DETECTING PHASECURVES IN THE K2 DATA

A Thesis

presented to

the Faculty of Natural and Applied Sciences

at Notre Dame University – Louaize

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

---

by

JANA HAJJ SLEIMAN

February 2019
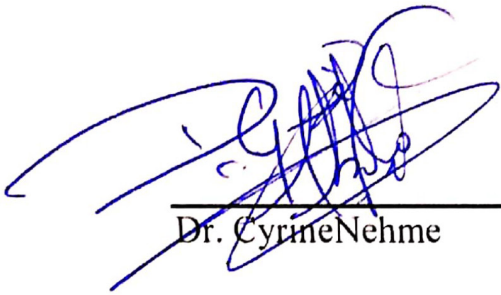
Notre Dame University – Louaize

Faculty of Natural and Applied Sciences

Department of Physics and Astronomy

We hereby approve the thesis of

Jana Hajj Sleiman

Candidate for the degree of Masters in Science

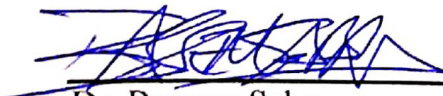Dr. CyrineNehme   Supervisor

Dr. Marwan Gebran   Committee Member, Chair

Dr. Bassem Sabra   Committee Member

Dr. Wehbe Farah   Committee Member

Abstract

Detecting Phase Curves in the K2 Data

by Jana Hajj Sleiman, February 2019

Exoplanets are planets orbiting stars other than our Sun. During their motion around their star, due to gravitational effects, and due to the planet's thermal and reflective radiation, small variations in the lightcuvre could be detected (Esteves et al. 2015). Phasecurves are small-scale variations in planetary flux that arise from an exoplanet's motion around its host star. If detected, the phasecurve can lead to determining the planet's mass, its albedo, and it can give us an idea about its atmosphere. In this work we aim to detect these these variations in the data from the K2 mission. To do this we first test two different models on two already published phasecurves detected in the K2 data (Malavolta 2018),(Močnik et al. 2018). We reveal the failure of one of the models (Serrano 2018) to recover a phasecurve for a planet with a relatively low mass having a lightcurve that posses a signal with a significant noise, while the other one (Faigler and Mazeh 2011) succeeds in recovering the phasecurve of the same planet. We also uncover significant modulations in the data of K2 31-b, which resulted in finding the planet's mass and its geometric albedo.

Key words: exoplanets: lightcurves-phasecurves techniques: model fitting - techniques: photometric

Résumé

Détection des courbes de phase dans les données K2

par Jana Hajj Sleiman, Février 2019

Avec plus de 3916 découvertes, les planètes extrasolaires représentent un champ d'étude majeur de lastronomie actuelle. Les exoplanètes sont des planètes en orbite autour d'autres étoiles que notre Soleil.

Au cours de l'orbite d'une planète autour de son étoile, et en raison des effets gravitationnels, du rayonnement thermique et réfléchi de la planète, des petites variations de lumière peuvent dès lors être détectées. Les courbes de phase, donc, sont des variations à petite échelle du flux planétaire résultant du mouvement dune exoplanète autour de son étoile hôte. Caractériser les parametres physique d'une planète extra solaire (I.e la masse, l'albedo et son atmosphere) est directement relié à la détection de sa courbe de phase.

Durant ce stage, nous nous sommes particulièrement interessés à détecter ces petites variations de flux dans les données de la mission K2. Pour ce faire, nous testons deux modèles différents, d'abord sur deux phases détectes et publiées dans la base de données K2 data. Nous révélerons l'échec de l'un des modèles à récupérer la phase pour une planète de masse relativement faible, tandis que l'autre reussit. Nous avons également découvert des variations de flux significatives dans les données de K2 31-b, qui ont permis d'estimer la masse de la planète et son albdo géométrique.

Mots clés: exoplanétes: techniques courbes de lumière - courbes de phase - techniques: ajustement de modèles - techniques: photométrique

Acknowledgments

I owe debts of gratitude to

My family, who believed in me;

Prof. Olivier Demangeon, for his patience, support, and for lighting up every dark avenue I encountered;

Dr. Cyrine Nehme, my advisor, for all the help she offered

To all my instructors, for their guidance throughout the masters program;

To Dr. Marrie Abboud, for her help in the Erasmus Grant.

To Bachar Wehbe, for his efforts towards the Erasmus Grant and going to Porto.

To all my colleagues, for making this journey pleasurable and memorable;

My special thanks goes to Elias Kammoun, for all the help he offered, and the remarks he gave, to Elie for all his support and encouragements. To my friends, I do not know how to thank you.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Our observable universe consists of billions of galaxies, which in turn is a collection of billions of stars. The stars are orbited by planets. Unlike stars, planets are relatively small and do not generate or radiate energy. They absorb energy and re-emit it at different wavelengths. Those two factors make them difficult to study and hard to directly image them even with the largest telescopes. The study of planets is rather worthy, as it can help us understand the planetary diversity as well as the formation of planetary systems.

1.1 History on the Search of Exoplanets

An exoplanet is a planet outside our solar system. It was only till 1992 that the first exoplanet was detected. The detection was made when studying pulsar PSR1257. A pulsar is a highly magnetized rotating neutron star that emits a beam of electromagnetic radiation. This radiation can be observed only when the beam of emission is pointing toward Earth, and is responsible for the pulsed appearance of emission. Neutron stars are very dense, and have short, regular rotational periods. Pulsar PSR1257 is a millisecond pulsar, with a rotation period of 6.22 milliseconds , and was found to have anomalies in the pulsation period. These anomalies led to the first confirmed discovery of planets outside our solar system. (Wolszczan and Frail 1992)

After that, from 1995 till 1998, several planets were detected using a spectrograph which can detect slight and periodic velocity changes in the spectral lines of a star. These changes are caused by a planet's gravitational effect on a star. This technique is known as the radial velocity method (see section 1.2.1). Later on in 1999 planet

HD 209458 b which was previously discovered with the radial velocity technique, was oberved eclipsing its parent star. (Charbonneau et al. 2000)

This method for detecting exoplanets is known as the transit method. It has eversince discovered thousands of exoplanets. With the big discoveries of exoplanets, several space missions projects were done with an objective to find exoplanets and characterize them.

COROT

COROT was launched in December 2006 [1] . It was the first spacecraft dedicated to the detection of transiting exoplanets. COROT was able to to announce 600 candidates and confirm 29. On November 2nd 2012 COROT suffered a computer failure and was announced retired by June 2013 as repair missions failed.

Kepler

After COROT came Kepler. Kepler was a space observatory launched by NASA in March 2009. The spacecraft's goal was to observe a large sample of stars and determine the fraction of Earth sized or larger planets lay in the habitable zone. The habitable zone is the orbital region around a star in which an Earth-like planet can possess liquid water on its surface and possibly support life. Kepler used the transit method to detect planets that are 30 to 600 times less massive than gas giants. This was a challenge as most of the previously detected planets by other projects were Jupiter sized planets or bigger. As of 2012, there were a total of 2,321 candidates. Out of which 207 are similar in size to Earth, 680 are super-Earth-size, 1,181 are Neptune-size, 203 are Jupiter-size and 55 are larger than Jupiter.[2]

---

[1]http://sci.esa.int/corot/
[2]http://archive.stsci.edu/

On May 15, 2013, NASA announced the spacecraft had been crippled by failure of a reaction wheel that keeps it pointed in the right direction. A second wheel had previously failed, and the spacecraft requires three wheels (out of four total) to be operational for the instrument to function properly. Developed over the months following this failure, the K2 mission represents a new concept for spacecraft operations that enables continued scientific observations with the Kepler space telescope. Smart engineers had devised a remarkable solution which is using the pressure of sunlight to stabilize the spacecraft so it could continue to do science, with only 2 functioning reaction wheels, and the spacecrafts thrusters. The fix worked, and Kepler was given a new mission as K2. [3] (Refer to figure 1.1)

---

[3]https://www.nasa.gov/
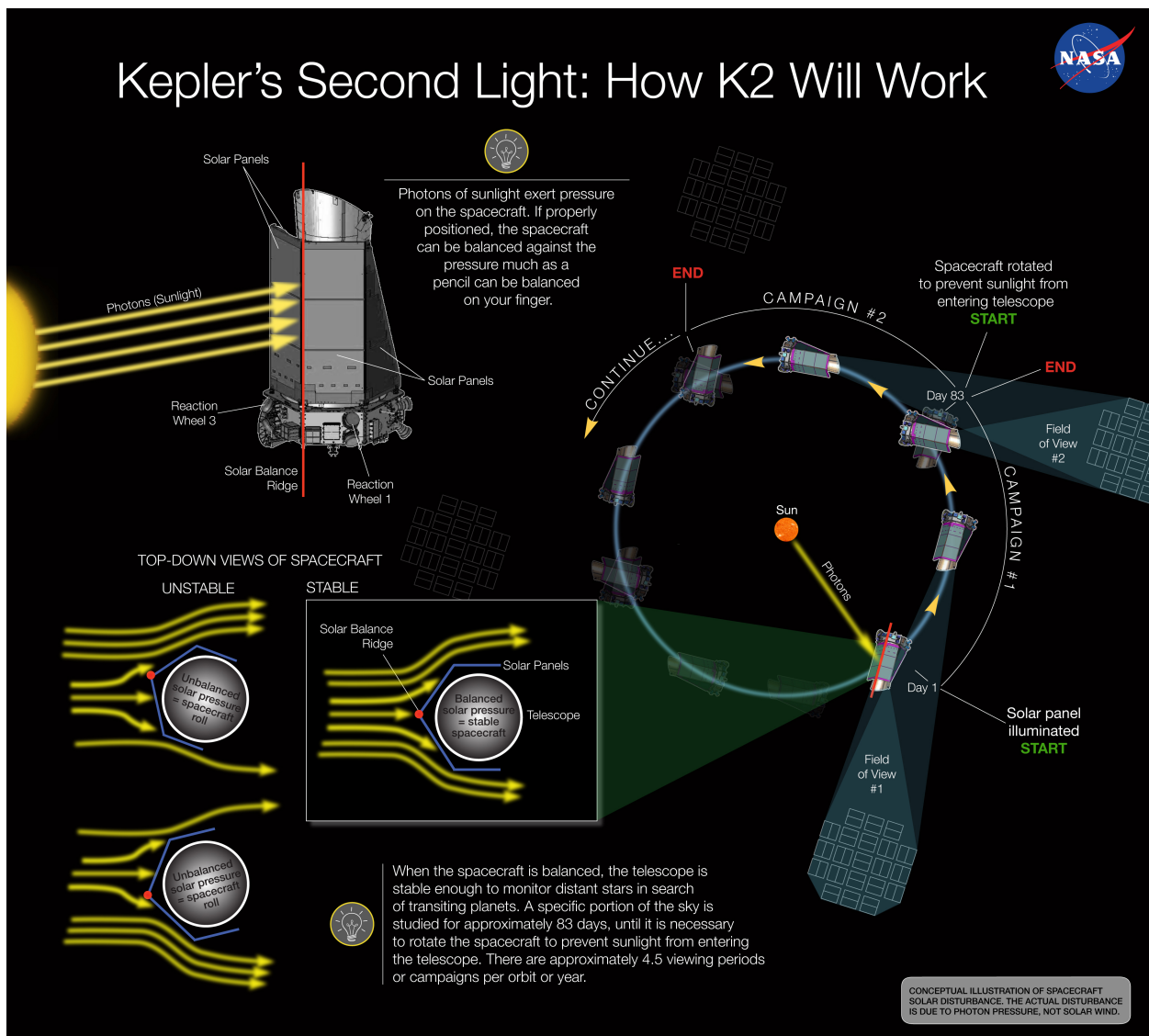
Figure 1.1: The proposed fixture of Kepler

Using the transit method to detect brightness changes, the K2 mission entailed a series of sequential observing "Campaigns" of fields distributed around the ecliptic plane and offered a photometric precision approaching that of the original Kepler mission. Operating in the ecliptic plane minimizes the torque exerted on the spacecraft by solar wind pressure, reducing pointing drift to the point where spacecraft attitude can effectively be controlled through a combination of thrusters and the two remaining reaction wheels. Each campaign is therefore limited by sun angle constraints to a duration of approximately 80 days.the number of confirmed exoplanets discovered by K2 extension mission is around 360(to date). Although kepler is no longer active(out of fuel), a lot of data still awaits to be studied.

1.2 Detection Methods

Many detection methods evolved with the development in the discoveries of exoplanets. While there are several methods, in this chapter we will briefly discuss two methods, which are the most common detection techniques. The transit method, and the radial velocity method.

1.2.1 Radial Velocity Method

The radial velocity method, also known as Doppler spectroscopy, is the most effective method for locating extrasolar planets with existing technology. The radial velocity method relies on the fact that a star does not remain completely stationary when it is orbited by a planet. It moves, ever so slightly, in a small circle or ellipse, responding to the gravitational tug of its smaller companion. When viewed from a distance, these slight movements affect the star's normal light spectrum, or color signature. If the star is moving towards the observer, then its spectrum would appear slightly shifted towards the blue; if it is moving away, it will be shifted towards the red. Using highly sensitive spectrographs, we can search for periodic shifts towards

Figure 1.2: Schematic representation of the radial velocity technique.
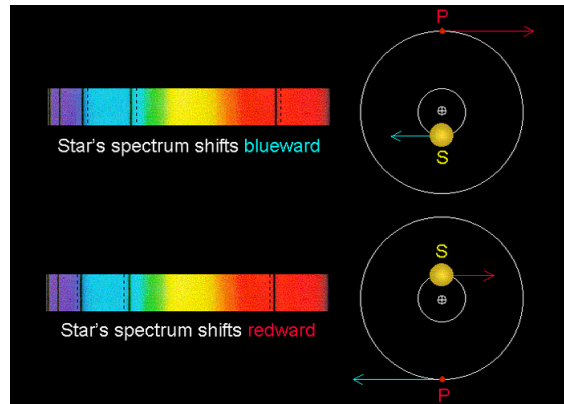
the red, blue, and back again. If the shifts are regular, repeating themselves at a certain period it means that the star is moving back and forth, towards the Earth and then away from it in a regular cycle. This, in turn, is almost certainly caused by a body orbiting the star, and if it is of a low enough mass it is called a planet.

Radial velocity observations covering all orbital phases are able to measure the orbital period P, the eccentricity e, and the Radial Velocity (RV) semi-amplitude K. From these observables, the so-called 'minimum mass' $M_p \sin i$ can be computed, provided the total mass of the system is known.

$$K = \frac{28.4329 m s^{-1}}{\sqrt{1-e^2}} \frac{m_p \sin i}{M_J} \left(\frac{M_* + M_P}{M_\odot} \right)^{-2/3} \left(\frac{P}{1yr} \right)^{-1/3}$$

Where: m_p is the mass of the planet, M_J is the mass of Jupiter, M_* is the mass of the star, P is the planet's orbital period, and e is the ecccentricity.

In practice, planetary masses are usually negligible compared to the mass of the parent star. The stellar mass can be obtained indirectly via spectroscopic analysis, photometry, parallax measurements and comparison with stellar evolutionary models. (Fischer et al. 2014)

Figure 1.3: Effects of system's inclination on the detection of radial velocity.
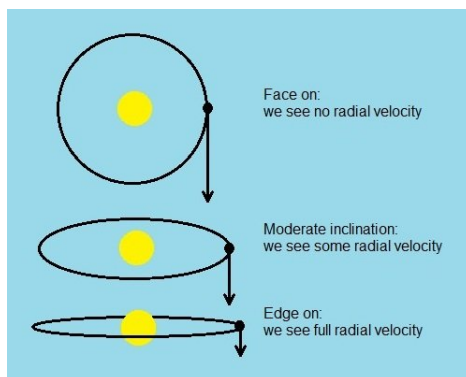
One disadvantage of this method is that it can only detect the movement of a star towards or away from the Earth. If the orbital plane of the system is "edge-on" when observed from the Earth, then the entire movement of the star will be towards or away from the Earth, and can be detected with a sensitive spectrograph. In this case the mass of the planet can be accurately derived. If, however, the orbital plane of the planet is "face on" when observed from the Earth, the entire movement of the star will be perpendicular to an observer's line of sight. No part of its movement will be towards or away from the Earth. No spectrum shift will be detected, and the observer will not know about the presence of a planet orbiting the star (See figure 1.3 for representation). In most cases a planet's orbital plane is tilted with respect to the line of sight at some angle, which is usually unknown. The mass of the suspected planet is directly proportional to the star's actual wobble. But since only a portion of this wobble is detected, then the measured mass will be lower than the true one and provide only a minimum value for the planet's mass.

Another drawback of this method is that it is most likely to find the types of planets that are the least likely to be habitable. Early on, most of the planets detected

by spectroscopy were of a type known among scientists as "hot Jupiters." These are giant planets composed mostly of gas, similar to our neighbor, Jupiter, but orbiting at dizzying speeds at a very short distance from their star. Their size, short periods, and close proximity to their star ensures that they produce the quick and relatively large stellar wobbles that are most easily detected by spectroscopy. Cooler planets orbiting further away produce more moderate wobbles in their home star, and take years to complete each orbit, factors which make them much harder to detect with spectroscopy. But while hot Jupiters are relatively easy to find, they are unlikely homes to any form of life as we know it. Even worse, their presence at the center of a planetary system makes it less likely that more Earth-like planets had survived in their neighborhood. In other words, while the discoveries made with spectroscopy established the presence and prevalence of planets outside our Solar System, most of the systems detected with this method are very unlikely abodes for life.

1.2.2   The Transit Method

Transit occurs when an exoplanet crosses infront of its host star relative to the observer's line of sight. When this happens, the planet blocks some of the received star's flux. This event will periodically occur at a period equal to the planet's orbital period around its host star. If the stellar radius is known, the depth of the eclipse can give us some information about the planet's radius by referring to the following formula (Seager and Mallén-Ornelas 2003)

$$\Delta F = \frac{F_{nontransit} - F_{transit}}{F_{nontransit}} = (\frac{R_p}{R_*})^2$$

Where $F_{nontransit}$ is the flux received by the star when the planet is not eclipsing it, $F_{transit}$ is the flux received by the system during transit, R_p is the planet's radius, and R_* is the star's radius.

When the planet passes behind the host star relative to the observer's line of sight, if the geometry is convenient, a secondary eclipse can be detected (Check figure

1.4 for more details). Although the majority of the flux received from the system is contributed by the star, the planet has a small contribution due to its thermal emission.
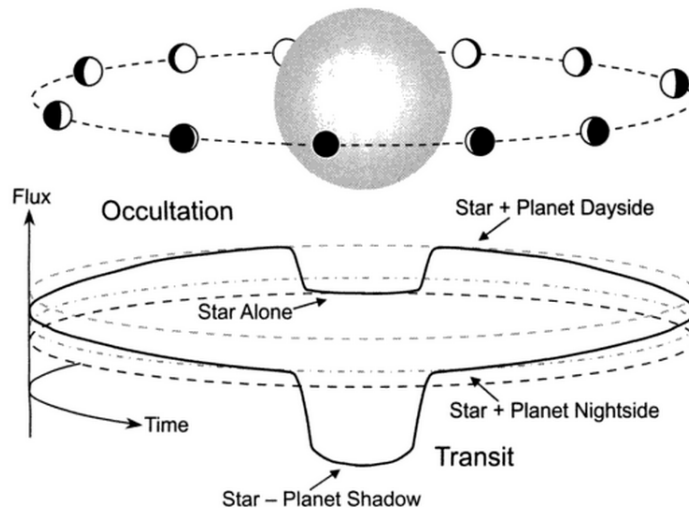
Figure 1.4: Schematic representation of the light curve of a transiting planet.

The planet's equilibrium temperature is calculated in López-Morales and Seager (2007). They showed that the equilibrium temperature of the planet depends on stellar radius, the distance between the star and the planet, and the planet's bond and geometric albedo.

$$T_{eq} = T_* \left(\frac{R_*}{a} \right)^{1/2} [f(1 - \frac{3}{2}A_g)]^{1/2} \qquad (1.1)$$

where T_* is the stellar temperature, R_* is the stellar radius, a is the semi-major axis, A_g is the geometric albedo and f is the normalized flux.

The bond albedo is the fraction of power in the total electromagnetic radiation incident on an astronomical body that is scattered back out into space. It accounts

for all of the light scattered from a body at all wavelengths and all phase angles, so it can determine how much energy an object absorbs, which is crucial for determining the equilibrium temperature of a body. The geometric albedo however, is the ratio of its actual brightness as seen from the light source to that of an idealized flat, fully reflecting, diffusively scattering (also known as Lambertian) disk with the same cross-section. Later on, a study related the eclipse depth with the geometric albedo. (Esteves et al. 2013) So detecting eclipses can constrain the thermal emission and the reflection by the planet.

1.3   Phasecurves

Exoplanet detection and characterization is important on so many levels. Foremost among these is its impact on our understanding of the nature and formation of planetary systems. The greater the sample of known planets, the better the conclusions we can draw about their common properties and formation mechanisms. Until the discovery of exoplanets, we had only eight major planets and several minor bodies endemic to one solar system to study and theorize. With thousands of new worlds to study, we are able to statistically analyze the overall planet population, which refines our understanding of planetary science. Naturally, a priority in the field of exoplanets is the discovery of another Earth-like planet and the existence of extraterrestrial life. Phase curve studies have the potential to significantly broaden the pool of known planets, thereby advancing the field generally.

Phase curves are small-scale, variable photometric effects that arise from an exoplanets motion about its host star. For a homogeneously emitting star, the phase curve would show the primary transit (when the planet passes in front of the stellar disk), the secondary transit (when the planet is hidden by the star), as well as three additional modulations. The first modulation is the Doppler boosting effect; it

consists of a modification of the stellar brightness proportional to the radial velocity variation induced by the planet (Barclay 2012) . The second effect is the ellipsoidal modulation, that is the gravitational attraction of the planet on the stellar surface, the star is deformed, with a surface tide that moves following the planet (Esteves et al. 2013). The last component is the planetary flux which mainly contains atmospheric reflection at optical wavelengths. The reflected light depends on the planetary albedo, which is the ability of a planet to reflect the star's light. (see Chapter 2 for further discussion). During the first stages of Kepler, most light curves analysis focused only on the transiting parts of the lightcurve. However, when studying the full phasecurve of the star-planet system, those small variations can be detected.

One rare advantage of analyzing phasecurves is that the study allows the full characterization of the physical and orbital parameters of an exoplanet. When all effects are taken into consideration, the planetary mass can be calculated (see later for more details). Morever parameters that indicate the reflectivity and temperature of the planet can be derived. With such parameters known, we can have a hint about the composition and dynamics of the atmosphere. These information usually require transit spectrospy, which is a complicated technique that studies the resolved spectra of planets as they transit their host star. If one is able to characterize the underlying stellar spectrum, the planetary atmospheric spectrum is derived by subtracting the stellar baseline from a spectrum taken during transit.

Another advantage of studying phasecurves is that they have the potential to study non-transiting exoplanets. Like the radial velocity technique, the magnitude of the phasecurve effects diminishes with decreasing inclination, and thus analysis is constrained solely by the sensitivity of the photometric detector. Thus studying phasecurves can fully characterize exoplanets over a greater range of inclinations, al-

lowing studies of the majority of exoplanets that are not fully characterized yet.

Esteves et al. (2013) previously studied planets observed by kepler. They were able to detect phasecurves in 8 planets. They detected ellipsoidal variations in 5 of these planets: planets Kepler-5, Kepler-6, Kepler-8, TrES-2, and HAT-P-7. From those variations, they were able to recover planetary masses that are within 2.5 σ of their published radial velocity measurements. Later on Esteves et al. (2015) did a similar study and they were able to detect phasecurves in 14 planets observed by Kepler (Kepler-5b, Kepler-6b, Kepler-7b, Kepler-8b, Kepler-10b, Kepler-12b, Kepler-41b, Kepler-43b, Kepler-76b, Kepler-91b, Kepler-412b, TrES-2b, HAT-P-7b, and KOI-13b). They were able to derive geometric albedos and estimate the planets' equilibrium temperatures.
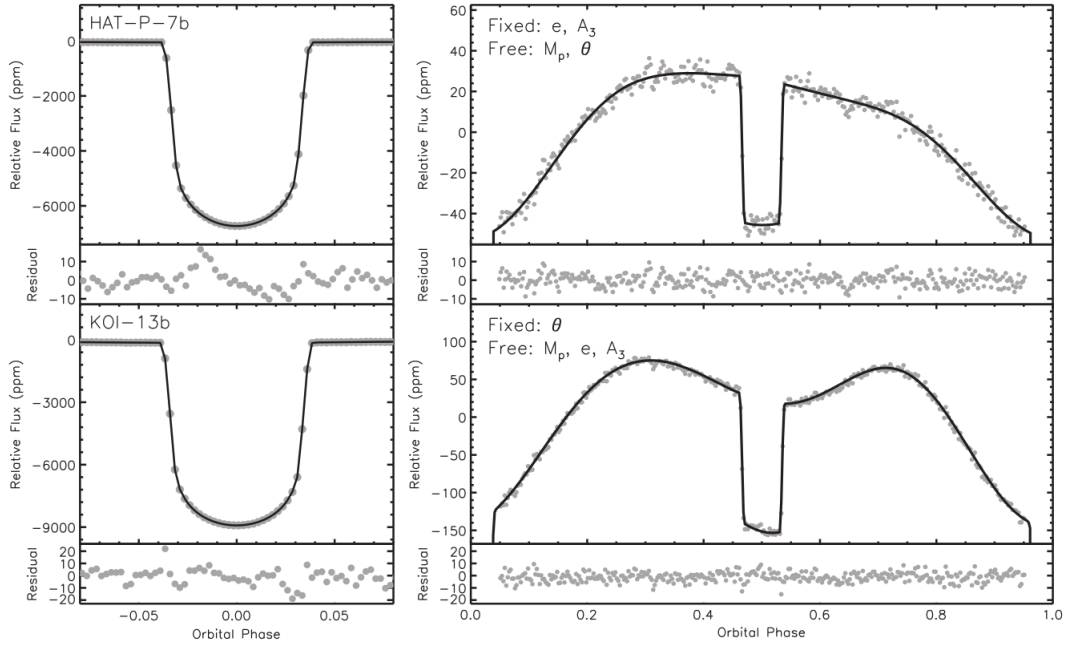
Figure 1.5: Phase-folded transit and phasecurves for Kepler-91b, Kepler-412b, and TrES-2b. (Esteves et al. 2015)

In this work, we aim to detect phasecurves in the K2 data. Similar work was previously done on Kepler's data (Esteves et al. 2013). But with the K2 data, there are only two detected phasecurves up to date. (Malavolta 2018), (Močnik et al. 2018)

The first part of the project consists of testing a phasecurve model on planets with published phasecurves. A model developed in (Serrano 2018) was created as a mean to simulate planetary signal, but it was never tested on actual data. In this project, we try to validate this model and test its limitations. We compare its results with a model that has been tested before (Faigler and Mazeh 2011). The second part of the project consists of searching through the confirmed planets in the K2 mission, in an attempt to detect a phasecurve, and characterize the planet.

In the next chapter, we present the physics behind the models that were used to fit the data. Then in chapter 3 we go through the data processing and the fitting methods. We proceed our work by comparing Serrano's model to Mazeh's model on previously published phasecurves and discover limatations for Serrano's model. Finally we present a new discovered phasecurve of planet K2 31-b.

Chapter 2

Light Curves of Transiting Exoplanets

2.1 Primary and Secondary Transit Model

The light curve of the star-planet system consists of the primary and secondary transits, as well as small variations that are due to three physical phenomena. The beaming effect, the ellipsoidal variations, and the planetary reflected light. These three effects combined build up the phasecurve. Throughout the work the transits are modelled using the Batman Transit Model. However for the modelling of the phasecurve we test two different models. In the following section we will describe each model used.

2.1.1 Batman Transit Model

For the transit fitting we use the python package batman. (Kreidberg 2015) This package models high precision light curves. The depth of the transit is determined by the ratio of the planets radius to its host star. However, the transit shape and depth can be highly modified due to the limb darkening effect. Limb darkening is an optical effect seen in stars, where the center part of the disk appears brighter than the edge or limb of the image. The limb darkening causes the apparent brightness of the stellar disk to decrease from center to edge. The stellar intensity profile can be fit with several functional forms, some of which are linear limb darkening law (Schwarzschild and Villiger 1906), quadratic (Kopal 1950), square-root(Diaz-Cordoves et al. 1995) and four-parameter nonlinear (Claret 2000).
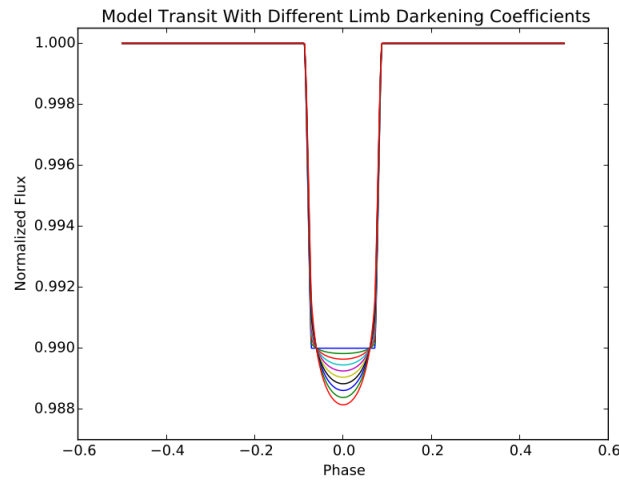
Figure 2.1: Effects of different limb darkening coefficients on the shape of the transit curve.

In order to calculate the transit light curve, we need to choose a a limb darkening model. whether linear, quadratic, square-root or nonlinear, we need to provide the model with the appropriate limb-darkening coefficients. These coefficients can be found in a study byClaret and Bloemen (2011), who computed the limb-darkening coefficients of different laws, for a wide range of stars with different effective temperature, gravity and metallicity.

After choosing the limb-darkening model and the appropriate coefficients, the batman transit model needs 7 more parameters to input.

Time of inferior conjunction "t0"

Orbital period of the planet

Planetary radius (in units of stellar radii)

Semi-major axis "a" (in units of stellar radii)

Inclination of orbital plane "i"

Eccentricity of planetary orbit "e"

Longitude of periastron "w"(in degrees)
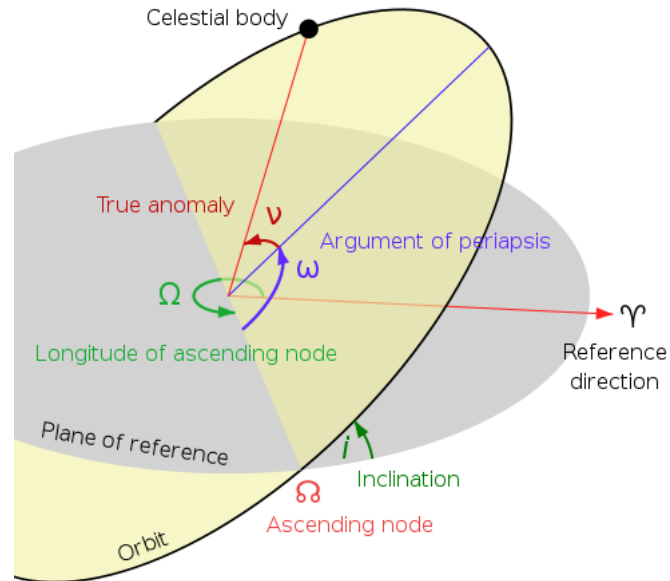
Figure 2.2: The stars spin axis is perpendicular to the orbital plane (yellow), which is tilted a certain number of degrees relative to the observer, called the inclination i. Inclination is 90 for an edge on orbit and 0 for a face on orbit. The celestial body, in this case an exoplanet, orbits elliptically on this orbital plane, following the stars spin. True anomaly (ν here) indicates the location of an object in its orbital ellipse relative the the foci of the systems center of mass. Argument of periapsis ω documents the position of the planet relative to its pericentre, which is the point in its orbit when its distance from the star is at minimum.

With those parameters provided the model can now calculate the light curve for a given time interval.

This package also allows us to model the secondary eclipse using the following formula:

$$f = 1 + f_p(1 - \alpha)$$

where f is normalized flux, f_p is the planet-to-star flux ratio, and α is the fraction of the planet disk that is occulted by the star.

2.2   Phasecurve Models

In our study we tested two models on previously published phasecurves (to be shown later). The first model we tested was a model developed by Serrano (2018), and the other model is a model developed by Faigler and Mazeh (2011). Both models take into account the three physical phenomena that produce the small variations known as phasecurves. The beaming effect, the ellipsoidal variations, and the planetary light.

2.2.1   Serrano Model

The first model, developed by Serrano (2018), is a model that computes the three modulations based on calculations done by Lillo-Box (2014). The phasecurve normalized flux is given by the following formula:

$$\frac{F_{pc}}{F_*} = \frac{F_{ellip}}{F_*} + \frac{F_{beam}}{F_*} + \frac{F_{refl}}{F_*} \tag{2.1}$$

Planetary Light

The first factor to take into account is the light coming from the planet. The light could be either reflected from the star or emitted by the planet. In their work, they described the planet as a Lambertian sphere as it was done in Esteves et al. (2013).

This means that the planet is assumed to be a perfect sphere with an atmosphere reflecting the stellar flux isotropically. The modulation of flux from the planetary light effects if modelled as follows:

$$\frac{F_p}{F_*} = A_g \left(\frac{R_p}{r}\right)^2 \sin i \cos \theta \qquad (2.2)$$

where A_g is the geometric albedo, r is the distance between the planet and the star, and θ is the angle between line of sight and star-planet direction. As mentioned earlier, these modulation are small scaled and require high precision to detect. The most detected phasecurves are the ones that correspond to planets that are close to their host stars. Such planets are tidally locked to their star, and the orbit remains circular. For this reason, from now on the distance r, between the planet and the star, will be considered as a, which the semi major axis. (Lillo-Box 2014)

In order to calculate the planetary light, the model needs 8 parameters as input:

- Planetary Radius "R_p"

- Orbital Period "P"

- Inclination "i"

- Semi-major axis "a"

- Eccentricity "e"

- Time of conjunction "t_0"

- Longitude of periastron "w"

- Geometric Albedo "A_g"

The Beaming Effect

The second factor to take into account is the beaming effect, also known as the doppler boosting. This phenomenom is a result of two effects. The first is the non-relativistic doppler boosting of the stellar light in the direction of the star's radial velocity, which is a function of the planet's mass and distance from its star (Esteves et al. 2015). The second one is a band-pass dependant effect. It is a periodic red/blue shift of the star's spectrum, which results in a periodic change of the measured brightness as part of the star's spectrum move in and out of the observed bandpass.(Barclay 2012) This effect is modeled by:

$$\frac{\Delta F_{beam}}{F_*} = (3 - \Gamma)\frac{K}{c}(\sin\theta + e\cos w) \tag{2.3}$$

Where Γ is the beaming factor that depends on the observed bandpass, K is the radial velocity semi-amplitude that depends on the planetary mass, and c is the speed of light.

In order to calculate the beaming effect, the model requires 8 parameters as input:

- Planetary Mass "M_p"

- Stellar Mass "M_*"

- Inclination "i"

- Effective temperature of the star "T_{eff}"

- Eccentricity "e"

- Time of conjunction "t_0"

- Longitude of periastron "w"

- Orbital Period "P"

The Ellipsoidal Effect

The third effect is due to ellipsoidal variations. These variations are periodic changes in the observed stellar light caused by tidal distortion. The star's visible surface area fluctuates in and out of the observer's view due to the gravitational tides between the planet and the star itself (Mislis and Hodgkin 2012).
This effect is modeled by:

$$\frac{\Delta F_{ellip}}{F_*} = -\alpha_e \frac{M_p}{M_*} \left(\frac{R_*}{a}\right)^3 \left(\frac{1 + e \cos \Psi}{1 - e}\right)^3 \sin^2(i) \cos 2\theta \tag{2.4}$$

Where α_e is a limb darkening coefficient, and Ψ is the true anomaly shown in figure 2.2 In order to calculate the ellipsoidal effect, the model requires 11 parameters as input:

- Planetary mass "M_P"

- Stellar mass "M_*"

- Stellar radius "R_*"

- Inclination "i"

- Semi-major axis "a"

- Eccentricity "e"

- Time of conjunction, "t_0"

- Longitude of periastron, "w"

- Orbital period "P"

- Limb darkening coefficient, "u"

- Gravitational darkening coefficient, "g"

Gravitational darkening is a phenomenom where a star rotates so rapidly that it has a detectably oblate spheroid shape. When a star is oblate, it has a larger radius at its equator than it does at its poles. As a result, the poles have a higher surface gravity, and thus higher temperature and brightness. Hence, the poles are gravity brightened, and the equator is gravity darkened.
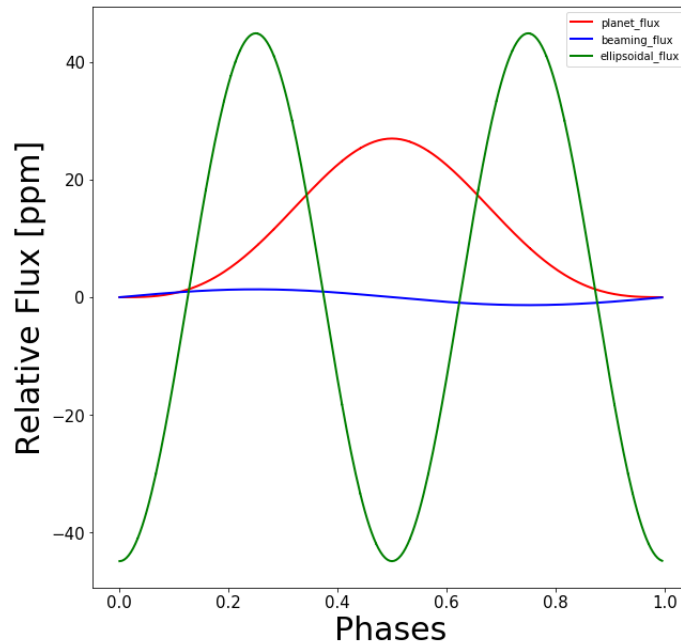
Figure 2.3: Representions of the three physical effects using Serrano's model over one ortbital period for a 4 M_J planet orbiting close to a sunlike star at a period of 1 day. Phase 0.5 represents the part where the planet is behind the star with respect to the observer. We note that ppm is a unit used for small variations. It stands for parts per million. (The actual value should be divdied by a million)

This model was previously created as a mean to simulate planetary signal. It was never applied on actual data, and never tested. In our work we aim to test this model, and try to set its limitations.

2.2.2 Faigler & Mazeh's Model

The other model we tested is a model developed by Faigler and Mazeh (2011). They approximated the ellipsoidal, beaming, and reflection effects by pure sine/cosine functions relative to the midtime of the first detected transit (time of conjuction Faigler and Mazeh 2011).

The Elliposidal Effect

$$\frac{\Delta F_{ellip}(t)}{F_*} = -A_{ellip} \cos\left(\frac{2\pi}{P_{orb}/2}t\right) \tag{2.5}$$

The ellipsoidal amplitude, A_{ellip}, is theoretically approximated to be:

$$A_{ellip} = \alpha_{ellip}\frac{m_p \sin i}{M_*}\left(\frac{R_*}{a}\right)^3 \sin i = 128\alpha_{ellip}\sin i \left(\frac{R_*}{R_\odot}\right)^3 \left(\frac{M_*}{M_\odot}\right)^{-2}\left(\frac{P_{orb}}{1day}\right)^{-2}\left(\frac{m_2 \sin i}{10M_J}\right) ppm \tag{2.6}$$

The ellipsoidal coefficient, $\alpha_{ellipsoidal}$, is theoretically approximated to be:

$$\alpha_{ellip} = \frac{0.15(15+u)(1+g)}{3-u} \tag{2.7}$$

Where u is the limb darkening coefficient and g is the gravitational darkening coefficient.

The Beaming Effect

$$\frac{\Delta F_{beam}(t)}{F_*} = A_{beam} \sin\left(\frac{2\pi}{P_{orb}}t\right) \tag{2.8}$$

The beaming amplitude, A_{beam}, is theoretically approximated to be:

$$A_{beam} = \alpha_{beam}\frac{K_{RV}}{c} = 27\alpha_{beam}\left(\frac{M_*}{M_\odot}\right)^{-2/3}\left(\frac{P_{orb}}{1day}\right)^{-1/3}\left(\frac{m_2 \sin i}{10M_J}\right) ppm \tag{2.9}$$

$\frac{4K_{RV}}{c}$ is the factor that takes into account the beaming effect for bolometric photometric obersvations, and α_{beam} takes into account the doppler shift which appears

when made in a specific bandpass. For F,G, and K soectral type stars observed by Kepler, α_{beam} ranges between 0.8 and 1.2.

Planetary Light

$$\frac{\Delta F_{refl}(t)}{F_*} = -A_{ref}\cos\left(\frac{2\pi}{P_{orb}}t\right) \tag{2.10}$$

A_{refl} is theoretically approximated to be:

$$A_{refl} = \alpha_{refl}0.1\left(\frac{r_p}{a}\right)^2\sin i \tag{2.11}$$

The reflection coefficient, α_{refl}, is theoretically approximated to be

$$A_{refl} = \rho_{geo}\left(\frac{Rp}{a}\right)^2\sin i \tag{2.12}$$

Where ρ_{geo} is the geometric albedo.

Final Phasecurve

The resulting phasecurve flux from this model is given by the following formula:

$$F_{pc} = a_{1c}\cos\left(\frac{2\pi t}{P_{orb}}\right) + a_{1s}\sin\left(\frac{2\pi t}{P_{orb}}\right) + a_{2c}\cos\left(\frac{2\pi t}{P_{orb}/2}\right) \tag{2.13}$$

Where:

$$a_{1c} = -A_{ref}$$

$$a_{1s} = A_{beam}$$

$$a_{2c} = -A_{ellip}$$

So in this model, the input parameters are 5: Orbital period, time of conjuction, and the three coefficients.
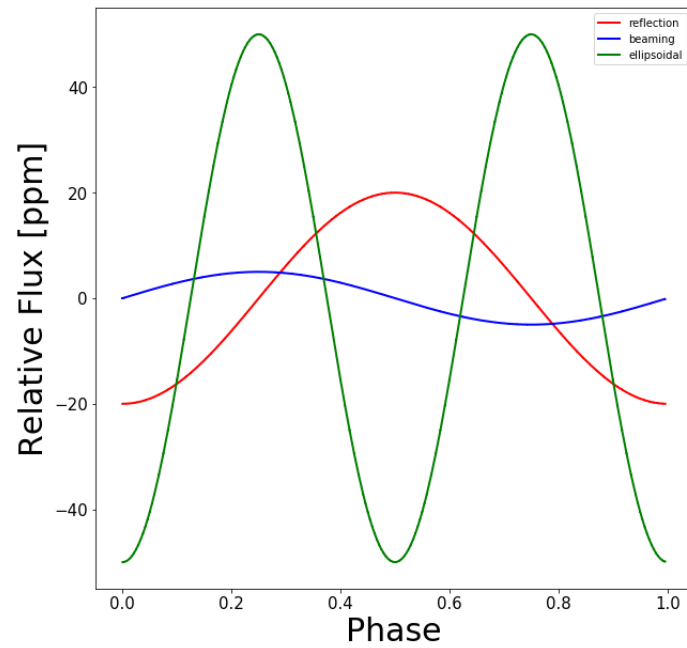
Figure 2.4: Representions of the three physical effects using the second model.

On the contrary to the first model, this model has been used before on actual data and it was able to detect signals comming from planetary phasecurves. ((Močnik et al. 2018), (Faigler and Mazeh 2011),(Shporer 2011)

CHAPTER 3

Data Reduction and Removing Stellar Activity

With the physics and models explained, in this chapter we will go through the steps of our work. As mentioned earlier, we aim to detect phasecurves modulation in the K2 data. The first step in this work is to test the models we have. To do this, we need try to fit the models to the light curves of 2 planets, that have a published phasecurve. Planet K2-141 (Malavolta 2018) and wasp 104 (Močnik et al. 2018). We then decide to stick to the model that gives result consistent with the published ones.

3.1   Data Reduction

In this work we use the data provided by the polar pipeline (Barros et al. 2016). In their network they download the calibrated pixel data (pixel files) from the Mikulski Archive for Space Telescopes (MAST). The first step of the polar pipeline is to produce the raw light curve. This is done by optimizing an aperture on the selected target, calculating the flux inside the aperture, and removing the background flux. The next step is to correct for the spacecraft's pointing stability. Owing to the particular pointing stabilisation mechanism of K2, the satellite slowly rolls around its line of sight and to correct for this, every six hours, the thrusters are fired returning the spacecraft close to its initial orientation. For each roll of the spacecraft, the target crosses a similar path on the CCD. But since pixels' sensitivity differ, they use the self-flat-fielding method, which calibrates the sensitivity variations with respect to the centroid position of the target by calculating the mean flux at each of a series of centroid position bins. Then the flux can be corrected from those sensitivity changes.
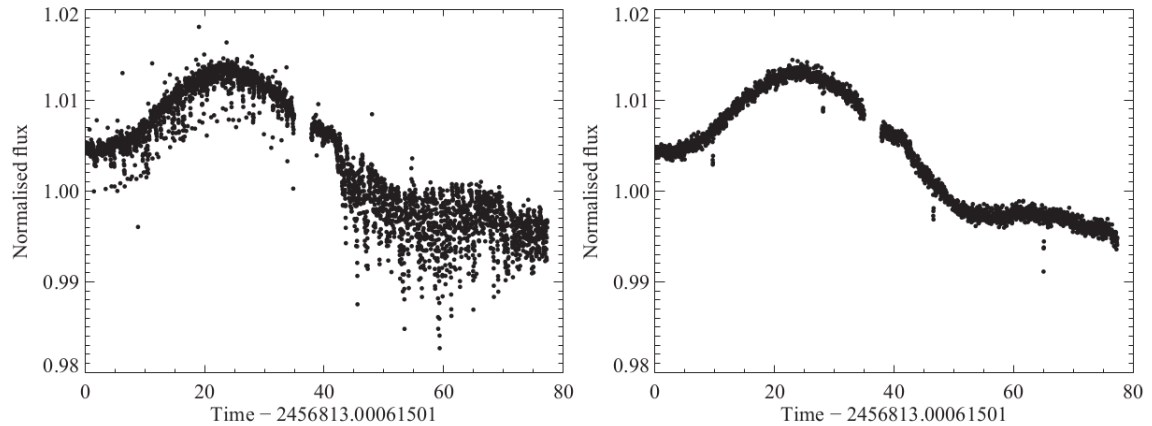
Figure 3.1: Left panel: the large systematic noise is clearly visible on the raw light curve of EPIC 201465501, a 14.95 magnitude star in Campaign 1. Right panel: same light curve after being position-decorrelated.(Barros et al. 2016)

3.2   Removing Stellar Activity

A variable star is a star whose brightness as seen from Earth fluctuates. When studying light curves of expoplanets, we have to take into account the variability that affects the photometric measures. The most frequent phenomenas that modify the transitting light curves are stellar rotation and star spots. These two effeccts are mostly present simultaneously.

The surface of the star is not uniformly bright, but has darker and brighter areas (like the sun's solar spots). Spots are regions of reduced surface temperature caused by concentrations of magnetic field flux that inhibit convection. Stars with sizeable sunspots may show significant variations in brightness as they rotate, and brighter areas of the surface are brought into view. The star's chromosphere may

also vary in brightness. As the star rotates we observe brightness variations of a few tenths of magnitudes. Stellar activity can prevent us from identifying the phasecurve variations and therefore prevent a possible characterization of the planet. The local brightness on the surface can either be decreased by dark spots or increased by bright faculae compared to the undisturbed photosphere. Spots (or faculae) located within the eclipsed section of the stellar surface lead to a decrease (increase) in the transit depth, and the true profile depends on the distribution of those structures across the planetary path. The overlap of a transiting planet and a stellar spot, for instance, can produce anomalies in the transit light curve that may lead to inaccurate estimation of the transit duration, depth, and timing.

Several methods are available to correct for stellar activities. In our work we use an easy and fast way by using the "spline function" provided by Python. This function fits a spline y = spl(x), that is a polynomial of degree k to the provided x, y data, and t which is internal knots of the spline. The knots are points in our time array. This fuction takes the provided knots, and slices the time array into intervals. The interval size is determined by the knots we provided. It then fits a polynomial of degree k to the data that lies in this interval.

Removing outliers from the baseline

In order to have a function that best describes the stellar variability, we need to remove any data points that don't correspond to the system, i.e., outliers. We have to be careful though, not to mark the transit points as outliers. To do so, we first mask the transit points. This is easy since we know the initial transit time, the period, and the transit duration. Python provides a simple masking function once these 3 parameters are known. Once the transit points are masked, we fit a spline function to the flux, and we normalize the flux to 1 by dividing it by the spline function. The

result will be a normalized baseline only, since the transit points are masked. We then perform "sigma clipping" on that baseline to identify the indices of the outliers. This function works by truncating the time array into intervals, calculating the mean flux in this interval, and then according to the threshold we choose, it will identify points that are lower or higher than this threshold. We choose a threshold of 3σ. The function then returns the indices of the outliers. These indices are then used to remove completely the outliers from the original light curve. This step is demonstrated in figures 3.3 and 3.4

Removing outliers from the transit

The previous step only removes outliers that are in the baseline. However, some outliers may be present even in the transit. To identify these outliers, we first phasefold the light curve. When we phasefold a lightcurve, we time shift all the data to a reference time (usually the initial detected time of transit), in order to mimic the scenario in which the observations were made in one period (meaning we fold the lightcurve onto itself for every orbital period). After that, we perform a "savitzky golay" filter on the phasefolded lightcurve. This filter is similar to the spline function, where it fits a polynomial to the data. However this filter acts on small time intervals and filters out small fluctuations while keeping the overall trend of curve. We only apply this to detect outliers in the transiting part of the lightcurve, as we do not want to remove the fluctuations that arise from the phasecurve. Hence, the filter is applied on the phasefolded lightcurve. The phasefolded lightcurve is then normalized by the filter, and another baseline is formed. Again sigma clipping is done to remove points that are above or below the mean by a certain threshold. We increase the threshold here up to 7σ as we do not want to remove points critical to the transit. Most of the outliers will be at the time of the transit, as we have already removed the non transitting outliers in the first step. The outliers' indices are obtained, and then the

data corresponding to those indices are also removed from the original light curve. This step is demonstrated in figures 3.5 and 3.6.

Final lightcurve

After removing all outliers, a spline function is applied to the data with outliers removed. This will result in a function that contains the trend of the flux with respect to time. To remove stellar activity, we normalize the flux by the obtained spline function. This is a typical smoothing technique. Few problems rise with this procedure however.

One of the problems we have to avoid is not to remove the transit points. So the spline fitting should be done on all points except the the transitting ones. To do this, we first remove all the points that are present in the transit. So we now have the data without any transitting points. The spline function is then applied to that data (data with no transits). As seen in figure 3.3 the spline fit ignores the transit points. The obtained fit is then used on the original data (with transit) to normalize. So the obtained normalized flux's formula is something like this:

$$F_{normalized} = \frac{F_{data-without-outliers}}{Spline_{without-transit}}$$

The other problem we have to avoid is over smoothing. In the spline function, if we choose the knots with a very short time interval, we will also be smoothing any variations comming from the phasecurve itself. To avoid this, we take the knots to be seperated by a few times the planet's orbital period. This will allow the function to smooth out patterns with a longer period (comming from the star) and keep variations that are on a shorter period (comming from the planet).

Removing stellar activity is the first step we do to test which model we will be choosing. We download the data of the 2 planets, K2-141 and wasp 104 (planets with published phasecurves from the K2 data) from the polar pipline, and we start the data reduction process discussed in this section on the data of each system.

3.3   Wasp 104-b

WASP-104 was observed by K2 during the observing Campaign 14, which covered a time-span of 80 days between 2017 June 1 and 2017 August 19. In the discovery paper Smith (2014), were able to constrain the stellar parameters of wasp 104 using photometric and spectroscopic analysis. The parameters are shown in Table 3.1

M_* ($M_\odot$)	1.011 $\pm$ 0.05
R_* ($R_\odot$)	0.940 $\pm$ 0.016
ρ_* ($\rho_\odot$)	1.2178 $\pm$ 0.007
T_{eff} (K)	5450 $\pm$ 130
Isochronal age estimate (Gyr)	3.5 $\pm$ 2.4

Table 3.1: Published stellar parameters for Wasp 104 (Smith 2014)

Later on the discovery of the phasecurve of Wasp 104-b (Močnik et al. 2018) yielded its planetary parameters which are presented in Table 3.2.
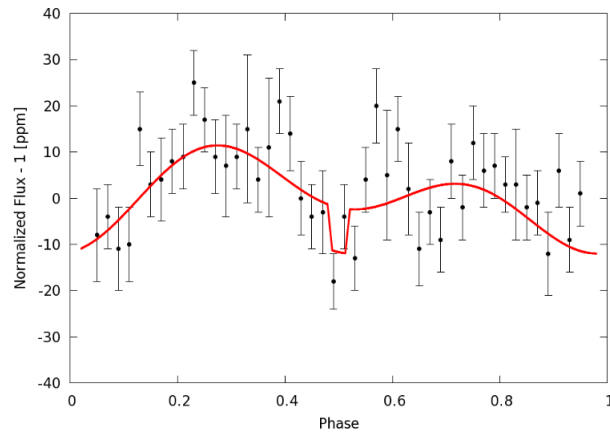
Figure 3.2: Published phasecurve of wasp 104-b.

t_0 (JD)	$57935.0702321 \pm 0.0000086$
per (days)	$1.75540636 \pm 0.00000014$
i (degrees)	83.612 ± 0.026
e (R_J)	0 (fixed)
a (au)	0.0286 ± 0.00047
M_p (M_J)	1.311 ± 0.053
R_p (R_J)	1.106 ± 0.019
Tp (K)	1.2178 ± 0.007

Table 3.2: Published planetary parameters for Wasp 104-b (Močnik et al. 2018)

In order to perform data reduction, remove outliers, and correct for stellar activity, we use the published orbital period and the time of the first detected transit (t_0). The procedure done is presented in the following figures.
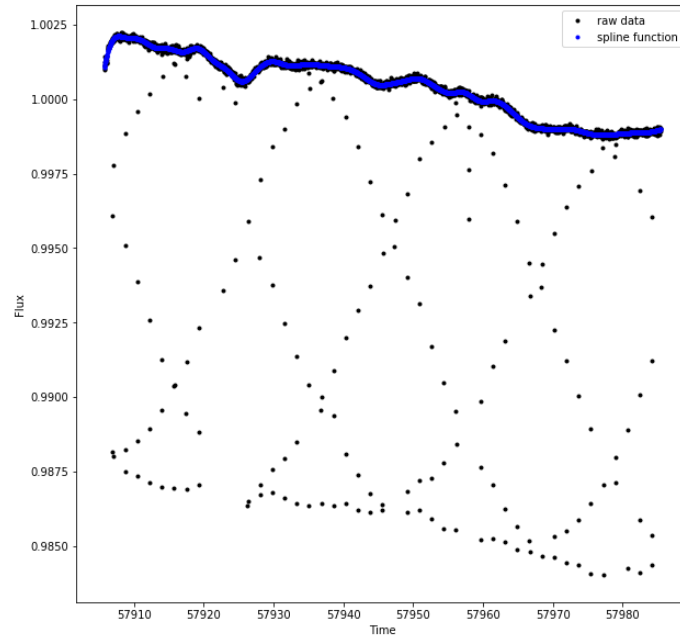
Figure 3.3: Spline Function fitted to the data of WASP 104 with transit points ignored.
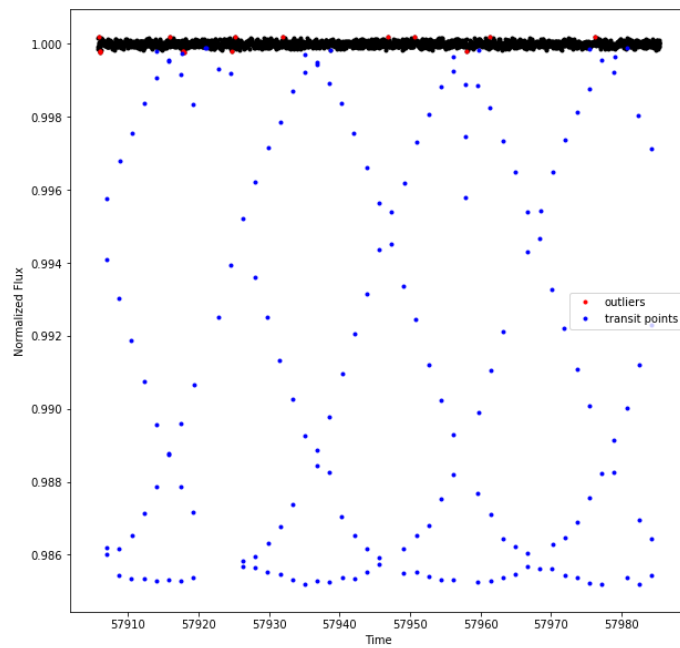
Figure 3.4: 3σ outliers in the normalized flux of wasp 104-b.

Figure 3.5: Outliers detected inside the transit in the phasefolded light curve of wasp 104-b
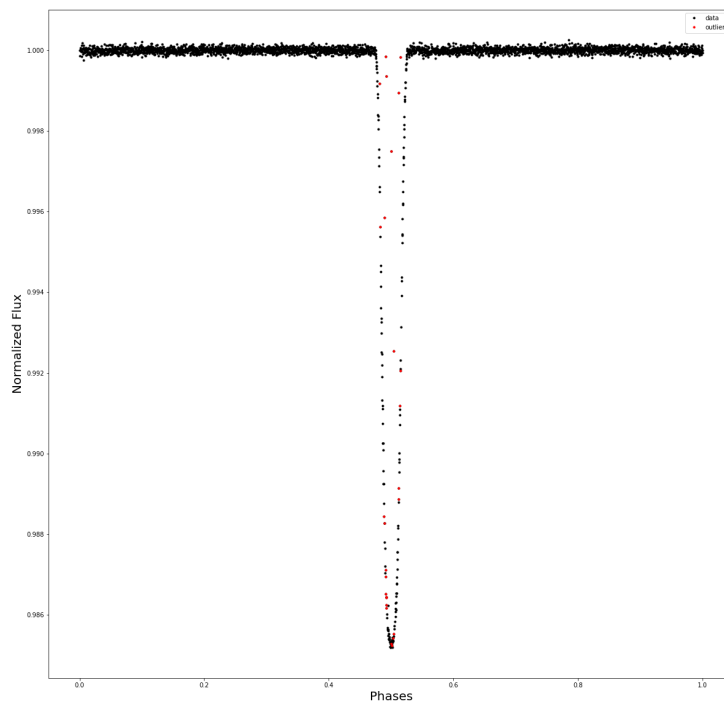
Figure 3.6: Final phasefolded lightcurve of wasp 104-b
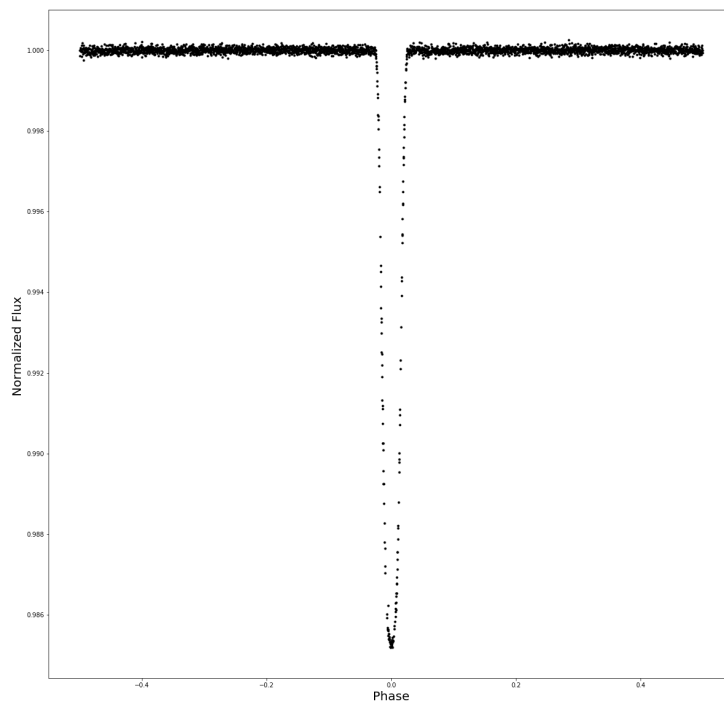
3.4   K2 141-b

K2-141 was observed by K2 for about 80 days between December 15 2016 and March 4 2017, with a loss of 5.3 days of data due to a safe mode state. In the discovery paper (Malavolta 2018), they were able to identify strong planetary signal at a period of only 6.7 hours (planet b), and a second planetary candidate signal with a period of 7.75 days (planet c). Their spectroscopic analysis reulted in obtaining the stellar parameters.

$M_* \ (M_\odot)$	0.708 ± 0.028
$R_* \ (R_\odot)$	0.681 ± 0.018
$\rho_* \ (\rho_\odot)$	2.244 ± 0.161
T_{eff} (K)	4599 ± 79
Age(Gyr)	6.3 ± 5

Table 3.3: Published stellar parameters for K2 141-b.

After removing the stellar activity signal from the K2 light curve and phase-folding the data to the orbital period of K2-141b, they also detected the signal of the secondary eclipse of this planet, and its phasecurve. This allowed them to obtain the planetary parameters of K2 141-b.
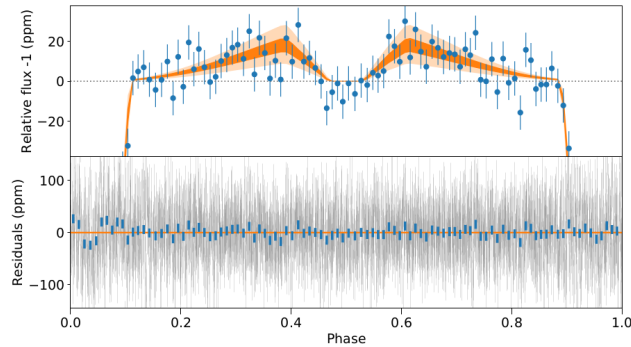
Figure 3.7: Published phasecurve of K2 141-b

t_0 (JD)	5774.0716 ± 0.0000022
per (days)	0.2803244 ± 0.000015
i (degrees)	86.3 ± 3
e (R_J)	0 (fixed)
a (R_*)	2.292 ± 0.05
M_p ($M_\oplus$)	5.08 ± 0.41
R_p ($R_\oplus$)	0.02037 ± 0.0046

Table 3.4: Published planetary parameters for K2 141-b.

The phasecurve shows modulation that results from planetary light. That is either reflection of the stellar light from the planet's surface, or the planet's thermal emission. The result of their model fit showed that this effect can be produced either by taking a geometric albedo (reflection effect) of 0.3 ± 0.06, or by thermal emission if the planet has a bond albedo of 0.37. This bond albedo corresponds to a temperature of 2100 K. This temperature doesnt belong to Kepler's observing band, but it is detected in the Infrared. So, in order to know for sure which effect is responsible for

this modulation, further observation should be done in the infrared to constrain one of the degenerate effects.

The same data reduction and stellar removal activity procedure was done on K2 141-b as it was done on wasp 104-b. The only thing we did in addition to the steps mentioned earlier, is that we masked points that correspond to the transit of planet c. Unlike wasp 104, K2 141 shows significant stellar activity. This made us face many problems in recovering the phasecurve later on. (to be shown later)
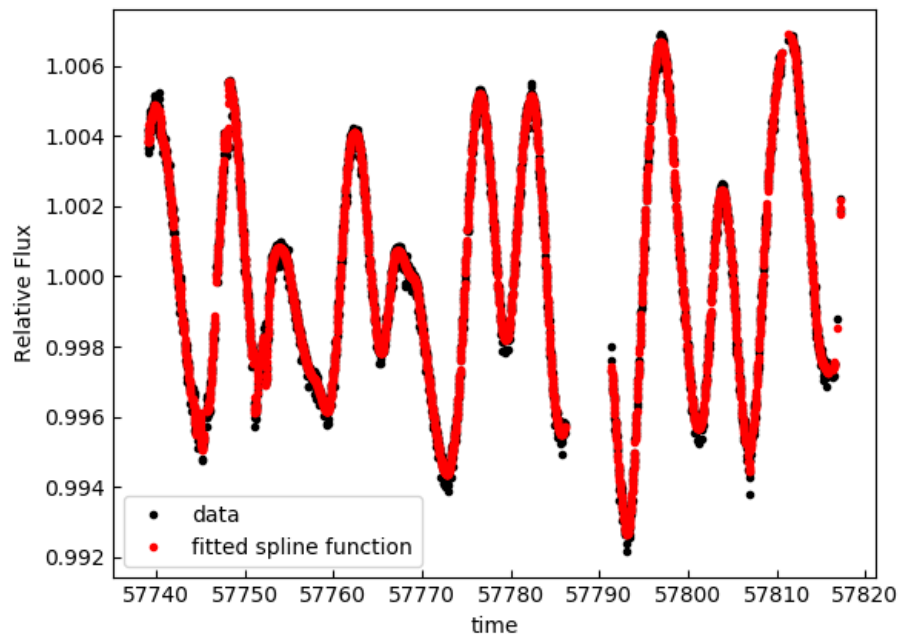
Figure 3.8: Spline functions fitted to K2 141 data.

As we can see in Figure 3.8, the gap corresponds to 5.3 days, where the spacecraft was in a safe mode state. The spline function we defined here was split into two parts, the first being to the data before the gap, and the second to the data after the gap. Each part of the data was then normalized with its corresponding spline function.
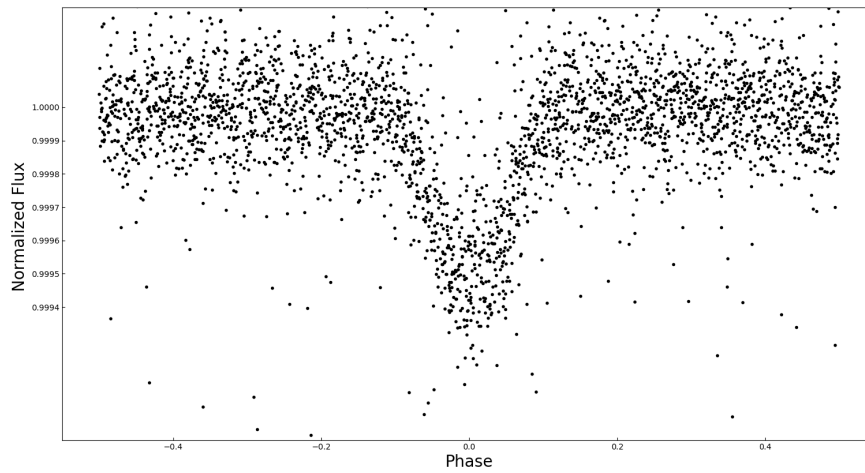
Figure 3.9: Obtained phasefolded lightcurve of K2 141-b after all the data reduction.

3.5   Fitting Technique

In choosing a method to fit our phase curve model to photometry, we need to consider several factors. Chief among these is how to cope with parameter degeneracy, which occurs when multiple parameters have a similar effect on the resulting model. For instance increasing the radius of the planet R p or increasing the day-side planet temperature T day both serve to increase the amplitude of the thermal contribution to the systems light curve. Temperature and albedo are degenerate parameters, as they can be used interchangeably to produce the same effect. The difference between a bright cool planet, and a dark warm planet, is difficult to determine.

In such a case, a fitting routine may have difficulty deducing the correct pair of parameter values. Instead, the routine may return one parameter value inflated and the other depressed, a combination that may give an overall thermal amplitude similar to that of the correct values. The addition of transit analysis to our assessment gives us another probe for the radius value, thereby eliminating the mentioned degeneracy, Without this additional transit constraint,planetary albedo, temperature, and radius would be triple degenerate parameters. Perhaps a convenient way of constraining this degeneracy would be to calculate a minimum expected temperature based on flux received from the star at the planets orbital distance. The temperature could potentially be warmer than this, assuming the atmosphere traps heat, but would not go beyond this lower bound. We will implement this constraint in future models.

An equally important issue is ensuring the fitting routine actually finds the best fitting set of parameters. Such algorithms aim to minimize the residuals, the absolute difference between a model and data. As with most optimization algorithms, fitting routines often times suffer from the tendency to return a best-fit model that actually corresponds to a local minimum in the residuals, rather than the global minimum. Once in a local minimum, many algorithms have a difficult time extracting them-

selves, since the surrounding residual gradient serves to keep them in the minimum. If the fitting routine cannot escape, it will output incorrect parameter values that correspond to a local minimum in residuals rather than a global minimum.

Monte Carlo Markov Chain Method

In an attempt to avoid these issues, we opt to implement a Markov chain Monte Carlo (MCMC) fitting routine (Crossfield (2015),Esteves et al. (2013)). MCMC populates a model parameter space with one or more walkers, which execute a random walk, evaluating the goodness of fit at each step. Theoretically, the walkers will converge to the best-fitting parameters. There are many implementations of the MCMC, varying in computational intensity and implementation of the random walk. In this work we use the Affine Invariant MCMC, which is an accessible tool in python that we used to fit the model into our data. This fitting procedure is further explained in Appendix A.

In the next chapters, we proceed our work by fitting the data of wasp 104-b and K2 131-b by the models of (Serrano 2018) and (Faigler and Mazeh 2011). We compare both models' results to the literature and proceed by finding more phasecurves by the appropriate model.

Chapter 4

Testing Models on Wasp 104b

In this chapter we present the results obtained from fitting the models of Serrano (2018) and the model of Faigler and Mazeh (2011) to the data of wasp 104. We compare our results to the values presented in the literature.

4.1 Serrano's Model

For the first model, the total number of parameters used for the lightcurve is 15, out of which we fix the stellar parameters that are previously published (R_*, M_*, T_{eff}). For planets with detected phasecurves, they are close in planets that are gravitationally locked to their host star, and thus remain in a circular orbit. So we fix the eccentricity to zero. We also fix the values of the limb and gravitational darkening coefficients using values from the tables provided by (Claret and Bloemen 2011). The fixed parameters and their corresponding values are presented in Table 4.1.

We set the prior distributions for the remaining parameters to be a normal distrution for the first detected time of transit (t_0) and the published orbital period. The median and the standard deviation of the distributions were chosen to be equal to the published values and their corresponding uncertainties. The prior distrbutions for the rest of the parameters are set to be uniform random distributions while taking into account the physical limitiations. The prior distributions for each parameter is presented in Table 4.2

The remaining free parameters, are 7: planetary radius, semi major axis, inclination, planetary mass, albedo, secondary eclipse depth, and time of first secondary eclipse detection. We choose the number of walkers to be 30 walkers , then the MCMC

is ran twice. The first run is done over 10,000 iterations. We disregard the first few iterations (15 % of the chain) as a burn in phase. In this phase, the walkers are randomly wandering through the parameter space and the "walks" they perform are quiet chaotic. After the first 10,000 iterations are done, we take the median value of each parameter, create a new initial normal distrubution around that median, with a standard deviation corresponding to the 68th percentile. The second run is then done over the new distrubiution for another 10,000 iterations. After the runs are done, the fitted values are taken to be the median value of every parameter's chain. The uncertainty of this value is the value of the 68th percentile,(1σ). The results of the fit are presented in Table 4.3

Fixed Parameters	Used Value	Source
eccentricity	0	adopted
R_*	0.94 $R_\odot$	(Smith 2014)
M_*	1.011 $M_\odot$	(Smith 2014)
T_{eff}	5450 K	(Smith 2014)
u_1 (linear limb dark. coef.)	0.6421	(Claret and Bloemen 2011)
y (gravitational dark. coef.)	0.4247	(Claret and Bloemen 2011)

Table 4.1: Fixed Parameters introduced to the MCMC fitting process of wasp 104.

Parameter	Prior Distribution
T_i	$\mathcal{N}(57935.0702321, 8.6 \times 10^{-6}$) Julian Days
Period	$\mathcal{N}(1.75540636, 1.4 \times 10^{-7}$) days
Planetary Radius	Random Uniform [0.1,0.3] R_*
Semi major axis	Random Uniform [5,7] R_*
inclination	Random uniform[80,90] degrees
Planetary Mass	Random Uniform [1,1.5] M_J
Geometric Albedo	Random Uniform[0,0.2]
Secondary Transit Depth	Random Uniform [0,100] ppm
Time of first secondary transit	$\mathcal{N}($ti+per/2, $\sigma_{t_i}+\sigma_{per})$

Table 4.2: Prior distributions used for parameters of wasp 104 in model 1.
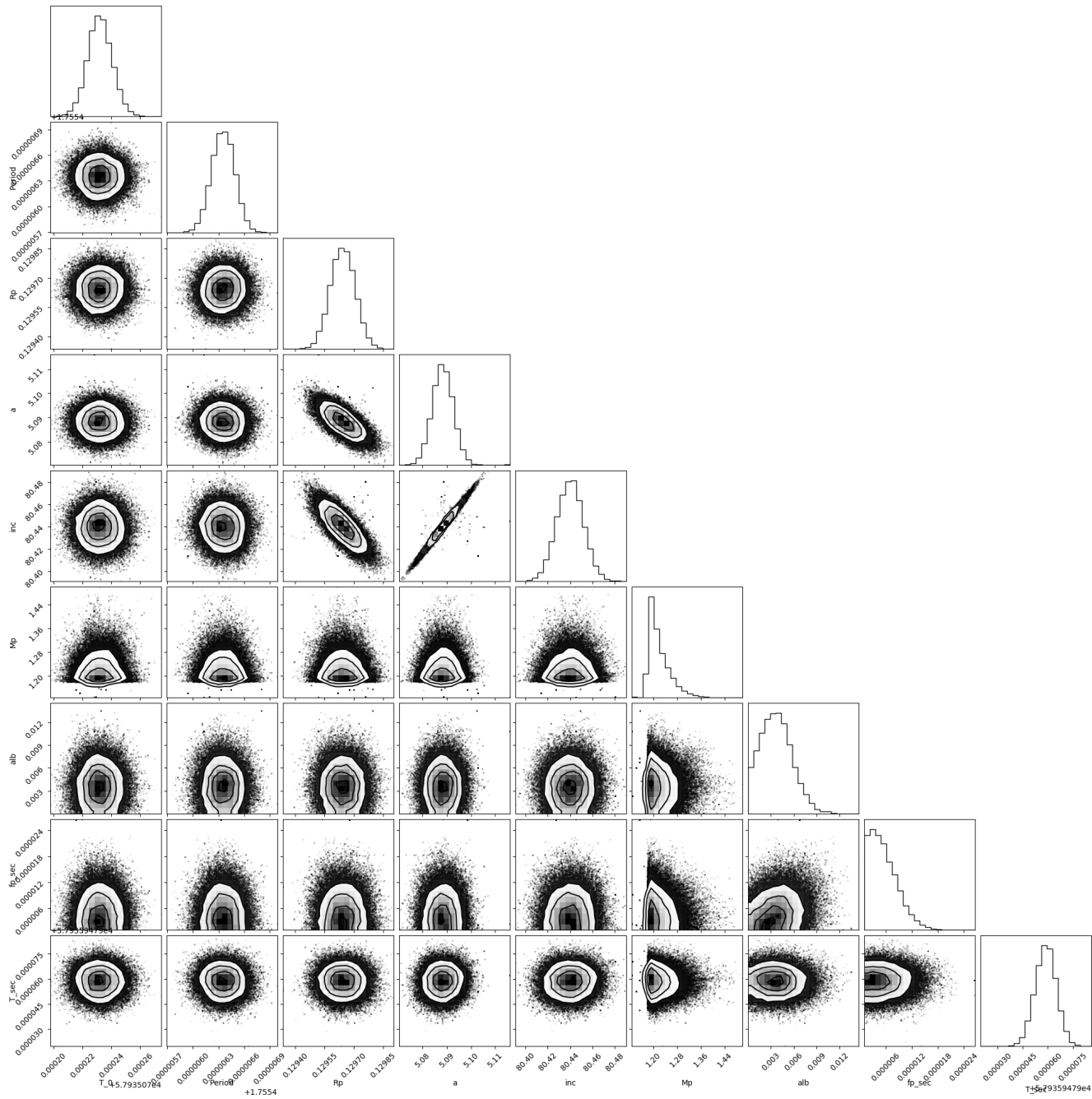
Figure 4.1: A triangle plot, showing the one-dimensional and two-dimensional posterior distributions for the MCMC run of model 1 to the data of wasp 104.

The fitted parameter are taken to be the median value of every parameter chain. As for the uncertainty,

Parameter	Fitted Value	Published Value
Intial time of transit (Julian Days)	57935.0702 ± 0.000017	$57935.0702321 \pm 0.0000086$
Orbital Period (days)	1.75540635 ± 0.0000003	$1.75540636 \pm 0.00000014$
a (au)	0.022 ± 0.00037	0.0286 ± 0.00047
inc (deg)	80.44 ± 0.023	83.612 ± 0.026
Rp (R_J)	1.21 ± 0.00125	1.106 ± 0.019
Mp (M_J)	1.21 ± 0.067	1.311 ± 0.053
Geometric Albedo	0.0036 ± 0.0041	Upper limit of 0.03

Table 4.3: Comparison between published values, and values obtained by fitting Serrano's model to wasp 104.
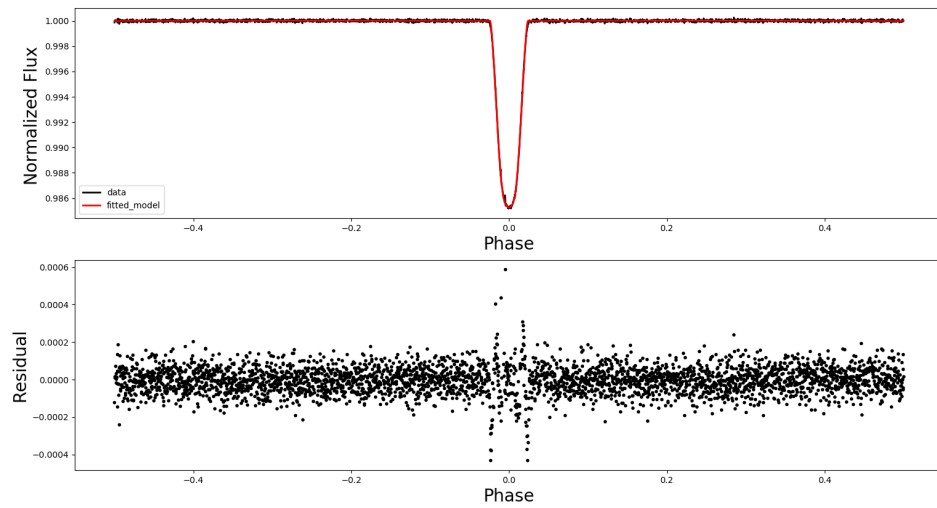
Figure 4.2: Phase-folded light curve of WASP-104. The red line is the fitting MCMC transit model. Shown in the lower panel are the residuals after fitting the transit model.
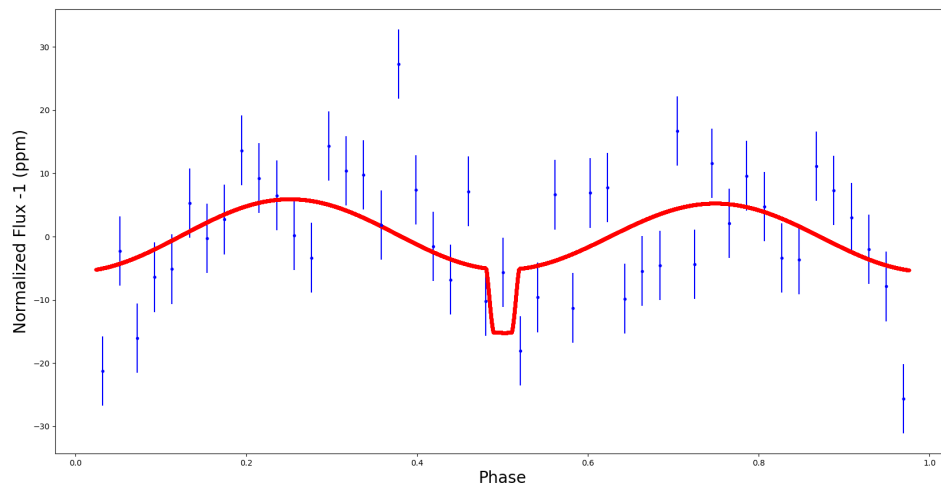
Figure 4.3: Phasefolded phasecurve of WASP-104, binned by a factor of 50. The red line is the fitting MCMC for the first phasecurve model.

4.1.1 Results

The fitting of the first model to the data of wasp 104 was successful as all the chains converged to one value (See FigureB.1). The convergence is reassured in the triangle plot(Figure 4.1, since all parameters have a peak in the posterior distribution around one value only, this indicates convergence and the absence of degeneracy for any parameter. Using the model in Serrano (2018) resulted in recovering the mass for wasp 104 b to be:

$$M_p = 1.21 \pm 0.067 M_J$$

. The obtained value is consistent with the published value within uncertainties. As for the geometric albedo, the fact that the uncertainty is greater than the value itself, this indicates that there was no significant detection of the reflection light comming from the planet, and thus the chain never converged. However, if we take the 98th percentile of the values taken by the walkers, we obtain an upper limit of 0.008. Močnik et al. (2018) obtained an upper limit of 0.03 for the geometric albedo. They argued that the emission phasecurve signal (from the thermal emission of the planet) would superimpose with reflection and increase its amplitude. Therefore, the true geometric albedo is likely to be significantly lower than the upper limit they recovered.

4.2 Faigler & Mazeh's Model

For the second model, the total number of parameters used for the lightcurve is 13, out of which we fix the eccentricity to 0. We also fix the values of the limb darkening coefficients using values from the tables provided by (Claret and Bloemen 2011). We set the initial distributions to be:

-Normal distribution for the orbital period and the time of the first transit.

-Uniform distribution for the remaining parameters:planetary radius, semi major axis, inclination, secondary eclipse depth, reflection coefficient(a_{1c}), beaming coefficient(a_{1s}),

and ellipsoidal coefficient (a_{2c}).

Also in this model the mcmc is ran twice. Each run is done over 10,000 iterations. Again, we choose the number of walkers to be 30 walkers. After the runs are done, the fitted values are taken to be the median value of every parameter's chain. The uncertainty of this value is the value of the 68th percentile.

Parameter	Prior Distribution
T_i	$\mathcal{N}(57935.0702321, 8.6 \times 10^{-6})$ Julian Days
Period	$\mathcal{N}(1.75540636, 1.4 \times 10^{-7})$ days
Planetary Radius	Random Uniform [0.1,0.3] R_*
Semi major axis	Random Uniform [5,7]R_*
inclination	Random Uniform [80,90] degrees
$F_{secondary}$	Random Uniform [0,100] ppm
reflection coefficient(a_{1c})	Random Uniform [0,100] ppm
beaming coefficient(a_{1s})	Random Uniform [0,100] ppm
ellipsoidal coefficient (a_{2c})	Random Uniform [0,100] ppm

Table 4.4: Prior distributions used for parameters of wasp 104 in the Mazeh & Faigler's model.
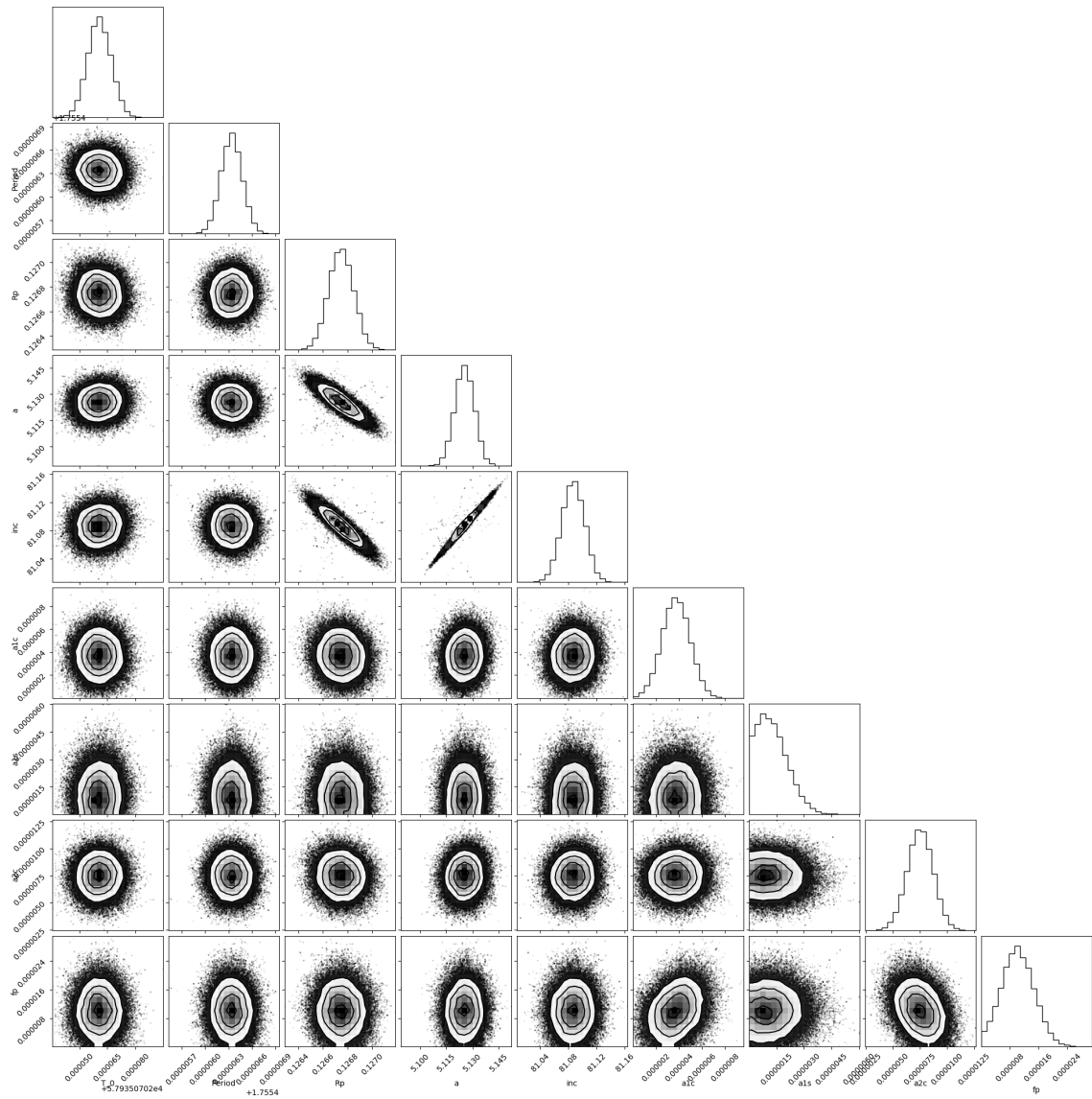
Figure 4.4: A triangle plot, showing the one-dimensional and two-dimensional posterior distributions for the MCMC run of model 2 to the data of wasp 104.

Parameter	Fittted Value	Published Value
A_{ellip} (ppm)	6.9 $\pm$2.2	7.48 $\pm$2.49
A_{beam} (ppm)	4.2$\pm$1.9	1.2 $\pm$1.76
A_{refl} (ppm)	4.8 $\pm$2.1	3.72 $\pm$2.43

Table 4.5: Comparison of obtained parameters for Wasp 104-b using model 2.
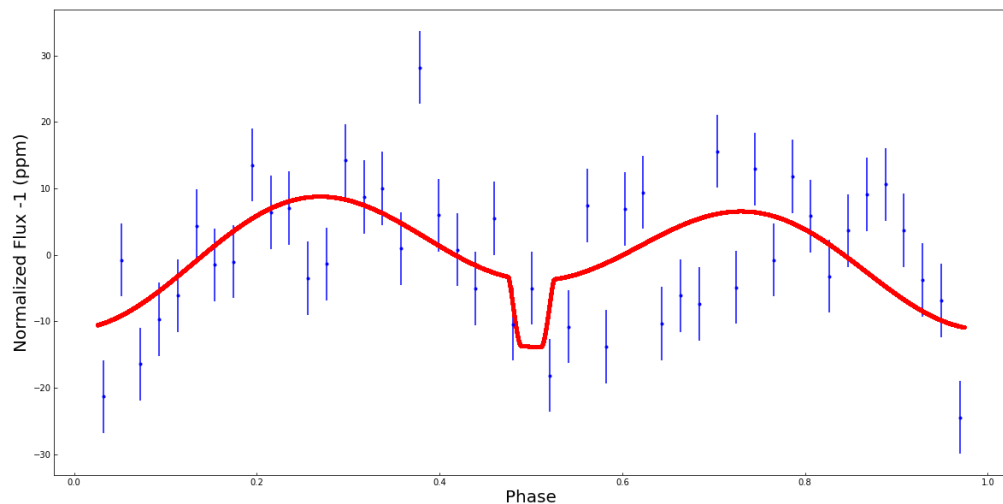
Figure 4.5: Binned phasefolded phasecurve of WASP-104. The red line is the fitting MCMC for the second phasecurve model.

4.2.1   Results

The fitting of the second model to the data of wasp 104 was also successful as all the chains converged to one value (figure(B.2). The convergence is reassured in the

triangle plot (figure 4.4), where no parameters showed degeneracy.

Using the second model we obtain an ellipsoidal amplitude of 7.48 ± 2.49 ppm, consistent with the published value within uncertainties. Using the limb and gravitational darkening coefficients from (Claret and Bloemen 2011), and substituting them in 2.7 we get $\alpha_{ellip} = 1.42$. Substituting this value along with the obtained elliptical ampitude in 2.6, and using the obtained orbital parameters from the MCMC fitting, we get a mass of:

$$M_p = 1.6 \pm 0.4 M_J$$

The obtained mass agrees with the published value ($1.311 \pm 0.053\ M_J$) within uncertainties. It also agrees, within uncertainties, with the mass obtained using the first model ($M_p = 1.21 \pm 0.067\ M_J$).

In this model the error on the beaming coefficient is greater than the value itself, indaicating that the effect is small and could not be constrained, and therefore an insignificant detection of this phenomenom.

As for the reflection coefficient, sunstituting it in Equation 2.11, and using the orbital parameters obtained from the MCMC fitting, we get a geometric albedo of:

$$Ag = 0.0062 \pm 0.004$$

As a conclusion, testing models 1 and 2 on wasp 104 gave consistent results. Both models were unabale to detect a clear reflection signal, but they were able to detect significant ellipsoidal modulations. They were also able to recover a planetary mass that is consistent with the published value. We proceed by further testing these models on another published phasecurve in the K2 data, it corresponds to the system of K2 141.

Chapter 5

Testing Models on K2-141 b

As seen in chapter 3, K2 141 is a very active star, the attempt of removing stellar activity by using a spline function without removing the planetary signal is tricky. In this chapter we present our attempts in fitting both models to the data of K2 141 after reducing the data, and removing the stellar activity.

5.1 Serrano's Model

Planet b rotates at a very short period, and at a very close distance to the star (Malavolta 2018), this close in planet is gravitationally locked to its host star, and thus remain in a circular orbit. Therefore we fix the eccentricity to 0. We also fix the stellar parameters that are previously published (R_*, M_*, T_{eff}). The values of the limb and gravitational darkening coefficients are fixed to the values obtained from the tables provided by (Claret and Bloemen 2011).

Fixed Parameters	Used Value	Source
eccentricity	0	adopted
R_*	0.681 $R_\odot$	(Malavolta 2018)
M_*	0.708 $M_\odot$	(Malavolta 2018)
T_{eff}	4599 K	(Malavolta 2018)
u_1 (linear limb dark. coef.)	0.693	(Claret and Bloemen 2011)
y (gravitational dark. coef.)	0.15	(Claret and Bloemen 2011)

Table 5.1: Fixed Parameters introduced to the MCMC fitting process of K2-141.

As for the rest of the parameters in this model, the prior distributions of each parameter is shown in Table5.2.

Parameter	Prior Distribution
T_i	$\mathcal{N}(5744.07160,\ 2.2 \times 10^{-4}$) Julian Days
Period	$\mathcal{N}(0.28,\ 1.5 \times 10^{-4}$) days
Planetary Radius	Random Uniform [0.1,0.2] R_J
Semi major axis	Random Uniform [1,5] R_*
Inclination	Random uniform[80,100] degrees
Planetary Mass	Random Uniform [0.001,0.1] M_J
Geometric Albedo	Random Uniform[0,0.5]
Secondary Transit Depth	Random Uniform [0,100] ppm
Time of first secondary transit	$\mathcal{N}(\text{ti+per}/2,\ \sigma_{t_i}+\sigma_{per})$

Table 5.2: Prior distributions used for parameters of K2 141 b in Serrano's model.

The obtained parameters of fitting Serrano's model to the reduced data of K2 141 are presented in Table 5.3.

Parameter	Fitted Value	Published Value
Intial time of transit (Julian Days)	57744.0716 ± 0.0004	57744.07160 ± 0.00022
Orbital Period (days)	0.28033 ±0.0000006	0.2803244 ± 0.0000015
a (R_*)	2.34420198 ± 0.009	2.292 ± 0.05
inc (deg)	90.027 ± 2.89	86.3 ± 3
Rp (R_*)	0.019 ± 0.00009	0.02 ± 0.00046
Mp (M_J)	0.075 ± 0.017	0.016 ± 0.0013
Geometric Albedo	0.006 ± 0.01	0.30 ± 0.06

Table 5.3: Comparison between published values, and values obtained by fitting model 1 to K2-141.

Using Serrano's phasecurve model on the data of K2 141 b, the planetary mass is obtained to be

$$M_p = 0.075 \pm 0.017 M_J$$

. This value does not agree with the published mass. The model also failed to detect a significant reflection modulation as was proposed by Malavolta (2018). This discrepancy raises two hypotheses. Either the model is unable to detect the phasecurve when the noise level is significant, or that the data processing did remove the phasecurve modulation. We try to resolve this issue by fitting the data of K2 141 by the model of Faigler and Mazeh (2011). In their work, they tested the model and were able to recover phasecurves of planets orbiting stars as active as K2 141 (K6222 and K8064). If we were able to recover a phasecurve using this model, we then know that the first model has limitations and that the data processing did not remove modulations that correspond to the phasecurve.
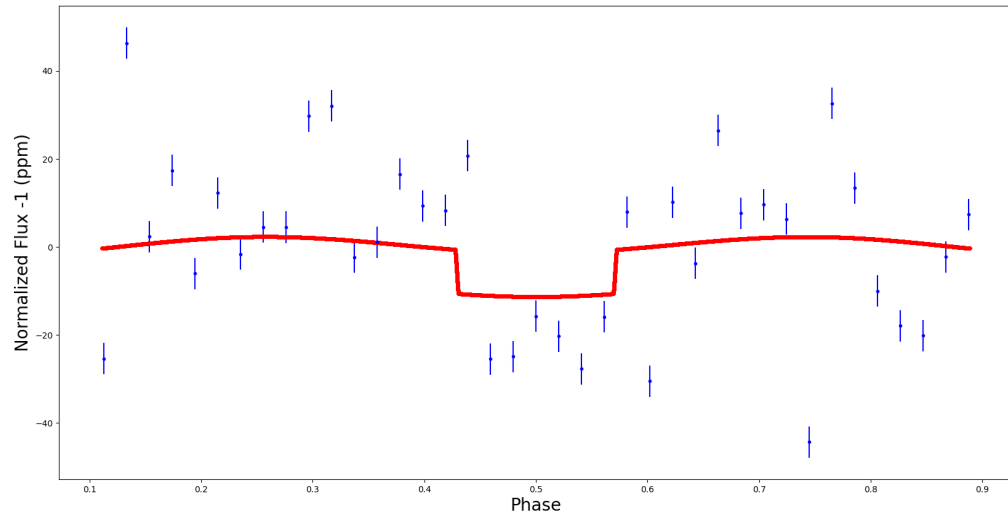
Figure 5.1: Binned phasefolded phasecurve of K2-141 b. The red line is the fitting MCMC for Serrano's phasecurve model.

One thing we didnt take into account in this specific case, is the low number of data points in the transiting phase. The exposure time of kepler spacecraft is 30 minutes per picture. This planet has a very short orbital period, the duration of the transit is around 1.5 hours (Malavolta 2018), this would mean we would have as few as 2 data points per transit. Over sampling should have been done on the data in order to achieve a better fit.

5.2 Faigler & Mazeh's Model

Our second attempt in detecting the phasecurve of K2 141 was done by fitting model 2 to the reduced data. (Faigler and Mazeh 2011) We also fix the same parameters as we did in model 1 (table 6.1). The obtained parameters are shown in the table below.

Parameter	Fitted Value
A_{ellip} (ppm)	5.8 ± 2.079
A_{beam} (ppm)	0.09 ± 0.1
A_{refl} (ppm)	26.44 ± 2.28

Table 5.4: Parameters obtained from the MCMC fitting of model 2 to the data of K2 141.

Similar analysis as in section 4.2.1 yields to a planetary mass of:

$$M_p = 0.047 \pm 0.016 M_J$$

The obtained mass agrees with the published value (0.016 ± 0.0013 M_J) within less than 2σ.

The reflection phenomenom is detected significantly with respect to the other two phenomenas (26.44 ppm). The analysis of the reflection amplitude yields to a geometric albedo of:

$$A_g = 0.4 \pm 0.03$$
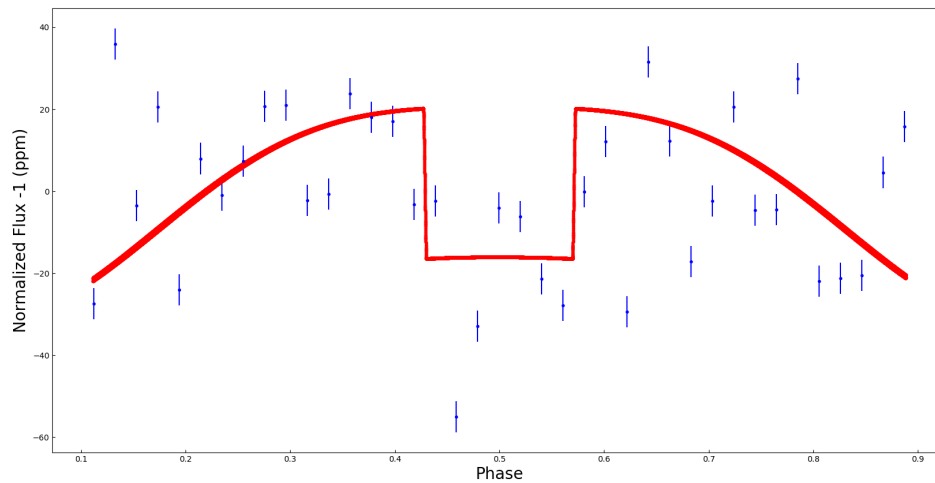
Figure 5.2: Binned Phasefolded phasecurve of K2-141 b. The red line is the best fitting MCMC for the second phasecurve model.

5.3 Analysis

Serrano's model failed to recover a mass that agrees with the published one, while Mazeh & Faigler's model found a mass that agrees with the published one within 2σ. This allows us to raise questions on the limitations of Serrano's model. Further testings should be done in order to test the model's effectiveness and limitations with different planetary masses. One could progress this work and find a lower limit for the detectable mass using this model.

The reflection phenomenom was succcessfully detected using Mazeh's model , when Serrano's model failed. This could indicate that model 1 fails to detect phase-curves when the data is as scattered as it is in the data of K2 141. One could also

proceed with testing Serrano's model and test its limitations in recovering phase-curves at different noise levels. Further studies should be done on this system. A different approach in removing stellar activity could be done, to get a better normalized lightcurve.

In the following chapter we proceed our work by searching through the data of K2, in the aim to find phasecurves that are not previously detected, and determine the planetary mass and geometric albedo.

Chapter 6

New Discoveries and Conclusions

In order to search for phasecurves in the K2 data, we head to The Extrasolar Planets Encyclopaedia (http://exoplanet.eu/). We filter the entries in order to choose confirmed planets from the K2 mission, detected via the transit method. We then scan through the entries and choose planets that have a short orbital periods and are in close orbits around their host star $(a/R_*) < 10$. Such planets are the ones that allow us to detect phasecurve modulations.(Esteves et al. 2013) We then head to the MAST archive (archive.stsci.edu/k2/search_retrieve.html) in order to check in which campaign was this planet observed. Once we have the planets name (known as EPIC ID) and the campaign in which it was observed in, we download the data that corresponds to that planet from the POLAR database at the center of astrophysics in university of Porto (CAUP). We download the file that has not been detrended (without the removal of stellar activity).

After retrieving the files, we start the data processing described in section 3.2 as an attempt to remove stellar activity. Once we have the normalized light curve, we phasefold the lightcurve onto itself using the published orbital period. We mask the transit, and bin the data and then visually inspect the phasefolded lightcurve to see if any modulation is present in the out of transit phases.

6.1   K2 31-b

One planet we came across to have a visually distinguishable phasecurve modulation is K2 31b. K2-31b is a short-period giant planet transiting a V=10.8 mag star at a period of 1.26 days. K2 31 was obseverd by K2 for about 80 consecutive days as part of Campaign 2 between August 23 and November 13 2014. (Grziwa 2016)
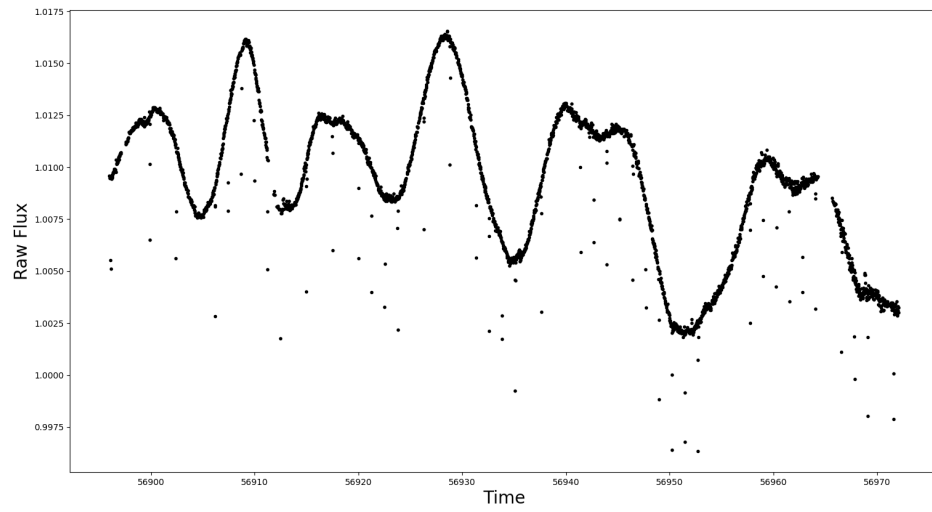
Figure 6.1: Raw Lightcurve of K2-31
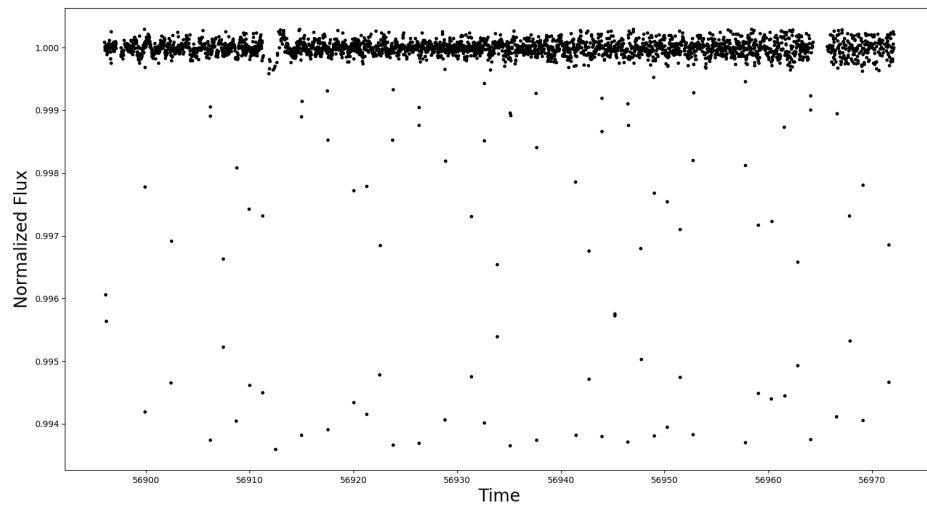
Figure 6.2: Normalized lightcurve of K2-31 after removing stellar activity and outliers.
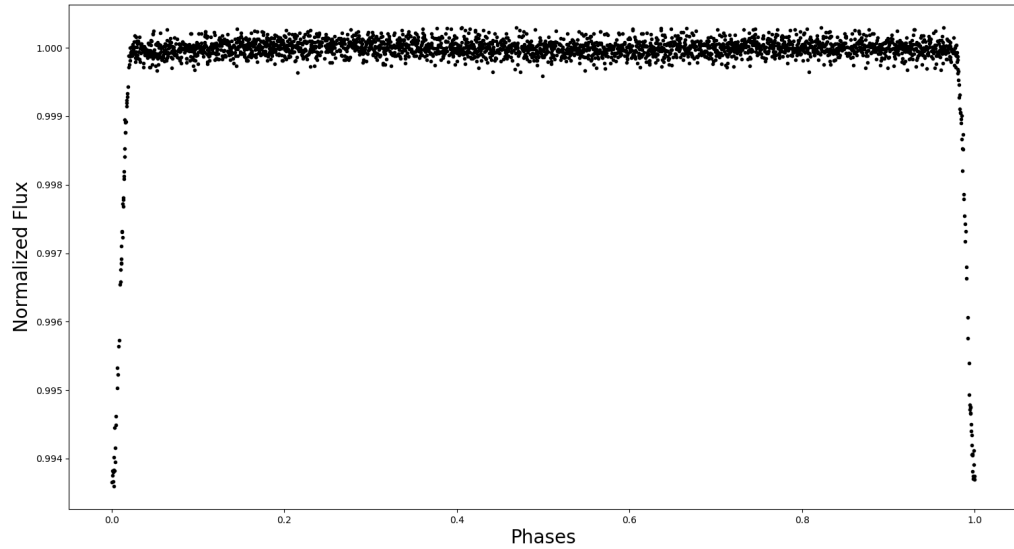
Figure 6.3: Phase folded lightcurve of K2 31-b.

The out of transit lightcurve seems to have some modulations. We preceed by fitting model 2 to the data of K2 31-b.

Fixed Parameters	Used Value	Source
eccentricity	0	adopted
R_*	0.78 $R_\odot$	(Grziwa 2016)
M_*	0.91 $M_\odot$	(Grziwa 2016)
T_{eff}	5280 K	(Grziwa 2016)
u_1 (linear limb dark. coef.)	0.514	(Claret and Bloemen 2011)

Table 6.1: Fixed Parameters introduced to the MCMC fitting process of K2-141.

Parameter	Prior Distribution
Period	$\mathcal{N}(1.257850 , 0.000002)$ days
Planetary Radius	Random Uniform $[0.01,0.5]$ R_*
Semi major axis	Random Uniform $[4,8]R_*$
inclination	Random Uniform $[60,90]$ degrees
reflection coefficient(a_{1c})	Random Uniform $[0,100]$ ppm
beaming coefficient(a_{1s})	Random Uniform $[0,100]$ ppm
ellipsoidal coefficient (a_{2c})	Random Uniform $[0,100]$ ppm

Table 6.2: Initial distributions used for MCMC fitting of model 2 to the data of K2 31 b.

Parameter	Fitted Value	Published Value
Orbital Period (days)	1.2578 ± 0.0000015	1.257850 ± 0.000002
a (R_*)	4.5402 ± 0.02	6.05 ± 0.5
inc (deg)	73.6 ± 0.9	79.9 ± 0.8
Rp (R_*)	0.3 ± 0.07	0.106 ± 0.035
Reflection Coefficient (ppm)	5.1 ± 1.49	-
Beaming Coefficient (ppm)	5.08 ± 1.4	-
Ellipsoidal Coefficient (ppm)	24.3 ± 1.5	-

Table 6.3: Comparison between published values, and obtained values from the MCMC fitting of Mazeh & Faigler's model to the data of K2-31 b.
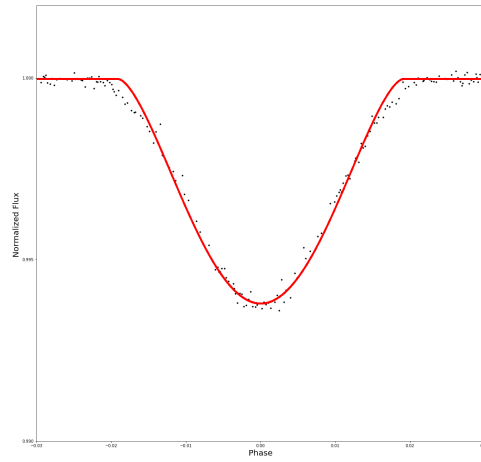
Figure 6.4: Phase-folded light curve of K2 31. The red line is the best-fitting MCMC transit model.
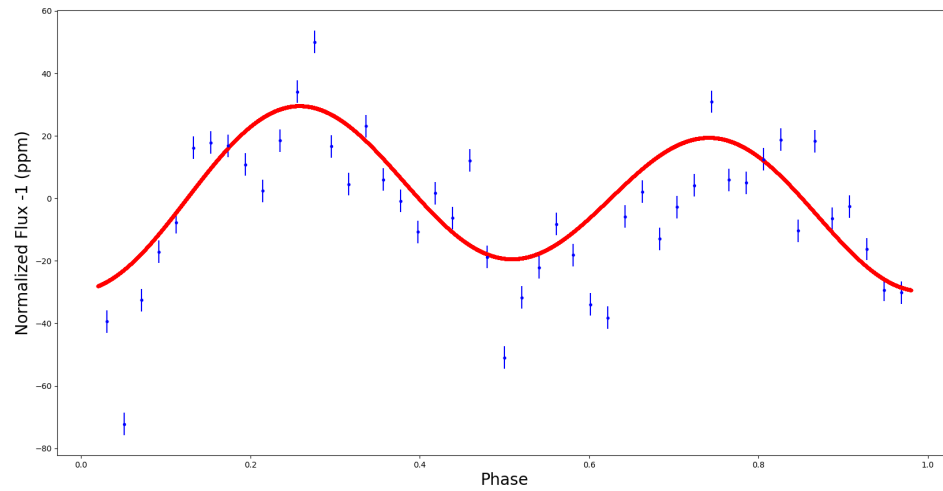

Figure 6.5: Binned out-of-transit lighcurve of K2-31b with the best fitting MCMC phasecurve model of Mazeh & Faigler.

The detection of the three physical contributions to the phasecurve is clearly successful. From them we try to recover the mass of K2-31b as well as its geometric albedo. Since the detections of both the beaming and the elliptical effects are successful, we can find the mass of the planet using both amplitudes independantly. From the ellipsoidal detection, we first calculate α_{ellip}. To do this we substitute the values of u and y presented in table (6.1) in equation (2.7).

$$\alpha_{ellip} = 1.37$$

Now that we have α_{ellip}, we substitute it in equation (2.6) to get a planetary mass of:

$$M_p = 2.02 \pm 0.18 M_J$$

To find the planetary mass from the beaming effect we use equation (2.9), where we substitute α_{beam} by 1, which corresponds to observations made by Kepler as proposed by (Faigler and Mazeh 2011).

$$M_p = 2.065 \pm 0.5 M_J$$

The published planetary mass obtained from the radial velocity study (Grziwa 2016) is 1.774 ± 0.079 M_J. The planetary mass obtained using both methods is consistent with the published values within uncertainties.

Next we use the detected reflection ampliude in an attempt to find the geometric albedo. This is done by substituting the reflection amplitude in equation (2.12) .

$$A_g = 0.001 \pm 0.0003$$

Our analysis yield that K2-31b has a mass of approximately $2M_J$ and a radius of $2.34R_J$, orbiting at a short distance from its host star (0.01 au). These properties are the properties of a hot Jupiter. Hot jupiters are tidally locked to their hosts stars and tend to have permanent day side and permanent night side. This means that

the day side is so hot clouds can't form, and clouds are typically very reflective, as Venus demonstrates. It's also too hot for ice the kind of surface that makes the moon Enceladus so bright. The fact that K2 31-b has a very low albedo means that it almost doesn't reflect much light in the visible band. It may indicate that it has a thick atmosphere which absorbs light in the visible spectrum, making the planet very dark on the day side. On the night side, away from the starlight, clouds may form, but that side never sees daylight, so there's no light nearby for it to reflect.(Demory and Seager 2011)

Another interesting property of this planet, is its low mean density. From its mass and radius, we can approximate its density to be 0.156 ρ_J. This could be explained by the location of the planet. Being so close to its star, it is receiving very intense levels of flux. The flux will heat up the planet and cause its atmosphere to inflate, resulting in a larger radius and lower density.

6.2   Conclusions

The goal behind this work is to detect phasecurve modulations in the data of K2. To do this we searched in the confirmed planets of K2, and chose the planets that are good candidates to produce a phasecurve (massive planets orbiting close to their stars). Once we selected those planets we had to go through several data processes, in order to remove stellar activity and normalize the light curves. Finally, we fit the light curves with models that best describe the physics behind the phasecurve, and see if we can detect any modulations. Detection of these modulations will allow us to determine the planetary mass, and estimate the geometric albedo of the planet. These parameters are important in finding the density of the planet, and having an idea of what its made up. Hence this method can achieve a good characterization of the planet, all done photometrically.

In this project we were able to test two different models that represents the three physical contributions to the phasecurve. The planetary reflection, the ellipsoidal modulations, as well as the beaming effect. The first model we tested, Serrano's model, was previously implemented only to simulate planetary signals. It was never tested or fitted on real data. Throughout our testing we showed that the model is able to represent the phasecurve modulations fairly well if the data is not affected by much noise (Wasp 104). The model fails to recover a phasecurve when the mass is relatively small, or when the data contains significant noise (K2-141). Further tests should be done in order to test the limitations of the model. For example one could assemble a data comming from different planetary masses, and set a lower limit on the mass that results in recovering the phasecurve. One could also sample different noise levels on a data set, and set a threshold on which the model can still recover the phasecurve

The second model we used, is a previously implemented model, and it has been used on actual data before. The fact that the second model was able to detect the phasecurves even when the data has some noise, assures us that the data processing was done correctly, and it didnt wash away any small scaled fluctuations. This model was able to recover both phasecurves that it was tested on. We procedeed our phasecurve hunting using this model.

Our search in the K2 data allowed us to discover a new phasecurve that hasn't been discovered before. The model recovered an amplitude for all three effets which allowed us to estimate the planet's mass to be twice that of Jupiters. The photometrically obtained mass matches with the previously published one, which was measured using the radial velocity method. This allows us to conclude that determining masses for some planets, to a relatively accurate value can be done photometrically with the already available data, without wasting time on the follow up spectorscopic observa-

tions.

We were also able to recover a small geometric albedo for K2-31 b which indicates its low reflectivity in the visible band. We note that previous spectroscopic studies showed that a big number of Hot Jupiters have an atmosphere that contains Sodium, Potassium, Titanium(II) oxide (TiO), and Vinadium(II) oxide (VO). The presence of these elements in the atmosphere of a planet will cause significant absorption in the optical wavelengths which will lead to a low geometric albedo. In order to test this hypothesis, we need a high resolution spectroscopic study to check for the presence of these absorption features in the atmosphere of K2 31-b. (Demory and Seager 2011)

As a final conclusion, this study offers a photometric way to detect planets around host stars even when they dont transit. The phasecurve modulations could be detected even if the planet does not transit between the star and the observer. It provides a mean to determine planetary masses and have an idea of what the planet's atmoshphere is made up of.

The next step in this project is to improve what we already have in order to continue. For example, different analysis on lightcurves of active stars should be done in order to remove the stellar activity. Several studies have used the Guassian Regression process in order to simulate a signal that corresponds to stellar activity.(Angus 2018) Gaussian process measures the similarity between points (the kernel function) to predict the value for an unseen point from the data. The prediction is not just an estimate for that point, but also has uncertainty information, which is a one-dimensional Gaussian distribution. In future work, we will try to simulate the stellar activity using guassian regression process, in an attempt to reduce the noise obtained when normalizing the data of active stars.

The work presented here is only a first step towards a bigger project. Our upcomming work consists of going through every possible candidate in the confirmed K2 planets and detect more phasecurves in an attempt to characterize more planetary system photometrically.

Bibliography

Angus, R. e. a. (2018). Inferring probabilistic stellar rotation periods using Gaussian processes. *MNRAS*, 474:2094–2108.

Barclay, e. a. (2012). Photometrically Derived Masses and Radii of the Planet and Star in the TrES-2 System. *ApJ*, 761:53.

Barros, S. C. C., Demangeon, O., and Deleuil, M. (2016). New planetary and eclipsing binary candidates from campaigns 1-6 of the K2 mission. *A&A*, 594:A100.

Charbonneau, D., Brown, T. M., Latham, D. W., and Mayor, M. (2000). Detection of Planetary Transits Across a Sun-like Star. *ApJ*, 529:L45–L48.

Claret, A. (2000). A new non-linear limb-darkening law for LTE stellar atmosphere models. *A&A*, 363:1081–1190.

Claret, A. and Bloemen, S. (2011). Gravity and limb-darkening coefficients for the Kepler, CoRoT, Spitzer, uvby, UBVRIJHK, and Sloan photometric systems. *A&A*, 529:A75.

Crossfield, I. J. M. a. (2015). A Nearby M Star with Three Transiting Super-Earths Discovered by K2. *ApJ*, 804:10.

Demory, B.-O. and Seager, S. (2011). The High Albedo of the Hot Jupiter Kepler-7 b. *ApJ*, 735:L12.

Diaz-Cordoves, J., Claret, A., and Gimenez, A. (1995). Linear and non-linear limb-darkening coefficients for LTE model atmospheres. *A&AS*, 110:329.

Esteves, L. J., De Mooij, E. J. W., and Jayawardhana, R. (2013). Optical Phase Curves of Kepler Exoplanets. *ApJ*, 772:51.

Esteves, L. J., De Mooij, E. J. W., and Jayawardhana, R. (2015). Changing Phases of Alien Worlds: Probing Atmospheres of Kepler Planets with High-precision Photometry. *ApJ*, 804:150.

Faigler, S. and Mazeh, T. (2011). Photometric detection of non-transiting short-period low-mass companions through the beaming, ellipsoidal and reflection effects in Kepler and CoRoT light curves. *MNRAS*, 415:3921–3928.

Fischer, D. A., Howard, A. W., Laughlin, G. P., Macintosh, B., Mahadevan, S., Sahlmann, J., and Yee, J. C. (2014). Exoplanet Detection Techniques. *Protostars and Planets VI*, pages 715–737.

Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J. (2013). emcee: The MCMC Hammer. *PASP*, 125:306.

Goodman, J. and Weare, J. (2010). Ensemble samplers with affine invariance. *Communications in Applied Mathematics and Computational Science, Vol. 5, No. 1, p. 65-80, 2010*, 5:65–80.

Grziwa, S. e. a. (2016). K2-31B, a Grazing Transiting Hot Jupiter on a 1.26-day Orbit around a Bright G7V Star. *AJ*, 152:132.

Kopal, Z. (1950). Detailed effects of limb darkening upon light and velocity curves of close binary systems. *Harvard College Observatory Circular*, 454:1–12.

Kreidberg, L. (2015). batman: BAsic Transit Model cAlculatioN in Python. *PASP*, 127:1161.

Lillo-Box, e. a. (2014). Kepler-91b: a planet at the end of its life. Planet and giant host star properties via light-curve variations. *A&A*, 562:A109.

López-Morales, M. and Seager, S. (2007). Thermal Emission from Transiting Very Hot Jupiters: Prospects for Ground-based Detection at Optical Wavelengths. *ApJ*, 667:L191–L194.

Malavolta, e. a. (2018). An Ultra-short Period Rocky Super-Earth with a Secondary Eclipse and a Neptune-like Companion around K2-141. *AJ*, 155:107.

Mislis, D. and Hodgkin, S. (2012). A massive exoplanet candidate around KOI-13: independent confirmation by ellipsoidal variations. *MNRAS*, 422:1512–1517.

Močnik, T., Hellier, C., and Southworth, J. (2018). WASP-104b is Darker Than Charcoal. *AJ*, 156:44.

Schwarzschild, K. and Villiger, W. (1906). On the Distribution of Brightness of the Ultra-Violet Light on the Sun's Disk. *ApJ*, 23:284.

Seager, S. and Mallén-Ornelas, G. (2003). A Unique Solution of Planet and Star Parameters from an Extrasolar Planet Transit Light Curve. *ApJ*, 585:1038–1055.

Serrano, L. M. e. a. (2018). Distinguishing the albedo of exoplanets from stellar activity. *A&A*, 611:A8.

Shporer, A. e. a. (2011). Detection of KOI-13.01 Using the Photometric Orbit. *AJ*, 142:195.

Smith, A. M. S. e. a. (2014). WASP-104b and WASP-106b: two transiting hot Jupiters in 1.75-day and 9.3-day orbits. *A&A*, 570:A64.

Wolszczan, A. and Frail, D. A. (1992). A planetary system around the millisecond pulsar PSR1257 + 12. *Nature*, 355:145–147.

APPENDIX A

Affine Invariant MCMC

In this section we further explain the implementation of the Affine Invariant MCMC. The affine-invariant (AI) MCMC (Goodman and Weare (2010); Foreman-Mackey et al. (2013)) attempts to resolve parameter degeneracy by using an ensemble of walkers, and proposing steps in multiple dimensions. To this end, the proposal position Y for each walker is selected from a proposal distribution connecting that walkers current position to that of another randomly selected walker. After the MCMC is well into its run, this proposal selection method will better enable any roaming walkers to join walker clumps surrounding a region of high probability. That the walkers make non-orthogonal steps also enables more efficient convergence, as the ensemble can more directly move through non-orthogonal, degenerate subspaces.

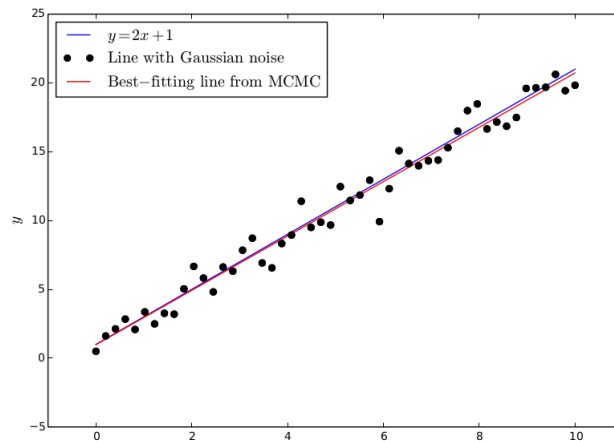To demonstrate the effectiveness of the AI MCMC, we consider the process of fitting

Figure A.1: Solid blue line: Linear function from which data was derived. Black dots: Data used in MCMC fitting routine, created by adding Gaussian random noise to evenly-sampled points from the linear function. Solid red line: Best-fitting linear function as given by the AI MCMC fitting routine.

a linear trend. The data set in figure A.1 was created by adding random Gaussian noise to points uniformly sampled from a line. An AI MCMC routine using 100 walkers was run for 500 steps. The results of a run are commonly visualized using step figures and triangle plots. Figure A.2 is a set of step figures, that is, visualizations of the walkers movements through each dimension of the parameter space. Each black line corresponds to the path of a single walker. Tracing a lines path through the steps
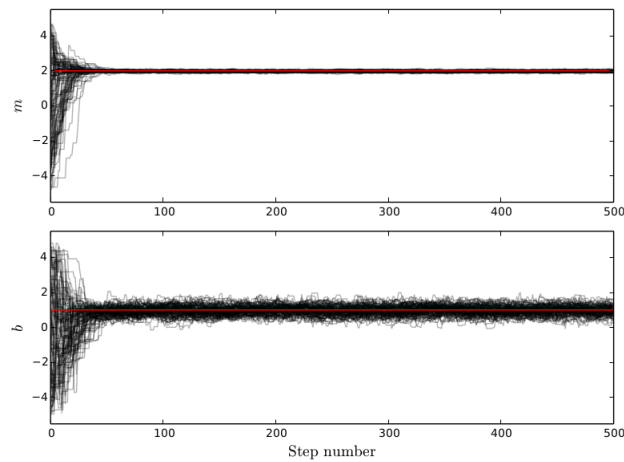
Figure A.2: Step figures for linear parameters m (top panel) and b (bottom panel), tracing the paths of the AI MCMC walkers through parameter space at each step in the fitting routine. The red overplotted lines indicate the true values m = 2 and b = 1.

of the run (x-axis) shows the various parameter values occupied by the walker; in this case, the slope m and y-intercept b. The walkers were initially uniformly distributed on the interval [5, 5] for both m and b. The step figure clearly demonstrates the convergence of the walkers to best fit values from this uniform initial distribution. While step figures assist in evaluating the convergence of an MCMC, a triangle plot provides a means to evaluate a fits accuracy. Figure A.3 is a triangle plot of the linear fit, with the first 39 steps from each walker discarded as a burn-in, so as to remove any influence of the caotic starting distribution. The histograms display the proba-

bility distribution (PD) for each parameter, all the positions occupied by the walkers following the burn-in phase. For constrained parameters, these one-dimensional PDs should take on a semi-Gaussian shape. Overlaid on the histograms are dashed lines that correspond to the 16th, 50th, and 84th percentiles, effectively showing the median value with 1σ errors. The solid blue line corresponds to the set of best-fitting parameters found by the routine. This ideally corresponds to the median value.

The scatter plot is a two-dimensional PD that demonstrates the level of correlation, or covariance, between pairs of parameters. Uncorrelated parameters will show circular scatter, or if the scaling of one parameter is dramatically different than the other, elongation in either the horizontal or vertical directions. Degenerate parameters manifest angled trends, reflecting that a change in one parameter can be well-compensated by altering the value of the other. Such is the case for the parameters m and b in our linear fitting example. The best-fit line given by the MCMC run is plotted in red in Figure 4.2, and agrees well with the original linear function.

The affine invariant sampling MCMC (Goodman and Weare (2010); Foreman-Mackey et al. (2013)) links together the position of two random walkers across the parameter space. Each step taken by one walker is selected from a proposed distribution connecting its position in parameter space to its partners. This is particularly useful for probing anisotropic probability distributions with high numbers of dimensions. The classic problem faced by many probability density sampling routines is commonly visualized as a highly anisotropic probability density, defined as (Foreman-Mackey et al. 2013)

$$\pi(x) \propto exp\left(-\frac{(x_1 - x_2)}{2\epsilon} - \frac{(x_1 + x_2)^2}{2}\right)$$
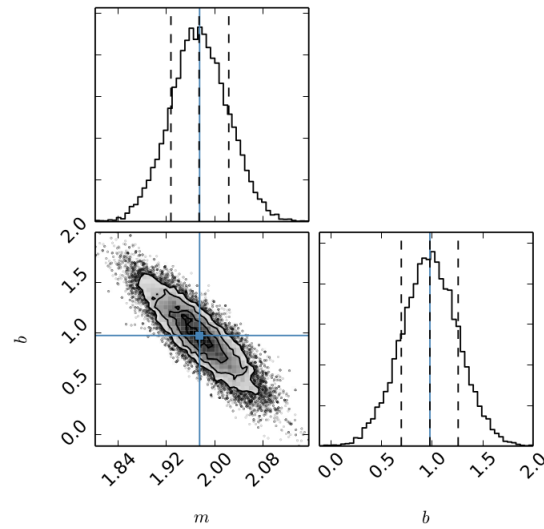
Figure A.3: A triangle plot, showing the one-dimensional and two-dimensional posterior distributions for the MCMC run. The best-fitting parameters are indicated by the solid blue lines, while the median values and 1 errors are given by the dashed black lines.

where ϵ is a proportionality constant. This definition of probability density is sensitive to parameter constraints of N^2 parameters,which makes this problem much worse with higher dimensions (Foreman-Mackey et al. 2013). The problem can be simplified by the affine transformation of $y_1 = \frac{x_1 - x_2}{\sqrt{\epsilon}}$, $y_2 = x_1 + x_2$. Hence invariant sampling is capable of performing well under all linear transformations, and therefore reducing degeneracy between parameters.

APPENDIX B
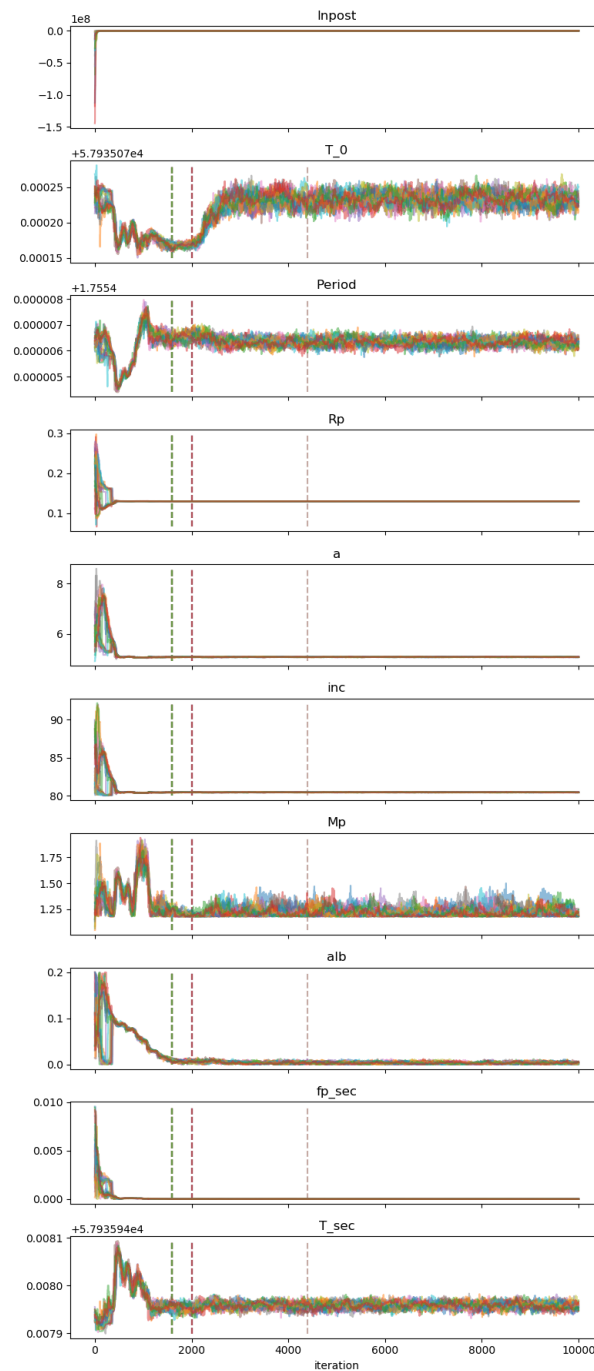
MCMC Plots

B.1 Step Figures

Figure B.1: Step figures that show the walkers through the parameter space at each step of the fitting routine of Serrano's model to the data of wasp 104.
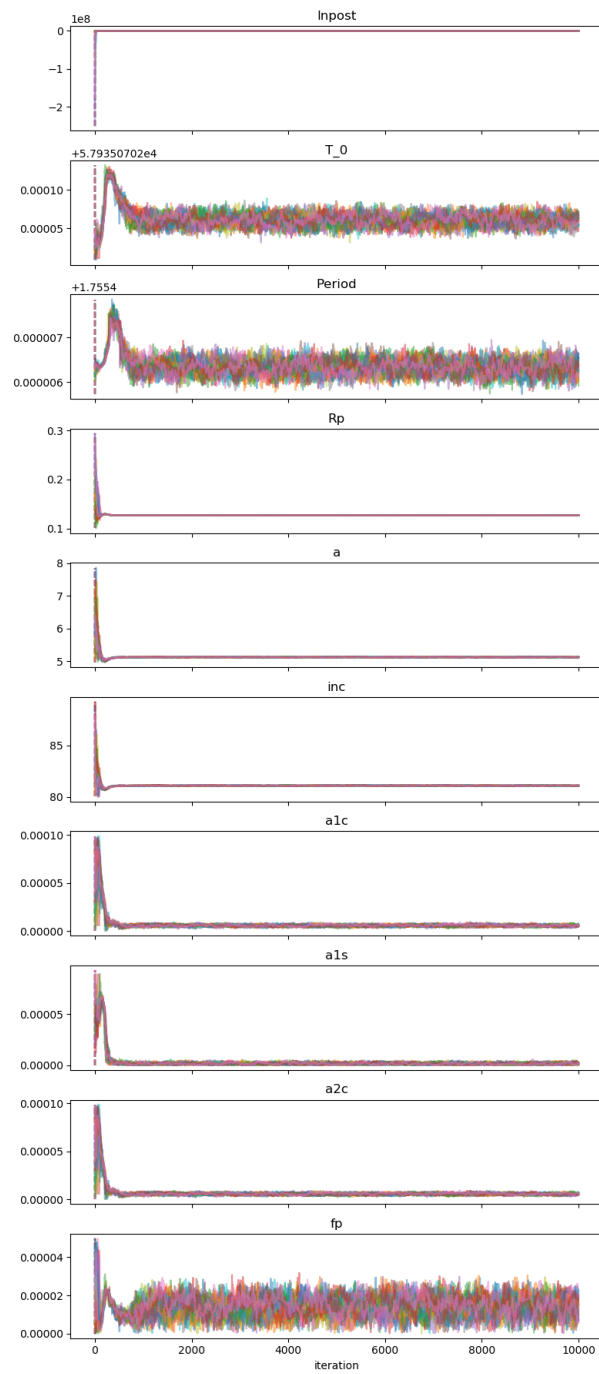
Figure B.2: Step figures that show the walkers through the parameter space at each step of the fitting routine of Mazeh & Faigle's model to the data of wasp 104.
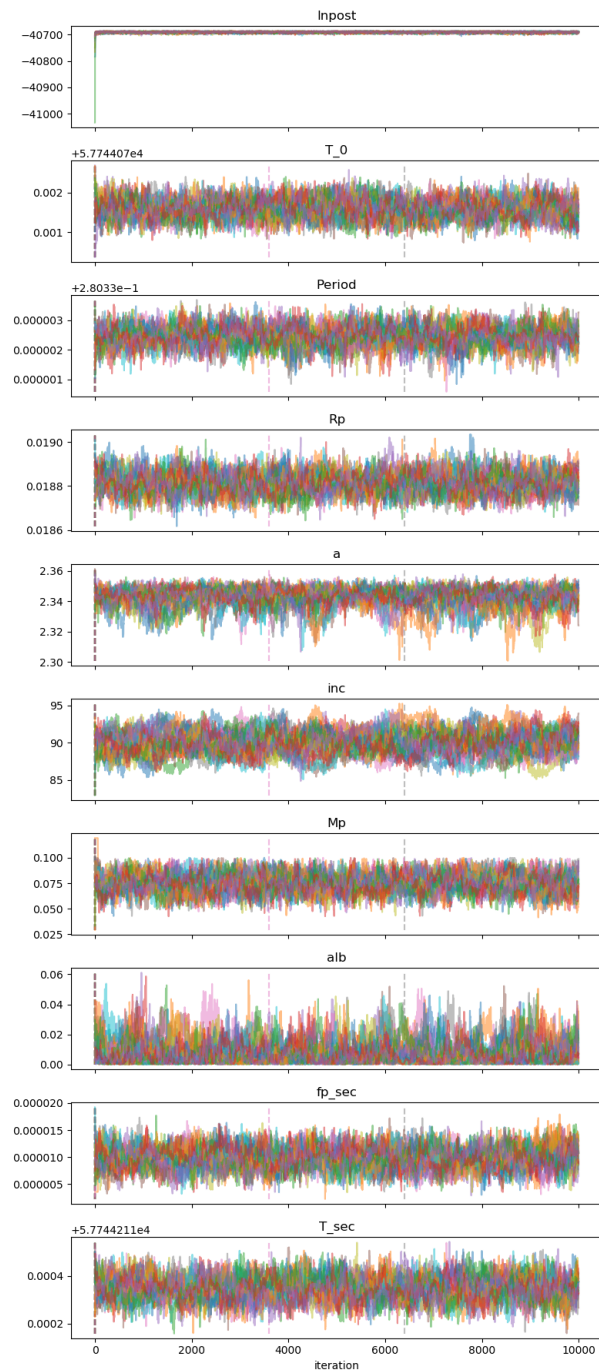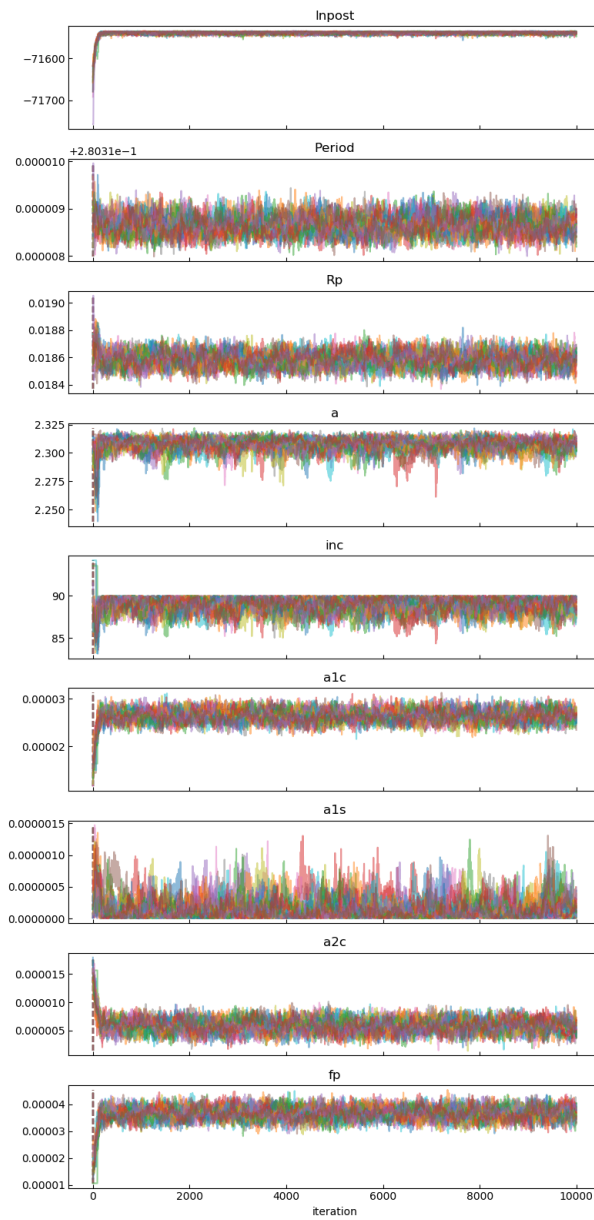
Figure B.3: Step figures that show the walkers through the parameter space at each step of the fitting routine of Serrano's model to the data of K2 141-b.

Figure B.4: Step figures that show the walkers through the parameter space at each step of the fitting routine of Mazeh & Faigle's model to the data of K2 141-b.
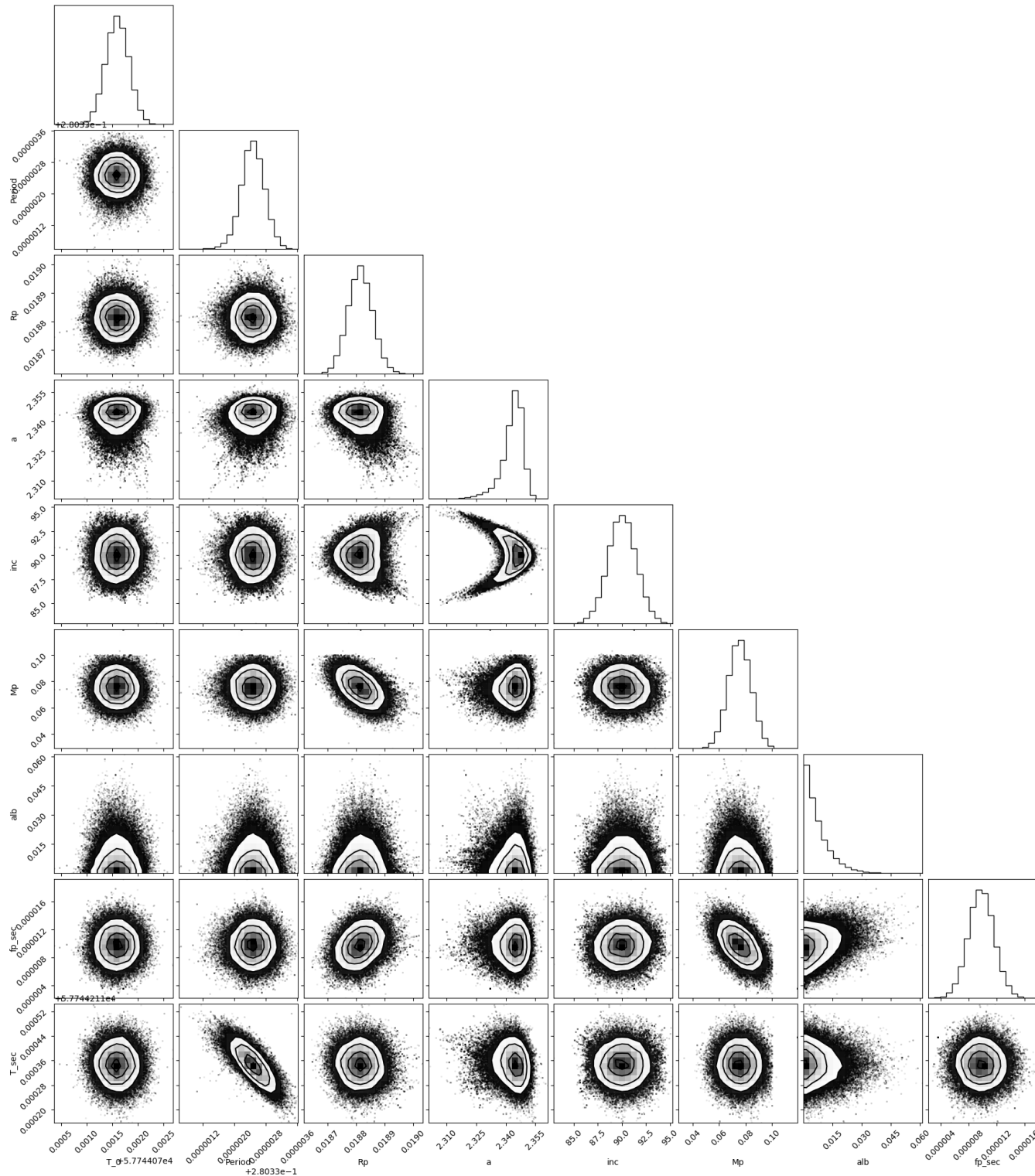
B.2 Triangle Plots

Figure B.5: A triangle plot, showing the one-dimensional and two-dimensional posterior distributions for the MCMC run of Serrano's model to the data of K2 141-b.
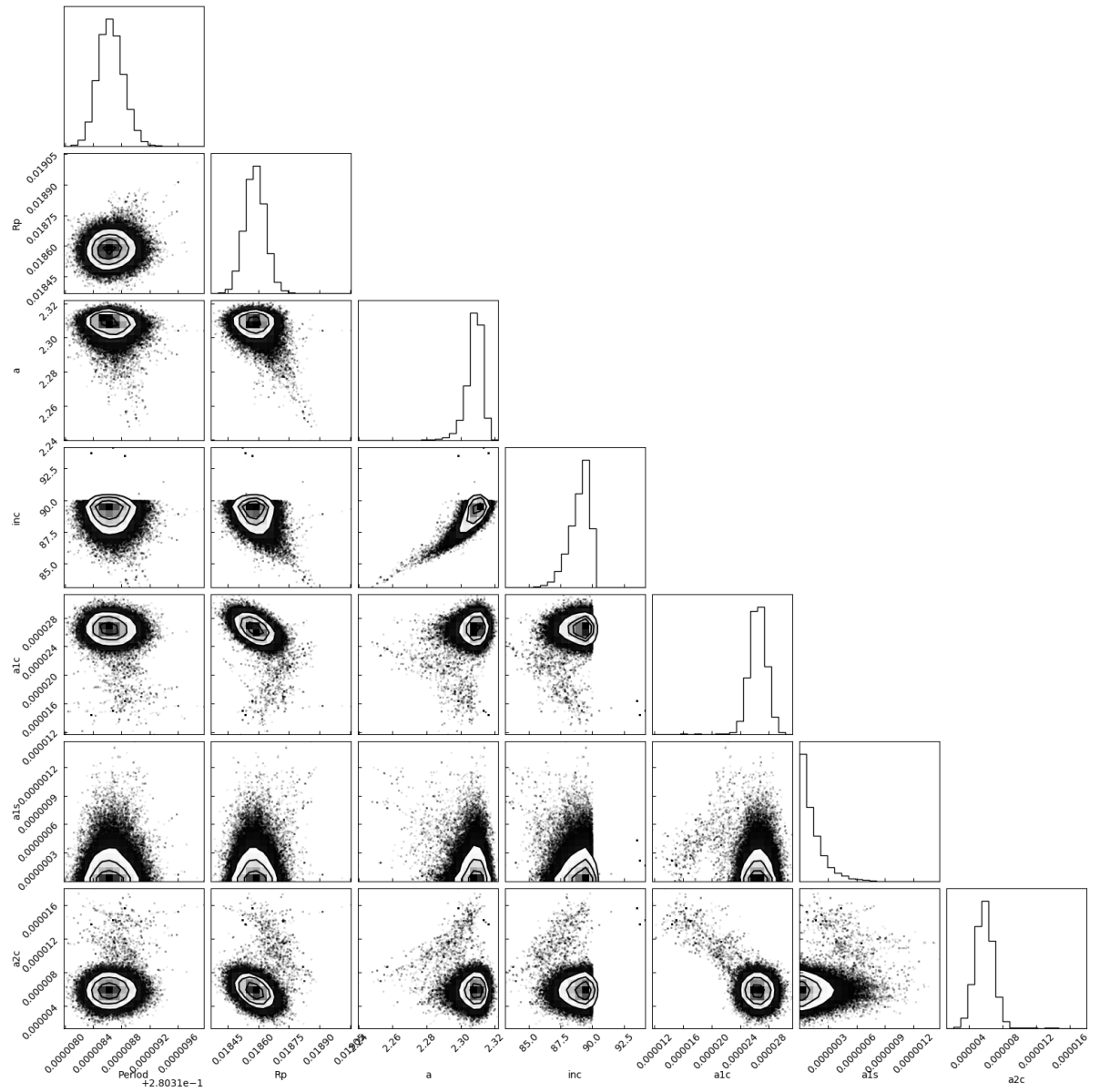
Figure B.6: A triangle plot, showing the one-dimensional and two-dimensional posterior distributions for the MCMC run of Mazeh & Faigler's model to the data of K2 141-b.