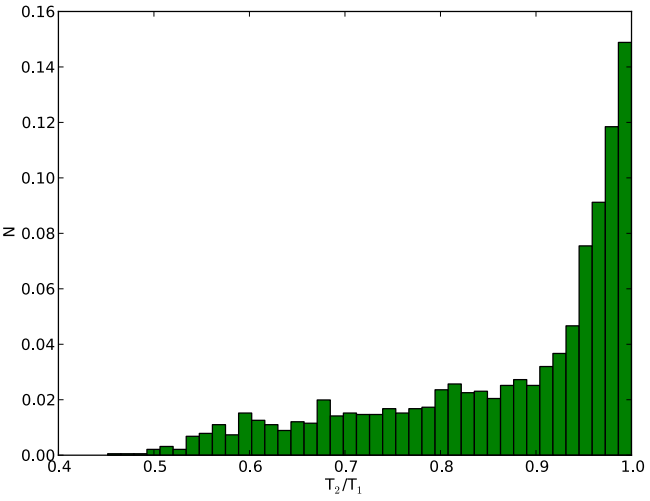

Figure 6.9: The normalised distribution of T_2/T_1 for solar type stars, total number 1908.

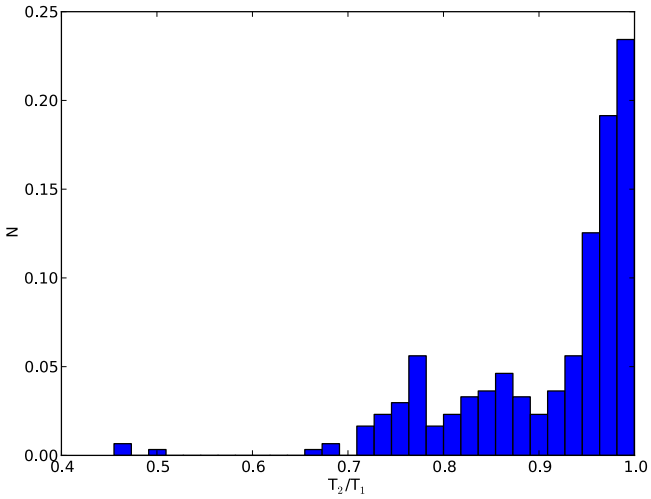

Figure 6.10: The normalised distribution of T_2/T_1 for stars cooler than 5200K, total number 303.

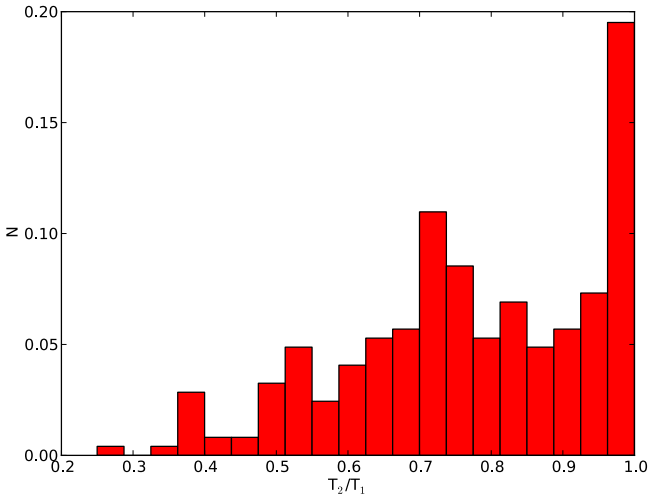

Figure 6.11: The normalised distribution of T_2/T_1 for stars hotter than 7500K, total number 246.

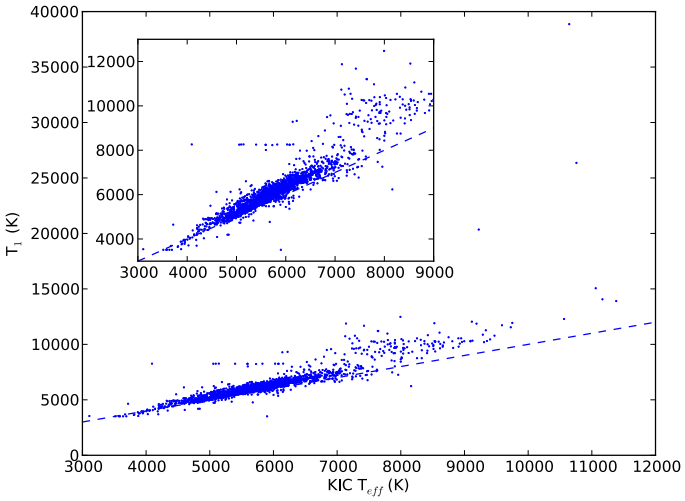

Figure 6.12: Comparison of KIC $T_{\rm eff}$ and final catalogue T_1 values. The dashed line represents a 1:1 match. The difference arises from the inclusion of two stars in our model as compared to the single star of the KIC, and also our higher temperature limit.

6.5 Discussion

6.5.1 Shortcomings

There are various issues involved in the production of this catalogue, the worst of which we summarise here. Some of these issues are in common with those of the KIC (Brown et al., 2011).

High/Low T

The CK stellar atmospheres used have a lower limit of 3500K. As such, temperatures which lie close to this value ($T < 3750K$) should not be taken as accurate. In terms of selecting cool stars however they can be used - while the temperature is not accurate, any object with a catalogue temperature around this level is unambiguously cool. At high temperatures ($T > 9000K$) a large systematic effect, visible in Fig. 6.1, means that our temperatures have much higher errors, and are likely underestimated. Again, these are still hot stars, but the exact temperature values should not be trusted.

Binary Solid Angle vs R_1/D and R_2/R_1

As described in Section 6.3, the R_1/D and R_2/R_1 values are generally poorly constrained. This is because for most of the systems in the KEBC, the temperatures of the two stars are close enough that little information is contained on the secondary star's relative contribution to the flux. Fig. 6.5 shows that only systems with $T_2/T_1 \leq 0.6$ constrain R_1/D and R_2/R_1, and in these cases there is still a systematic underestimate of ~0.2 in R_2/R_1 which has not been adjusted for. In the other systems, which represent the majority of those here, the binary solid angle, equal to $(R_1^2 + R_2^2)/D^2$ and representing a combined measure of the level of flux incoming at the Earth is constrained instead, and can be recovered from the catalogue values as $(R_1/D)^2(1 + (R_2/R_1)^2)$.

Contamination

As has been mentioned previously, our model assumes only two stellar components, and does not take account of additional sources of flux. The proportion of KEBC objects with additional companions is ~20% (Rappaport et al., 2013), meaning they represent a significant part of our catalogue. Data flagged as contaminated in the various photometric surveys has been removed, and objects which are confirmed as or possible 3+ body systems marked (Rappaport et al., 2013; Carter et al., 2011; Derekas et al., 2011; Conroy et al., 2013; Gies et al., 2012). These systems can still be of use - for example, KIC5897826 is a triple star and is tested successfully against our two star model in Table 6.5. In cases where more than two stellar components are present in a system, the derived physical properties will be affected to the extent that the extra companions contribute to the flux.

Extinction and Reddening

The use of the KIC A_V values allows us to reduce our already large parameter space while still using reasonable extinction values. Attempts to fit these values ourselves were very poorly constrained. Extinction is not generally well constrained - typical errors on the KIC E(B-V) values are of 0.1 mag, implying a 0.31 mag typical error on A_V (Brown et al., 2011), which is of the order of the A_V values. The KIC paper itself notes the problems involved in its extinction parameters, including no account of small scale structure in the interstellar extinction. These errors are incorporated here, and will affect the systems in our catalogue. We have attempted to include the systematic effects of bad extinction in our errors (see Section 6.4.2), but anomalously high extinctions will produce too high temperatures, and the reverse for low

extinctions.

6.5.2 Distributions

While our individual stellar temperatures are not constrained by spectroscopy, and so have a larger potential for biases and non-physical effects to show themselves, they offer a particularly large sample size - 303 cool stars, 1908 solar type, and 246 hot stars. However, as the stars under consideration are drawn from an unconstrained sample of ages (and the expected stellar temperature varies with evolution), the temperature distributions are generally uninformative by themselves. The temperature ratio distributions split by spectral type show slight differences, however due to the lower number of samples in each bin of the cool and hot distributions there is not a strong case for any statistically significant difference. This implies that systematic effects in the fitting of different temperature stars were not strong, supporting the robustness of our results across the different temperature regimes. All of the presented distributions have potential biases present, in the effect of the particular combination of colour bands which were utilised, as a residual effect of the KEBC recovery of eclipse depths, or as a sampling effect in the eclipsing binaries which Kepler selects.

6.5.3 Performance on known objects

A limited number of these eclipsing binary systems, analysed in depth for other reasons, are available to compare with our results. While this sample is far from ideal (by definition these are systems which were selected as 'interesting' for a variety of reasons, and hence unlikely to be typical) we would still hope to predict their temperatures reasonably well. Eleven systems were found where detailed analysis had been done. Five of these represent interesting stellar objects (e.g. sdB+white dwarves, triple systems) whereas the other 6 were circumbinary planet hosts. In the stellar cases the primary star often dominates the flux, and in these circumstances this star should fit reasonably well. The spectroscopically derived individual temperatures are compared to ours in Table 6.5. The comparison was contained to temperatures, as in the majority of cases no distance information was available. Encouragingly both the primary and secondary temperatures fit well, within the errors derived. For the three hot star entries in the table, the fits are generally worse, in all cases lying outside 3σ for at least one star. They are however, unambiguously high, representing some of the highest temperatures in the catalogue. This highlights that while 'hot' stars are selected reasonably well, precise temperatures at values greater

Table 6.5: Performance on known systems

KIC ID	Description	T1 (Actual, K)	T1 (This Work, K)	T2 (Actual, K)	T2 (This Work, K)	Reference
12644769	Kepler-16	4450 ± 150	4013 ± 350	N/A	3984 ± 550	Doyle et al. (2011)
8572936	Kepler-34	5913 ± 130	5937 ± 380	5867 ± 130	5943 ± 610	Welsh et al. (2012)
9837578	Kepler-35	5606 ± 150	5902 ± 370	5202 ± 100	4913 ± 620	Welsh et al. (2012)
6762829	Kepler-38	5623 ± 50	5834 ± 350	N/A	3539 ± 540	Orosz et al. (2012a)
10020423	Kepler-47	5636 ± 100	5881 ± 350	3357 ± 100	3985 ± 680	Orosz et al. (2012b)
4862625	PH1	6407 ± 150	6705 ± 390	3561 ± 150	4626 ± 860	Schwamb et al. (2013)
10661783	δ-Scuti	8000 ± 160	9197 ± 480	N/A	7714 ± 810	Southworth et al. (2011)
5897826	Triple System	5875 ± 100	5798 ± 380	N/A	5951 ± 750	Carter et al. (2011)
6889235	A+WD binary	14500 ± 500	10874 ± 470	9500 ± 250	9519 ± 1000	Bloemen et al. (2012)
7975824	sdB+WD binary	34730 ± 250	32604 ± 2190	15900 ± 300	10753 ± 1460	Bloemen et al. (2010)
9472174	sdB+dm binary	29564 ± 106	38868 ± 387	N/A	16350 ± 2090	Østensen et al. (2010)

than ~9000K can have larger errors. This does not affect the selection of 'hot' stars used in the above distributions.

6.6 Summary

The Kepler Eclipsing Binary Catalogue has been combined with information from the HES, KIS and 2MASS photometric surveys to produce spectral energy distribution fits to over 2600 eclipsing binaries in the catalogue over a wavelength range of 0.36 to 2.16Å. Primary (T_1) and secondary (T_2) stellar temperatures are presented, plus information on the stellar radii and system distance ratios. The derived temperatures are on average accurate to 370K in T_1 and 620K in T_2. These results improve on the similarly derived physical parameters of the Kepler Input Catalogue through consideration of both stars of the binary system rather than a single star model, and inclusion of additional U band photometry. The results are expected to aid future uses of the Kepler Eclipsing Binary data, both in target selection and to inform users of the extremely high precision light curves available. Surface gravities and system metallicities were not included, as these were found to have an insignificant effect on the observed photometric bands.

Chapter 7

Circumbinary Rates of Occurrence

Previous chapters have laid out a number of developments which make possible the detailed study of the occurrence rates of circumbinary planets. The first observationally based determination of this rate is presented here. This is derived from the publicly available Kepler data, using the search algorithm described in Chapter 4 and a debiasing process to produce occurrence rates implied by the known transiting systems. These rates depend critically on the planetary inclination distribution: if circumbinary planets are preferentially coplanar with their host binaries, as has been suggested, then the rate of occurrence of planets with $R_p > 6R_\oplus$ orbiting with $P_p < 300$ d is $10.0^{+18}_{-6.5}$% (95% confidence limits), higher than but consistent with single star rates. If on the other hand the underlying planetary inclination distribution is isotropic, then this occurrence rate rises dramatically, to give a lower limit of 47%. This implies that formation and subsequent dynamical evolution in circumbinary disks must either lead to largely coplanar planets, or proceed with significantly greater ease than in circumstellar disks. The investigation also demonstrates that giant planets ($>10R_\oplus$) are significantly less common in circumbinary orbits than their smaller siblings, and confirms that a proposed shortfall of circumbinary planets orbiting the shorter period binaries in the Kepler sample is a real effect. This work has been published in Armstrong et al. (2014), which forms the majority of the Chapter.

7.1 Overview

The process for obtaining the desired occurrence rates proceeds as laid out in Chapter 2. The key addition to this here is the debiasing process. This is where the complete sample of Kepler EBs is tested, in order to find which of these binaries we could observe a planet in if such a planet existed. This is described below, and results in defined samples of binaries which form the basis of the population synthesis subsequently performed. In this way, all binaries which contribute no information to the occurrence rate (through yielding no planet to the search algorithm regardless of a planet's presence or absence) are removed from the process, allowing greater efficiency. The larger the sample of useable binaries is, the better the statistical noise on the resulting rates becomes, and so it is desirable to have the most efficient search algorithm possible. This must be weighed against becoming too targeted in searching - while it may be possible to obtain a highly effective algorithm by targeting specific forms of transit for example, there is a danger of missing more unusual but still planetary signals, which would cast doubt on the derived occurrence rates.

The search algorithm used has been described in Chapter 4. Here its application to this specific purpose is given. Occurrence rates are defined here to mean the number of binaries with one (or more) planets as a fraction of the total binary number, leaving the question of multiple planets per binary to future work.

7.2 Data Processing

7.2.1 Data Source

Targets were selected from the Kepler Eclipsing Binary Catalogue (KEBC) (Slawson et al., 2011). The version of the catalogue as found online[1] on 18th September 2013 was used, yielding 2610 objects. Of these only non-overcontact binaries with morphology parameter <0.7 (see Matijevič et al. (2012) for detail) were used, as the history of planets in highly evolved over-contact binaries is likely to be significantly different from those in other systems, and these binaries present different challenges to systematic planetary detection. The initial sample then comprised 1735 binaries. Light curves from Quarters 1 through 16 were downloaded, comprising a baseline of approximately 4 years for most objects. Where some Quarters were unavailable for some objects, partial light curves were used with as much data as possible.

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

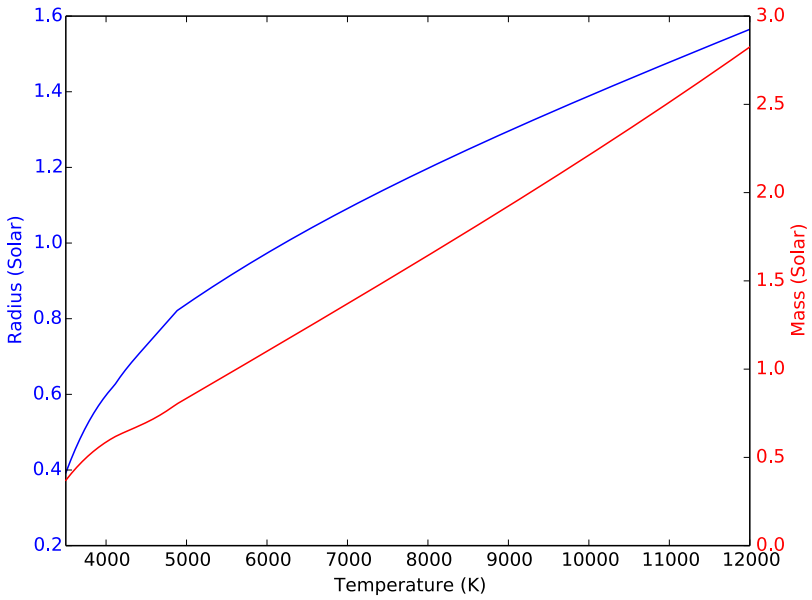

Figure 7.1: The calibrations used to obtain stellar radii and masses from the known temperatures, which assume main sequence stars. As larger evolved stars are more likely to eclipse, this represents a potential source of bias.

7.2.2 Parameter Source

Period and eccentricity parameters were taken from the KEBC as described previously in Chapter 1. Where insufficient information was available on the binary eclipses (generally due to non-detection of the secondary eclipse) the binary eccentricity was set to zero. Temperature information on the sample was obtained as described in Chapter 6. This allowed main-sequence calibrated stellar radii and masses to be generated from the stellar temperatures, using the calibration of Torres et al. (2010) with surface gravity 4.5 and solar metallicity. The lower mass limit for this process was $0.6M_{\odot}$ - below this the calibrations of Boyajian et al. (2012) were used. Between 0.6 and $0.8M_{\odot}$, in the valid range of both calibrations, we interpolated between them to ensure no discontinuity. The mapping of temperature to radius and mass under this regime are shown in Figure 7.1. Using these calibrations implicitly assumes main sequence stars. For the detached stars studied here, this is a reasonable approximation for the majority of the sample (for the contact stars of the KEBC, where significant post-main-sequence evolution is liable to have occurred, it would not be).

7.2.3 Data Preparation

This was carried out as described in Chapter 4. In summary, covariance basis vectors were used to mitigate instrumental systematics. After this, quarters of data were joined to create single light curves for each object, and the eclipsing binary signal then removed.

7.3 Search Algorithm

The search algorithm as described in Section 4.4 was used. This consisted of a search for individual transit-like events, followed by a periodicity test allowing for the quasi-periodic nature of circumbinary transits.

The candidate system KIC6506534 was included in calculating occurrence rates. This was due to the strong transit timing variations seen. Also supporting the hypothesis that this is a real planet is that the candidate period shown by its transit signals corresponds to $\sim 6P_{\rm bin}$, similar to the currently known planets and outside the inner stability limit for this system. The other two strong planetary candidates had periods outside the 300d threshold (KIC5473556) or did not transit consecutively (KIC9632895). Using our calibrated stellar radii (see Section 7.2.2), for KIC6506534 the transit depth of ~0.2% would represent a planet radius of 4.3 $R_\oplus$. This defines the radius bin used for this object.

7.3.1 Detection Threshold

To test the efficacy of the search algorithm a minimum significance threshold had to be set (the details of forming the output statistic itself are given in Section 4.5). This was done using the significances of recovered test transit injections (see Section 7.4.1) and is shown in Figure 7.2. Note the large number of unrecovered injections at periods below 60 d – due to the increased box size at short periods (caused by increasing maximum TTV), both the number and significance of false detections is increased in this region. As shorter period real signals have their significance increased for the same reason, as well as having more transit events, the threshold is allowed to rise at low periods to exclude this additional noise. This led to 308 out of the 1735 scanned systems showing signals with significance over the threshold, which were in each case examined transit by transit by eye. Further to this, every periodogram and light curve was checked by the author and another independent researcher. Both flagged all of the currently known planets clearly, with the exception of the Kepler-47 system, which was only weakly detected due to

both small transits and stellar noise. Two strong candidate planetary systems were also detected within our period threshold of 300 d as well as various other signals. These were described in Section 4.7.2.

7.4 Debiasing

7.4.1 Transit Injections

To determine what the true sample of searched light curves was through testing the search algorithm, simulated planetary signals were injected into the light curves, which were then rescanned. Simulated transit times were found using an N-body integration code, as described in Section 2.4. The planets were put in circular orbits coplanar with the binaries on periods of 10.2 $P_{\rm bin}$ and 300 d. These were made slightly eccentric (e=0.05, with a uniform distribution of argument of periapsis between 0 and 2π) so that the additional TTVs which eccentricity may bring, via for example precession, were not excluded. Systems which were unstable (typically very long period binaries where 300d proved to be within the inner stability limit) were dropped, providing an implicit stability check on our sample. Exact resonances with the binary were avoided, due to the possibility of localised stability effects (Doolin & Blundell, 2011). Times and durations of transits were extracted.

These transits were injected into the light curves of each binary using U-shaped transits of the recorded duration, centred on the transit times. Only transits of the primary star were used, as these dominate the detectability of a planet. Transit depths were set using stellar radii derived as described in Section 7.2.1, for planets of radius $4R_\oplus$, $6R_\oplus$, $8R_\oplus$ and $10R_\oplus$. Dilution from the secondary star was included, along with quarter-by-quarter contamination ascertained from the Kepler data archive (Fraquelli & Thompson, 2012, typically a few percent). Note that contamination by unknown tertiary stars in the system is not included, as no information is available as to the extent or magnitude of this. Although approx. 20% of the KEBC binaries are thought to have stellar tertiary companions (Rappaport et al., 2013) the amount that these will dilute transits of the primary binary star is unknown. Each planetary radius and period combination was injected and searched separately. The detected output statistics for a typical injection group are shown in Figure 7.2, and led to the threshold shown. Injections where the maximum periodogram peak was not at a harmonic of the injected signal, or where the detected significance was below the threshold, are shown as blue dots and represent the background noise distribution. Detections were allowed for any harmonic down to $P_p/10$. Signals on shorter periods show more events and as such generally have

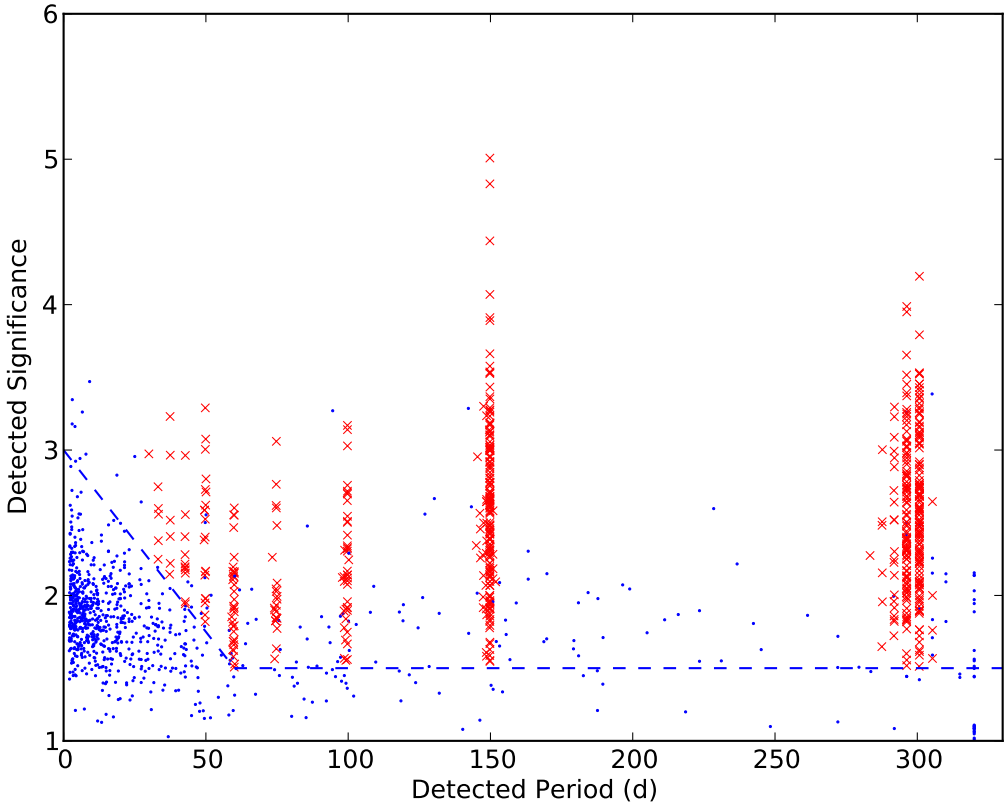

Figure 7.2: Detected significance of the whole binary sample with injected planetary signals, for $P_p = 300$d, $R_p = 10R_\oplus$. Red crosses represent successful detections at period or harmonic, while blue dots represent the detected significance of the highest peak in the periodogram in systems where the injection was not recovered. Dashed line represents the significance threshold used. Some blue dots fall above this line due to chance periodicity in the noise (be it astrophysical or instrumental) of those system's lightcurves.

Table 7.1: Planets and candidates within bins for which test transit injections were successful. Note that the 8-10 bin is equivalent to > 8 here.

Period	Radius ($R_\oplus$)	Number	Included Planets
10.2 $P_{\rm bin}$	> 10	0	
	$8 - 10$	2	K-16b, K-35b
	$6 - 8$	1	PH1
	$4 - 6$	2	K-38b, KIC6504534
	$6 - 10$	3	K-16b, K-35b, PH1
	$4 - 10$	3	K-16b, K-38b, KIC6504534
300 d	> 10	0	
	$8 - 10$	3	K-16b, K-34b, K-35b
	$6 - 8$	1	PH1
	$4 - 6$	2	K-38b, KIC6504534
	$6 - 10$	3	K-16b, K-34b, PH1
	$4 - 10$	3	K-16b, K-38b, KIC6504534

higher detection significances. The effect was particularly strong for signals under 60 d, which is reflected in the threshold.

Similarly, light curves containing known planets or candidates were subjected to testing, after removal of the known transits. This allowed us to probe the sensitivity of the search algorithm in these systems. Table 7.1 shows the number of planets or candidates contained within each period or radius bin for which these test injections were recovered successfully. Note that some planets which would be expected to appear in bins do not because we did not successfully recover transit injections at those minimum radii in these systems. An example is Kepler-34b in the 4–10$R_\oplus$ group, where a 4$R_\oplus$ transit injection was not recovered and so the system is not included, as in the Kepler-34 system a planet could not be detected over the whole bin range. On the other hand Kepler-16b is included in the $4 - 10$ bin, as a test 4$R_\oplus$ planet was successfully recovered for this system and the real planet radius lies within the bin.

7.4.2 Test Results

The number of systems where recovery was successful according to the stated threshold is shown in Table 7.2, split into each radius and period group. This includes systems where the detected period was a harmonic of the injected period. The recovery rate varied between ~10% of the total stable sample for the most difficult 300d, 4$R_\oplus$ case, and ~55% for the 10.2$P_{\rm bin}$, 10$R_\oplus$ case. To check these surprisingly

Table 7.2: N systems with successful transit injection recovery, out of 1735 total.

Period	Radius ($R_\oplus$)	$N_{recovered}$
10.2 P_{bin}	10	857
	8	757
	6	597
	4	322
300 d	10	581
	8	490
	6	328
	4	143

low sample sizes, a subsample of the failed systems were examined to determine the cause of the recovery failure. In ~50% of $R_p = 10R_\oplus$ cases this was light curve noise or stellar activity dominating the transit signal depth. A further sixth of the failed cases were due to remnants of binary eclipses, with another sixth due to short light curves (generally under 1 year) which were not long enough to show multiple transits. The remainder were due to transits falling in gaps in the light curve, with a few percent finding the correct injected period but at too low significance. It should be noted that dilution by the secondary star in general had a large effect on the transit depths, resulting in transits significantly shallower than would be expected for e.g. $10R_\oplus$ planets around single stars. In the following, it is assumed that a system which tested successfully at a given planet period and radius would also be successful at any shorter period or larger radius, as both of these changes make detection easier. Note that this method allows us to use the specific sample of binaries in which we could detect planets, so that we are finding the implied occurrence rates truly given by this sample. As such while the recovery percentages give a good idea of completion rates, it is the specific binaries which make up each sample, and moreover the parameter space of each within which planets would be observable as found in Section 7.5, that are the most important outcome of this debiasing process.

7.5 Population Synthesis

7.5.1 Overview

Converting the sample size and number of observed planets that we have into useable statistics requires some synthesis of circumbinary planet populations. The method for this was described in Section 2.2. Planets orbiting our sample binaries were simu-

lated using a Keplerian approach. While this ignores the more complex dynamics of circumbinary planetary systems, including rapid precession, period and eccentricity oscillations (e.g. Farago & Laskar, 2010b; Doolin & Blundell, 2011; Leung & Lee, 2013), the approximation must only hold for the time baseline of 4 years used. Furthermore, the produced planet count would only be sensitive to systematic offsets caused by these effects, which would be expected to be small, rather than orbital element variation which would be taken account of when distributing the orbital elements. This approximation allows us to rapidly sample many possible combinations of orbital elements, something that would be both time and computationally expensive using a full N-body simulator.

7.5.2 Planet Distributions

For planets the crucial distributions are those of inclination and period. There are theoretical indications that planetary inclinations should be preferentially coplanar with the binary (Foucart & Lai, 2013). The actual distribution is largely unknown, with influences possible from protoplanetary disk alignment, planet-planet scattering (e.g. Chatterjee et al., 2008) and other sources of orbital evolution (e.g. Kley & Haghighipour, 2014b). If all circumbinary planets were near perfectly aligned with their binary orbital planes, the detected numbers would represent a significantly different underlying abundance than if the planets were uniformly distributed. As such a variety of inclination distributions are tested, and occurrence rates are given as a function of these. All inclinations are measured relative to the binary plane. Gaussian distributions with means of zero and standard deviations ranging from 5 to 40 degrees are trialled. These are simple functions which can easily be made 'more misaligned', and so without better knowledge of the true distribution represent a good test case. Each of these is convolved with the standard isotropic uniform in cos i distribution (i.e. convolved with sin i at the probability distribution stage), to avoid the bias towards values near zero which would result from using the gaussian distributions directly. An isotropic distribution and fully coplanar distribution are also tested, representing the extreme cases. The injected distributions are shown in Figure 7.3.

In terms of planetary period, the underlying distribution is again poorly known. Using the justification that far from the central binary planet formation and evolution can be expected to proceed as if the host was a single star, the distribution of periods found from the Kepler objects of interest was taken (cut off above 300 d, and corrected for the reduced probability of long period planets transiting), and without further knowledge assumed to hold down to the inner stability limit

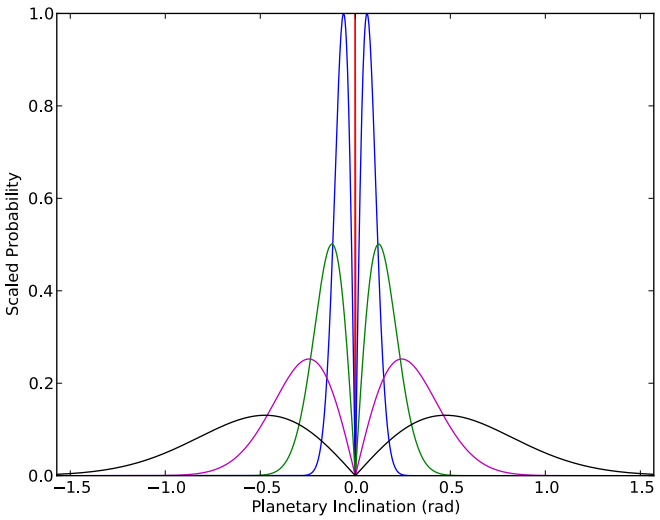

Figure 7.3: Probability density functions for the synthesised planet population inclinations. Distributions are (from centre out) Coplanar, then Gaussian 5, 10, 20 and 40 ° (See text). These are normalised such that the Gaussian 5° distribution peaks at unity.

(Holman & Wiegert, 1999; Dvorak et al., 1989) of each binary, below which planets are taken to be ejected or absorbed by a host star. There have been indications (Welsh et al., 2013) of a potential 'pileup' of planets close to this stability limit - for example through the halting of inward migration at the disk boundary. As such, a distribution whereby 50% of the planets located within the inner stability limit are 'recovered' and placed randomly between 1.1 and 1.4 multiples of that limit was also trialled. The results proved to be generally insensitive to this, and so final results are presented without this pile-up.

7.5.3 Binary Distributions

Many of the necessary binary parameters are already known and were used as described in Chapter 1. Binary inclinations were drawn uniformly across the range within which they would still at least partially eclipse. It is critical to include the binary inclination variation, as for preferentially coplanar planets on much larger semimajor axes a change of a few degrees can have significant consequences for observability.

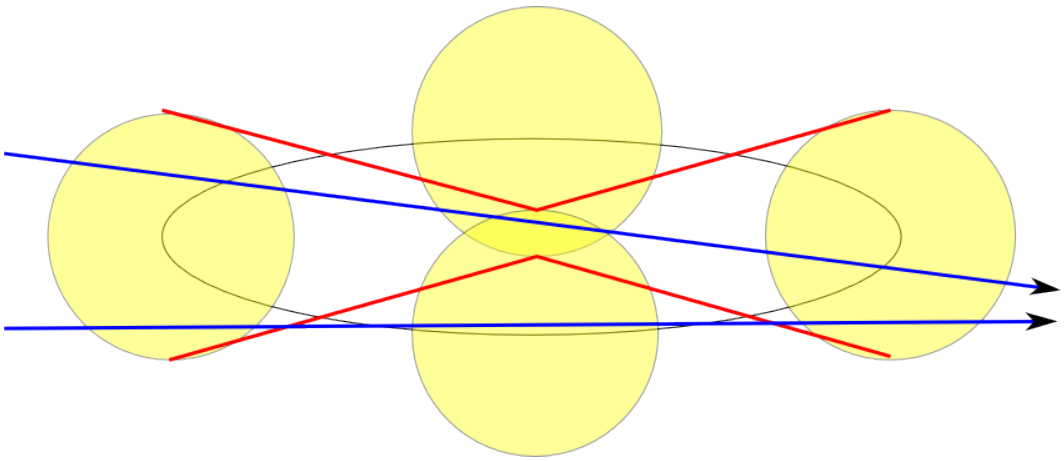

Figure 7.4: Consecutive transit threshold. Four projected positions of the primary star of a binary are shown. The threshold is defined by the red lines. A projected planet trajectory must pass between the red lines. Two possible planet trajectories are shown, one meeting the threshold (top) and one not (bottom).

7.5.4 Output

In each iteration, a proportion of binaries are assigned a planet randomly based on a tested occurrence rate. At this point the relevant binary and planet parameters are drawn, and planets then checked for observability. The test is limited to planets transiting consecutively, i.e. on every orbit, as while the sample has been shown complete for consecutive transits, the effect of occasional missed transits is difficult to quantify. This constraint requires a high degree of alignment with the binary plane, to a degree commensurate with the stellar radii. Within the simulation, this consecutive transit constraint must be quantified by a threshold, the value for which is not immediately obvious. For binaries with inclination $\pi/2$, where the stars only move in one axis, it is clear - the radii of the stars. However, for slightly inclined binaries, where the projected motion of the primary star follows a tight ellipse, it is less so. Planet parameters need to be restricted to a region which ensures consecutive transits. This required the constraint shown in Figure 7.4, which places a tighter restriction on the planet's position at conjunction. At conjunction, the planet must pass through an area which is covered by the primary star on every orbit. As such, the overlapping circles of the projected primary star radius at binary phase 0 and 0.5 (primary and secondary binary eclipse) define this area.

Continuing along the method detailed in Section 2.2, we obtain probability density functions for the occurrence rate, for each of the tested parameter ranges.

7.6 Results

Using these detections and the debiased sample of Section 7.4, we can obtain probability density functions of the implied circumbinary planet rate of occurrence in the Kepler sample, around non-contact binary stars. Typical such distributions are shown in Figure 7.5. These are non-Gaussian, and so maximum likelihood values along with 50 and 95% confidence limits are given. The specific values and errors were found to be only moderately sensitive to the presence of a pile up in planet periods near the inner stability limit. Without full confirmation of its existence values are presented without this pileup, but including one (through recovering 50% of unstable planets into the pileup region as described in Section 7.5.2) leads to occurrence rates which are ∼10% lower for the 300 d period, coplanar group, and unchanged for the 10.2 P_{bin} group. These further reduce in significance for more uniform inclination distributions, and are well within the 50% confidence limits.

The occurrence rates are however critically dependent on the input planetary inclination distribution. As such results are shown as a function of this, and are summarised in Figures 7.6–7.9. The full list of values and confidence limits can be seen in Tables 7.3 and 7.4. (note that modal values are typically accurate to ∼0.5%, unless higher precision is given). The rates in Table 7.3 are lower (and more precise) than for Table 7.4 as 10.2 P_{bin} is generally lower than 300 d in the Kepler sample. This improves transit detection, increasing the sample size of binaries while not increasing the planet count, as nearly all planets are still detectable at 300 d and Kepler-34b no longer lies within the period window. This concentration of the known transiting circumbinary planets at periods close to the binary has been discussed in Section 7.5.2. The varying rates are then a consequence of the window on parameter space one uses to look at the sample.

A number of planet radius bins are shown, both large and small, so that readers may use whichever is most useful for their science. For the periods below 10.2 P_{bin} results are given both with and without Kepler-34b. Strictly Kepler-34b lies at $10.4P_{bin}$, just above the period threshold. In the case of CBs it seems plausible however that a more suitable boundary would be defined by multiples of the binary inner stability limit. In the Kepler-34 case, this limit is particularly large, at ∼190 d, due to the high eccentricity of the binary. Under this definition, Kepler-34b would clearly lie within a similarly defined period boundary. As such both results are given where relevant.

Although a large range of planetary inclination distributions was tested, previous work suggests that some are more likely than others, and that a strong pref-

Table 7.3: Percentage rates of occurrence for planets within $10.2P_{\rm bin}$. Values are maximum likelihoods, with the rates of occurrence corresponding to 50 and 95% confidence intervals shown as super and subscripts.

$R_{\rm planet}$ ($R_\oplus$)	Planetary Inclination Distribution					
	Coplanar	Gauss $\sigma = 5°$	Gauss $\sigma = 10°$	Gauss $\sigma = 20°$	Gauss $\sigma = 40°$	Isotropic
> 10	$0^{0.67\ 2.8}$	$0^{1.1\ 4.8}$	$0^{1.9\ 8.0}$	$0^{3.7\ 15.9}$	$0^{9.0\ 39}$	$0^{26\ 84}$
$8-10$	$2.0^{4.2\ 7.6}_{1.8\ 0.57}$	$3.7^{7.6\ 13.9}_{3.3\ 1.1}$	$7.1^{12.9\ 24}_{5.7\ 2.0}$	$15.0^{26\ 48}_{11.7\ 4.1}$	$33^{60\ 92}_{28\ 10.1}$	$100^{}_{70\ 27}$
$6-8$	$1.5^{3.6\ 7.5}_{1.3\ 0.25}$	$2.5^{6.7\ 13.9}_{2.4\ 0.59}$	$5.0^{11.4\ 24}_{4.1\ 1.0}$	$9.0^{8.4\ 48}_{23\ 2.1}$	$23^{53\ 90}_{19.9\ 5.1}$	$75^{79\ 97}_{39\ 10.9}$
$4-6$	$5.0^{9.4\ 17.3}_{4.2\ 1.4}$	$9.5^{18.0\ 33}_{7.9\ 2.8}$	$15^{31\ 56}_{13.7\ 4.9}$	$31^{60\ 91}_{27\ 10.0}$	$86^{100\ 100}_{66\ 24}$	$100^{}_{76\ 33}$
$6-10$	$4.2^{6.8\ 11.7}_{3.4\ 1.4}$	$7.5^{12.7\ 22}_{6.3\ 2.7}$	$13^{22\ 37}_{10.7\ 4.5}$	$25^{44\ 73}_{22\ 9.4}$	$68^{80\ 97}_{46\ 21}$	$100^{}_{79\ 40}$
$4-10$	$7.1^{12.2\ 21}_{6.1\ 2.6}$	$13.5^{24\ 40}_{11.7\ 5.0}$	$25^{40\ 68}_{20\ 8.6}$	$48^{72\ 95}_{39\ 17.2}$	$100^{}_{75\ 35}$	$100^{}_{81\ 44}$
$8-10$ (inc Kepler-34b)	$3.1^{5.4\ 9.3}_{2.7\ 1.1}$	$5.8^{9.9\ 17.0}_{4.9\ 2.1}$	$10.2^{16.8\ 29}_{8.4\ 3.5}$	$20^{34\ 58}_{17.1\ 7.3}$	$51^{73\ 95}_{39\ 17.0}$	$100^{}_{78\ 38}$
$6-10$ (inc Kepler-34b)	$5.5^{8.3\ 13.5}_{4.5\ 2.1}$	$10.0^{15.5\ 25}_{8.3\ 4.0}$	$17.0^{27\ 43}_{14.3\ 6.9}$	$35^{54\ 82}_{29\ 14.2}$	$87^{86\ 98}_{56\ 29}$	$100^{}_{83\ 49}$

Table 7.4: Percentage rates of occurrence for planets within 300 d. Values are maximum likelihoods, with the rates of occurrence corresponding to 50 and 95% confidence intervals shown as super and subscripts.

R_{planet}	Planetary Inclination Distribution					
($R_\oplus$)	Coplanar	Gauss $\sigma = 5°$	Gauss $\sigma = 10°$	Gauss $\sigma = 20°$	Gauss $\sigma = 40°$	Isotropic
> 10	$0\,{}^{1.4\ 5.9}$	$0\,{}^{3.4\ 14.8}$	$0\,{}^{7.1\ 31}$	$0\,{}^{16.8\ 67}$	$0\,{}^{33\ 89}$	$0\,{}^{45\ 93}$
$8-10$	$6.4\,{}^{11.6\ 19.7}_{5.7\ 2.5}$	$17.5\,{}^{31\ 53}_{15.4\ 6.6}$	$38\,{}^{62\ 91}_{32\ 13.7}$	$100\,{}_{71\ 32}$	$100\,{}_{80\ 42}$	$100\,{}_{83\ 44}$
$6-8$	$3.5\,{}^{8.5\ 17.6}_{3.1\ 0.75}$	$10\,{}^{25\ 51}_{8.9\ 2.2}$	$21\,{}^{18.5\ 88}_{50\ 4.7}$	$47\,{}^{75\ 96}_{34\ 9.1}$	$100\,{}_{65\ 18.4}$	$100\,{}_{69\ 21}$
$4-6$	$14\,{}^{26\ 48}_{11.7\ 4.2}$	$51\,{}^{73\ 96}_{36\ 13.4}$	$100\,{}_{70\ 27}$	$100\,{}_{76\ 32}$	$100\,{}_{77\ 34}$	$100\,{}_{78\ 35}$
$6-10$	$10.0\,{}^{16.3\ 28}_{8.2\ 3.5}$	$30\,{}^{47\ 77}_{23\ 10.0}$	$58\,{}^{78\ 96}_{44\ 19.7}$	$100\,{}_{77\ 37}$	$100\,{}_{81\ 43}$	$100\,{}_{83\ 47}$
$4-10$	$20\,{}^{34\ 58}_{17.2\ 7.4}$	$67\,{}^{82\ 98}_{49\ 22}$	$100\,{}_{78\ 38}$	$100\,{}_{81\ 44}$	$100\,{}_{82\ 44}$	$100\,{}_{85\ 50}$

erence for coplanarity is probable (Foucart & Lai, 2013). Using the coplanar results as an indicative case, there is a 95% confidence upper limit on the occurrence rate of giant ($>10R_\oplus$) planets within 10.2 $P_{\rm bin}$ of 2.8%. Making comparisons to the single star rate of occurrence (Fressin et al., 2013) is difficult, as we do not use the same period ranges. However, looking at their largest two ranges, 0.8–245 d and 0.8–418 d, these rates are ~5% for planets with $R_p > 6R_\oplus$ and ~8% for planets with $R_p > 4R_\oplus$, the latter derived by summing the appropriate radius bins in their paper. Both of these are consistent with the coplanar results, although our modal values are higher. It is worth noting that if it is assumed that the single star rate of occurrence holds in the circumbinary case, for the $6-10R_\oplus$, within 300 d bin the 10° Gaussian inclination distribution would be excluded with probability >99.9%, along with all more misaligned distributions. As such, should a large very misaligned population of circumbinary planets exist, it would imply that circumbinary planets exist in significantly greater numbers than planets with single stellar hosts.

The derived probability density functions also allow us to investigate differences between planetary radius groups (see Section 2.3 for more detail). It has been proposed that giant (Jupiter like) planets should be less common in coplanar circumbinary orbits than Saturn-like or smaller equivalents, due to increased chances of ejection for higher mass planets (Pierens & Nelson, 2008). The Kepler sample supports this, with the rate of occurrence for planets $>10R_\oplus$ within 300 d being significantly lower than the other radius groups. In the coplanar case the probability of this difference is 99.8% ($4-10R_\oplus$), 98.4% ($6-10R_\oplus$), and 96.4% ($8-10R_\oplus$). This finding becomes less significant for distributions more misaligned than the 10° Gaussian case.

It has also been proposed that there is a preference for CBs to have longer period binary hosts (Welsh et al., 2013). All of the known planets so far orbit binaries with periods greater than 7d, despite these longer period binaries being significantly undersampled in the Kepler dataset. It is possible to test whether this effect is due to a sampling bias or represents a real trend using the debiased sample. The sample was split into short and long period binaries, using a period cut of 10 d. For coplanar CB planets with periods less than $10.2P_{\rm bin}$, the probability that the occurrence rate is lower around shorter period binaries was 96.3%(4-10$R_\oplus$), 97.7%(6-10$R_\oplus$) and 95.6%(8-10$R_\oplus$). This becomes more significant for more misaligned inclination distributions, rising to 99.9% for the 5° Gaussian case and higher. Using a binary period cutoff at seven days (below all published CB planets) reduces the significance of the result, to a 92.6% probability for the 6-10$R_\oplus$ sample. This again becomes more significant for more misaligned distributions.

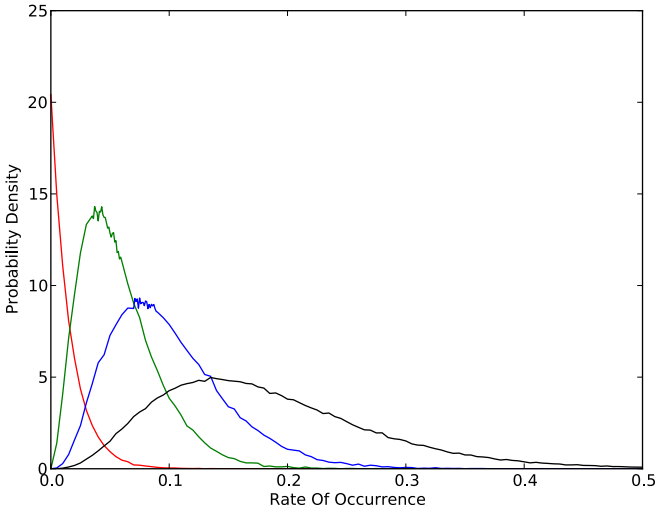

Figure 7.5: Probability density functions for the rate of occurrence of CB planets following a Gaussian inclination distribution with $\sigma = 5°$, within $10.2P_{\rm bin}$. The distributions are shown for (from left to right) planets with radii $>10R_\oplus$, 8-$10R_\oplus$, 6-$10R_\oplus$, and 4-$10R_\oplus$. The $>10R_\oplus$ density function has been scaled down by a factor of three for clarity, and takes a different form to the others due to the zero detections of planets within this group.

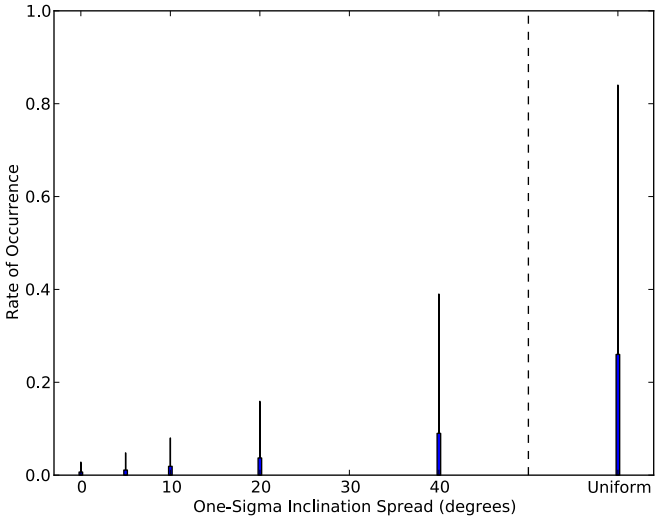

Figure 7.6: Rates of occurrence for a range of Gaussian planetary inclination distributions, for planets within $10.2P_{\rm bin}$ with $R_p > 10R_\oplus$. The large boxes show 50% confidence limits, with the thin 'whiskers' extending to 95% limits.

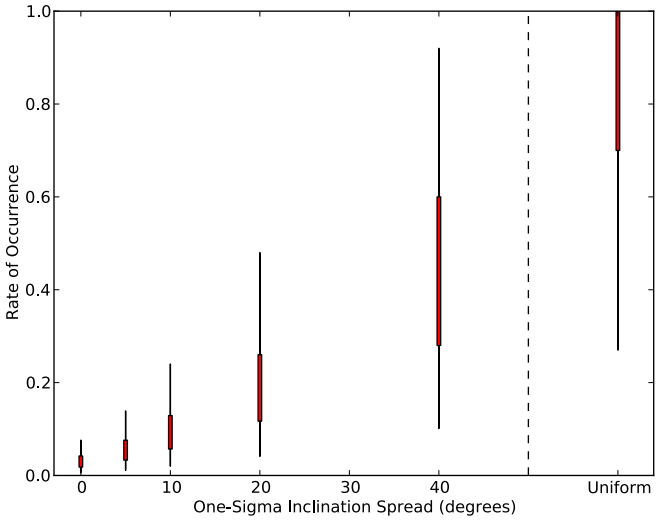

Figure 7.7: As Figure 7.6 for $8 < R_p < 10R_\oplus$

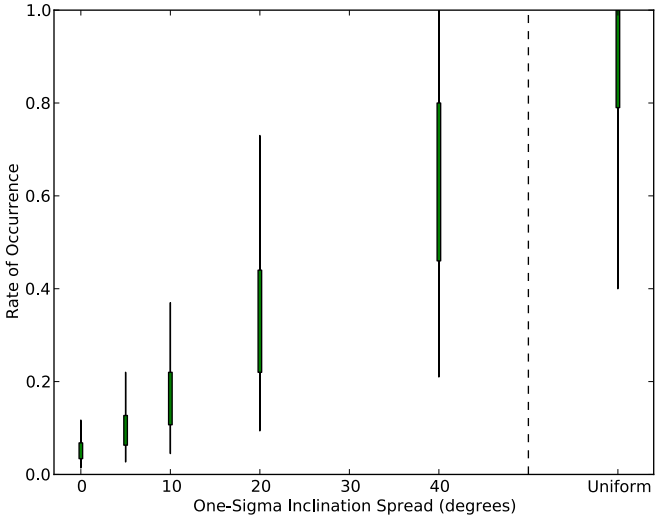

Figure 7.8: As Figure 7.6 for $6 < R_p < 10R_\oplus$

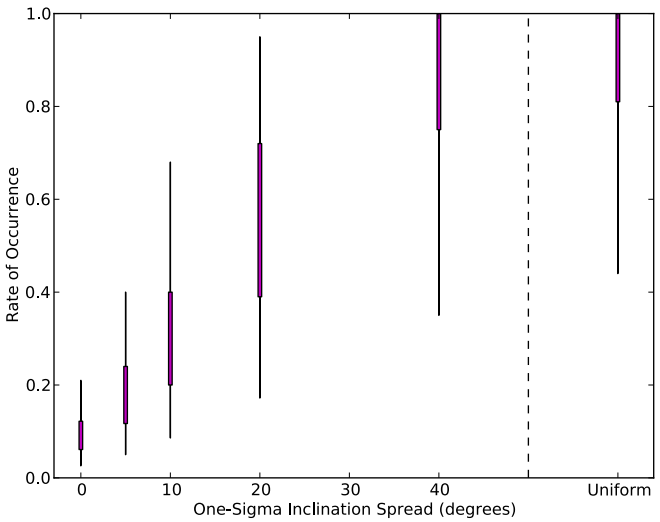

Figure 7.9: As Figure 7.6 for $4 < R_p < 10R_\oplus$

7.6.1 Highly Inclined Distributions, and Multiple Planets/System

As said earlier, the use of rates of occurrence here is defined as the number of binaries with one or more planets as a proportion of the total binary number. This leads naturally to a maximum occurrence rate of 100%. However, as can be seen in Tables 7.3 and 7.4, some tested cases run into this limit, particularly the highly inclined planetary inclination distributions. This has some effect on the results. A large potential area under the probability density function curve in these cases can be found above 100% (i.e. representing multiple planets per binary) and is excluded from the presented values and analysis due to this definition of the occurrence rate. While including multiple planets formally at this time would be a significant extension (note the additional search algorithm, planet parameter correlations, and dynamical questions which would need to be answered), it is informative to investigate the effects of these unused areas of the probability curves.

The particularly high occurrence rates required by highly inclined distributions, such as the isotropic case, have been noted previously. Allowing multiple planets per binary, the full extent of this issue can be demonstrated. This was tested by allowing the occurrence rate to rise above unity in the population synthesis model (keeping all planet parameters independent). In a typical high inclination case ($P < 10.2P_{\rm binary}$, 8-10$R_\oplus$, Isotropic, without Kepler-34b) the results rise to 113^{418}_{36}% (with the values corresponding to 95% confidence limits super and subscripted), showing a strong preference for more than one planet per system. In the

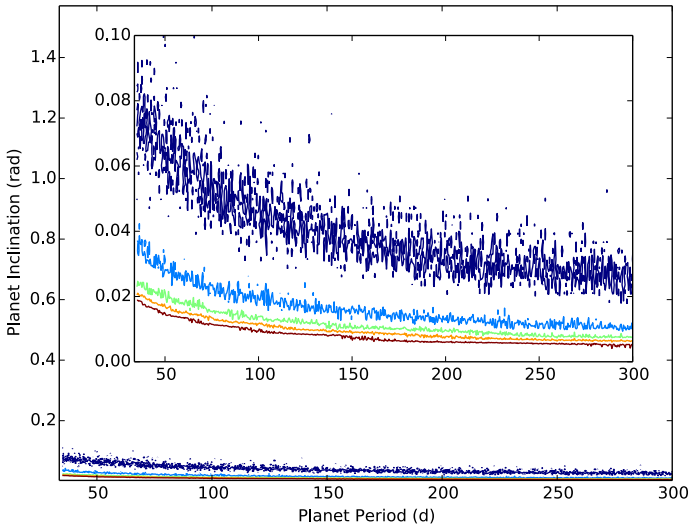

Figure 7.10: Contours of the proportion of planets showing consecutive transits, drawn from a uniform sample of planets with the shown range of inclinations and periods orbiting a binary (itself with inclination $\pi/2$, period 8.5d, and solar radii and mass stars). Contours are plotted at 20% intervals, starting at 10%. for the uppermost line.

most extreme case ($P < 300$ d, 4-10$R_{\oplus}$, Isotropic) the results rise dramatically to a modal value of near 50 planets per binary, a number which would presumably lead to serious stability issues within this relatively tight period bound. Note that due to the change in definition of occurrence rate implied here these numbers cannot be considered a direct extension of the previous results, and are merely indicative. When values are needed, those given in Tables 7.3 and 7.4 should be used with the earlier definition of the occurrence rate. In the light of this however, it is worth reemphasising that should the true inclination distribution of circumbinary planets be particularly misaligned with respect to their host binaries, their formation must be abundant, common and in essence very hard to avoid.

7.7 Discussion

7.7.1 Rate of Occurrence and Errors

The errors associated with the presented rates of occurrence are particularly large compared with those for single stars; this is a function of both the reduced sample size and moreover the constraint of consecutive transits. The region of parameter space within which consecutive transits occur is decidedly smaller than for single

stars, reducing the sensitivity of a given sample. Both these occurrence rates and their errors increase sharply for increasingly uniform planetary inclination distributions. This behaviour is expected, as in the uniform case many more planets must exist in order to produce the few we see transiting. The errors increase as the possible parameter space of planetary orbits becomes largely unprobed by our consecutive transit requirement, which is only sensitive to nearly coplanar planets. To illustrate how this situation comes about, the sensitivity of this method as a function of planetary inclination is plotted in Figure 7.10. The region of parameter space where consecutive transits are possible is shown. The tiny area of the total parameter space this represents is striking - that a reasonable number of planets should be found within it (as they have been) largely explains the qualitative form of these results. Unbiased searches for misaligned CB planets, for example on non-eclipsing binary stars (Martin & Triaud, 2014), will be essential to constraining the CB inclination distribution and the implied rates of occurrence. Interestingly, the occurrence rate estimated in that work is compatible with our values, despite being based on a different method and involving no analysis of the Kepler light curves themselves.

It should be noted that the presented values test planets after significant periods of evolution. It is unlikely that any of the known transiting CB planets formed where they are currently located (Meschiari, 2012). As such these rates include both planet formation and subsequent dynamical evolution, through disk migration, scattering or otherwise. Furthermore the starting point of this history is not fully understood - the abundance of circumbinary disks is not yet well known, although it has been shown that they should be common (Alexander, 2012). If these disks occur more or less readily than circumstellar disks then it impacts the formation rates implied by the presented rates of occurrence.

7.7.2 Biases and Approximations

Any statistical study is subject to various potential biases, which are summarised here. The first is in the sample chosen, of Kepler eclipsing binary stars. This is not a general sample of binaries, with a study of the full effects of the Kepler pipeline well beyond this work. There is also a bias towards shorter period binaries, the usual geometric bias associated with selecting eclipsing objects. As such our rates of occurrence are skewed towards these shorter period binaries. Given that the currently known transiting CBs are found orbiting generally longer period binaries ($P_{\rm bin} \gtrsim 5$d) this may be significant, and the effects of this will be the target of future work. There is also a bias towards larger, evolved stars as they are more

likely to eclipse, which may be present in the binary sample (although the removal of overcontact binaries may ameliorate this).

There is also a bias against more active stars (with noisier light curves) due to the difficulty in detecting planetary transits, especially where the timescale of that noise becomes shorter than ~3 d (see Section 4.3). This will preferentially reject closer binaries, as they are more likely to have activity induced by the companion, and so leads to a sample bias towards longer period binaries within the dataset. Stars with particularly sharp binary eclipses may also be affected, although the effect of these eclipses has been mitigated as far as possible (as described in Section 4.2.3). Similarly, planets with orbits on very close integer resonances with the host binary are more likely to be rejected as noise, or to have their transits removed with the binary signal.

In checking for consecutive transits, a Keplerian approximation was made as to the planetary orbits. This will become important for planets orbiting on short enough periods that their precession timescales become comparable to the data baseline (~4 yr). Using the formula of Doolin & Blundell (2011), derived from Farago & Laskar (2010b), it is possible to determine where this region typically begins: for a moderately eccentric $e_p = 0.2$ coplanar planet at the inner stability limit, orbiting an $e_{\rm bin} = 0.1$ binary, the binary must have a period under ~0.06 d for the planetary precession period to fall below 4 years. As such, this will not be a problem here. The precession timescales of CB planets are however fairly short, on the order of decades (see Chapter 3). This means that objects which consecutively transit through the dataset may well not do in several years time, as is the case for Kepler-16b (Doyle et al., 2011). This is accounted for by the statistical nature of our method – a planet on a slightly misaligned orbit will consecutively transit for a fraction of the iterations, and only be counted for that fraction. As detailed in the above section, the consecutive transit requirement also impacts our sensitivity to inclined planets. This is included in the presented errors, but shows that the information leading to these results comes from a narrow region of parameter space in terms of planetary inclination.

Tertiary stellar companions have also not been accounted for, which will dilute planetary transits and reduce the chance of detection. This contamination has been estimated to be potentially as high as 20% (Rappaport et al., 2013). Without further detail it is impossible to estimate how strongly such tertiary companions would dilute transits of the primary star, so rates of occurrence were produced without this. However should the contamination rate be this high and the dilution be significant it will raise all of the presented rates by effectively reducing the sample

size. In the worst case (if in 20% of the sample the dilution was always strong enough that we could not in fact detect planets) the true sample size would be reduced by this 20%. This would have the effect of increasing the presented occurrence rates by ~20% of their present values. This would not affect the conclusions made regarding the various trends evident.

Similarly in terms of the injected transits, transits of the secondary star were not included. We do not expect these to contribute significantly to the detection. Dilution by the primary star means that the relative depth of transits of the secondary star compared to those on the primary goes as $(T_2/T_1)^4$, implying that for all but particularly equal temperature binaries ($T_2/T_1 > \sim 0.92$, corresponding to a transit depth ratio of ~0.7) transits of the secondary would not contribute significantly to the detection. From Raghavan et al. (2010), their figure 16, it is possible to estimate how many binaries this applies to. This estimate is somewhat rough (as the samples are by no means the same, and it involves converting mass ratio to temperature) but leads to ~10-15% of the sample having significant secondaries. As several of these binaries will already be successful detections, including secondaries would increase the detectable binary sample by at most a few percent, decreasing the derived abundance rates by a few percent of their present values.

We have relied on an element of human eyeballing of the search algorithm results, introducing potential subjectivity. This was mitigated through using two independent checkers, and the results supported in that every known planet host (excepting Kepler-47, which the algorithm did not detect) was marked by both. The use of defined significance thresholds (see Section 7.4.1) also constrained the sample to a size amenable to finely detailed checking.

Finally, there is a possible effect from errors on the temperatures of Chapter 6; these are ~400K for the primary stars and ~600K for the secondaries, which would affect the derived radii used to produce transit depths and check for consecutive transits. As our results are statistical, errors on individual binaries will not have a large effect, the important factor being whether systematic offsets are found in the temperatures. We cannot check for this, but there is no indication that they should be present.

7.8 Summary

The rates of occurrence of circumbinary planets orbiting close ($P_{\rm bin} <\sim 60$ d) non-contact binary stars were investigated using the Kepler sample of eclipsing binaries. This produced a number of interesting results:

1. The most significant controlling distribution is that of planetary inclination - whether these planets lie preferentially coplanar with their host binaries, or in a more uniform pattern. The results show that if such a uniform or close to uniform distribution is the norm, then the rate of occurrence of CBs must be exceptionally high, significantly more so than analogous rates for single stars. While not formally excluding very uniform, misaligned planetary inclination distributions, these results show that to exist such distributions need planetary formation rates at levels very difficult to explain.

2. Conversely, if coplanarity is preferred, to the level implied by a Gaussian distribution with standard deviation ~5° or tighter (although the distribution by no means must be Gaussian, and may even be bimodal) then the rate of occurrence of CBs is consistent with that of single star planets. Evidence suggests that circumbinary planets orbiting sub-AU binaries should be preferentially coplanar due to alignment of the protoplanetary disk, supporting this option (Foucart & Lai, 2013; Kennedy et al., 2012).

3. CB giant planets (defined as $>= 10R_{\oplus}$) are significantly less common than their smaller equivalents. There remains the possibility of a non-coplanar giant CB population at any rate of occurrence, formed for example by dynamical evolution, but a coplanar CB giant population on the same order as planets with $R < 10R_{\oplus}$ is excluded, at least within the tested period range. Given that proto-planetary disk masses scale with the mass of the central object (Andrews et al., 2013), and that more massive disks produce more gas giants (Mordasini et al., 2012) this supports the finding of Pierens & Nelson (2008), that CB Jupiter mass planets if present will likely orbit at larger distances from the central binary due to increased scattering.

4. CB planets are less common in coplanar orbits around shorter period binaries ($P_{\rm bin} < \sim 5-10$ d) than around binaries of longer period. We have shown that this trend is not the result of detection bias, with 99.9% confidence for all tested misaligned planetary inclination distributions and 97.7% for a coplanar distribution. The observed difference could be explained through a significantly different orbital distribution between planets orbiting shorter and longer period binaries (such as a more misaligned shorter population, so that we do not observe them) or by an effect of the formation of these binary systems. If shorter period binaries form through secular interactions with a tertiary stellar companion, planets in these systems would either be disrupted, or if present difficult to see due to dilution by the companion. If such close binaries have

evolved to their current orbit via angular moment loss (through e.g. stronger magnetic braking) then this process may influence the protoplanetary disk and impact planet formation. This remains a promising area of future work.

Chapter 8

Conclusion

8.1 Summary

Through this thesis, the aim has been to develop the area of circumbinary (CB) planets from one of a few specific objects to a more general understanding. When it was begun, all that was known observationally was that they existed. Over the last few years, several more planets have been discovered, providing more material for investigation. It is the general studies, taking the now available sample of planets and deriving what can be derived from them, where the discoveries into the properties lying at the heart of this class of objects lie. Each Chapter of the preceding work develops its own topic, all with the eventual goal of an in depth study of this type. This was realised in Chapter 7, where all of the previous work was united in investigating the abundances of circumbinary planets.

Initially, Chapter 3 used the general geometric principles of a CB orbit to derive analytic limits on the transit timing variations (TTVs) of the planet's transits. This was then extended to take into account precession of the planet's argument of periapsis. With both of these effects, the TTVs of planets became limitable to an accuracy of ~1%, providing a significant boost to survey efficiency and design. Moreover, the expressions derived are amenable to approximating some binary parameters if they are not known, making them particularly versatile.

These limits were used to form the base of a search algorithm targeting CB planets specifically. While this step in itself is not new - see Ofir (2008) - the use of TTV limits incorporated into the algorithm is, and is here proved as a viable method which can be added to future algorithms. Moreover, at that time no easily useable algorithms were available, Carter & Agol (2013, the QATS algorithm) only being published later. QATS is, interestingly, amenable to being upgraded with

these TTV limits, via a setting of the maximum shift allowed. The search algorithm developed here was applied to both the WASP and Kepler datasets. Later results, in particular involving the longer period host binaries of CB planets, showed that the WASP dataset, which generally holds binaries with periods less than ~5 days, was not an ideal target for a CB planet search. However, it may prove useful in other ways, see below. This precipitated the move onto the Kepler light curves, which allow both a longer baseline and importantly, continuous observations. The result was a number of new candidate systems, both planetary and stellar, as well as the ability to statistically debias the sample of binaries, something particularly useful later. The newly discovered candidates will go on to add to the small sample of CB systems, while the stellar multiples can each be useful in studying n-body dynamics (some show very rapid orbital evolution, on the timescale of months) as well as the benefits to stellar evolution study of having several stars formed coevally.

The target at this point was the occurrence rate work of Chapter 7. However, this proved impossible with the stellar parameters available at the time - in particular stellar radii were needed, because without them the transit depths used to test binaries would be heavily approximated. Stellar radii from the KIC (Brown et al., 2011) could have been used, but these were already under some doubt (Farmer et al., 2013), potentially subjective to temperature bias and poorly recovered radii, something unacceptable for a statistical investigation. As such it was necessary to remove this bias, and so a catalogue was produced based on spectral energy distribution fits. Fortunately, a number of surveys of the Kepler field had already been undertaken, providing the necessary data for the task. This catalogue, as well as allowing the later work presented here, will be useful to anyone studying binaries in the Kepler field. As said, the only available parameters before came from the KIC, and suffered a bias in temperature. The catalogue developed allows for more efficient target selection (cool or hot stars for example), better initial guesses when studying systems in more detail, and generally serves as a guide to the binaries of the field. There are already several tens of papers focused on these systems, with many more expected in future.

In the context here, the catalogue is most important for what it, along with the TTV limits and search algorithm developed previously, allows. It was now possible to take the sample of binaries and planets and investigate the abundance of CBs, something which had only been attempted theoretically before. This first observationally determined abundance would be a very useful constraint on the large number of studies looking at CB formation and evolution - all must as their end point produced the observed abundance rate. With this as the objective, the

Kepler sample of binaries was debiased, producing a set of binaries within which planets could be detected. With this and the known planets and candidates found earlier, population synthesis was performed to find what underlying distributions of CB planets could have produced the observed sample. The results of this form the main conclusion of this thesis. Specific occurrence rates were found, which show the CB planets are as common as those around single stars. If they follow a significantly more misaligned than coplanar distribution, they must in fact be more common. It was shown that there is a significant lack of Jupiter-sized and larger planets in coplanar orbits, in agreement with previously suggested theory. Pierens & Nelson (2008) predicted that beyond a rough size threshold, larger planets would preferentially scatter out of the system or onto more misaligned orbits, raising the intriguing possibility that a population of Jupiter-sized planet exists at more misaligned and as-yet unobserved inclinations. Furthermore, this work shows that CB planets are strongly favoured orbiting binaries with host periods greater than ~7 days. This may be a first look at a threshold in binary formation processes - however shorter period binaries form, be it via magnetic braking, three body interactions or another method, it seems to be relatively inimical to planet formation.

This work has taken place as the field of circumbinary planets has developed, from the discovery of the first transiting CB, Kepler-16b, at the beginning, to a state where the underlying distributions of these planets can be begun to be investigated. It has contributed to this development, through TTV limits which allow for more efficient search algorithms, a catalogue to guide and inform users of the still to be fully utilised Kepler binary sample, and the first observational determination of the CB occurrence rate. This rate, and the conclusions drawn while investigating it, will drive theoretical work in the subject which until now has proceeded informed mainly by individual planets. This development goes beyond CB planets, and into the more general area of planet formation, including how our own solar system formed. Planet formation, a subject under much scrutiny but as yet not fully understood, must proceed under all manner of conditions. One of those conditions, and a very different one to the single star case, is in binary systems. Formation under 'S-Type' orbits has been studied for several years, and we can now hope that the CB, 'P-Type' case will add its own understanding to the investigation. Eventually we will discover how our own solar system, and the huge numbers of others now being found throughout the Galaxy, formed from the clouds of dust and gas from which they began.

8.2 Future Work

This field is only beginning to enter maturity, and much remains to be done. The key next stage is in investigating the CB inclination distribution. Up until now, we have been limited to planets coplanar with their host binaries, largely due to observational constraints - they transit more regularly and more often. At present the misaligned planets known are only misaligned on the order of a few degrees, leaving a large parameter space as yet unexplored. This space is important: the inclinations of planets are a key indicator of their past dynamical history. Are for example planets perturbed from initially coplanar CB disks, or are the disks even initially coplanar? Are as predicted Jupiter-size planets more likely to be found in inclined orbits? This distribution of inclinations also feeds back into the occurrence rate of CB planets, allowing it to be calibrated much more tightly.

There are a number of possibilities to detect inclined CB planets. Proposals have been made to target non-eclipsing binaries, looking for planetary transits. While such transits will be irregular, there may be enough planets to make it a plausible route (Martin & Triaud, 2014), if the transits are clear enough and the binary nature of the host can be determined through for example light curve variations (which may be possible for longer period binaries due to the highly precise Kepler data, Faigler et al., 2012). Sparse, irregular transits of inclined planets can also be searched for in eclipsing binary systems, although there will always be a bias to coplanar planets with more transits. While radial velocities are not biased towards coplanar planets, they do not pin down the planet's inclination, making them of only limited use in this.

However, more discoveries increasing the small sample size of CBs are very desirable. From this perspective, both radial velocity and direct imaging surveys should provide valuable new CB systems in the near future. Each has been ongoing for several years now, but with radial velocities specifically the baseline obtained should reach the point at which discoveries can be expected. In the realm of transits, future missions will provide large numbers of new systems. The planned TESS mission is unlikely to find more than a handful due to its shorter observing baseline on each field - combining the preference for longer period binaries with an inner stability limit of ~5 times the binary period means ~150 days of baseline is likely needed for three planetary transits, at a minimum. This does not preclude it from further investigating the lack of CB planets orbiting shorter period binary hosts. In the long term, the recently selected PLATO mission with its longer time baseline and high precision should find a larger number. Moreover, the PLATO targets

are bright stars, leading to the exciting possibility of doing planetary atmosphere work on CBs (how would equilibrium atmospheric chemistry change under a high amplitude temperature forcing cycle?). In the more medium term from the ground, the NGTS observatory will observe a large portion of the sky over a long baseline. While being ground based it will suffer from the same non continuous data issues as WASP in finding CB planets, the increased precision will lead to the ability to detect individual transits of the largest planets, solving the issue of trying to bin transits with high and unknown TTVs. Despite having these problems, WASP itself may still prove useful - the exceptionally large sample it holds of ~7000 eclipsing binaries may allow for significantly improved statistics on the jupiter sized CB planets. There is also the SOLARIS project, which using four 0.5m telescopes aims to detect CBs using eclipse timing and precision radial velocities.

Finally, theoretically the effects of multiple planets in CB systems is only beginning to be explored. With only one transiting system known of this type (Kepler-47), it is a fledgling subject. Higher mass planetary CB companions are known (e.g. NN Ser) but are in a very different regime, both in mass and orbital period. The effects of the binary on resonances between planets, and on the dynamics between planets in these cases, will be the subject of future investigation. In all, the study of CB planets promises much development, both in the short and long term, leading to new insights into their formation, evolution, and how these impact on the study of planetary systems of all kinds.

Bibliography

Adams E. R., Dupree A. K., Kulesa C., McCarthy D., 2013, The Astronomical Journal, 146, 9

Alexander R., 2012, The Astrophysical Journal, 757, L29

Andrews S. M., Rosenfeld K. A., Kraus A. L., Wilner D. J., 2013, The Astrophysical Journal, 771, 129

Applegate J. H., 1992, The Astrophysical Journal, 385, 621

Armstrong D. et al., 2012, Astronomy and Astrophysics, 545, L4

Armstrong D. J. et al., 2013a, Monthly Notices of the Royal Astronomical Society, 437, 3473

Armstrong D. J. et al., 2013b, Monthly Notices of the Royal Astronomical Society, 434, 3047

Armstrong D. J., Osborn H., Brown D., Faedi F., Gómez Maqueo Chew Y., Martin D., Pollacco D., Udry S., 2014, Monthly Notices of the Royal Astronomical Society, 444, 1873

Artymowicz P., Lubow S. H., 1994, The Astrophysical Journal, 421, 651

Baines E. K., McAlister H. A., Brummelaar T. A. t., Turner N. H., Sturmann J., Sturmann L., Ridgway S. T., 2008, The Astrophysical Journal, 682, 577

Batalha N. M. et al., 2010a, The Astrophysical Journal Letters, 713, L109

Batalha N. M. et al., 2013, The Astrophysical Journal Supplement Series, 204, 24

Batalha N. M. et al., 2010b, The Astrophysical Journal, 713, L103

Bergfors C. et al., 2012, Monthly Notices of the Royal Astronomical Society, 428, 182

Beuermann K. et al., 2010, Astronomy and Astrophysics, 521, L60

Bloemen S. et al., 2012, Monthly Notices of the Royal Astronomical Society, 422, 2600

Bloemen S. et al., 2010, Monthly Notices of the Royal Astronomical Society, pp no–no

Boss A. P., 2006, The Astrophysical Journal, 641, 1148

Bowler B. P., Liu M. C., Kraus A. L., Mann A. W., 2014, The Astrophysical Journal, 784, 65

Boyajian T. S. et al., 2012, The Astrophysical Journal, 757, 112

Brown T. M., Latham D. W., Everett M. E., Esquerdo G. A., 2011, The Astronomical Journal, 142, 112

Bryson S. T. et al., 2010a, The Astrophysical Journal, 713, L97

Bryson S. T. et al., 2010b, The Astrophysical Journal, 713, L97

Campbell B., Walker G. A. H., Yang S., 1988, The Astrophysical Journal, 331, 902

Cardelli J. A., Clayton G. C., Mathis J. S., 1989, The Astrophysical Journal, 345, 245

Carter J. A., Agol E., 2013, The Astrophysical Journal, 765, 132

Carter J. A. et al., 2011, Science, 331, 562

Castelli F., Kurucz R. L., 2004, eprint arXiv:astro-ph/0405087

Chaplin W. J. et al., 2013, The Astrophysical Journal Supplement Series, 210, 1

Chatterjee S., Ford E. B., Matsumura S., Rasio F. A., 2008, The Astrophysical Journal, 686, 580

Claret A., 2004, Astronomy and Astrophysics, 428, 1001

Cohen M., Wheaton W. A., Megeath S. T., 2003, The Astronomical Journal, 126, 1090

Collier Cameron A. et al., 2006, Monthly Notices of the Royal Astronomical Society, 373, 799

Conroy K. E., Prsa A., Stassun K. G., Orosz J. A., Fabrycky D. C., Welsh W. F., 2013, eprint arXiv:1306.0512

Conroy K. E. et al., 2014, The Astronomical Journal, 147, 45

Debosscher J., Blomme J., Aerts C., De Ridder J., 2011, Astronomy and Astrophysics, 529, A89

Delfosse X. et al., 2004, Spectroscopically and Spatially Resolving the Components of the Close Binary Stars, 318, 166

Delorme P. et al., 2013, Astronomy and Astrophysics, 553, L5

Derekas A. et al., 2011, Science, 332, 216

Devor J., 2005, The Astrophysical Journal, 628, 411

Doolin S., Blundell K. M., 2011, Monthly Notices of the Royal Astronomical Society, 418, 2656

Doyle L. R. et al., 2011, Science, 333, 1602

Duquennoy A., Mayor M., 1991, Astronomy and Astrophysics, 248, 485

Duquennoy A., Mayor M., 2005, Binaries as Tracers of Stellar Formation. Cambridge University Press

Dvorak R., Froeschle C., Froeschle C., 1989, Astronomy and Astrophysics, 226, 335

Everett M. E., Howell S. B., Kinemuchi K., 2012, Publications of the Astronomical Society of the Pacific, 124, 316

Faigler S., Mazeh T., Quinn S. N., Latham D. W., Tal-Or L., 2012, The Astrophysical Journal, 746, 185

Farago F., Laskar J., 2010a, Monthly Notices of the Royal Astronomical Society, 401, 1189

Farago F., Laskar J., 2010b, Monthly Notices of the Royal Astronomical Society, 401, 1189

Farmer R., Kolb U., Norton A. J., 2013, Monthly Notices of the Royal Astronomical Society, 433, 1133

Fischer D. A., Marcy G. W., 1992, The Astrophysical Journal, 396, 178

Foucart F., Lai D., 2013, The Astrophysical Journal, 764, 106

Fraquelli D., Thompson S. E., 2012, Kepler Archive Manual (KDMC-10008-004)

Fressin F. et al., 2013, The Astrophysical Journal, 766, 81

Gies D. R., Williams S. J., Matson R. A., Guo Z., Thomas S. M., Orosz J. A., Peters G. J., 2012, The Astronomical Journal, 143, 137

Girardi L., Bressan A., Bertelli G., Chiosi C., 2000, Astronomy and Astrophysics Supplement Series, 141, 371

Greiss S. et al., 2012a, The Astronomical Journal, 144, 24

Greiss S. et al., 2012b, eprint arXiv:1212.3613

Haas M. R. et al., 2010, The Astrophysical Journal Letters, 713, L115

Hadden S., Lithwick Y., 2014, The Astrophysical Journal, 787, 80

Halbwachs J. L., Mayor M., Udry S., Arenou F., 2003a, Astronomy and Astrophysics, 397, 159

Halbwachs J. L., Mayor M., Udry S., Arenou F., 2003b, Astronomy and Astrophysics, 397, 159

Hastings W. K., 1969, Biometrika, 57, 97

Hatzes A. P., Cochran W. D., Endl M., McArthur B., Paulson D. B., Walker G. A. H., Campbell B., Yang S., 2003, The Astrophysical Journal, 599, 1383

Hellier C. et al., 2014, Monthly Notices of the Royal Astronomical Society, 440, 1982

Hinse T. C., Horner J., Lee J. W., Wittenmyer R. A., Lee C.-U., Park J.-H., Marshall J. P., 2014, Astronomy and Astrophysics, 565, A104

Holman M. J., Wiegert P. A., 1999, The Astronomical Journal, 117, 621

Horner J., Wittenmyer R. A., Hinse T. C., Marshall J. P., Mustill A. J., Tinney C. G., 2013, Monthly Notices of the Royal Astronomical Society, 435, 2033

Howard A. W. et al., 2010, Science, 330, 653

Howell S. B. et al., 2014, Publications of the Astronomical Society of the Pacific, 126, 398

Ibe O. C., 2009, Markov Processes for Stochastic Modelling. Elsevier Academic Press

Irwin J. M. et al., 2011, The Astrophysical Journal, 742, 123

Jaime L. G., Pichardo B., Aguilar L., 2012, Monthly Notices of the Royal Astronomical Society, 427, 2723

Jenkins J. M. et al., 2010, in Radziwill N. M., Bridger A., eds, SPIE Astronomical Telescopes and Instrumentation: Observational Frontiers of Astronomy for the New Decade. SPIE, pp 77400D–77400D–11

Kallrath J., Milone E., 2009, Eclipsing Binary Stars: Modelling and Analysis, 2 edn. Springer

Kennedy G. M., Wyatt M. C., Sibthorpe B., Phillips N. M., Matthews B. C., Greaves J. S., 2012, Monthly Notices of the Royal Astronomical Society, 426, 2115

Kley W., Haghighipour N., 2014a, Astronomy and Astrophysics, 564, A72

Kley W., Haghighipour N., 2014b, eprint arXiv:1401.7648

Koch D. G. et al., 2010, The Astrophysical Journal Letters, 713, L79

Konacki M., Muterspaugh M. W., Kulkarni S. R., Hełminiak K. G., 2009, The Astrophysical Journal, 704, 513

Kopal Z., 1959, Close Binary Systems. New York: Wiley

Kostov V. B. et al., 2014, The Astrophysical Journal, 784, 14

Kostov V. B., McCullough P. R., Hinse T. C., Tsvetanov Z. I., Hébrard G., Díaz R. F., Deleuil M., Valenti J. A., 2013, The Astrophysical Journal, 770, 52

Kovacs G., Zucker S., Mazeh T., 2002, Astronomy and Astrophysics, 391, 369

Kraus A. L., Ireland M. J., Cieza L. A., Hinkley S., Dupuy T. J., Bowler B. P., Liu M. C., 2013, The Astrophysical Journal, 781, 20

Lee J. W., Kim S.-L., Lee C.-U., Lee B.-C., Park B.-G., Hinse T. C., 2013, The Astrophysical Journal, 763, 74

Leung G. C. K., Lee M. H., 2013, The Astrophysical Journal, 763, 107

Lillo-Box J., Barrado D., Bouy H., 2012, Astronomy and Astrophysics, 546, A10

Lines S., Leinhardt Z. M., Paardekooper S., Baruteau C., Thébault P., 2014, The Astrophysical Journal, 782, L11

Mandel K., Agol E., 2002, The Astrophysical Journal, 580, L171

Marsh T. R. et al., 2013, Monthly Notices of the Royal Astronomical Society, 437, 475

Martin D. V., Triaud A. H. M. J., 2014, Arvix

Martin R. G., Armitage P. J., Alexander R. D., 2013, The Astrophysical Journal, 773, 74

Matijevič G., Prsa A., Orosz J. A., Welsh W. F., Bloemen S., Barclay T., 2012, The Astronomical Journal, 143, 123

Mayer L., Wadsley J., Quinn T., Stadel J., 2005, Monthly Notices of the Royal Astronomical Society, 363, 641

Meschiari S., 2012, The Astrophysical Journal, 761, L7

Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, The Journal of Chemical Physics, 21, 1087

Morales J. C. et al., 2009, The Astrophysical Journal, 691, 1400

Mordasini C., Alibert Y., Benz W., Klahr H., Henning T., 2012, Astronomy and Astrophysics, 541, A97

Morgan W. W., Harris D. L., Johnson H. L., 1953, The Astrophysical Journal, 118, 92

Mugrauer M., Ginski C., Seeliger M., 2014, Monthly Notices of the Royal Astronomical Society, 439, 1063

Mugrauer M., Neuhäuser R., 2009, Astronomy and Astrophysics, 494, 373

Murphy S. J., Bedding T. R., Shibahashi H., Kurtz D. W., Kjeldsen H., 2014, Monthly Notices of the Royal Astronomical Society, 441, 2515

Musielak Z. E., Cuntz M., Marshall E. A., Stuit T. D., 2008, Astronomy and Astrophysics, 480, 573

Nataf D. M., Stanek K. Z., Bakos G. A., 2010, Acta Astronomica, 60, 261

Nesvorny D., Kipping D. M., Buchhave L. A., Bakos G. A., Hartman J., Schmitt A. R., 2012, Science, 336, 1133

Ofir A., 2008, Monthly Notices of the Royal Astronomical Society, 387, 1597

Orosz J. A. et al., 2012a, The Astrophysical Journal, 758, 87

Orosz J. A. et al., 2012b, Science, 337, 1511

Østensen R. H. et al., 2010, Monthly Notices of the Royal Astronomical Society: Letters, 408, L51

Patil A., Huard D., Fonnesbeck C. J., 2010, Journal of Statistical Software, 35

Pelupessy F. I., Portegies Zwart S., 2013, Monthly Notices of the Royal Astronomical Society, 429, 895

Perryman M. A. C. et al., 2001, Astronomy and Astrophysics, 369, 339

Pfahl E., Arras P., Paxton B., 2008, The Astrophysical Journal, 679, 783

Pierens A., Nelson R. P., 2007, Astronomy and Astrophysics, 472, 993

Pierens A., Nelson R. P., 2008, Astronomy and Astrophysics, 483, 633

Pilat-Lohinger E., Funk B., Freistetter F., Dvorak R., 2002, Exo-Astrobiology, 518, 547

Pollacco D. et al., 2006, Astrophysics and Space Science, 304, 253

Prsa A. et al., 2011, The Astronomical Journal, 141, 83

Prsa A., Guinan E. F., Devinney E. J., DeGeorge M., Bradstreet D. H., Giammarco J. M., Alcock C. R., Engle S. G., 2008, The Astrophysical Journal, 687, 542

Prsa A., Pepper J., Stassun K. G., 2011, The Astronomical Journal, 142, 52

Rafikov R. R., 2012, The Astrophysical Journal Letters

Rafikov R. R., 2013, The Astrophysical Journal, 764, L16

Raghavan D. et al., 2010, The Astrophysical Journal Supplement Series, 190, 1

Rappaport S., Deck K., Levine A., Borkovits T., Carter J., El Mellah I., Sanchis-Ojeda R., Kalomeni B., 2013, The Astrophysical Journal, 768, 33

Rappaport S. et al., 2012, The Astrophysical Journal, 752, 1

Reid I. N., Hawley S. L., 2005, New Light on Dark Stars. Praxis Publishing Ltd

Roell T., Neuhäuser R., Seifahrt A., Mugrauer M., 2012, Astronomy and Astrophysics, 542, A92

Schleicher D. R. G., Dreizler S., 2014, Astronomy and Astrophysics, 563, A61

Schwamb M. E. et al., 2013, The Astrophysical Journal, 768, 127

Schwarzenberg-Czerny A., 1989, Monthly Notices of the Royal Astronomical Society, 241, 153

Schwarzenberg-Czerny A., 1997, The Astrophysical Journal, 489, 941

Seager S., D'Angelo G., Durisen R., Lissauer J., 2010, in , Exoplanets. The University of Arizona Press

Seager S., Murray C., Correia A., 2010, in , Exoplanets. The University of Arizona Press, p. 19

Seager S., Winn J., 2010a, in , Exoplanets. The University of Arizona Press

Seager S., Winn J., 2010b, in , Exoplanets. The University of Arizona Press, pp 57–58

Shibahashi H., Kurtz D. W., 2012, Monthly Notices of the Royal Astronomical Society, 422, 738

Sigurdsson S., 2003, Science, 301, 193

Skrutskie M. F. et al., 2006, The Astronomical Journal, 131, 1163

Slawson R. W. et al., 2011, The Astronomical Journal, 142, 160

Smith J. C. et al., 2012, Publications of the Astronomical Society of the Pacific, 124, 1000

Southworth J. et al., 2011, Monthly Notices of the Royal Astronomical Society, 414, 2413

Stellingwerf R. F., 1978, The Astrophysical Journal, 224, 953

Still M., Barclay T., 2012, Astrophysics Source Code Library, -1, 08004

Stumpe M. C. et al., 2012, Publications of the Astronomical Society of the Pacific, 124, 985

Sybilski P., Konacki M., Kozłowski S., 2010, Monthly Notices of the Royal Astronomical Society

Thalmann C. et al., 2013, European Planetary Science Congress 2013, held 8-13 September in London, UK. Online at: http://meetings.copernicus.org/epsc2013, id.EPSC2013-1020, 8, 1020

Thébault P., Marzari F., Scholl H., 2006, Icarus, 183, 193

Thompson S. E. et al., 2012, The Astrophysical Journal, 753, 86

Torres G., Andersen J., Giménez A., 2010, The Astronomy and Astrophysics Review, 18, 67

Torres G. et al., 2010, The Astrophysical Journal, 727, 24

Tran K., Levine A., Rappaport S., Borkovits T., Csizmadia S., Kalomeni B., 2013, The Astrophysical Journal, 774, 81

Wang J., Xie J.-W., Barclay T., Fischer D. A., 2014, The Astrophysical Journal, 783, 4

Welsh W. F., Orosz J. A., Carter J. A., Fabrycky D. C., 2013, arXiv.org

Welsh W. F. et al., 2012, Nature, 481, 475

Wittenmyer R. A., Horner J., Marshall J. P., 2013a, Monthly Notices of the Royal Astronomical Society, 431, 2150

Wittenmyer R. A., Horner J., Marshall J. P., 2013b, Monthly Notices of the Royal Astronomical Society, 431, 2150

Zorotovic M., Schreiber M. R., 2013, Astronomy and Astrophysics, 549, A95

Zsom A., Sándor Z., Dullemond C. P., 2011, Astronomy and Astrophysics, 527, A10