# Exoplanet transit modelling: three new planet discoveries, and a novel Artificial Neural Network treatment for stellar limb darkening

Kirstin Hay

University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of

Doctor of Philosophy (PhD)

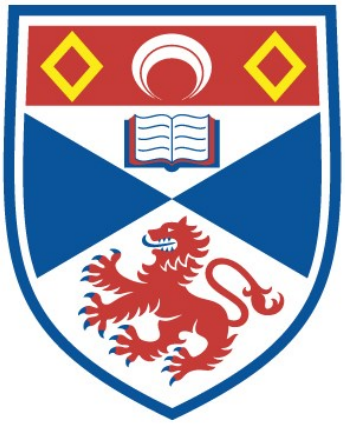
at the University of St Andrews

March 2018

# EXOPLANET TRANSIT MODELLING: THREE NEW PLANET DISCOVERIES, AND A NOVEL ARTIFICIAL NEURAL NETWORK TREATMENT FOR STELLAR LIMB DARKENING

Kirstin Hay

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



2018

# Exoplanet transit modelling:

**Three new planet discoveries, and a novel Artificial Neural Network treatment for stellar limb darkening**

*by*

**Kirstin Louise Hay**

*Submitted for the degree of Doctor of Philosophy in Astrophysics*

26th March 2018

# Declaration

I, Kirstin Hay, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 37,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews in September 2014.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date Signature of candidate

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date Signature of supervisor

# Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Kirstin Hay, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

**Printed copy**: No embargo on printed copy

**Electronic copy**: No embargo on printed copy

Date Signature of candidate

Date Signature of supervisor

# Underpinning Research Data or Digital Outputs

**Candidate's declaration**

I, Kirstin Hay, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date Signature of candidate

# Abstract

This first part of this thesis concerns the discovery and parameter determination of three hot Jupiter planets, first detected with by the SuperWASP collaboration, and their planetary nature is confirmed with the modelling of radial velocity measurements and further ground-based transit lightcurves. WASP-92b, WASP-93b and WASP-118b are all hot Jupiters with short orbital periods – 2.17, 2.73 and 4.05 days respectively. The analysis in this thesis finds WASP-92b to have $R_p = 1.461 \pm 0.077\ R_J$ and $M_p = 0.805 \pm 0.068\ M_J$; WASP-93b to have $R_p = 1.597 \pm 0.077\ R_J$ and $M_p = 1.47 \pm 0.029\ M_J$, and WASP-118b to have $R_p = 1.440 \pm 0.036\ R_J$ and $M_p = 0.514 \pm 0.020\ M_J$.

The second part of this thesis presents three novel approaches to modelling the effect of stellar limb darkening when fitting exoplanet transit lightcurves. The first method trains a Gaussian Process to interpolate between pre-calculated limb darkening coefficients for the non-linear limb darkening law. The method uses existing knowledge of the stellar atmosphere parameters as the constraints of the determined limb darkening coefficients for the host star of the transiting exoplanet system.

The second method deploys an artificial neural network to model limb darkening without the requirement of a parametric approximation of the form of the limb profile. The neural network is trained for a specific bandpass directly from the outputs of stellar atmosphere models, allowing predictions to be made for the stellar intensity at a given position on the stellar surface for values of the $T_{\rm eff}$, $\log g$ and [Fe/H]. The efficacy of the method is demonstrated by accurately fitting a transit lightcurve for the transit of Venus, and for a single transit lightcurve of TRES-2b. The final limb darkening modelling method proposes an adjustment to the neural network model to account for the fact that the stellar radius is not constant across wavelengths. The method also allows the full variation in light at the edge of the star to be modelled by not assuming a sharp boundary at the limb.

# Acknowledgements

The work contained in this thesis would not have been possible without the support and assistance of many people, and I would like to take this opportunity to thank some of those who have been part of the network of people whose help over the last 3.5 years both scientifically and personally have made completing this work a reality.

Thank you to Andrew Collier Cameron, who as my supervisor has brought inexhaustible enthusiasm and continuous support even when things are not going exactly to plan.

A special mention to Yvonne Unruh at Imperial College London, who whilst supervising my MSci project inspired me with enough confidence to believe that I could do a PhD, and enjoy it as well.

Thank you to Annelies Mortier for providing invaluable guidance throughout my PhD, and for being a sensible and steady influence. To the rest of the planet hunters gang – Raphie, Nicole and Duncan, I will miss our cake-fuelled journal clubs and camaraderie.

My time in St Andrews would not have been the same without the community of Astronomers I have had the pleasure of sharing friendships, coffee breaks, and pub trips with. They are too numerous to name here, but thank you all.

Thank you to my family for their support. And to Claire, thank you for everything.

## Funding

This work was supported by the Science and Technology Facilities Council (STFC) through the doctoral training grant ST/M503812/1.

# Collaboration statement

This thesis is the result of my own work carried out at the University of St Andrews between September 2014 and March 2018. In all cases the text in the Chapters has been written entirely by me. All Figures, unless explicitly stated in the Figure caption have been produced by me.

The work in Chapter 2 was published as: "WASP-92b, WASP-93b and WASP-118b: three new transiting close-in giant planets", K. L. Hay, A. Collier-Cameron, A. P. Doyle, G. Hébrard, I. Skillen, D. R. Anderson, S. C. C. Barros, D. J. A. Brown, F. Bouchy, R. Busuttil, P. Delorme, L. Delrez, O. Demangeon, R. F. Díaz, M. Gillon, Y. Gómez Maqueo Chew, E. Gonzàlez, C. Hellier, S. Holmes, J. F. Jarvis, E. Jehin, Y. C. Joshi, U. Kolb, M. Lendl, P. F. L. Maxted, J. McCormac, G. R. M. Miller, A. Mortier, E. Pallé, D. Pollacco, J. Prieto-Arranz, D. Queloz, D. Ségransan, E. K. Simpson, B. Smalley, J. Southworth, A. H. M. J. Triaud, O. D. Turner, S. Udry, M. Vanhuysse, R. G. West and P. A. Wilson, 2016, *Monthly Notices of the Royal Astronomical Society*: **463**, 3276–3289.

The text of the chapter was written by me, the MCMC code used for the main analysis was written primarily by A. Collier-Cameron, and used by me. A. P. Doyle performed the stellar parameter analysis described, and I. Skillen performed the adaptive optics imaging analysis described. All other authors participated in the collection of the observational data analysed; are involved in the SuperWASP collaboration, or offered comments of the manuscript of the paper.

# Contents

# List of Figures

# List of Tables

# 1

# The exoplanet transit

From the first discovery of a planet orbiting a main-sequence star other than the sun in 1995[1], the prospect of detecting a planet analogous to our own has seemed tantalising. Whilst the capability of techniques to detect such planets have improved, a detection of a planet like our own Earth remains elusive.

Directly observing an exoplanet is not a trivial act, since they are such small objects so far away, so much of the research into exoplanets has focused on the study of the effects that the planet has on its host star. The work in this thesis focuses primarily on the transits of exoplanets in front of their host stars.

## 1.1 Exoplanet orbits and notation

When trying to understand exoplanet systems, an understanding of the orbits involved is key. A brief introduction to the geometry and notation used throughout this thesis is described

[1]The first detected planet was 51 Peg b, discovered by Mayor & Queloz (1995) using the ELODIE spectrograph at Observatoire de Haute-Provence in France.

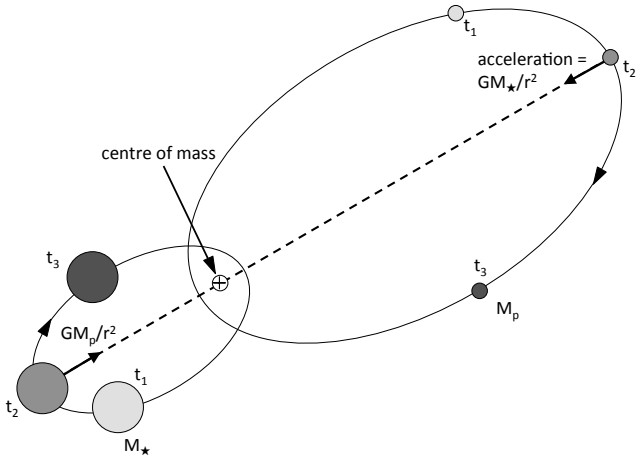



**Figure 1.1:** Schematic of the mutual elliptic orbits of a host star and its orbiting planet. Figure adapted from Figure 2.3 in Perryman (2011)

here.

At the core of a star-planet system, is the elliptic orbit each component follows around a shared centre of mass (also known as the barycentre), an example of which is shown (not to scale) in Figure 1.1. The schematic shows the orbits of the planet with mass $M_p$, and the star with mass $M_\star$. Each component has the same orbital period $P$, and orbital eccentricity ($e$), but a semi-major axis ($a$) dependent on the mass of the object. The mutual orbital positions for this example orbit are shown for three time steps to demonstrate the timing of the orbit, which can be defined with some reference time ($t_p$) for a given position in the orbit, and an orbital period ($P$).

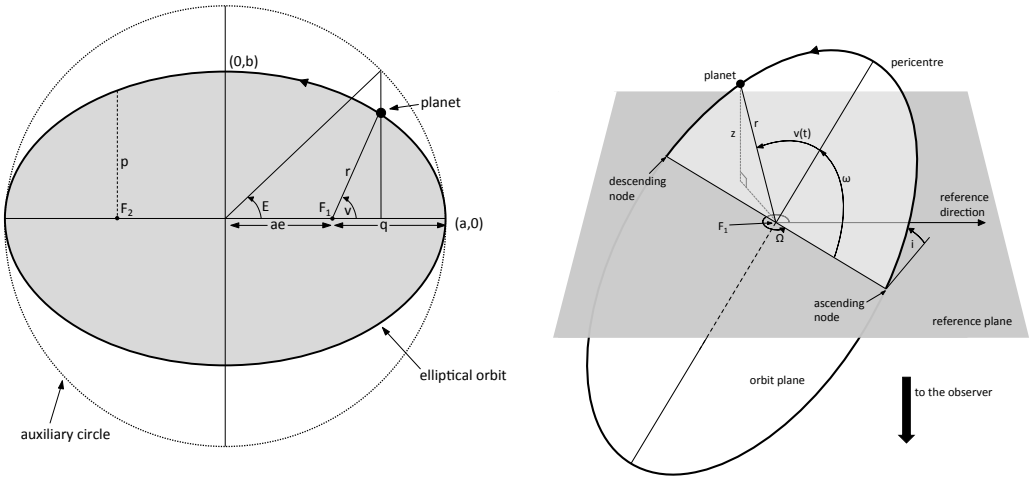
Considering just the orbit of one of the bodies in the system allows for a further description of the parameterisation of the orbit. A schematic of an individual elliptic orbit is shown in 1.2a. The shape of this orbit is defined as:

$$b^2 = a^2(1-e^2)\,, \tag{1.1}$$

where $a$ is the semi-major axis; $b$ is the semi-minor axis, and $e$ is the eccentricity of the orbit[2]. The shape and size of the orbit can be described with just $a$ and $e$.

Of course this orbit is in some real alignment in space, so cannot just be thought of as its

[2]Note that any further use of $b$ as a symbol in this work will not refer to the semi-minor axis, rather the transit impact parameter, which is defined later in this section




**(a)** Schematic of the orbital parameters of an elliptic orbit projected onto 2 dimensions.

**(b)** Schematic of the orbital parameters of an elliptic orbit projected onto 3 dimensions.

**Figure 1.2:** Figures adapted from Figure 2.1 and 2.1 respectively in Perryman (2011).

2D projection, and therefore the orbit as observed from Earth is defined with some position relative to a reference frame, an illustration of which is shown in Figure 1.2b. The orientation of the orbit is defined with the Euler angles: $i$, $\Omega$ and $\omega$ defined as in Figure 1.2b.

Following the definitions given above, a Keplerian orbit can be fully defined with seven parameters: $a$, $e$, $P$, $t_p$, $i$, $\Omega$ and $\omega$; the timings of the orbit described with $P$ and $t_p$, and the relative position described with $i$, $\Omega$ and $\omega$. Some part of studying exoplanet systems aims to build up an understanding of these fundamental parameters describing the orbit, despite not being able to measure all of them them directly.

A particular orientation of an orbit results in the planet passing in front of the star as seen from Earth, this occurs when $i \approx 90°$ within a few degrees depending on orbital separation. For the planet to be transiting, the following condition must be met:

$$a \cos i < R_\star + R_p \ , \tag{1.2}$$

where $R_p$ and $R_\star$ are the radii of the planet and star respectively. The value of $a \cos i$ will only be $< R_\star + R_p$ for values of $i$ close to 90°, as $a \gg R_\star > R_p$. Because of the role of $a$ in this expression, the wider the orbit, the smaller the range of $i$ for which a transit will be observable from Earth. The geometry of such a transit is shown in Figure 1.3.

The Keplerian orbital parameters are not the only useful information that can be gleaned

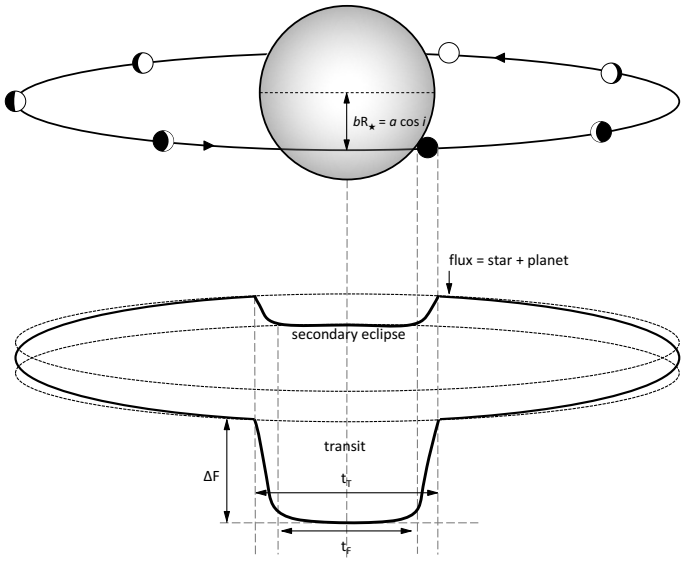



**Figure 1.3:** Schematic of a transit showing the relation between physical parameters and transit lightcurve shape. Figure adapted from Winn (2009)

from an exoplanet system. The shape and size of a planetary transit can also provide more information about the composition of the star-planet system, the size of the planet in particular. The total transit duration, denoted $T_T$, is the time from where the edge of the planet begins to occult the limb of the star, blocking some of its light, until no part of the planet is overlapping with the surface of the star. The flat portion of the transit ($T_F$) describes the duration for which the planet and star are completely overlapping.

In Figure 1.3, $\Delta F$ is the transit depth, which relates to the relative sizes of the star and the planet:

$$\Delta F = \frac{R_p^2}{R_\star^2} \, . \tag{1.3}$$

The transit depth is simple to measure from a transit lightcurve to a low precision, and is a useful initial indicator of the type of transiting system observed.

Another useful parameterisation used for a transiting system is the impact parameter ($b$), which is a representation of the angle at which the planet's orbit is inclined relative to the observer. This is configured as:

$$bR_\star = a \cos i \, . \tag{1.4}$$

For a transiting planet, $b$ will be between 0 (transit passes along centre of stellar surface) and $1 + \frac{R_p}{R_\star}$ (edge of planet grazes the edge of the star). The planet will only fully transit the star

where $b < 1 - \frac{R_p}{R_\star}$, and any value of $b$ larger than that, but $< 1 + \frac{R_p}{R_\star}$ is a grazing transit. It is useful to note that for higher values of $b$, less information about the stellar surface is encoded into the transit lightcurve, as only part of the full radius of the star is occulted by the planet during the transit, assuming radial symmetry of the stellar surface.

Figure 1.3 also shows that the system brightness baseline is not constant throughout the entire orbit. This is due to the planet reflecting light from the star. At different points in the orbit, more or less of the surface of the planet facing the star is visible, which would make the system appear brighter or fainter. The amplitude and shape of this phase curve is one source of information about exoplanetary atmospheres through the reflectivity of the atmosphere, as demonstrated by Knutson et al. (2009) for HD 189733b.

When planets are discovered by observing the radial velocity variation that the planet exerts on its host star throughout its orbit, a measurement of the mass of the planet can be made. The method is known as the radial velocity method. This mass measurement is however dependent on the inclination of the orbit. If the planet is also transiting, then the orbital inclination can be ascertained, as well as a measurement of the radius of the planet, which allows the bulk composition of the planet to be modelled. Discovering and characterising transiting planets is therefore vital for building a full picture of planet formation and migration.

## 1.2 The search for planet transits

The first exoplanets detected by the radial velocity method (RV) were completely unexpected based on existing understanding of the solar system – there was no expectation that a planet similar in mass to Jupiter could exist so close to its host star as 51 Peg b is (Mayor & Queloz, 1995). If a planet like 51 Peg b exists, then it was very likely that there were more systems similar to it to find. Suddenly existing capabilities to detect exoplanet transits from the ground could realistically be used for the transits of this hitherto unimagined class of planets. The searches for transit signatures began in earnest. The transit of HD 209458 b was the first exoplanet transit observed Charbonneau et al. (2000), after it had been discovered with the RV method.

The first exoplanet discovered by its transit alone was OGLE-TR-56-b (Udalski et al., 2002; Konacki et al., 2003). Since hot giant planets have a significant transit depth, and a high transit frequency, the prospects of finding more transits for favourably aligned systems using small

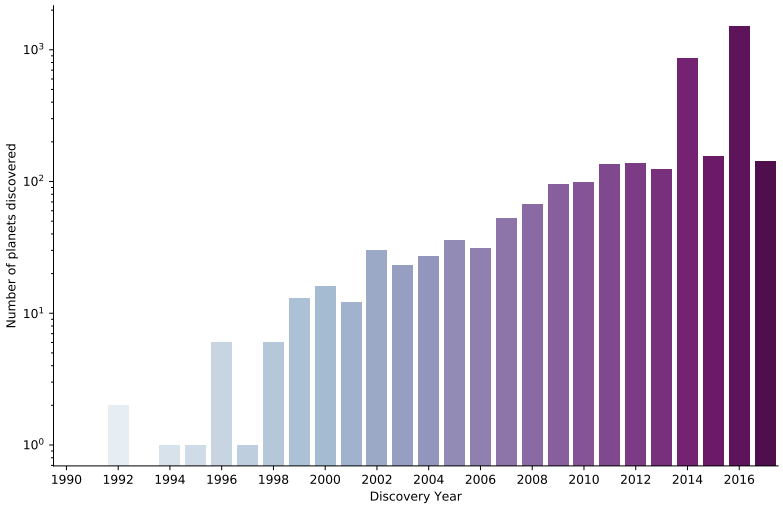



**Figure 1.4:** Plot of number of planets discovered each year. Data accessed from NASA Exoplanet Archive on 8th January 2018.

telescopes were good. As the possibility of discovering the transits of planets from ground-based telescopes became a reality, new whole-sky surveys with dedicated instruments focusing on detection of transits came online, ushering in a swathe of new discoveries. The first of these transit surveys to publish a successful detection was the Trans-Atlantic Exoplanet Survey (TrES) network with the detection of TRES-1b (Alonso et al., 2004). The Hungarian Automated Telescope Network (HATNet) (Bakos et al., 2002) also saw early success with the detection of HAT-P-1b (Bakos et al., 2007). The WASP project closely followed with the publication of the detection of WASP-1b and WASP-2b (Collier Cameron et al., 2007). More details about the SuperWASP collaboration are described in Section 1.3. TrES, HATNet and SuperWASP are not the only surveys focusing on the detection of giant exoplanets, KELT (Pepper et al., 2007) and other ground-based surveys have also found detection success.

The *N*ext *G*eneration *T*ransit *S*urvey (NGTS) (Wheatley et al., 2018) is a more recent addition to the ground-based transit searching efforts. NGTS has harnessed the insights gained from earlier surveys (particularly SuperWASP) to be able to achieve better photometric precision, and thus provide the capability to discover planets with much shallower transits than previous whole sky transit surveys, looking to find many Neptune and some super Earth sized planets. Operational since 2015 at Paranal in Chile, the project announced its first exoplanet discovery, NGTS-1b, in October 2017 (Bayliss et al., 2018).

Naturally there is a far larger diversity of extra-solar planets than solely the hot giant planets readily detected from the ground – the *Kepler* space mission (Borucki et al., 2010) changed

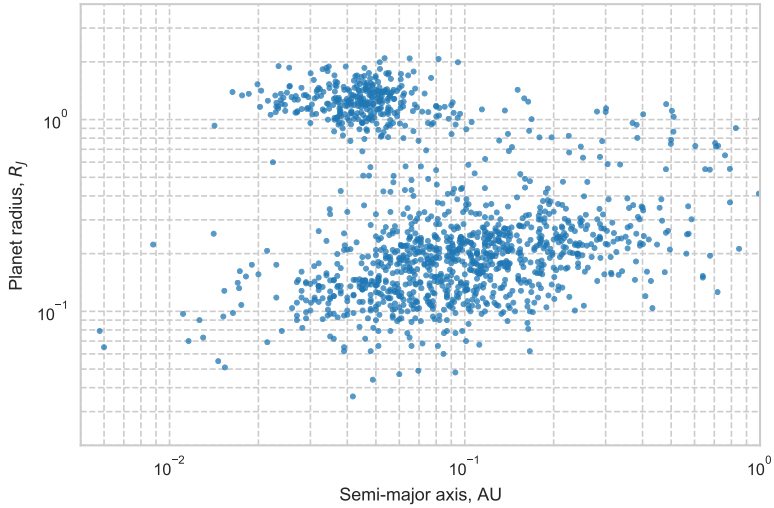



**Figure 1.5:** Plot of planets discovered by transit method for planet size, and orbital semi-major axis. Data accessed from NASA Exoplanet Archive on 8th January 2018.

the landscape of discovered planets by surveying a single field in the sky continuously for just under four years. It was able to observe much shallower transits and maintain a stable photometric baseline impossible from the ground, and thus found smaller planets and transiting bodies with longer orbital periods. Since the telescope was launched in 2009, thousands of planet discoveries have been associated with the mission[3].

Figure 1.4 shows the number of exoplanet discoveries per year – the y-axis scale is logarithmic, since the rate of increase of discoveries in recent years has been too fast to display fully on a linear axis. The rate of discoveries is a strong sign of how fast-moving the field of exoplanetary research is, and the pace of discovery is likely to remain high with the launch of TESS (*T*ransiting *E*xoplanet *S*urvey *S*atellite) (Ricker et al., 2014) in 2018, and the continued operation of NGTS.

Figure 1.5 highlights the diversity of planets discovered via their transits – demonstrated through the distance the planet is from its host star in AU, and the size of the planet in radii of Jupiter ($R_J$). Most of the smaller planets ($R_p \lesssim 0.5R_J$) shown in this figure were detected by *Kepler* – the vast majority of the planets found by *Kepler* are small planets orbiting very close to their host star.

An initial look at the exoplanet distribution in Figure 1.5 would indicate that there are very many hot giant planets (upper left) relative to other planet types. It is however important to

[3]Catalogue of planets maintained at the NASA Exoplanet Archive https://exoplanetarchive.ipac.caltech.edu/ (Accessed 8th January 2018)

note that this plot only shows discovered planets, and not the true full exoplanetary population. The top left of the plot where large close-orbiting planets are located is the region of the displayed parameter space that is relatively easy to detect from the ground. The number of planets plotted in that region is more representative of the true population of that type of planet for the whole sky compared to other parts of the parameter space. For larger orbital separations and shallower transits, only *Kepler* was able to detect the transits, and the detection efficiency of the *Kepler* pipeline also degrades for systems that are more challenging to observe (Christiansen et al., 2016). The likelihood of an exoplanet being aligned favourably for observing a transit is also dependent on the orbital separation:

$$p \sim \frac{R_\star + R_p}{a} , \tag{1.5}$$

which is derived from Equation 1.2. This probability expresses that using solely the transit method, observational results are not only biased by the difficulties of detecting shallower and less frequent transits, but also that larger orbital separations mean that the planet is very unlikely to be transiting.

## 1.3 The WASP project

As mentioned in Section 1.2, the WASP project collaborators were early pioneers of ground-based transit search surveys. WASP is an acronym for *W*ide *A*ngle *S*earch for *P*lanets, and consists of two survey sites – one at Observatorio del Roque de los Muchachos (ORM) on La Palma in the northern hemisphere, and a second in the south at Sutherland in South Africa. The details of the technical set up of the instrumentation are described in Pollacco et al. (2006).

The key aim of the project is to take wide angle images of the sky at a high enough cadence to detect the periodic signature of the transit of a hot Jupiter planet in stellar lightcurves using several small telescopes at each of the two observing sites. Each observing night produces several photometric measurements of thousands of stars, which are then computationally searched for appropriate periodic signals. The best candidates are then vetted by eye and further observations are organised to validate the status of the candidate. An example of the confirmation process for planet detections is described in Chapter 2 for hot Jupiters WASP-92b, WASP-93b and WASP-118b.

The success of the WASP project has led to the discovery of more than 100 exoplanets,

most of which can be described as hot Jupiters, although also some planets which are the size of Neptune or Saturn.

## 1.4 Transit parameter fitting

A vital part of the process for discovering an exoplanet is having a strong grasp of the parameters of the system. Many other astrophysical effects can mimic an exoplanet transit or bias the fitting of the transit parameters, therefore it is important to model the system carefully to provide a reliable picture of the observed system.

Once the true nature of the detected transit has been verified as an exoplanet system, the careful fitting of the parameters of the system can begin. The model used for the shape of the transit is based on the geometries of the orbits introduced in Section 1.1, and the timings defined by the Keplerian orbital relations with the laws of gravitation. The stellar surface must also be modelled, as the brightness of the light occulted by the planet at a given part of the transit is dependent on the distance of that position from the centre of the star. This brightness variation from the centre to the limb of the star is known as limb darkening, and is introduced in much more detail in Chapter 3.

Any fit of a transit lightcurve should also include consideration for the instrumental or weather effects present in the data, as well as any stellar variability, which if not appropriately included could bias the final system parameters extracted. The systematics signal observed in a transit lightcurve cannot be assumed to be white noise, and so proper consideration should be made of the form of the noise signals when modelling the lightcurve (Pont et al., 2006).

## 1.5 This thesis

The work described in this thesis has been split into several chapters. Chapter 2 describes the discovery and parameter determination of WASP-92 b, WASP-93 b, and WASP-118 b. Chapter 3 introduces the concept of stellar limb darkening and describes different modelling approaches that have been deployed for stellar atmospheres and limb darkening. Chapter 4 develops a new method for interpolating limb darkening coefficients using Gaussian processes. Chapter 5 introduces the use of Artificial Neural Networks in astrophysics. Chapter 6 shows how these neural networks can be deployed to interpolate limb darkening directly from stellar atmosphere models. Chapter 7 shows how this neural network model can be adjusted to account

for the stellar radius not being constant at all wavelengths. Finally the results and analysis presented in this thesis are concluded in Chapter 8.

# 2

# The discovery of WASP-92b, WASP-93b and WASP-118b

This chapter is based on work published as Hay et al. (2016), presenting the discoveries of WASP-92b, WASP-93b and WASP-118b.

## 2.1 Introduction

Hot Jupiters are not common objects. They are found around $\sim$ 1% of main sequence stars (Howard et al., 2012; Wright et al., 2012; Wang et al., 2015), with some variation on the exact value depending on the selection of the sample studied. To be able to detect these rare objects when they are in favourable alignments to transit, thousands of stars must be studied. Surveys such as SuperWASP (as introduced in Section 1.3) have focussed on covering enough of the sky to make significant numbers of detections, with lots of success.

Since there is no analogue to hot Jupiters in the solar system, the question of how these

planets could have formed was quickly posed. Gas giant planets cannot form close to their host stars, as the amount of gas and dust found in that region would not be enough to stably accrete and form such a large planet before the proto-planetary disk dissipates. The giant planets instead form beyond the snow line of the system, where the cooler temperatures allow ice mantles to develop, which accelerates the dust accretion to the planetary core. There is also a bigger gas supply at larger distances from the host star, which allows for the formation of the large gaseous envelopes that are observed on hot Jupiters (Mizuno, 1980; Boss, 1995).

For the hot Jupiter planets to reach the close-in orbital separations at which they are observed, they must migrate closer to the host star than the location of their formation. The mechanism responsible for this migration remains unclear. The migration is thought to originate from two possible types of processes.

The first proposal is for the giant planet to undergo tidal interactions with the proto-planetary disk (Goldreich & Tremaine, 1980; Ward, 1997), which is known as Type I migration.

The second migration process is for the giant planet to interact with another body, inducing changes in its orbit. The Jupiter sized planet can be transferred from its large orbital semi-major axis through a perturbation from another body onto a highly eccentric orbit, which then sheds angular momentum through tidal forces with the primary host star to circularise at a close-in orbital separation – this is known as high eccentricity migration (HEM). The perturbations could be caused by Kozai cycling with a binary stellar companion (eg. Wu & Murray (2003)) or another planet in the system (eg. Rasio & Ford (1996)), or through planet-planet scattering (Chatterjee et al., 2008).

It is expected that HEM would not just result in highly eccentric intermediate orbits, but also the interactions could induce misalignments between the rotational axis of the star and the orbital plane of the planet. The scale of this misalignment projected onto the observation angle can be measured through the Rossiter-McLaughlin (RM) effect (Rossiter, 1924; McLaughlin, 1924). The RM effect refers to the perturbation in the RV curve during the transit caused by different parts of the stellar surface being occulted by the the planet. Since the star is rotating, the starlight blocked by the planet will be red or blue shifted depending on the path the planet takes across the surface, and fitting the form of the RV anomaly during the transit allows the relative alignments of the planetary orbit and stellar rotation to be mapped. The effect was first observed and modelled for an exoplanet transit by Queloz et al. (2000b).

For a faster rotating star ($v \sin i \gtrsim 10$ km s$^{-1}$), the RM perturbation can be readily seen as a bump in the wider average absorption line profile of the spectra during the transit. The location (in velocity), and size of that bump at a given time during the transit can be modelled to determine the path the planet is observed to take across the rotating surface of the star. This technique is known as doppler tomography and was first deployed for an exoplanet transit by Collier Cameron et al. (2010).

Several studies of orbital misalignments have resulted in a plethora of measurements of the misalignment, which noted that there was a higher likelihood of observing a significant misalignment for systems with host stars hotter than 6250 K (Winn et al., 2010; Albrecht et al., 2012). This result indicates that Kozai and scattering interactions are likely to play a part in the description of the migration processes of hot Jupiters with hotter host stars, indicating that the migration of hot Jupiters with cooler host stars is a result of disk migration. Statistical analysis of the observed Jupiter populations by Dawson et al. (2015) found a paucity of intermediate Jupiters with high eccentricities, which suggests that HEM cannot account for a large proportion of the migration. The existence of the WASP-47 system (Becker et al., 2015) with several smaller inner companions to the hot Jupiter cannot be explained with HEM, rather provides strong evidence for disk migration.

The whole picture of hot Jupiter migration processes remains unclear – explaining the lack of intermediate systems, as well as the population of misaligned hot Jupiters, requires more than than one prevailing migration process. The orbits of more systems for a variety of host stars must be studied to populate the understanding of why certain processes happen in each situation. This chapter describes the steps leading to the discovery and parameter determination for three hot Jupiter planets.

## 2.2 Observations

The discovery of WASP-92b, 93b and 118b resulted from the use of several telescopes around the world, the details of the observations obtained are presented below.

### 2.2.1 WASP photometry

The first detections of the transit signals for these planets were made with the WASP telescopes. SuperWASP in La Palma at the Observatorio del Roque de los Muchachos (ORM) was

**Table 2.1:** Details of the photometric observations made by the WASP telescope survey operations for each system.

| System | Interval | $N_{points}$ |
|---|---|---|
| WASP-92 | 2007-03-30 – 2007-08-04 | 11674 |
| | 2008-03-28 – 2008-08-03 | 8551 |
| | 2009-03-31 – 2009-08-03 | 10358 |
| | 2010-03-28 – 2010-08-03 | 11693 |
| WASP-93 | 2007-07-26 – 2007-12-25 | 8614 |
| WASP-118 | 2008-07-13 – 2008-12-13 | 7750 |
| | 2009-07-17 – 2009-12-09 | 9297 |
| | 2010-08-18 – 2010-12-13 | 4906 |

used to detect WASP-92b and WASP-93b, and WASP-118b was detected jointly with WASP-S at Sutherland in South Africa. The WASP observing windows for each host star are detailed in Table 2.1.

The datasets were detrended using the WASP pipeline, following Cameron et al. (2006). The WASP Collaboration identified each of the systems as candidate planets through searching for a periodic signal in each of the datasets using an adapted Box-Least Squares (BLS) search algorithm (Kovács et al., 2002) as described for the WASP implementation in Cameron et al. (2007). The BLS folds the data on multiple trial periods and searches for where the folded data fits well to a square transit window. The search includes various trial transit depths and durations, and the strongest peaks in the likelihood of the search parameters are identified as potential transit signals. Once candidate signals are detected, the lightcurves and other information for the star are inspected by eye to further assess the likelihood of detection being as a result of a planetary transit. Candidate transits around the stars of WASP-92, WASP-93 and WASP-118 were detected with periods of 2.175, 2.732, and 4.046 days respectively with this method.

In order to confirm the planetary nature of the signals detected, the transits must be observed with a higher spatial resolution and a higher cadence than the WASP telescopes. The WASP observations have low resolution (13.7″), so lightcurves are susceptible to blending with nearby stars. The blending can dilute the observed transit depth in the WASP data, and background binary star systems blended with the foreground object can also mimic the transit of a planetary-sized object on the primary star.

**Table 2.2:** Details of photometric follow up observations with telescope/instrument used, and the date at start of observation

| System | Obs start | Instrument | Filter | Exp time (s) | $N_{points}$ | Defocus? | Full? |
|---|---|---|---|---|---|---|---|
| WASP-92 | 2012-07-24 | OverSky | r | 90 | 179 | N | N |
| | 2013-06-26 | NITES | none | 20 | 751 | Y | Y |
| | 2013-07-09 | RISE | V+R | 20 | 743 | Y | Y |
| | 2013-07-22 | NITES | none | 20 | 630 | Y | Y |
| WASP-93 | 2008-07-10 | RATCam | Z | 10 | 144 | Y | Y |
| | 2009-01-23 | JGT | R | 60 | 156 | N | N |
| | 2011-09-28 | JGT | R | 15 | 627 | N | N |
| | 2011-10-31 | PIRATE | R | 45 | 142 | N | Y |
| | 2012-01-21 | JGT | R | 30 | 458 | N | Y |
| | 2012-07-22 | PIRATE | R | 120 | 87 | N | Y |
| | 2012-08-21 | NITES | none | 15 | 785 | Y | Y |
| | 2013-10-05 | RISE | V+R | 3.5 | 3717 | Y | Y |
| WASP-118 | 2013-10-27 | EulerCam | z' | 60 | 226 | Y | Y |
| | 2014-10-11 | TRAPPIST | I+z | 7 | 1029 | N | N |
| | 2014-10-18 | EulerCam | z' | 50 | 264 | Y | Y |
| | 2015-10-01 | EulerCam | I | 60 | 215 | Y | N |

**Note:** Details of the filters used for each observation are described in more detail with the descriptions of observing procedures for each telescope, where values for amount of defocussing used are also noted. Transit coverage can be seen in the lightcurves for each observation in the second panels of Figures 2.2–2.4

### 2.2.2 Photometric follow-up

Several telescopes were used to collect follow-up transit lightcurves of WASP-92b, WASP-93b and WASP-118b – the details of when the observations were made; what filters were used, and whether the images were defocused are included in Table 2.2. Further details of how these lightcurves were observed and the data reduced are included below.

#### Oversky observations

La Palma plays host to Oversky, which is a 0.36 m robotic telescope, operated primarily by amateur astronomers. A partial transit of WASP-92b was obtained with this telescope on the night beginning 2012 July 24th. The telescope has a $19.9' \times 19.9'$ field of view (FOV), and is operated entirely remotely. The camera used in this set up is an SBIG STL-1001E, which has a $1024 \times 1024$ pixels Kodak KAF-1001E CCD. The guiding of the telescope was performed internally throughout the observations. Each image was in focus and had an exposure time of 90 s, with 7 s in between each for CCD readout. The filter used was a Sloan r filter.

The images were bias, dark and dusk flat calibrated, and then the lightcurves were extracted using the Munipack Muniwin 2.0 software tools. The diferential photometry was performed with two reference stars and one check star in the field.

**NITES observations**

NITES (*N*ear *I*nfra-red *T*ransiting *E*xoplanet*S*) is a 0.4 m (f/10) Meade LX200GPS Schmidt-Cassegrain telescope, which is situated at the ORM in La Palma. The telescope is operated with targets and observing patterns requested remotely. The CCD installed on the telescope is a Finger Lakes Instrumentation Proline 4710, which has 1024 × 1024 deep-depleted pixels providing an $11' \times 11'$ FOV. The telescope is operated without a filter, and the peak efficiency of the CCD is at 800 nm (McCormac et al., 2014).

Defocussing was used during the transits observed, the resulting full width at half maximum (FWHM) of the point spread function (PSF) of the 2013 June 26 WASP-92 observation was $3.3''$; 2013 July 22 was $2.8''$, and $6.6''$ on 2012 August 21. PyRAF[1] and the wrapped IRAF[2] routines were used to subtract the bias frames, and to perform a flat-field correction on the images.

DAOPHOT (Stetson, 1987) was used to perform aperture photometry, with 8 comparison stars used in the field of WASP-92, and 5 stars for WASP-93 data. The aperture radius used for the first WASP-92 observation data was $5.9''$; $5.0''$ radius for the second WASP-92 transit, and $10.2''$ for the WASP-93 data. Aperture size selection was made by minimising the out of transit root mean square (RMS) noise. The target's electron noise, combined with the readout and sky noise for the aperture were used to determine the uncertainty values for each measurement. A first order line fit was also applied to each set of observations' out of transit data to normalise the lightcurves.

**RISE observations**

The 2 m Liverpool Telescope in La Palma is mounted with RISE (*R*apid *I*maging *S*earch for *E*xoplanets), which is an optical camera controlled robotically like the rest of the telescope operations (Gibson et al., 2008; Steele et al., 2008). The images collected for this work were observed with a fixed V+R filter, for which the passband has > 50% efficiency in the range

[1]PyRAF is a product of the Space Telescope Science Institute, which is operated by AURA for NASA.
[2]IRAF is distributed by the National Optical Astronomy Observatories, which are operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation.

500–700 nm. The FOV of the instrument is 9.4′ × 9.4′, and this is captured with a 1024 × 1024 pixel e2v CCD. IRAF routines were utilised to perform the debiasing and flat-fielding with twilight flats. Differential photometry was performed with the target star, and some nearby non-variable stars as comparison using PyRAF and DAOPHOT.

#### RATCam observations

RATCam is an optical imaging instrument, which is now decommissioned, but was installed on the Liverpool telescope in February 2014 (Mottram et al., 2004; Steele et al., 2004). The EEV CCD had 2048 × 2048 pixels, which when installed saw a 4.6′ × 4.6′ FOV. 10 s exposures were used for the images taken during the transit of WASP-93b on the night beginning 2008-07-10. IRAF routines were used to perform the bias subtraction, and the flat-fielding with a recent twilight flat image. IRAF was again used with DAOPHOT for the aperture photometry. Three nearby bright stars were used as comparison, once they had been confirmed as non-variable. Because of the short exposure time, the lightcurve was binned with each bin calculated from three adjacent lightcurve values.

#### PIRATE observations

PIRATE (*P*hysics *I*nnovations *R*obotic *A*stronomical *T*elescope *E*xplorer) is a robotic 0.43 m telescope, which was situated at the Observatori Astronomic de Mallorca for the observations in this work (Holmes et al., 2011; Kolb, 2014). The telescope has an SBIG STX-16803 camera, which contains a 4096 × 4096 pixel Kodak KAF-16803 CCD. The pixel scale is 0.63″ pixel$^{1}$ for the 43′ × 43′ FOV of the instrument. All of the transits observed with PIRATE were observed with a broad R-band filter, and with the telescope in focus. The images were collected without pixel binning, but for during the transit observation on the night beginning 2011-10-31, where a 2×2 pixel binning was applied. Holmes et al. (2011) describes how the images were reduced – bias, dark and dusk flat calibrated, and how differential ensemble photometry was performed using DAOPHOT.

#### JGT observations

The University of St Andrews Observatory houses the James Gregory Telescope (JGT), which has a 0.94 m aperture. The images from this telescope are taken with a 1024 × 1024 e2v CCD, which has a useable FOV of 15′ diameter. The lightcurves obtained with the JGT were

completed with the telescope in focus and a broad R-band filter. Routines from the STARLINK[3] software packages were used for the bias and flat-field corrections, and then GAIA from STARLINK was used for the aperture photometry. The lightcurves were constructed from relative photometry with a single nearby comparison star with a similar magnitude to the target star.

#### EulerCam observations

EulerCam is installed on the Swiss Euler Telescope at La Silla, Chile (Lendl et al., 2012). The 1.2 m telescope when combined with the 4k × 4k e2v deep-depletion silicon CCD has a 15.68′ × 15.73′ FOV. The telescope tracking was performed by a precise exposure to exposure movement tracking system, to minimise the movement of the stellar images across the CCD pixels. The images were also overscan, bias and flat-field corrected. Ensemble photometry was performed with anywhere between 3–10 comparison stars of similar colour and magnitude to the target star. The exact number of reference stars was determined iteratively so that RMS scatter in the lightcurve was minimised.

#### TRAPPIST observations

TRAPPIST (*TRA*nsiting *P*lanets and *P*lanetes*I*mals *S*mall *T*elescope) is a 0.6 m robotic telescope situated at La Silla, Chile (Jehin et al., 2011). The imaging CCD has 2048 × 2048 pixels, which when mounted on the telescope see a 22 × 22 ′ FOV. The TRAPPIST lightcurve of the transit of WASP-118 was obtained with an I+z filter, which is has an approximate FWHM in the region 750 – 1100 nm. The telescope is fitted with a high precision tracking system, which uses astrometric solutions of the imaged field to correct the pointing for every exposure, which minimises the movement of the stellar image relative to the CCD pixels. The images were bias, dark and flat-field corrected. Aperture differential photometry was then performed using IRAF/DAOPHOT2 routines with comparison stars selected to be non-variable and of a similar brightness to the target star.

### 2.2.3 Radial velocity follow-up

Radial velocity (RV) data for each of the star-planet systems was also obtained, where the effect of the planet on the centre of mass of the star was recorded in order to characterise the orbits. The data included in this work were obtained from three different instruments:

[3]STARLINK is maintained by the East Asian Observatory, and the routines are published as open source at http://starlink.eao.hawaii.edu/starlink

**Table 2.3:** Details of the observing intervals of the RV follow up

| System | Interval | Instrument | $N_{points}$ |
|---|---|---|---|
| WASP-92 | 2012-05-23 – 2012-07-26 | SOPHIE | 9 |
| WASP-93 | 2008-07-15 – 2010-11-28 | SOPHIE | 13 |
| | 2010-08-17 – 2010-08-18 | SOPHIE | 34 |
| | 2011-08-18 – 2011-08-19 | SOPHIE | 27 |
| WASP-118 | 2010-10-16 – 2010-10-19 | SOPHIE | 4 |
| | 2010-07-19 – 2014-08-26 | CORALIE | 23 |
| | 2011-07-29 – 2011-07-31 | SOPHIE | 2 |
| | 2015-07-23 – 2015-11-07 | CORALIE | 4 |
| | 2015-10-01 – 2015-10-02 | HARPS-N | 25$^{\star}$ |
| | 2015-10-04 – 2015-10-06 | HARPS-N | 3 |

**Note:** Observations have been grouped according to the instrumental configurations used. $^{\star}$Data set not included in the global fit for the determination of the final parameters presented.

SOPHIE, CORALIE and HARPS. Details of the observing windows of the datasets for each instrument are shown in Table 2.3, and the data themselves are plotted in the lower plots of Figures 2.2-2.4 folded on the best fit transit phases presented in Section 2.5.

SOPHIE (Perruchot et al., 2008; Bouchy et al., 2009) is the spectrograph at the Observatoire de Haute-Provence (OHP), installed on the 1.93 m telescope. The spectrograph is fed with two fibres – one pointed to the target star, and the other to the sky. Using the second fibre in this way allows for observation specific sky and moon subtraction from the stellar spectral data. The targets were observed in two different modes with SOPHIE – high resolution and high efficiency, which have different baseline offsets in the RV measurements (Bouchy et al., 2009). Due to the potential offsets between datasets, observations made with each mode were treated as from different instruments, for which an individual offset value was included in the model fit.

WASP-93 was observed on two occasions at a high cadence over individual nights with SOPHIE. The collection of observations were designed to match with the predicted transit time to quantify the Rossiter-McLaughlin (RM) effect for the system.

The CORALIE spectrograph (Baranne et al., 1996; Queloz et al., 2000a) was also used to observe WASP-118. The instrument is installed on the Swiss Euler 1.2 m telescope in La Silla. In contrast to SOPHIE, the second fibre for CORALIE is pointed to a calibration lamp rather than the sky, which means that the spectral observations are only taken with CORALIE when the moon is not present to prevent contamination. November 2014 saw upgrades to the

instrument, which may have introduced an offset in the RV measurements, thus data taken before and after the upgrade are treated as separate in the analysis.

HARPS-N (Cosentino et al., 2012) was also utilised for observations of WASP-118. The HARPS-N spectrograph is mounted on the Telescopio Nazionale Galileo (TNG) on La Palma, and uses the same technique as SOPHIE for sky and moon subtraction – the second fibre is pointed to the sky during observations. An attempt was made to observe the RM effect during the transit of WASP-118 on the night beginning 2015 October 1st, however poor weather adversely affected the collected spectra.

To transform the observed stellar spectra from each instrument into an RV measurement, the data were convolved using a weighted cross-correlation function (CCF) with a numerical G2-spectral template (Pepe et al., 2002). Whilst all of the host stars in this work are hotter than a G2 spectral type, this template is the hottest available for each instrument, and is regularly used for F type stars. The value for each RV is determined by a Gaussian fit to the mean CCF output for each of the spectral orders. The HARPS-N data were reduced using the data reduction software (DRS) version 3.7.

Other features of the average spectral line profiles, which is represented by the CCF output, were determined for each measurement. The features were: full width at half maximum (FWHM), line bisector inverse slope (BIS) and the line contrast. The uncertainties for each of these values were scaled from the factors presented in Santerne et al. (2015) and the RV uncertainties determined with each of the instruments' data reduction pipelines. The features described above were not obtainable for the observations made by SOPHIE of WASP-93, since the CCF is particularly broad.

The William Herschel Telescope (WHT) was also used with the red arm of the Intermediate dispersion Spectrograph and Imaging System (ISIS) to observe an H$\alpha$ spectrum of WASP-93. The observation was made on 2012 December 23, and covers 6161–6953 Å. The PAMELA and MOLLY software (Marsh, 1989) were used to flux-calibrate and reduce the spectrum[4]

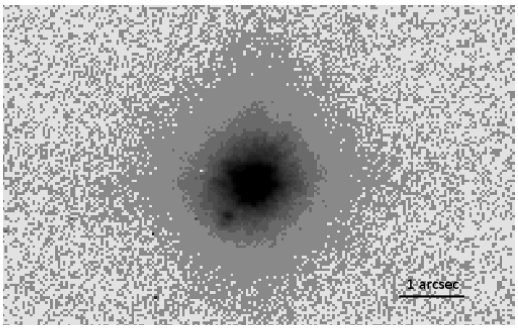



**Figure 2.1:** H-band image of WASP-93 with correction to 0.14 arcsec, which was taken with INGRID on the WHT in 2011. This image was observed in telescope sky position angle 0 ° North is up, and East is to the right.

### 2.2.4 Adaptive optics imaging follow-up

Since WASP-93 is such a fast rotating star, it is difficult to obtain high precision RV measurements to confirm the planetary nature of the transit signal. As an extra test of whether the observed transits could have been introduced by a background eclipsing binary, high resolution near-infrared images of the star were obtained. On the night of 2011 August 13, the near-infrared camera (INGRID) fed by the NAOMI adaptive optics (AO) system installed on the 4.2 m WHT was used to take high resolution images in the J, H and Ks bands. These images are a useful tool for observing whether there are any objects blended with the primary star, which would be otherwise unresolved with lower resolution imagery.

The images were obtained at three sky position angles (0, 20, 35 ° from zenith), which ensured that any inhomogeneity in the PSF observed in every image can be attributed to a real effect, rather than a false positive introduced by features of the PSF at particular angles. On the night of the observations, the natural seeing in the H band was between 0.6 and 0.8 ″. With the corrections provided by the AO system, the observed FWHM for the H and Ks bands was 0.15 ″, and 0.20 ″ for the J band.

The H-band image is shown in Figure 2.1, where a nearby stellar companion can be seen 0.69 ± 0.01 ″ from the central star. The companion is at position angle 220.89 ± 0.60 °, and is fainter by 3.70 ± 0.18 magnitudes in the J band; 3.45 ± 0.10 magnitudes in the H band, and 3.37 ± 0.13 magnitudes in the Ks band. The position of the faint companion relative to North is unchanged in all images at each sky position angle, which verifies that this companion is

[4]Details of the software packages can be found at
http://www2.warwick.ac.uk/physics/research/astro/people/marsh/software .

not an artefact of the PSF. From the 2MASS survey (Skrutskie et al., 2006), the magnitudes of WASP-93 observed with the large apertures of the 2MASS survey can be used to determine that the companion has the following magnitudes: J=13.93 ± 0.18, H=13.45 ± 0.10 and Ks=13.31 ± 0.13. If associated with WASP-93, the companion must be an early to mid K dwarf.

As is indicated in Figure 2.1 and verified in conjunction with the other images, there are no other sources detected in the region of WASP-93. INGRID has a 40 $''$ FOV, where no other objects are observed down to J $\sim$ 19.0 ± 0.35, H $\sim$ 19.0 ± 0.35 and Ks $\sim$ 18.2 ± 0.35 at distances $>$ 4 FWHM from the centre of the corrected stellar profile. No sources are detected outwith 1.5 FWHM from the stellar core to magnitudes of J $\sim$ 17.0, H $\sim$ 17.0 and Ks $\sim$ 16.2. All of these sets of limits are 5$\sigma$ limiting magnitudes for detections.

## 2.3 Stellar parameters from spectra

For each of the systems presented here, several high quality spectral observations were made for the primary purpose of measuring the RV signal, however the obtained spectra can also be utilised to understand more about the host stars.

### 2.3.1 WASP-92 stellar parameters

The SOPHIE spectra for WASP-92 collected without moon pollution were combined to a single spectrum, which had a signal to noise ratio (SNR) of around 25:1 across the range of wavelength for the combined spectrum. The spectral analysis presented here made use of the methods set out in Doyle et al. (2013). The derived spectral parameters described below for WASP-92 are presented in the upper section of Table 2.5.

The excitation balance of the Fe I lines was used to define the effective temperature ($T_{\rm eff}$). A measure of surface gravity ($\log g$) was determined from the Na I D lines and the ionisation balance of Fe I and Fe II. The iron abundance ([Fe/H]) was measured from the equivalent widths of several unblended Fe lines. Microturbulence ($\xi_{\rm t}$) was quantified from Fe I via the method of Magain (1984).

The uncertainties quoted are propagated through measurement and atomic data uncertainties. The presence of interstellar Na D lines allow an estimation of the extinction. The interstellar lines had equivalent widths of $\sim$0.02 Å, which relates to an extinction of $E(B-V) =$

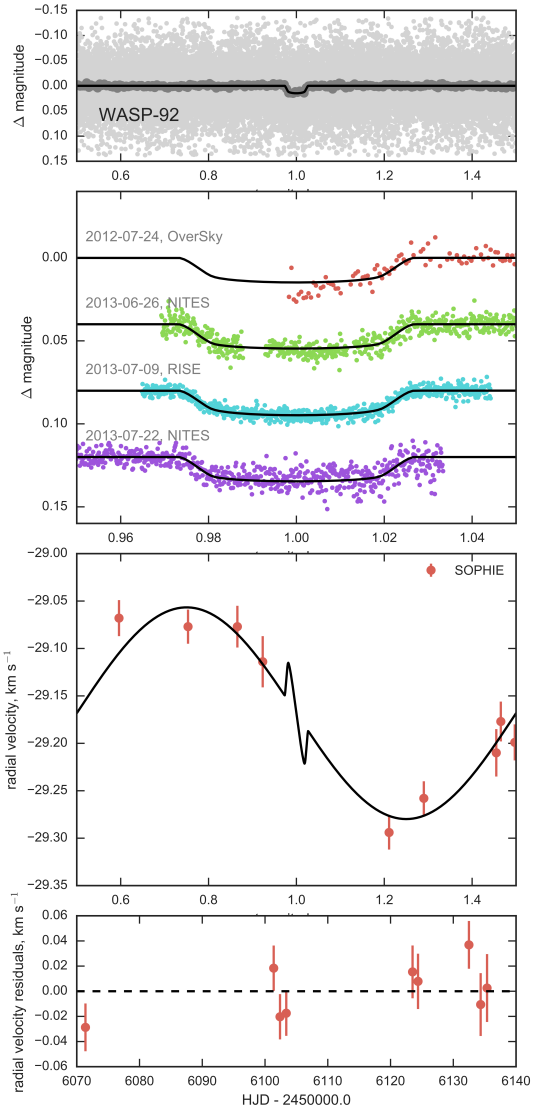



**Figure 2.2:** Figure showing the observations collected for WASP-92 and used in the global fit described in Section 2.5. The upper plot is of the photometry in collected by the WASP cameras, shown in light grey, which is folded onto best-fit orbital ephemeris presented in this work. The folded data are overlaid in dark grey with the data binned in 1/200th of the phase for clarity in the transit shape, and the transit model is shown in black. The upper middle plot shows the follow up photometry with reference to transit phase, the order presented is the same as for Table 2.2. The lower middle plot shows the RV points with reference to transit phase, and the residuals of the fit with reference to time are shown in the lower plot.

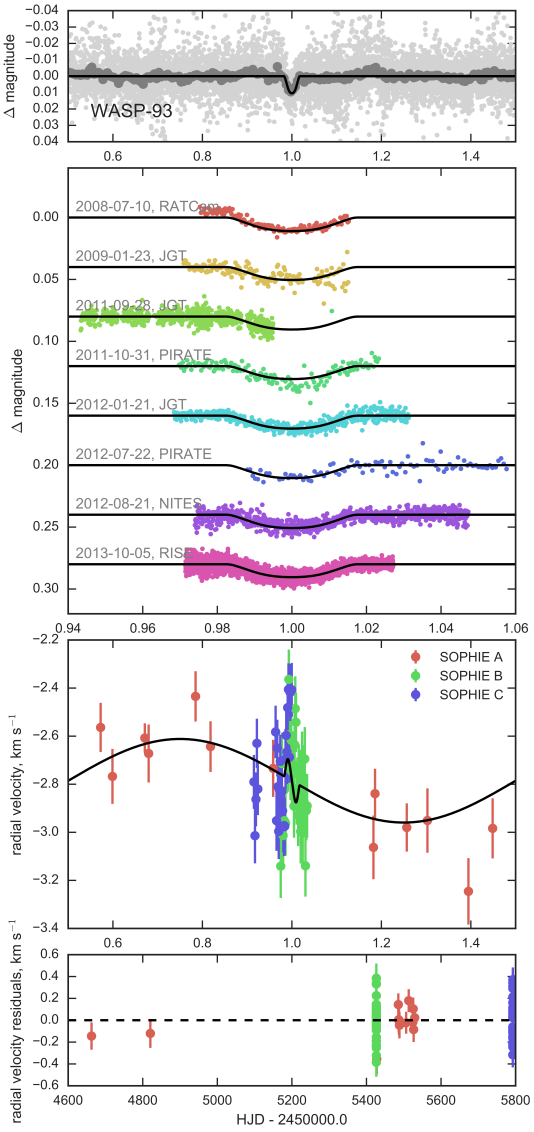



**Figure 2.3:** Figure showing the data collected for the WASP-93 system. The plots are presented in the same way as for WASP-92 in Figure 2.2.

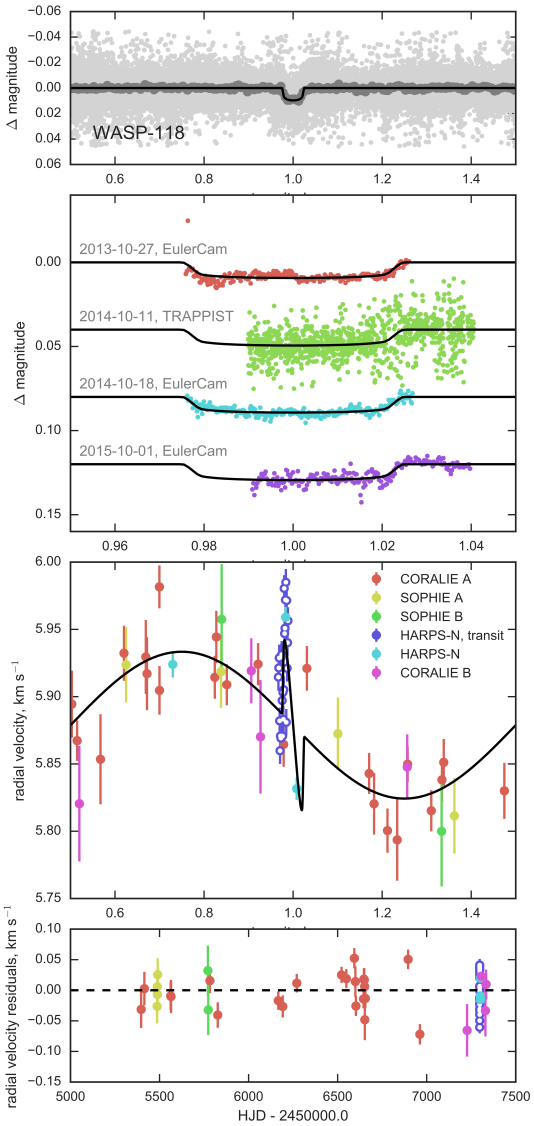



**Figure 2.4:** Figure showing the observations made of WASP-118, which are presented in the same way as described in the caption of Figure 2.2. The HARPS-N data taken as a series during a transit (shown in purple) have open circle markers to denote that the observations were not included in the global fit.

$0.005 \pm 0.001$ through the calibration of Munari & Zwitter (1997).

The instrumental FWHM of $0.15 \pm 0.01$ Å was measured from the profiles of the telluric lines around 6300 Å, which was then utilised to calibrate the measures of the projected stellar rotation velocity ($v \sin i$) and macroturbulence ($v_{\rm mac}$). The $v \sin i$ was determined from fitting the average line profiles of several unblended Fe I lines, and $v_{\rm mac}$ for WASP-92 was estimated from the calibration in Doyle et al. (2014).

The non-detection of lithium in the spectrum with an upper limit on the equivalent width of 52 mÅ, leads to an abundance upper limit of $\log A({\rm Li}) = 2.70 \pm 0.09$, and a stellar age estimate of several Myr from the calibration of Sestito & Randich (2005).

The $v \sin i$ combined with the estimate of the stellar radius implied by the values of $T_{\rm eff}$ and $\log g$ can provide a constraint on the stellar rotation period. In this case, the rotation is estimated to be $10.07 \pm 2.58$ days. The Barnes (2007) relation for gyrochronology estimates the age of the star to be $\sim 2.29^{+6.80}_{1.51}$ Gyr.

### 2.3.2 WASP-93 stellar parameters

The methods for determining the stellar parameters for WASP-93 were the same as for WASP-92, where in this case the co-added spectrum used for the analysis had an average SNR of 50:1. The broad lines found in the spectrum of WASP-93 due to fast stellar rotation presented some difficulties in measuring sufficient Fe lines to find a value for the $T_{\rm eff}$, so the ISIS spectrum of the H$\alpha$ region (mentioned in Section 2.2.3) was used to measure the Balmer lines and estimate the $T_{\rm eff}$. The broad lines also prevented a determination of a value for $\xi_{\rm t}$. The interstellar Na D lines in the spectrum allowed the extinction of WASP-93 to be measured as $E(B-V) = 0.07 \pm 0.02$.

A value for $v_{\rm mac}$ was found for WASP-93, however the effect of $v_{\rm mac}$ on line broadening is negligible when the $v \sin i$ is so high.

Lithium was not detected in the spectrum, with an upper limit on the equivalent width of lithium of 0.01 mÅ, which corresponds to an upper limit on its abundance of $\log A({\rm Li}) = 1.14 \pm 0.09$. The same relation as above indicates this implies a stellar age of several Gyr.

The $v \sin i$ and stellar radius estimate suggest a stellar rotation period of $1.45 \pm 0.4$ days, which indicates a gyrochronological age of $\sim 0.70 \pm 0.65$ Gyr.

### 2.3.3 WASP-118 stellar parameters

Whilst WASP-118 was observed with SOPHIE, there are significantly more spectra of the star from the CORALIE instrument, thus that data was used to create a co-added spectrum in the same way, which had an average SNR of 50:1. The same techniques to derive the stellar parameters for WASP-92 were also used for WASP-118, with the final parameters presented in Table 2.5.

For WASP-118, the interstellar Na D lines allowed the estimation of the extinction to be $E(B-V) = 0.10 \pm 0.03$. The instrumental FWHM was found to be 0.11 ± 0.01 Å for CORALIE, using the same technique as for the SOPHIE data.

As for WASP-93, there was no significant detection of lithium in the combined spectrum of WASP-118. This corresponds to an upper limit on the equivalent width of 4 m Å, and an upper limit on the abundance of $\log A(\mathrm{Li}) = 1.21 \pm 0.09$, which relates to a stellar age of several Gyr. The rotation rate of the star is estimated to be 6.12 ± 1.11 days from the values of $v \sin i$ and the derived stellar radius, which implies a gyrochronological age of $\sim 1.17^{+5.72}_{-0.75}$ Gyr.

### 2.3.4 Stellar activity analysis

All of these systems have been observed to have stable planetary transit signals, as well as a periodic RV signal in phase with the transits, thus the risk that the signals detected could be mimicked by activity on the host star is very low, particularly as the detected planets are large. It is however advisable to perform checks for correlations between the detection parameters and any indicators of stellar activity.

The WASP photometric datasets were searched for indications of stellar rotation through fitting sinusoids to the data with the method presented in Maxted et al. (2011). No such signals were detected to 3 mmag precision in any of WASP-92, 93 or 118. This is also a strong indication that stellar spot activity is minimal on all of the host stars, which is further evidence in favour of the planetary scenario in each case.

The photometric transit signals are not the only evidence of planets that can be mimicked by stellar activity, periodic signals in the RV data can be introduced by plagues or spots on the surface of the star. The effects of stellar surface phenomenon on the shape of the average absorption line profile of the stellar spectrum are different to those introduced by the pres-

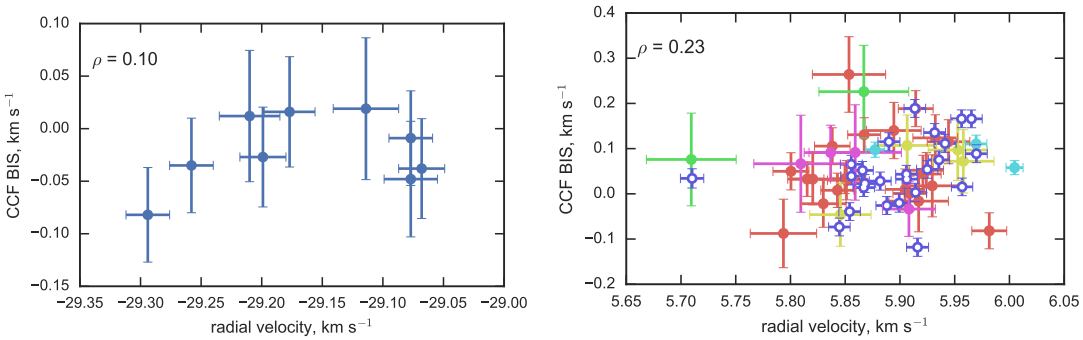



**Figure 2.5:** The distribution of CCF BIS with measured RV for WASP-92 (upper plot) and WASP-118 (lower plot), where the colours are equivalent to those used for the RV data in Figures 2.2 and 2.4. The Spearman's rank correlation coefficients ($\rho$) are shown in the upper left of the plots.

ence of a planet (Santerne et al., 2015), thus we can look for correlations between the RV measurements and the observed BIS (the absence of BIS measurements for WASP-93 data was discussed in Section 2.2.3). Figure 2.5 demonstrates the distribution of RV against BIS for WASP-92 and WASP-118. There is low correlation measured between the RV and BIS measurements, as can be seen by eye in Figure 2.5 and with the low values of the Spearman's rank correlation coefficient. The lack of strong correlation adds no evidence against the proposal that the observed periodic RV signals and transit lightcurves are originated from planetary sources.

## 2.4 Planetary system parameter determination

Determining the fundamental parameters associated with a star-planet system is a challenging statistical operation. The observational data are related to the parameters indirectly, each is non-linearly related to the other, and non-linear correlations exist between some of the parameters. Fitting the required model to the data, as well as estimating the uncertainty for each parameter is therefore best suited to Markov Chain Monte Carlo (MCMC) methods. Using an MCMC to fit the model allows not just the optimisation of the parameters, but also an exploration of the likelihood space surrounding the optimal parameter values. This ensures that estimates of the uncertainties are robust in the case of correlation between parameters. The methods used for the parameter determination of WASP consortium discoveries have made use of MCMC methods, the primary methods used were presented in Cameron et al. (2007) and expanded upon in Pollacco et al. (2008).

The core concept of an MCMC is that the selected parameters make a random walk through

**Table 2.4:** Details of jump parameters used in final global MCMC analysis

| Core parameters | | |
|---|---|---|
| Parameter | Definition | Prior shape |
| $T_0$ | Transit epoch, HJD - 2450000.0 | Uniform |
| $P$ | Orbital period, days | Uniform |
| $\Delta F$ | Transit depth, magnitudes | Uniform |
| $t_T$ | Transit duration, days | Uniform |
| $b$ | Impact parameter, $R_\star$ | Uniform |
| $T_{\rm eff}$ | Stellar effective temperature, K | Gaussian |
| [Fe/H] | Stellar abundance | Gaussian |
| $K_1$ | RV semi-amplitude, km $s^{-1}$ | Uniform |
| Optional parameters | | |
| $\sqrt{e}\cos\omega^a$ | Eccentricity proxy | Uniform |
| $\sqrt{e}\sin\omega^a$ | Eccentricity proxy | Uniform |
| $\sqrt{v\sin i}\cos\lambda^b$ | Misalignment proxy | Uniform |
| $\sqrt{v\sin i}\sin\lambda^b$ | Misalignment proxy | Uniform |
| $FWHM_p{}^c$ | FWHM of planet profile, km $s^{-1}$ | Gaussian |
| $v\sin i^d$ | Stellar rotation speed, km $s^{-1}$ | Uniform |
| $\lambda^d$ | Spin-orbit misalignment angle, ° | Uniform |

**Note:** [a] parameters used only when a non-circular orbit is allowed. [b] parameters used when fitting the RM effect. [c] parameter used when fitting tomographic transit. If $v\sin i >> 0$ parameters denoted [d] are used as jump parameters for tomographic transit fitting.

randomly generated values, where the randomness of the path is constrained by the likelihood generated from each set of parameter values related to the observed data from the previous step. The parameter values generated are referred to as trial values, and the set of selected parameters are known as jump parameters. The method to control the randomness of the walk is the Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970), which defines whether proposed trial values should be accepted. If the likelihood for the trial values is higher than that for the previous set, then the step is accepted. In the case of a lower likelihood for the new values, the Metropolis-Hastings algorithm determines whether the step is accepted with probability $\exp(\frac{\Delta Q}{2})$, where $\Delta Q = Q_i - Q_{i-1}$, and $Q$ refers to the log of the likelihood for each trial.

The parameters used to characterise the system as jump parameters in the MCMC are shown in Figure 2.4, where the shape of prior probability distribution applied for each parameter is also listed. A uniform prior means that no likelihood penalty is applied at whichever value that parameter is for that trial, unless it is outside some pre-defined bounds, where the likelihood is set to a very low value. A Gaussian prior makes the assumption that the parameter must sit near a value with some standard deviation, setting the probability distribution in this way is reserved for parameters for which we have prior information and are not strongly constrained by the transit and RV data.

The transit data was modelled using the methods set out by Mandel & Agol (2002), which analytically define the amount of starlight blocked by the intersection of two overlapping disks. The larger of the disks is limb darkened, and for this case the non-linear limb darkening law as presented in Claret (2000) is used. The jump parameters used to represent limb darkening were $T_{\rm eff}$ and [Fe/H], which were transformed into the four coefficients of the non-linear limb darkening law by interpolating through the tabulated coefficients in Claret (2000).

To model the RV data a simple Keplerian orbit was applied, where each separate set of RV data was assigned a zero-point floating offset value, which was optimised at each MCMC step through inverse variance weighted averaging of the RV residuals.

Selecting the jump parameters to use for the global fit is not a trivial choice, since each can have correlations with other parameters, or introduce biases in the direction that parameters jump. The approach to parameter choice used for characterising WASP collaboration planet discoveries has been developed since the original presentation of the fitting method

in Cameron et al. (2007). Enoch et al. (2010); Anderson et al. (2011) ensured that when including eccentricity in a fit, the choice of jump parameters did not bias towards very small eccentricities, so eccentricity is now parameterised as $\sqrt{e}\sin\omega$ and $\sqrt{e}\cos\omega$, where $e$ is eccentricity and $\omega$ is the longitude of periastron.

For a fit including the RM effect, the Ohta et al. (2005) model was applied, where $\sqrt{v\sin i}\sin\lambda$ and $\sqrt{v\sin i}\cos\lambda$ were used as jump parameters. $\lambda$ refers to the projected misalignment between the rotation axis of the star and the path of the planet's orbit. The choice of the form of the jump parameters is dictated by the same logic as for eccentricity, and to mitigate against the effects of the correlation between the $v\sin i$ and $\lambda$.

The MCMC fitting was performed for each of the systems with 5,000 steps of burn-in, to ensure that all of the jump parameters are initialised to the optimal values, and followed by 10,000 steps of a production chain. The initial values for the jump parameters for the burn-in chain were the parameters found by the initial WASP survey detection (method described in Section 2.2.1), and the stellar parameters from the spectral analysis (described in Section 2.3). The median parameter values for the final 1,000 steps of the burn-in phase were then used to initialise the jump parameters for the production chains.

The final parameters found for the fits are presented in Table 2.5. Circularity of the the orbits was assumed when initially fitting to reduce the number of parameters to optimise, but the possibility of eccentricity should also be investigated. To test for a significant detection of a non-circular orbit, the burn-in phase was applied with a circular orbit, and followed by another run of the same length without the same assumption of circularity. The production run was then completed with the inclusion of the parameters relating to eccentricity.

To determine whether including the parameters for eccentricity is statistically justified, the Bayesian Information Criterion (BIC) (Schwarz et al., 1978; Kass & Raftery, 1995) was used. The BIC offers a means to select between fits of different models. The likelihood of a fit to some data can be improved by adding more parameters, however more parameters can lead to overfitting – the BIC adds a penalty for the number of parameters used. A lower value for the BIC implies a more justified fit. According to Kass & Raftery (1995), $\Delta$BIC $> 10$ is strong evidence for the selection of the model with a lower value of the BIC. For all of the systems presented here, the $\Delta\mathrm{BIC} = \mathrm{BIC}_{\mathrm{ecc}} - \mathrm{BIC}_{\mathrm{circ}} > 10$, so the circular solutions were accepted.

**Table 2.5:** Table of planetary and stellar parameters found in the global system fit for WASP-92, 93 and 118.

| Parameter | Symbol | WASP-92 | WASP-93 | WASP-118 | Units |
|---|---|---|---|---|---|
| General information | | | | | |
| RA | | 16h 26m 46.08s | 00h 37m 50.11s | 01h 18m 12.12s | J2000 |
| Dec | | +51° 02' 28.2" | +51° 17' 19.5" | +02° 42' 10.2" | J2000 |
| V magnitude | | 13.18 | 10.97 | 11.02 | |
| K magnitude | | 11.52 | 9.94 | 9.79 | |
| Spectroscopic analysis results | | | | | |
| Effective temperature | $T_{\rm eff}$ | 6280 ± 120 | 6700 ± 100 | 6410 ± 125 | K |
| Stellar surface gravity | $\log g$ | 4.40 ± 0.12 | 4.5 ± 0.20 | 4.30 ± 0.10 | |
| Projected stellar rotation speed | $v \sin i$ | 5.73 ± 1.15 | 37 ± 3 | 9.68 ± 1.14 | km s$^{-1}$ |
| Macroturbulence | $v_{\rm mac}$ | 4.97 ± 0.73 | 6.95 ± 0.73 | 5.77 ± 0.73 | km s$^{-1}$ |
| Microturbulence | $\xi_{\rm t}$ | 1.05 ± 0.10 | – | 1.00 ± 0.06 | km s$^{-1}$ |
| Metallicity | [Fe/H] | 0.00 ± 0.14 | 0.07 ± 0.17 | 0.16 ± 0.11 | |
| Lithium abundance | $\log A$(Li) | 2.70 ± 0.09 | < 1.14 ± 0.09 | < 1.21 ± 0.09 | |
| Stellar mass | $M_*$ | 1.20 ± 0.10 | 1.30 ± 0.11 | 1.03 ± 0.08 | $M_\odot$ |
| Stellar radius | $R_*$ | 1.14 ± 0.18 | 1.06 ± 0.26 | 1.17 ± 0.16 | $R_\odot$ |
| Spectral Type | | F7 | F4 | F6 | |
| Distance | d | 530 ± 90 | 250 ± 60 | 250 ± 35 | pc |
| MCMC jump parameters | | | | | |
| Transit epoch (HJD-2450000.0) | $T_0$ | 6381.28340 ± 0.00027 | 6079.56420 ± 0.00045 | 6787.81423 ± 0.00062 | days |
| Orbital period | $P$ | 2.1746742 ± 0.0000016 | 2.7325321 ± 0.0000020 | 4.0460435 ± 0.0000044 | days |
| Planet/star area ratio | $(R_{\rm p}/R_*)^2$ | 0.01254 ± 0.00029 | 0.01097 ± 0.00013 | 0.00755 ± 0.00019 | |
| Transit duration | $t_T$ | 0.1153 ± 0.0012 | 0.0931 ± 0.0010 | 0.2002 ± 0.0019 | days |
| Impact parameter | $b$ | 0.608 ± 0.043 | 0.9036 ± 0.009 | 0.16 ± 0.10 | $R_*$ |
| Stellar effective temperature | $T_{\rm eff}$ | 6258. ± 120 | 6696. ± 101. | 6420. ± 121. | K |
| Metallicity | [Fe/H] | 0.00 ± 0.14 | 0.06 ± 0.17 | 0.16 ± 0.11 | |
| Stellar reflex velocity | $K_1$ | 0.1116 ± 0.0092 | 0.174 ± 0.034 | 0.0546 ± 0.0019 | km s$^{-1}$ |
| Deduced stellar parameters | | | | | |
| Stellar density | $\rho_*$ | 0.495 ± 0.055 | 0.376 ± 0.028 | 0.271 ± 0.012 | $\rho_\odot$ |
| Stellar surface gravity | $\log g_*$ | 4.259 ± 0.031 | 4.197 ± 0.019 | 4.100 ± 0.012 | (cgs) |
| Projected stellar rotation speed | $v \sin i$ | 5.73 (fixed) | 37.0 (fixed) | 9.68 (fixed) | km s$^{-1}$ |
| Stellar radius | $R_*$ | 1.341 ± 0.058 | 1.524 ± 0.040 | 1.696 ± 0.029 | $R_\odot$ |
| Stellar mass | $M_*$ | 1.190 ± 0.037 | 1.334 ± 0.033 | 1.320 ± 0.035 | $M_\odot$ |
| Scaled stellar radius | $R_*/a$ | 0.1790 ± 0.0067 | 0.1684 ± 0.0037 | 0.1445 ± 0.0022 | |
| Centre-of-mass velocity | $\gamma$ | −29.1683 ± 0.0067 | −2.785 ± 0.029 | 5.8826 ± 0.0038 | km s$^{-1}$ |
| Deduced planet parameters | | | | | |
| Orbital separation | $a$ | 0.03480 ± 0.00036 | 0.04211 ± 0.00035 | 0.05453 ± 0.00048 | AU |
| Orbital inclination | $i$ | 83.75 ± 0.69 | 81.18 ± 0.29 | 88.70 ± 0.90 | ° |
| Orbital eccentricity | $e$ | 0 (fixed) | 0 (fixed) | 0 (fixed) | |
| Argument of periastron | $\omega$ | 0 (fixed) | 0 (fixed) | 0 (fixed) | ° |
| Spin-orbit misalignment angle | $\lambda$ | 0 (fixed) | 0 (fixed) | 0 (fixed) | ° |
| Planet radius | $R_{\rm p}$ | 1.461 ± 0.077 | 1.597 ± 0.077 | 1.440 ± 0.036 | $R_{\rm J}$ |
| Planet mass | $M_{\rm p}$ | 0.805 ± 0.068 | 1.47 ± 0.29 | 0.514 ± 0.020 | $M_{\rm J}$ |
| Planet surface gravity | $\log g_{\rm p}$ | 2.939 ± 0.056 | 3.120 ± 0.093 | 2.757 ± 0.040 | (cgs) |
| Planet density | $\rho_{\rm p}$ | 0.260 ± 0.044 | 0.360 ± 0.084 | 0.175 ± 0.018 | $\rho_{\rm J}$ |
| Planetary equilibrium temperature | $T_{\rm P}$ | 1871. ± 56. | 1942. ± 38. | 1729. ± 36. | K |
| Roche limit | $a_R$ | 0.0171 ± 0.0010 | 0.0157 ± 0.0015 | 0.0204 ± 0.0009 | AU |
| Roche separation | $a/a_R$ | 2.038 ± 0.026 | 2.685 ± 0.245 | 2.671 ± 0.099 | |

**Note on spectroscopic parameters:** Mass and Radius estimate using the Torres et al. (2010) calibration. Spectral Type estimated from $T_{\rm eff}$ using the table in Gray (2008). Abundances are relative to the solar values obtained by Asplund et al. (2009). **Note on MCMC parameters:** The Roche limit is $a_R = 2.16 R_{\rm p}(M_*/M_{\rm p})^{1/3}$, as defined in Ford & Rasio (2006).

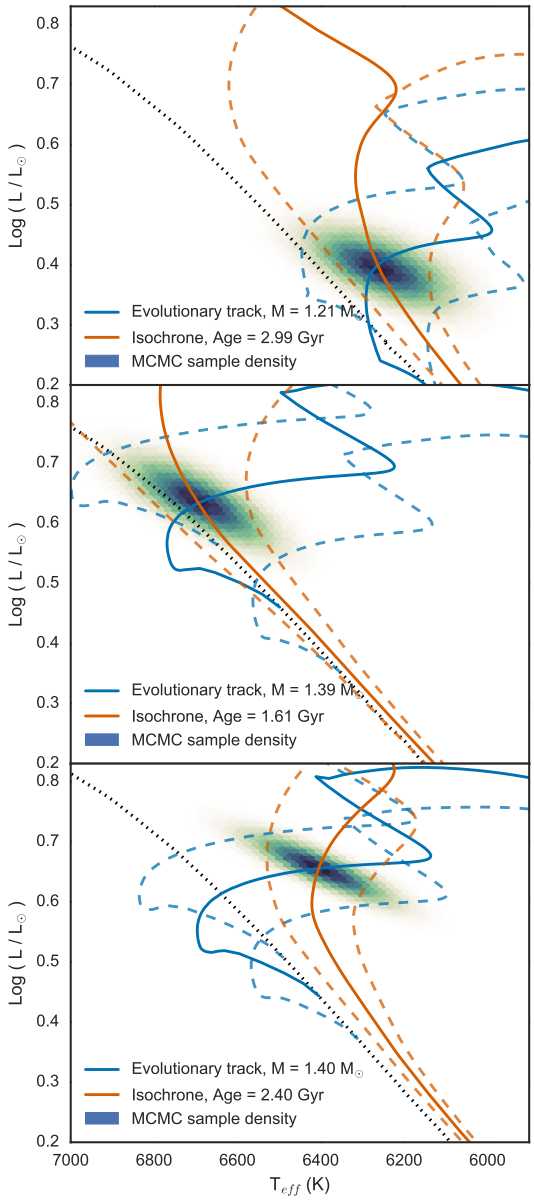



**Figure 2.6:** Results of the BAGEMASS MCMC analysis for WASP-92 (upper plot); WASP-93 (middle), and WASP-118 (lower). For each of the plots, the dotted black line is the ZAMS. The solid blue line is the evolutionary track for the mass found, and the dashed tracks either side are for the 1-$\sigma$ error of the mass. The solid orange line is the isochrone for the stellar age found, with the 1-$\sigma$ error denoted by dashed lines in the same colour. The density of MCMC samples is shown in the colour scale of the posterior distribution plotted.

## 2.5 Results and discussion

### 2.5.1 WASP-92 system

As presented in Table 2.5, WASP-92b is a ∼0.81 $M_J$ planet, which orbits its F7 type star every 2.17 days. When fitting the orbit of WASP-92b, the possibility of a non-circular orbit was

investigated. With the inclusion of the eccentricity parameters, a value of $e = 0.084^{+0.118}_{-0.060}$ was found. Including the eccentricity parameters effected the other transit parameters by <5%, and a BIC for the eccentric fit was found to be 19.1. In comparison, the circular fit resulted in a BIC of 5.7. The $\Delta$BIC is therefore 13.4 in favour of not including the two extra parameters required to include eccentricity, so the circular fit was used for the final quoted parameters.

To test that the observed stellar parameters and those derived from the MCMC analysis are consistent with stellar evolution models, the BAGEMASS tool from Maxted et al. (2015) was utilised. The tool takes observations with uncertainties of [Fe/H], $T_{\rm eff}$ and $\rho_\star$ as input to estimate the age and mass of the host star. BAGEMASS makes use of the GARSTEC code (Weiss & Schlattl, 2008) to create a large grid of stellar evolutionary models to interpolate between and find the optimal mass and age of the star using MCMC methods. For WASP-92, the optimal age found was $2.99 \pm 1.03$ Gyr and a mass of $1.21 \pm 0.06\ M_\odot$. The upper plot of Figure 2.6 shows the posteriors produced for this star, as well as the best fit evolutionary tracks and isochrones. The age found by BAGEMASS is consistent with the gyrochronological age presented in Section 2.3.1, but significantly older than the age suggested by the spectral lithium abundance. The masses found in the spectral analysis, the MCMC analysis and with BAGEMASS are also broadly consistent. The stellar mass determined with the MCMC fit of a circular orbit has been used to calculate the final quoted planetary mass.

**Blend scenario**

WASP-92 has not been observed with AO imaging, which could be used rule out the possibility of the planetary signal being mimicked by a blended eclipsing binary. Without this extra data, the planetary nature of WASP-92b can still be confirmed. If there were an associated binary system producing the transit signals, the RV amplitude observed would be significantly higher than is observed. If a background blended object were mimicking the transits, a periodic RV signal phased with the transits would not be expected. For a blended object to produce such a deep transit signal, signs of the spectral lines of the object may be seen in the spectra of WASP-92, which is not the case.

The extent to which WASP-92 is blended with background objects could be investigated with AO imaging, which dilute the transit depth and RV amplitude. The fact that WASP-92b has an inflated radius implies that any dilution is minimal.

### Tidal evolution

Since hot Jupiters orbit close to their host stars, tidal interaction between the star and the giant planet is expected. This interaction means that the orbits of hot Jupiters are not stable over long timescales (Matsumura et al., 2010). Tidal interactions between planet and star result in angular momentum transfer between the two objects. A Darwin stable system is defined as the objects orbiting with coplanarity, circularity and co-rotation, which is stable as it is the minimal energy state of the orbital system (Darwin, 1879) constrained by the conservation of angular momentum. This stable orbit can only exist if there is enough angular momentum in the system – if the total angular momentum of the system ($L_{\text{tot}}$) is larger than some critical angular momentum, $L_c$,

$$L_c = 4\left(\frac{G^2}{27}\frac{M_*^3 M_p^3}{M_* + M_p}\left(C_* + C_p\right)\right)^{\frac{1}{4}}, \tag{2.1}$$

where $C_*$ and $C_p$ are the moments of inertia of the star and planet respectively (Counselman, 1973; Hut, 1980). The total angular momentum is defined as

$$L_{\text{tot}} = L_{orb} + C_*\omega_* + C_p\omega_p\ , \tag{2.2}$$

where $\omega_*$ and $\omega_p$ are the angular velocities of rotation of the star and planet respectively, and $L_{orb} = M_* M_p \sqrt{\frac{Ga(1-e^2)}{M_* + M_p}}$. For a hot Jupiter, the planetary rotation period will match the orbital period of the planet, as the tidal dissipation timescale for the planet is very short, so it synchronises very quickly with the orbital period, the planet is then tidally locked. The system is referred to as dual-synchronous when the rotation period of the star is also synchronised with the orbital period of the planet.

As can be seen in the upper plot in Figure 2.7, for WASP-92 $L_{\text{tot}}$ (shown in green) is less than $L_c$ (the minimum of the blue line), so the orbit can never reach a stable state as there is not enough angular momentum in the system. $L_{tot}/L_c \sim 0.67$, which means than the planet is migrating towards the star, towards the Roche limit, as is the case for most hot Jupiters (Levrard et al., 2009; Matsumura et al., 2010). These calculations have made the assumption that the planetary orbit is circular, which is justified by the observations (as discussed above). The assumption has also been made that the stellar spin is co-planar with the planetary orbit, as there is no observational evidence that the orbit is misaligned. If the orbit were misaligned, the conclusion that the WASP-92 system does not have enough angular momentum to support

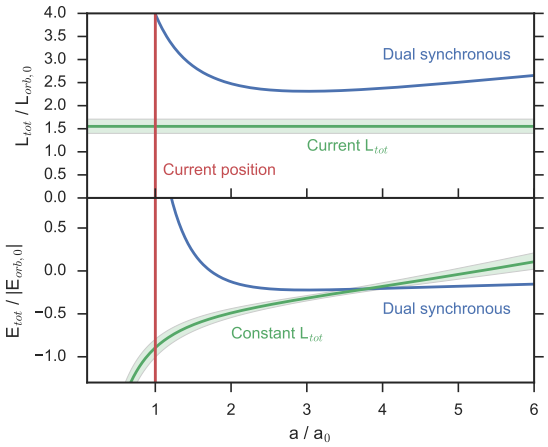



**Figure 2.7:** Plot showing the tidal equilibrium curves for WASP-92. In the upper plot, the blue line shows the total angular momentum of the system when dual-synchronised for the range of semi-major axis; the green line shows the current total angular momentum with an assumption of spin-orbit alignment, and the red line shows the current separation of the star and planet. The angular momenta are scaled with the current total orbital angular momentum, $L_{orb,0} = 1.398 \times 10^{42}$ kg m$^2$/s. The lower plot shows the curves for the total energy in the system, where the blue line shows the total orbital and rotational energy for the system when dual synchronised for the range of orbital separation; the green line shows the total energy when angular momentum is conserved, and the red line shows the current separation. The energies are scaled with the current orbital energy, $E_{orb,0} = -2.338 \times 10^{37}$ kg m$^2$/s$^2$. The 1-$\sigma$ uncertainties plotted are calculated from the output chains of the global MCMC analysis of the system.

a stable orbit does not change.

The rate at which the planet is tidally migrating towards its host star can be estimated with the following equation from Brown et al. (2011) for relatively slowly rotating stars:

$$t_{\text{remain}} = \frac{2Q'_{*,0}}{117n}\frac{M_*}{M_p}\left(\frac{a}{R_*}\right)^5 , \tag{2.3}$$

where $Q'_{*,0}$ is the current tidal quality factor for the star, and $n$ is the orbital frequency. If $Q'_{*,0}$ is set to $10^8$ (as suggested by Penev & Sasselov (2011)), the spiral in time, $t_{\text{remain}}$, is $\sim$ 16 Gyr, which is significantly longer than the remaining lifetime of the host star. The true value of $Q'_{*,0}$ is not well understood from observations of exoplanets, however for values lower by a couple of orders of magnitude, the value of $t_{\text{remain}}$ remains too large to detect.

Equation 18 from Matsumura et al. (2010):

$$\frac{da}{dt} = \frac{9}{Q'_{*,0}}\frac{M_p}{M_*}\frac{R_*^5}{a^4}\left(\frac{a_0}{a}\right)^{3/2}(\omega_{*,0}\cos\epsilon_{*,0} - n) , \tag{2.4}$$

where $\epsilon_{*,0}$ is the current spin orbit misalignment, can be integrated between the current orbital

separation ($a$) and the orbital separation which results in a change in the orbital period of 60 s to find the timescale for which the migration is detectable through orbital period shortening. For this case, as above, spin-orbit alignment is assumed. For $Q'_{*,0} = 10^8$, a period change of >60 s would be observed after ∼20 Myr, if orbital period changes were only resulting from tidal decay.

### 2.5.2 WASP-93 system

WASP-93b is a ∼1.47 $M_J$ planet, orbiting an F4 star every 2.73 days. In the same way as for WASP-92, BAGEMASS was used to justfiy the stellar mass and age determined, by verifying the values with stellar evolutionary models. The age and mass estimates were $1.61 \pm 0.48$ Gyr and $1.39 \pm 0.08$ $M_\odot$ respectively. The stellar mass values found in Section 2.3.2, and in the MCMC parameter determination are consistent with the mass found with BAGEMASS. The BAGEMASS age estimate is just above the 1-$\sigma$ bound for the gyrochronological age of the star, and significantly lower than the age estimated from the lithium abundance measured from the spectra. The final quoted planetary mass has been determined from the stellar density which was a jump parameter in the MCMC analysis.

The MCMC analysis was tested with the inclusion of the parameters for eccentricity. The optimal value of $e$ found for this fit was $0.012^{+0.035}_{-0.008}$, for with the BIC was 168.3 compared with 147.4 for the circular fit. Since the $\Delta$BIC >10, and the value found for $e$ is small, the orbit is treated as circular.

The two sets of observations mentioned in Section 2.2.3 where WASP-93 was observed spectroscopically at a high cadence in an attempt to measure the RM effect were not well centred on the mid-transit time, which was less certain when those observations were made. Due to the lack of full coverage of a transit, using the data to quantify the RM effect is challenging, but the attempts to do so are explored below. The CCFs for each observation were median normalised, and then the mean CCF profile across all the observations was subtracted, which should leave the effect of the planet blocking some part of the CCF during the transit. This process was completed separately for both data sets. The CCF residual data is shown in Figure 2.8, where an indication of the transit of WASP-93b is visible in the blue-shifted wing of the CCF residuals.

To attempt to fit for the RM effect and measure the projected spin-orbit alignment an-

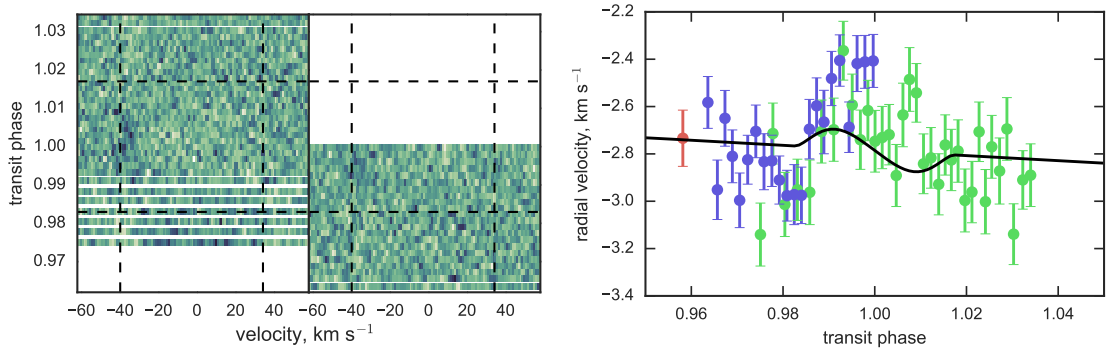



**Figure 2.8:** Left hand plot showing the time series of CCF residuals for WASP-93 after the subtraction of the average CCF profile with reference to transit phase for both attempts to observe spectra during transit. The vertical dashed lines show the width of the CCF profile centred on the centre-of-mass velocity of the system. The horizontal dashed lines show beginning of transit ingress and egress. The trail of the transit of WASP-93b is visible near the blue-shifted limb of the CCF. Right hand plot showing the RV measurements around the transit phase. The colours used correspond to those in the lower plot of Figure 2.3, and the green and blue RV points match the CCFs shown in the left and right of the upper plot respectively. The RV and RM effect model overplotted in black is for the case that the planetary orbit and the stellar spin are aligned.

gle, the parameters for fitting the RV measurements ($\sqrt{v \sin i}\cos\lambda$ and $\sqrt{v \sin i}\sin\lambda$) were included in the MCMC fit, however fitting this effect without the CCF data was not an improvement on fitting the data without the inclusion of the RM effect due to the fast rotation of the star. The residual CCF data was then included in the fit, and the parameterisation of the effect of the transit on the spectral data was adjusted to FWHM$_p$, $v \sin i$ and $\lambda$. The MCMC code did not converge on a solution for the value of $\lambda$, even when the value of $v \sin i$ was fixed to 37.0 km s$^{-1}$, as found in Section 2.3.2.

The maximum amplitude of the RV anomaly produced by the RM effect is:

$$\Delta RV_{RM} = \sqrt{1-b^2}\left(\frac{R_p}{R_\star}\right)^2 v \sin i \,, \tag{2.5}$$

which is ∼0.18 km s$^{-1}$ for the observed parameters of WASP-93. Since the true value of the RV anomaly is also dependent on the orbital obliquity, that value will likely be lower than the maximum value. For this data, the RM effect is difficult to detect, since the average 1-$\sigma$ uncertainty value for each RV observation is 0.13 km s$^{-1}$ and 0.11 km s$^{-1}$ for the first and second sets of observations respectively. From the just the RV data during transit (shown in the second plot in Figure 2.8), the exact shape of the RM anomaly is unclear.

The signature of the planetary transit is clear in the CCF residual data in Figure 2.8 in

the blue-shifted wing of the expected CCF profile. Observing this signature at the expected time, and inside the velocity window expected by the observed $v \sin i$ is very strong evidence in favour of the planetary scenario. Despite not finding a fit for the misalignment parameters with the global MCMC analysis, the fact that the planetary signal in the CCF residuals is only in the blue-shifted wing provides a strong expectation that the orbit of WASP-93b is almost entirely misaligned with respect to the spin axis of the star.

**Blend scenario**

Given the near v-shape to the observed transit profile (as shown in Figure 2.3), investigating the possibility that the detected signal is not caused by a planetary transit is particularly important. The two key scenarios which have the potential to imitate the signal of a planet are: a background blended eclipsing binary pair, or that the primary star is part of a hierarchical triple system.

The case of a blended background eclipsing binary pair, is already well ruled out by the periodicity in the photometric transit data being matched in phase and period by modulations in the RV data. To provide a further test against this hypothesis, the AO imaging presented in Section 2.2.4 searched for companions to WASP-93 which were unresolved in lower resolution images. A stellar object was observed with a low angular separation from WASP-93, however it is fainter than WASP-93 by more than three magnitudes. If this magnitude difference were to remain constant into the visible passbands where the transits were observed, a complete occultation of the object by a non-emitting body would not be enough to reproduce the observed transit signal. If the object is associated with WASP-93, then the magnitude difference between it and WASP-93 would be larger in visible light passbands, as the star would be a K type dwarf. It is also not possible for there to exist an object large enough to fully occult the star and be non-emitting. The AO images show no other stellar objects within 0.3 $''$, so the region of space remaining for which a chance aligned binary system bright enough to mimic the transits would have to exist is so small as to render the possibility extremely unlikely.

The prospect of WASP-93 being part of a hierarchical triple system, which is producing the transit and RV signal is more difficult to rule out, as the companion stars would not appear in AO images. Were this scenario to be the correct characterisation of this system, the CCFs obtained during the transits would have looked significantly different – a modulation would have been seen in the residual CCF data, but it would have been outside the velocity window

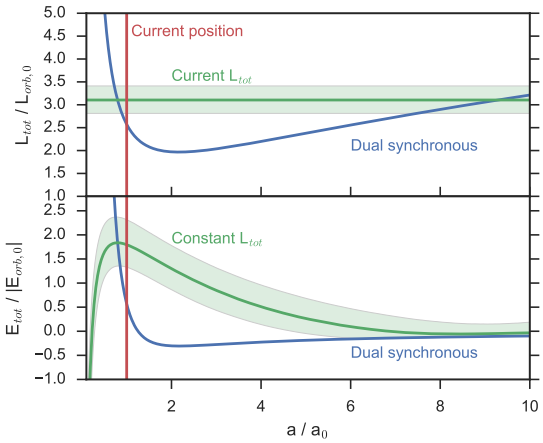



**Figure 2.9:** Plot showing the tidal equilibrium curves for WASP-93, which are presented in the same way as Figure 2.7. The upper plot is scaled with orbital angular momentum, $L_{orb,0} = 2.986 \times 10^{42}$ kg $m^2/s$, and the lower plot is scaled with orbital energy, $E_{orb,0} = -3.976 \times 10^{37}$ kg $m^2/s^2$.

of the observed CCF, since the binary pair which could have been mimicking the transit signal would not occult any part of the WASP-93 CCF. The fact that the observed signal in the residual CCF is entirely within the transit window further supports the conclusion that the WASP-93 system is formed of a planet orbiting a main sequence star. Even if it were assumed that the signal seen in the CCF residuals was a fluke in the noise, some sign of another set of spectral lines at a different systemic velocity would be expected to be seen. The CCF window was also calculated out to ± 100 km $s^{-1}$ of the systemic RV of WASP-93, no sign of another stellar object was seen.

### Tidal evolution

In terms of tidal stability, the WASP-93 system sits in an interesting position, as $L_{\rm tot} > L_c$, which is unlike most other hot Jupiter systems (Levrard et al., 2009; Matsumura et al., 2010). In this case $L_{\rm tot}/L_c = 1.57 \pm 0.16$, which means that the system has enough angular momentum to potentially eventually reach a stable orbit, rather than for the planet to continue to head towards the Roche limit of its host star.

The upper plot of Figure 2.9 is like that of Figure 2.7 for WASP-92, where the green line shows the current total angular momentum of the system, and the blue line shows the angular momentum required for a dual synchronous orbit at different orbital separations. The red line shows the current separation of the WASP-93 and WASP-93b. The green and blue lines clearly intersect, indicating that the system has enough angular momentum for a Darwin stable orbit.

Tidal stability is characterised by the minimisation of the total orbital and spin energies, as well as the conservation of total angular momentum, so to investigate the evolution of the system due to tidal interaction, the energies must be considered. The lower plot in Figure 2.9 shows the how the orbital and spin energies match with each of the lines plotted with reference to angular momentum. The energy plotted has been defined as:

$$E_{\rm tot} = -\frac{GM_*M_p}{2a} + \frac{1}{2}C_*\omega_*^2 + \frac{1}{2}C_p\omega_p^2\,, \tag{2.6}$$

where the first term is the orbital energy, and the remaining terms are the rotational energies of the star and planet respectively. Since hot Jupiters are understoon to be tidally locked, the value of $\omega_p = n$, which to find the value of $E_{\rm tot}$ leaves only $\omega_*$ to be determined by conservation of angular momentum or through being synchronised with $\omega_p$ and $n$.

It is clear from the lower plot in Figure 2.9 for the intersections of conserved angular momentum and a dual synchronous orbit, the inner equilibrium point is at an energy maximum, and the outer equilibrium is a minimum, thus the single point where a stable orbit can exist. For the case plotted here, the tidal interactions would be pushing WASP-93b away from the host star towards the stable orbital position along a path of conserved $L_{\rm tot}$.

Since the $v \sin i$ of WASP-93 is high (37 km s$^{-1}$), and the planet is potentially not migrating inwards, the estimation of the remaining time before the planet spirals in to its host star from Brown et al. (2011) (shown in Equation 2.3) is not appropriate, so a numerical integration of Equation 18 from Matsumura et al. (2010) (shown here as Equation 2.4) is used to estimate the timescale over which the planet is potentially migrating outwards. Using this method, the timescale for the stable orbit to be reached is $\sim 2\times10^7$ Gyr for $Q'_{*,0} = 10^8$, which is significantly longer than the remaining stellar lifetime even if the tidal interaction is allowed to be more efficient with a lower value of $Q'_{*,0}$. To detect the elongation of the orbital period through this migration (change of $>$ 60 s) would take $\sim$50 Myr when $Q'_{*,0} = 10^8$.

As for WASP-92, the analysis above assumed that the spin of the star was aligned with the orbital plane of the planet. The CCF residuals plotted in Figure 2.8 indicate that the assumption of alignment is not valid. Misalignment beyond $\lambda \sim 30^\circ$ (as is predicted in the tomographic data) would reduce the amount of angular momentum in the plane of the stellar rotation enough so that the current orbital separation of the star and planet would be inside the inside the inner intersection of the angular momentum curves. This would mean that WASP-93b

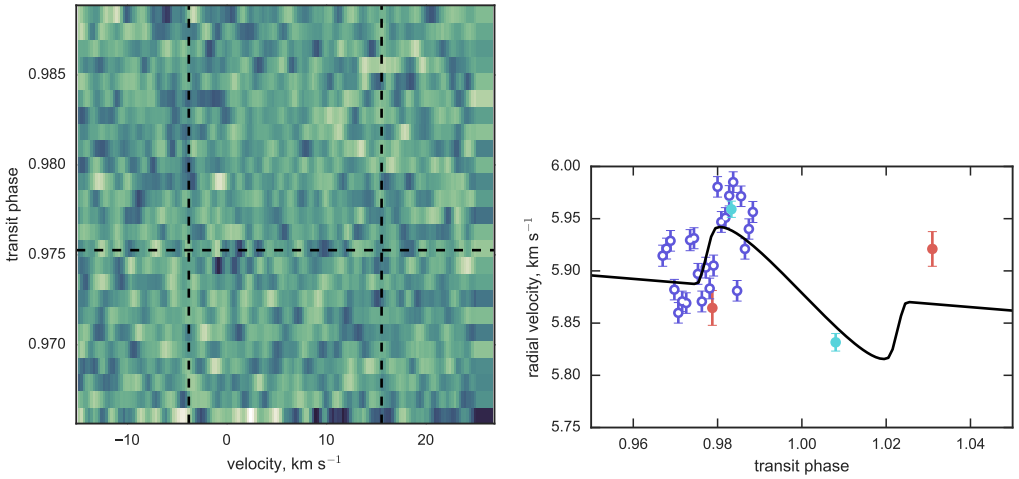



**Figure 2.10:** Upper plot showing the time series of CCF residuals for WASP-118 after the subtraction of the average CCF shape with reference to transit phase for the attempt to observe spectra throughout a transit . The trail of the transit of WASP-118b is not visible. Lower plot showing the RV measurements around the transit phase. The colours used correspond to those in the lower plot of Figure 2.4, and the purple RV points match the CCFs shown in the upper plot.

would be tidally migrating towards its host star, which is in line with the expected situation for tidal interactions in hot Jupiter systems.

### 2.5.3 WASP-118 system

WASP-118b is an ∼0.55 $M_J$ planet orbiting an F6 spectral type star with an orbital period of 4.04 days. As for WASP-92 and WASP-93, BAGEMASS was used to estimate the mass and age of WASP-118 from stellar evolutionary models. The age estimate was $2.38 \pm 0.38$ Gyr and $1.40 \pm 0.05$ $M_\odot$ for the mass, and the results are shown in the lower plot of Figure 2.6, which show that the star is slightly evolved off the ZAMS. The age found is in agreement with the determinations made from the spectral analysis in Section 2.3.3. There is an inconsistency between the mass found with BAGEMASS and the MCMC analysis, and the mass found in the spectral analysis, which is presented in Table 2.5. In the spectral analysis, the stellar mass was estimated from the spectroscopically determined surface gravity of the star, which is difficult to determine from spectra for stars hotter than the sun (Mortier et al., 2014). The stellar density is well defined from the shape of the observed transits, thus the larger mass estimate found from the MCMC analysis and BAGEMASS are treated as more reliable.

The weather was poor for the spectral observations on the night beginning 2015 October 1st using the HARPS-N instrument on La Palma, which resulted in fewer observations collected than expected, as well as a significant reduction in the quality of the data that was obtained.

There is significant scatter in the RV points from that set of observations (shown in the second plot of Figure 2.10), which is not reflected in the uncertainties derived by the DRS. Stellar activity would not be expected to create such large and real scatter in the RV values on such short timescales, so it can be assumed that the uncertainties have been significantly underestimated by the DRS for the observations that were adversely affected by poor weather.

The time series of CCF residuals from those observations, which were reduced in the same manner as the tomographic observations for WASP-93, are shown in the first plot of Figure 2.10. The residuals show a significant amount of noise for such a precise spectrograph, which was not only due to the bad weather, but also the bright moon (79% illuminated and on average 42° from the target) will have increased the noise level of the collected spectra. There is no sign of a planet signal in the CCF residuals, which is likely due to the low quality of the data.

Since the reliability of the RV measurements and uncertainties from the 2015 October 1st HARPS-N data is in question, the MCMC analysis was performed with and without that dataset. Without the transit series of HARPS-N data, the MCMC fit for a non-circular orbit found a value of $e = 0.089^{+0.057}_{-0.047}$. The BIC for the non-circular fit was 67.0, and 20.3 for a circular orbit. When all the data is included, the best fit non-circular orbit found $e = 0.132^{+0.105}_{-0.080}$ with a BIC of 520.4, in comparison with BIC = 50.3 for a circular fit. In each case $\Delta$BIC is significantly >10, so there is very strong evidence to support excluding the two free parameters required to include eccentricity in the global fit.

The underestimation of the uncertainties for the transit series of RV points are likely to place too high a weight on those data points during a global fit, which could bias the final parameters determined, therefore the values quoted for the WASP-118 system in Table 2.5 are excluding that data. The distribution of those RV points are indicative of an orbit aligned with the stellar spin axis, but without full transit coverage and higher quality data, a conclusive answer on the spin-orbit alignment is not possible.

#### Blend scenario

As for the WASP-92 system, there is no AO imagery of the WASP-118 system. The discussion in Section 2.5.1 for WASP-92 is also relevant for WASP-118.

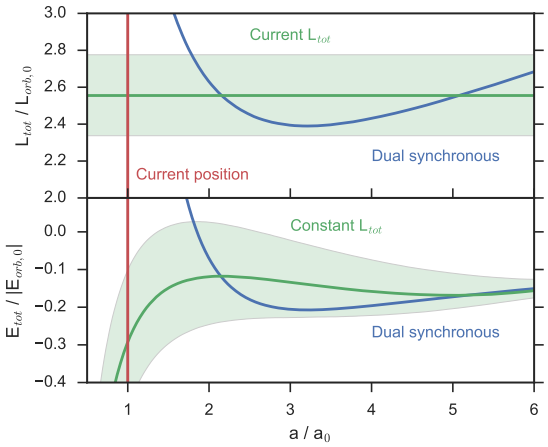



**Figure 2.11:** Plot showing the tidal equilibrium curves for WASP-118, which are presented in the same way as Figure 2.7. The upper plot is scaled with orbital angular momentum, $L_{orb,0} = 1.164 \times 10^{42}$ kg m$^2$/s, and the lower plot is scaled with orbital energy, $E_{orb,0} = -1.046 \times 10^{37}$ kg m$^2$/s$^2$.

### Tidal evolution

Just as for WASP-93, the WASP-118 system probably has enough angular momentum for a stable orbital solution to exist. In this case the value of $L_{\rm tot}/L_c = 1.07 \pm 0.08$. The plots in Figure 2.11 show the existence of the intersection between angular momentum conservation and a dual synchronous orbit, for the majority of space occupied by the 1-$\sigma$ uncertainty.

Unlike WASP-93, the current position of the planet relative to the star places the planet interior to the inner stable point, which means that the energy gradient for angular momentum conservation (shown in the lower plot of Figure 2.11) is towards the Roche limit of the star rather than towards the stable orbit position.

The timescale for the migration of WASP-118b towards its host star can be estimated using the same numerical integration as for WASP-93, since Equation 2.3 is not appropriate for a star rotating with $v \sin i \sim$10 km s$^{-1}$, such as WASP-118. Equation 2.4 was integrated from the current orbital separation to the surface of the star. As before, the system was assumed to have a circular orbit, as well as for the planetary spin to be tidally locked, both of which are valid for hot Jupiter systems. The planet was also assumed to be orbiting in the same plane as the stellar rotation, for which there is currently no evidence against, and is tentatively implied by the the RV measurements of the system taken during transits. This gives a value of $t_{\rm remain}$ $\sim$150 Gyr for $Q'_{*,0} = 10^8$. The spiral in timescale is therefore too long for the orbital period change to be observable, even with lower values of $Q'_{*,0}$.

## 2.6 Conclusions

In this chapter, I have presented results that describe the discovery of three new hot Jupiter planets. Each of the planets populate a new data point of in the distribution of known planets, and provide a further constraint on the understanding of the whole population of extra-solar planets.

In the 23 years since the first extra-solar planet was discovered, thousands more planets and hundreds more hot Jupiters have also been detected, which begets the question of whether it is still important to detect more planets, such as those presented above. Whilst hundreds of hot Jupiters are known, the understanding of how they form remains an open question. A key mechanism for differentiating between proposed hot Jupiter formation pathways is in the distributions of star-planet orbital misalignment that the planetary systems exist in, which is particularly important for stars with $T_{\rm eff} > 6250$ K.

The three planets presented in this chapter represent a range of host star $T_{\rm eff}$ values, which is important not just in understanding giant planet formation, but also the true underlying distribution of giant planets, not just those that are readily detectable.

# 3

# An introduction to stellar atmospheres and limb darkening

## 3.1 Introduction

From the earliest observations of the Sun, it became clear that the brightness of the Sun was not the same at the centre compared to the extrema. This is known as centre to limb intensity variation (CLIV[1]). The first quantification of the effect was made by Bouguer in 1748[2], who noted that at 0.75 $R_\odot$ from the centre of the Sun, the brightness was 0.73 of the central intensity. The measurement was made with the Heliometer he invented, which observed the limb and the centre of the solar surface simultaneously with identical optical systems with the images of each juxtaposed. By changing the aperture size of the system observing the centre of the Sun, the measurement could be made of the intensity difference of the two positions based on the different aperture sizes required for the images to be of the same brightness. The

[1] CLIV – centre to limb intensity variation is not to be confused with the roman numeral for 154.
[2] The results were published posthumously in Bouguer (1760)

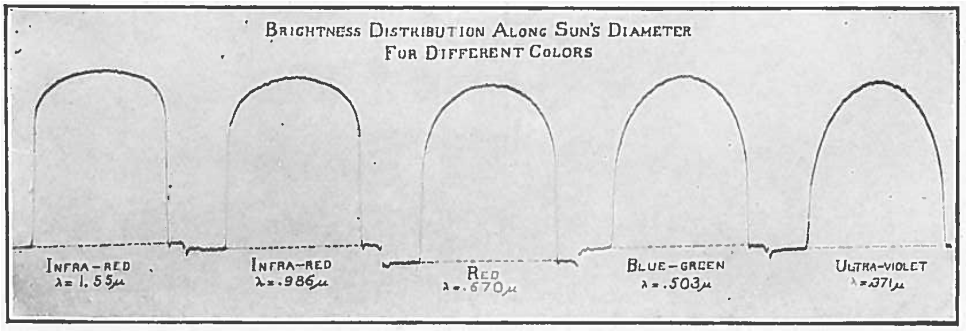



**Figure 3.1:** Image from Figure 23 in Abbot et al. (1923), which shows the different profiles of solar limb darkening measured for several wavelengths across the diameter of the Sun, with left to right showing long to short wavelengths.

result achieved was within a few percent of the current accepted measurements for the Sun.

As instrumentation improved, further attempts were made to quantify the effect of limb darkening by trying to compare the brightness of solar images in different positions from the centre of the Sun. Work was conducted by Arago (1858), finding that the light observed at the edge of the Sun is less bright than that at the centre by 2.4% at the same position that Bouguer's measurements were taken from. Liais (1866) observed a 10% reduction at the edge. Secchi (1870, 1875) proposed that for the same position at the edge of the Sun, the light would be reduced by 78% from the central light. The vastly discrepant estimates for the intensity drop-off can be attributed to each experiment assuming that the optical conditions for all of the images were consistent, when this was not the case.

Langley developed a more robust method of observing the difference in the intensity of the light from the centre of the Sun to the limb, where the optical arrangement ensured that the compared images came from exactly the same optical conditions. The method came with the added benefit of being able to investigate more specifically the colour dependence of the effect of limb darkening (Langley, 1875a,b). Concurrently with Langley's work, and the advancing development of photographic techniques, Vogel (1872, 1877) was also able to make measurements of limb darkening from the interaction of the light from the Sun with photographically active chemicals in resolved images of the solar surface.

Figure 3.1[3] shows the work of Abbot et al. (1913, 1923) to quantify solar limb darkening. The figure shows the increasing importance of limb darkening as wavelength of light decreases for the Sun, and highlights how the measurements of the solar CLIV had improved since the

[3]With thanks to the Harvard Smithsonian Library for providing a high quality scan of this figure to include.

earliest attempts to compare the brightness at two positions on the solar surface. Current instrumentation makes observing limb darkening for the Sun in the continuum trivial. The work by Neckel & Labs (1994) to quantify the effect for a range of wavelengths continues to be used as a benchmark for solar limb darkening observations.

Whilst the Sun's surface can be readily resolved from Earth, before the invention of CCDs obtaining accurate and precise direct CLIV measurements was a difficult task. Julius (1906) proposed a novel way to map the solar limb darkening by measuring the change in total visible sunlight during a solar eclipse. As the moon occulted different portions of the solar surface, the intensity variation could be modelled. Several expeditions were made to test this technique during eclipses, such as Julius (1913); Wesselink (1940); Kopal (1946) with varying success. The eclipse method has been shown to be less precise than direct observations (Evans, 1947; Hubenet & de Jager, 1956; Mädlow, 1962), predominantly because the lunar surface is inhomogeneous, thus the true portion of the solar surface occulted at any given time during the eclipse is challenging to model. This approach, whilst not optimal for the resolvable surface of the Sun, does bear some similarity to the way that limb darkening can be quantified for stars for which their surfaces cannot be resolved.

The lightcurves of binary star system eclipses can be used to quantify the CLIV of other stars, and measurements of this were first reliably achieved by Kron (1939). Now that there are an abundance of lightcurves of exoplanet transits, measurements of the form of limb darkening for a variety of stars can be made more precisely than from eclipsing binary lightcurves as a smaller portion of the stellar surface is occulted at any given time in the transit.

With the development of interferometry techniques, surface properties of other bright and nearby stars can also be resolved. This led to the first direct quantification of the limb darkening of a star other than the Sun by Hanbury Brown et al. (1974) for the surface of Sirius. Interferometric measurements continue to provide direct observations of limb darkening, although this remains challenging given the very small angular diameters of the stars (Wittkowski et al., 2004; Kervella et al., 2017).

## 3.2 Limb darkening theory

The underlying theory of what physical processes result in the effect of stellar limb darkening rests in the understanding of stellar atmospheres. The primary feature of stellar atmospheres

that gives rise to limb darkening is the temperature gradient in the atmosphere. When observing at an angle to the centre of the star, the same length of optical path (distance before optical depth equals unity) will be at a shallower point in the atmosphere than looking directly towards the centre of the star. Since a stellar atmosphere has a temperature gradient, being able to see less deep in the atmosphere means that the effective temperature at that depth is lower. A lower effective temperature means that flux of the emitted light is lower, and that the distribution of emitted light is skewed towards redder colours.

The temperature gradient of a stellar atmosphere is not always decreasing further from the centre of the star – for particular wavelengths, the temperature is increasing, which results in limb brightening. Limb brightening is observed in the UV and X-ray for the Sun, and for specific spectral lines.

The first steps in building a model to understanding (although now understood to be incorrect) were taken by marquis de Laplace (1805) in his opus, Traitè de mècanique céleste, where he modelled the early solar limb darkening observations made by Bouguer. It took until Schwarzschild (1906) produced the first consideration of radiative equilibrium at the edge of a star for the beginnings of stellar atmosphere theory to develop. The work resulted in a model for the temperature profile in the solar atmosphere and therefore produced the first parametric law of limb darkening. The model and law was generalised for stars other than the Sun by Milne (1921).

To predict the effect of limb darkening for a given star and viewing angle, a model of the stellar atmosphere structure must be constructed, followed by the modelling of the radiative transfer processes for that viewing angle.

## 3.3 Grids of stellar atmosphere models

The ATLAS stellar atmosphere model was first presented in Kurucz (1970) for atmospheres in radiative and convective equilibrium. The code makes approximations for the atmosphere's geometry and for some physical processes, so that modelling the atmosphere is computationally tractable. The form of the atmosphere is constructed as a set of finite plane-parallel layers with constant energies and abundances at each. Opacity distribution functions (ODFs) are used to provide a statistical means to describe the opacity structure to calculate the radiative equations throughout. The ATLAS model continued to be developed, and Kurucz (1979) pre-

sented the ATLAS9 version, which included more advanced treatment of some of the other atmospheric physics, and allowed the modeller to include ODFs for a large number of the line opacities. The ODFs include a tabulation of the line opacity for different temperatures and pressures for a range of wavelengths. A different ODF is required for different metallicities. This treatment allows the atmosphere model to be calculated very efficiently once the ODF for that metallicity has been generated. Because of this efficiency, the ATLAS9 models are used to create very large grids of stellar atmosphere models, such as Castelli & Kurucz (2003); Kirby (2011); Mészáros et al. (2012). A version of the ATLAS9 model for spherical rather than plane-parallel geometry was developed by Lester & Neilson (2008).

Hauschildt & Baron (1999) introduced a new version of the PHOENIX stellar atmosphere model that did not rely on the assumption of local thermodynamic equilibrium that the ATLAS9 model does, and that the original PHOENIX model did (Allard & Hauschildt, 1995). This means that the atmosphere model does not rely on tabulated ODFs. The opacity is instead directly sampled from a large list of lines. The model geometry was also upgraded from the original plane-parallel to a spherically symmetric structure (Hauschildt et al., 1999). The PHOENIX model has been used to generate a large grids of atmospheres for a range of $T_{\rm eff}$, $\log g$ and [Fe/H] including for the GAIA (Gaia Collaboration et al., 2016) mission by Brott & Hauschildt (2005). The most recent grid of PHOENIX atmospheres, Husser et al. (2013), provides several thousand model atmospheres.

Computational methods for the 3D modelling of stellar atmospheres were first instigated by Nordlund (1982), providing a framework with realistic radiative transfer and a solution of the hydrodynamic equations. Nordlund et al. (2009) provide an example of how the 3D model much better represents real stellar atmospheres (particularly the Sun) than the 1D ATLAS9 and PHOENIX models. Hayek et al. (2012) show the same for the constraints on a stellar atmosphere derived from an exoplanet transit lightcurve. Magic et al. (2013) present a grid of atmosphere models calculated with the STAGGER code. The code is a 3D hydrodynamic treatment that calculates solutions to the conservation of mass, momentum and energy with time dependence, and then solves for radiative transfer where optically thin. Using the 3D hydrodynamic approach removes the need to invoke parameterisations for convection and other time-dependent processes relevant to the stellar atmosphere, as they emerge from the model itself. Due to the increased computation time for this type of model, the grid consists of ∼200 atmospheres, rather than the $\sim 10^5$ in some grids created with 1D models.

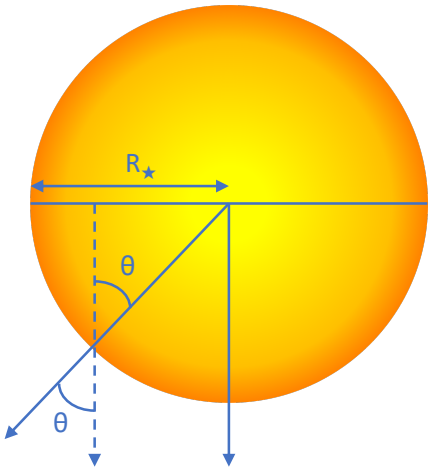



**Figure 3.2:** Schematic of the angle, $\theta$, to the viewing direction and the stellar surface. The location of the observer is looking up from the bottom of the schematic. $R_\star$ is the observer defined stellar radius, and $\theta$ is the angle between the line of sight and the normal to the given point on the stellar surface.

## 3.4 Limb darkening parameterisation

Limb darkening is parameterised via its relation to the angle, $\theta$. $\theta$ is the angle between the line of sight and the normal to the stellar surface at the given position, which is shown in Figure 3.2. To relate the angle $\theta$ to the projected radius, the cosine of this angle is then used:

$$\mu = \cos\theta \ . \tag{3.1}$$

$\mu$ is used throughout the rest of this work to describe the position for which a limb darkened intensity is defined.

The first law for parameterising limb darkening was proposed by Schwarzschild (1906), which was based on his physical model of the solar atmosphere in radiative equilibrium (and formalised by Milne (1921)). The law takes the simple form:

$$\frac{I(\mu)}{I(1)} = 1 - c(1-\mu) \ , \tag{3.2}$$

where $I(\mu)$ is the stellar intensity for a given $\mu$, $I(1)$ is the intensity at the centre of the star ($\mu = 1$), and $c$ is the coefficient of limb darkening for this law. For the Sun, $c$ is approximately 0.6 for visible light passbands. This formulation has been used extensively for the modelling of eclipsing binary lightcurves (eg. Kron (1939)), as the amount of information about the limb darkening of each star in the observed lightcurve is limited by the relative sizes of the stars,

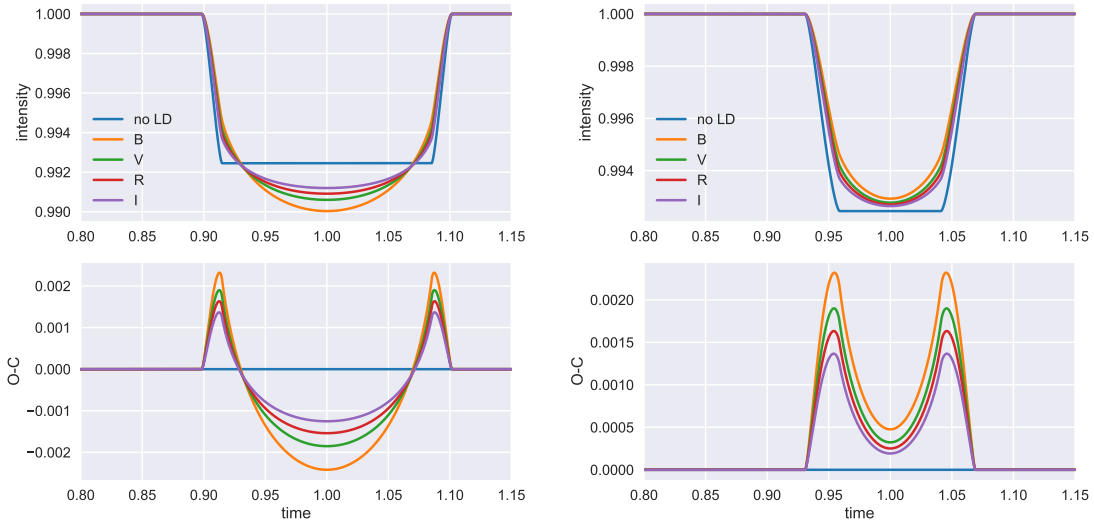



**Figure 3.3:** Figure showing simulated transit lightcurves for a hot Jupiter on a 4 day orbit around a solar type host star. Lightcurves are shown for the B, V, R and I passbands, as well as for a star with no limb darkening. The upper plots show the lightcurves, and the lower plots show the difference between the lightcurve with no limb darkening and the other lightcurves. The plots to the left are for a simulated transiting system with $b = 0$, and $b = 0.8$ for the plots to the right.

and other effects such as gravity darkening.

Once a rigorous solution of the radiative transfer equation was established by Wiener & Hopf (1931), Kopal (1950) was able to show that the linear law was insufficient to describe the true CLIV, and so proposed the quadratic law as a more representative description:

$$\frac{I(\mu)}{I(1)} = 1 - c_1(1-\mu) - c_2(1-\mu)^2 \,, \tag{3.3}$$

where $c_1$ and $c_2$ are the coefficients of the law. As observations of eclipsing binary system lightcurves improved, it became clear that a single parameter law would not be sufficient to represent the CLIV for stars other than the Sun, thus the quadratic law was used for these approaches when the data quality was good. The quadratic law was also preferred for hotter stars where the difference from the linear law and the true profile was larger (Grygar, 1965). This law is widely used in exoplanet transit fits, as it reasonably well represents the CLIV for visible light passbands and for stars in a similar region of atmospheric parameter space as the Sun, as well as only requiring two fitting parameters. The quadratic law is currently the most commonly used parameterisation of limb darkening for exoplanet transit analysis, as it only has two parameters, and provides a good enough estimation of the model limb profiles for many purposes.

Figure 3.3 demonstrates the strength of the effect of different observational bandpasses on the shape of an exoplanet transit lightcurve modelled with the quadratic law for two different impact parameters – a transit across the centre of the star ($b = 0$), and a near grazing transit ($b = 0.8$).

Several other two parameter laws have since been proposed, such as the logarithmic law (Klinglesmith & Sobieski, 1970),

$$\frac{I(\mu)}{I(1)} = 1 - c_1(1-\mu) - c_2\mu\ln(\mu)\ , \tag{3.4}$$

and the square-root law (Diaz-Cordoves & Gimenez, 1992),

$$\frac{I(\mu)}{I(1)} = 1 - c_1(1-\mu) - c_2(1-\sqrt{\mu})\ . \tag{3.5}$$

Each proposed two parameter law offers a good fit to the CLIV for certain regions of the HR diagram, however none of the laws can represent the whole range required for exoplanet studies. These laws, unlike those before them were proposed based on how well they could represent the data (from eclipsing binary observations, and model atmospheres), rather than derived from a physical process.

Another two parameter law, the "power-2" law was proposed by Hestroffer (1997) for interferometric measurements of limb darkening, and constructed by Morello et al. (2017) for exoplanet transit lightcurve analysis. The law takes the form:

$$\frac{I(\mu)}{I(1)} = 1 - c_1(1-\mu^{c_2})\ , \tag{3.6}$$

with $c_1$ and $c_2$ as the two coefficients. The justification given for the use of this law over the more commonly used quadratic law for exoplanet transits is that this form of law is more flexible, and can fit the model atmospheres better for the types of stars studied with exoplanet transits.

Claret (2000) proposed the non-linear law with four coefficients:

$$\frac{I(\mu)}{I(1)} = 1 - c_1(1-\mu^{\frac{1}{2}}) - c_2(1-\mu) - c_3(1-\mu^{\frac{3}{2}}) - c_4(1-\mu^2)\ , \tag{3.7}$$

where $c_1$–$c_4$ are the coefficients for this law. This law was proposed as it was flexible enough

to well represent the limb profiles of most stars, rather than have to utilise different limb darkening laws for different types of stars. A major disadvantage in using this law is that there are four coefficients, which can cause difficulties in fitting lightcurves if not fixed to values found from model stellar atmospheres.

Sing et al. (2009) introduced a variation on the four parameter law in Equation 3.7, which dropped the $\mu^{\frac{1}{2}}$ term:

$$\frac{I(\mu)}{I(1)} = 1 - c_1(1-\mu) - c_2(1-\mu^{\frac{3}{2}}) - c_3(1-\mu^2)\,, \tag{3.8}$$

where $c_1$–$c_3$ are the coefficients of the law. The justification for making this adjustment was primarily because the $\mu^{\frac{1}{2}}$ term affects mostly the shape of the limb profile at low $\mu$ values, and it is well known that the variation for this part of the limb profile is smooth and well approximated by a linear function, so the term is not required to provide a good representation of the CLIV. The second reason to remove the term is that it reduces the number of coefficients in the model, which is then easier to fit.

## 3.5 Tables of LDCs

Tabulated LDCs are the result of fitting the outputs of a grid of stellar atmosphere models for the coefficients of the chosen law for a chosen photometric passband. Several tables of LDCs have been calculated for different grids of stellar atmosphere models. Each of the grids of stellar atmosphere models mentioned here for fitting LDCs are introduced in Section 3.3. Claret & Bloemen (2011) made use of grids of the ATLAS9 plane-parallel models, and also the PHOENIX spherically symmetric models. Neilson & Lester (2013) made use of a version of the ATLAS9 model adapted to function with a spherically symmetric geometry rather than plane-parallel (Lester & Neilson, 2008). Sing (2010) also made use of a grid of ATLAS9 models. The table of LDCs presented in Magic et al. (2015) were generated from the STAGGER grid of 3D models, although the number of models included in this grid is significantly smaller than for those based on grids of 1D atmosphere models.

The basic process for generating a table of LDCs is briefly described in the following steps. The atmosphere model outputs take the form of specific intensities for a range of wavelengths and $\mu$ angles for each model. The specific intensities at each $\mu$ are then convolved with a photometric bandpass, and integrated over wavelength, resulting in an intensity for each angle

for each passband used. The grid of coefficients may be calculated using the least squares method, which optimises the fit of the parametric law for each atmosphere model across the calculated intensities.

The tabulated LDCs in Claret & Bloemen (2011) also include LDCs found using a second method to determine their values for each atmosphere model – the flux conservation method (FCM). Intuitively basing the determination of the LDCs on conserving flux appears the optimal method to utilise. The FCM does not however constrain the LDCs to well represent the form of the atmosphere model output for any non-linear limb darkening formulation, only that the integrated fluxes match. Another condition must be used in conjunction with flux conservation to allow a solution to be found for laws with more than one coefficient, and to force the fit to use some information about the form of the limb profile. Despite the desire for parameterised limb darkening to conserve the flux of the atmosphere model output, using the FCM leads to significantly worse fits of the LDCs than using a least squares fitting procedure (Claret, 2000).

# 4

# Limb darkening coefficient interpolation

## 4.1 Limb darkening and exoplanet transits

This chapter introduces a new approach for determining the limb darkening coefficients for a transit lightcurve, interpolating between tabulated values of limb darkening coefficients using a Gaussian Process (GP).

The forms of the parameterisations for limb darkening are discussed in Section 3.4. In order to apply these parameterisations to an exoplanet transit fit, some approach must be chosen to implement the model into the fitting procedure. When choosing which limb darkening law to utilise, a balance needs to be struck between accurately mapping the stellar surface variation with a higher order law, and having too many ill-constrained free parameters to fit. Once a limb darkening law has been selected, what fitting procedure to use must be determined – the coefficients can be freely fit, constrained only by the transit observation data; some existing knowledge of the stellar parameters can be used to apply a prior to some or all of the coefficients, or some or all of the coefficients can be fixed to the values expected for that host

star.

If choosing not to apply an informative prior to the coefficient fit, the coefficients for certain laws can be reparameterised to allow uniform sampling of the coefficients. The adjustment reduces the correlation between the parameters, and means that the parameters used must sit within the region [0,1] where the probability of each value is equal (a flat prior is appropriate). This reduces the parameter space that any MCMC fit must explore[1], and prevents unintended or non-physical non-uniformity in the prior. This form of reparameterisation has however only been found for the quadratic law (Kipping, 2013), and for the Sing et al. (2009) 3-parameter law (Kipping, 2016).

Many studies of exoplanet transit lightcurves have used these uninformative priors on the limb darkening coefficients for the quadratic law (eg. Barragán et al. (2016); Foreman-Mackey et al. (2016)), or for the 3-parameter law (eg. Sandford & Kipping (2017)). This method is most effective when the lightcurve is of a good SNR, and the impact parameter is low enough, so that enough information about the limb darkening for the system is encoded into the transit lightcurve. With a low SNR transit lightcurve, freely fitting the coefficients risks offering too much flexibility in the fit, and can result in over-fitting or biasing the transit parameters (Espinoza & Jordán, 2016).

The LDCs can be fixed to the expected values for the host star (eg. Adams et al. (2017); Wilkins et al. (2017)), which ensures that the modelled CLIV is physically justified for that star, and prevents the fit for the limb darkening over-fitting any other signal in the transit lightcurve. This approach does assume that the measurements of the stellar atmospheric parameters used to find the fixed coefficients are accurate and precise, and the atmosphere models and the form of the parametric law used to determine the coefficients reflect the physics of the atmosphere of the host star. The uncertainties in the determination of the stellar atmospheric parameters are also not propagated through this fitting procedure.

Instead of fitting the LDCs, prior information about the stellar atmosphere parameters can be used to inform the fitting of the LDCs. This approach provides some flexibility in the LDCs, as it allows the uncertainties in the stellar parameters to be propagated through the fit. Southworth (2008) proposes JKTLD[2], which performs bi-linear interpolation between tabulated LDCs from a choice of tabulations. The method interpolates between the $T_{\rm eff}$ and

[1] Markov Chain Monte Carlo (MCMC) methods are introduced in Section 2.4
[2] The JKTLD code is available from ASCL (Southworth, 2015), and http://www.astro.keele.ac.uk/jkt/codes/jktld.html

$\log g$ grid provided in the tables for a given set of input stellar parameters. This method can then be called during an MCMC transit lightcurve fit to sample the limb darkening coefficients with stellar parameters as jump parameters.

LDTK (Parviainen & Aigrain, 2015)[3] generalises the interpolation between tabulated coefficients to produce fits for LDCs for arbitrary photometric passbands from the Husser et al. (2013) grid of stellar atmosphere models. A prior on the stellar parameters ($T_{\rm eff}$, $\log g$, and [Fe/H]) is then sampled to produce a multivariate prior on the LDCs to be applied to a transit fit.

### 4.1.1 Biases from incorrect or insufficient fitting of limb darkening

Insufficiently or incorrectly modelling the stellar limb darkening when fitting an exoplanet transit can introduce biases in the transit parameters determined. The scales and extents of the biases are explored by several authors.

Work by Southworth (2008) found that in general fixing the limb darkening coefficients to those tabulated did not bias the transit parameters, but would result in an underestimate of the uncertainties. Howarth (2011) showed more significant differences between limb darkening determined from the lightcurve shape and from tabulated coefficients, although the scales of the differences showed some dependence on the stellar parameters of the host star. Müller et al. (2013) demonstrated that the predicted LDCs from theoretical atmosphere models systematically over-predicts the value of the second quadratic coefficient, which would induce biases in the transit parameters if fixed to that value. The study also highlights that fixing the LDCs to the tabulated values (or tightly constraining the coefficients) is the only appropriate approach for systems with high impact parameters, as not enough of the stellar surface is occulted during the transit to provide enough information about the limb darkening to determine the LDCs solely from the transit lightcurve.

Hayek et al. (2012) were able to show that fixing LDCs to values determined from the outputs of 3D stellar atmosphere models produced far lower residuals in transit fits than LDCs determined from 1D stellar atmosphere models. Espinoza & Jordán (2015) quantify some of the differences seen between LDC modelling approaches for transit lightcurves, which are significant, particularly when fixing the coefficients to tabulated values for the quadratic law

[3]The LDTK code is available from ASCL (Parviainen, 2015) and https://github.com/hpparvi/ldtk

– up to a 3% bias in $R_p/R_\star$. They recommend freely fitting the LDCs as part of a transit fit to prevent this bias.

Neilson et al. (2017) found significant differences (up to 300 ppm in the ingress/egress) between the transit shapes for limb darkening determined directly from stellar atmosphere models with plane-parallel geometry and those with spherically symmetric geometry. They did not however fully take into account the differences in definition of the stellar radius between the two geometries, so the scale of the true effect will be much smaller.

Ensuring that limb darkening is precisely and accurately modelled is particularly important when fitting for third order effects, which are potentially observable in the ingress and egress of an exoplanet transit lightcurve. The scale of these effects can be $\sim$ 100 ppm, which is comparable to the scales of biases introduced by ineffective modelling of limb darkening. A few examples of third order transit features, which will be more readily observable in transit lightcurves with a strong understanding of the stellar limb darkening – exo-rings (Brown et al., 2001), stellar granulation (Chiavassa et al., 2017), and planetary thermal emission (Kipping & Tinetti, 2010).

## 4.2 Tabulated limb darkening coefficients

In cases where some prior information about limb darkening is to be included in the model, tables of pre-calculated LDCs are used to inform the prior deployed. Examples of these tabulations and how the LDC values are calculated are introduced in Section 3.5. This work utilises the Claret & Bloemen (2011) table, however any sufficiently well sampled grid of coefficients would be suitable to use for the procedure described in the subsequent sections. Because of the quality of the fits using the FCM, this work only uses LDCs calculated from least squares fits.

Figure 4.1 shows the distributions of each of the four coefficients of the non-linear limb darkening law with reference to the other coefficients fitted from individual atmosphere models as calculated by Claret & Bloemen (2011). To ensure that the LDCs displayed represent a uniform sample of stellar parameters, only the ATLAS9 models are used and this plot is only for the R passband. The correlations shown demonstrate the best fit coefficients for variations in the stellar atmosphere parameters – $T_{\rm eff}$, $\log g$, and [Fe/H]. Whilst the ATLAS9 grid used by Claret & Bloemen (2011) advertises the inclusion of $\xi_{\rm t}$ as a varying parameter, all of the

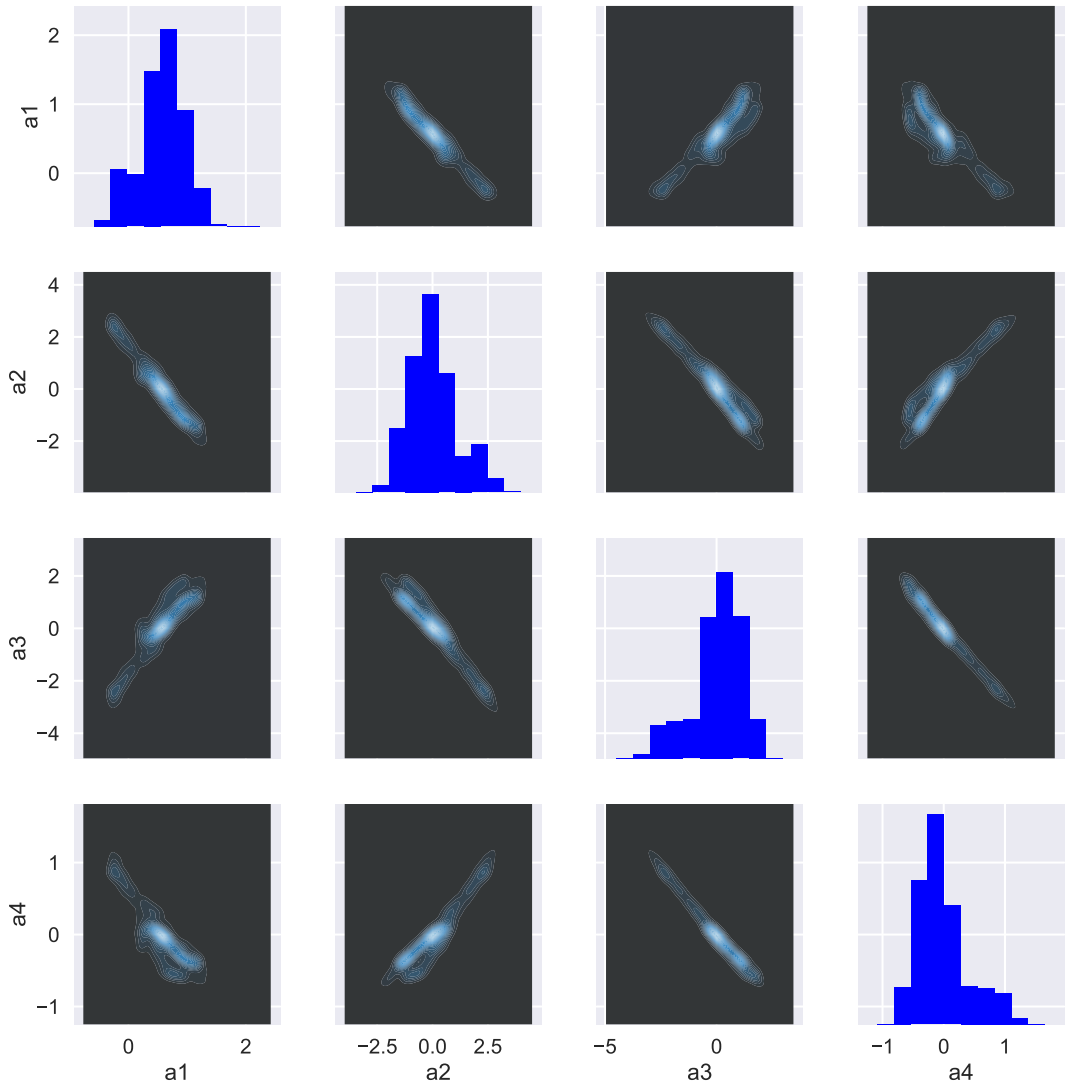



**Figure 4.1:** Figure showing the distributions of each of the values of the four non-linear law coefficients from the Claret & Bloemen (2011) tabulated LDCs for ATLAS9 atmosphere models (which are sampled uniformly in each of $T_{\rm eff}$, $\log g$ and [Fe/H]) and the R passband. The colour scale of the two-parameter plots represents the density of values found at that position.

models have $\xi_{\rm t}$=2.0 km s$^{-1}$ but for those with [Fe/H] = 0.0, so only models with $\xi_{\rm t}$ = 2.0 km s$^{-1}$ have been included in the analysis.

The distributions of the values of each of the coefficients show in Figure 4.1 are clearly correlated with the other coefficients, and the values of each do not sit in well defined ranges. Uniform and uninformative coefficient sampling schemes have been determined for the quadratic law and the 3-parameter law, but not for the non-linear law. Based on the distributions of coefficient values shown along the central spine of Figure 4.1, a potential approach to provide uninformative priors on these coefficients may be to sample from a Gaussian centred on the

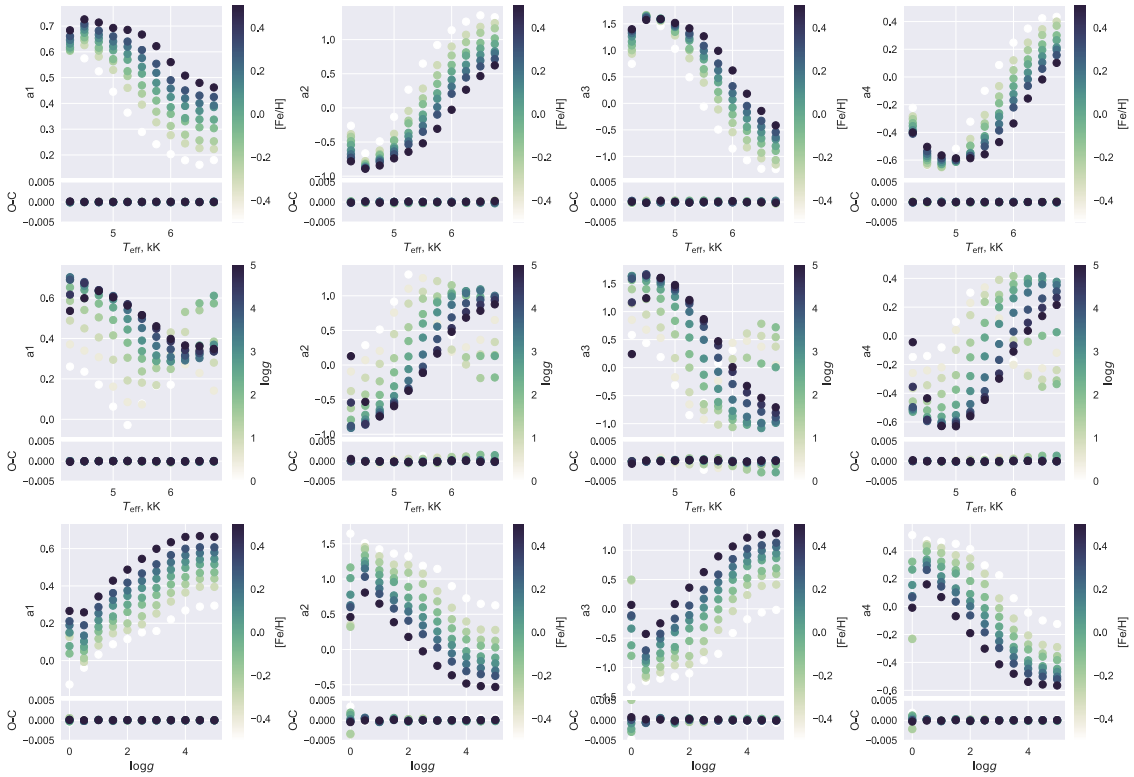



**Figure 4.2:** Figure shows how each of the four LDCs for the non-linear law vary with a subset of stellar parameters from the Claret & Bloemen (2011) table of LDCs. The LDCs included are: for the V band; with 4000 K $< T_{\rm eff} <$ 7000 K; $-1.0 \leqslant$ [Fe/H] $\leqslant +1.0$; $\xi = 2$ km s$^{-1}$; fitted using the least squares algorithm, and only for the ATLAS stellar atmospheres. Each column is for one of the four coefficients. Each row shows a different variation: the upper row is for models with $\log g = 4.5$; the middle row shows models with [Fe/H] $= 0.0$, and the lower row shows models with $T_{\rm eff} = 5500$ K. The lower plots in each window show the residuals from the fit to the tabulated LDCs plotted above.

peak value shown in the Figure, however this is only loosely supported by the shapes of the distributions, and does not account for the correlations between the coefficients. It is therefore useful to look at the variations in the calculated coefficients as the inputs to the stellar atmosphere models vary, which is shown in Figure 4.2.

It can clearly be seen in Figure 4.2 that each of the coefficients varies smoothly with the stellar atmosphere parameters. Given that these parameters are the input to the atmosphere models, it should not be surprising that this smooth variation exists in the parameter range displayed. As deployed by the JKTLD and LDTK schemes, the stellar atmosphere parameters can be used as parameters to sample the form of the limb darkening for an exoplanet transit fit. The variation shown is also clearly non-linear, so linearly (or bi-linearly) interpolating between adjacent tabulated coefficients will not entirely represent the underlying functional form of the coefficient variation with all of the atmospheric parameters.

Whilst deploying an uninformative prior on limb darkening allows the observed lightcurve to provide the sole constraint on the form of the limb darkening, the host star atmospheric parameters are not completely unknown. Not applying this existing knowledge of the system risks misrepresenting the posterior of the LDC, or biasing the posteriors found for other system parameters. The form of the limb darkening is correlated with the value of $R_\star/a$, and the transit impact parameter through the shape of the observed transit, and as such these parameters are most strongly affected by biased modelling of limb darkening.

## 4.3 Gaussian Process interpolation

Other efforts to map stellar atmospheric parameters to limb darkening coefficients, such as LDTK and JKTLD, have relied on linear interpolation between the tabulated values, which is susceptible to under-representing the full form of the multi-dimensional surface that the the tabulated coefficients sample from. This method proposes the use of a Gaussian Process (GP) to model the parameter space with reference to the stellar atmosphere parameters used in the grid of LDCs. The model can then be sampled using arbitrary stellar parameters and included in a fit of an exoplanet lightcurve.

A GP is particularly well suited to this problem, as it is known that the output surface is smooth and non-linear when parameterised with $T_{\rm eff}$, $\log g$ and [Fe/H], as shown in Figure 4.2.

### 4.3.1 Gaussian Processes Introduction

Gaussian Processes (GPs) are extremely useful tools for the probabilistic modelling of data without knowledge of the underlying function beforehand – the data determines the form of the model. The crux of the method defines a multivariate Gaussian distribution, where each data point is a dimension in the distribution. The GP is estimating the underlying function for the data with a mean function and a covariance matrix, which when determined allows the prediction of output values for arbitrary inputs without the need for a parametric model of the functional form of the intrinsic function (Rasmussen & Williams, 2006).

If there are $n$ data points in the input data, then the covariance matrix ($\mathbf{K}$) has the shape, $n \times n$, which defines the correlations between the data points. The matrix elements along the diagonal of $\mathbf{K}$ inform how an individual data point is correlated with itself, and as such

information about the white noise would contribute to the values of only those elements. The other elements of the covariance matrix provide the information about how a data point is correlated with another – for instance if the input data were a set of observations, with input: $x_1$, $x_2$, ..., $x_n$ and output: $y_1$, $y_2$, ..., $y_n$, then the covariance matrix element $\mathbf{K}_{i,j}$ would describe the covariance between $y_i$ and $y_j$.

Covariance functions are the descriptors for the domain of the function that the data should represent, and are the prior on the form of the GP. The form of covariance function used in this Chapter is the radial basis function (RBF), which is defined for each element in the covariance matrix:

$$k(x_i, x_j) = \theta_1^2 \exp\left[-\frac{(x_i - x_j)^2}{\theta_2^2}\right], \tag{4.1}$$

where $\theta_1$ is the amplitude hyperparameter describing the scale of the variability in the data, and $\theta_2$ is a length scale hyperparameter describing how well correlated the data is to nearby data. A larger value of $\theta_2$ means that the correlation between data points extends to data further away than a smaller value of $\theta_2$. Another commonly used covariance function is the white noise kernel, which as the name suggests provides a constraint on the white noise in the data. The white noise covariance function takes the form:

$$k(x_i, x_j) = \sigma_w^2 \delta_{i,j}\,, \tag{4.2}$$

where in this case $\sigma_w$ is the magnitude of the white noise, and $\delta_{i,j}$ is a Dirac delta function, that is 1 where $i = j$ and 0 everywhere else. Covariance functions are often combined to provide the appropriate prior on the form of the model.

Once the covariance function for the data set has been defined, the full covariance function for trial hyperparameters can be constructed for the data. The likelihood of the model for the selected hyperparameters can then be calculated using the following expression:

$$P(\underline{y}|\mathbf{K}) = \mathscr{L} = \left(\frac{1}{2\pi}\right)^{n/2} \frac{1}{\sqrt{\det \mathbf{K}}} \exp\left[-\frac{1}{2}(\underline{y} - \underline{\mu})^T \mathbf{K}^{-1} (\underline{y} - \underline{\mu})\right], \tag{4.3}$$

where $\underline{y}$ is the vector of the output data, and $\underline{\mu}$ is the mean model for the data. An optimisation procedure is then used to find the best hyperparameters to describe the data by maximising the value of the likelihood – the logarithm of $\mathscr{L}$ is often used for this step. Given the $\mathbf{K}^{-1}$ term in the likelihood calculation, for large datasets, the matrix inversion can be slow to compute

– it scales with $n^3$.

Having found the optimal hyperparameters, and calculated $\mathbf{K}$ for the training data, the GP is said to be conditioned, and can now be used to predict outputs for arbitrary inputs. To find the outputs of a set of test inputs, $x_{\star 1}$, $x_{\star 2}$, ..., $x_{\star m}$, the predicted outputs ($\underline{y}_\star$) are given by:

$$\underline{y}_\star = \mathbf{K}_\star \mathbf{K}^{-1} \underline{y} \,, \tag{4.4}$$

where $\mathbf{K}_\star$ is the covariance function for the cross terms between the training and test data with the same covariance function as $\mathbf{K}$, and has the shape $m \times n$. Since $\mathbf{K}^{-1}$ is already calculated from determining the optimal hyperparameters, predicting the output values for test inputs is computationally simple.

GPs have found many uses in exoplanet and stellar research, such as stellar activity variation in time, modelling red noise in transit lightcurves, or for determining stellar rotation periods; see Haywood et al. (2014); Parviainen et al. (2016); Dittmann et al. (2017); Angus et al. (2018) for recent examples.

### 4.3.2 Application and training

In order to construct the GP model for interpolation, a set of coefficients are required to train the GP hyper-parameters. This approach has been tested using the tabulated LDCs from Claret & Bloemen (2011). Only the LDCs for $4000 < T_{\rm eff} < 7000$ K; $-0.5 \leqslant$ [Fe/H] $\leqslant +0.5$, and $0.0 < \log g < 5.0$ were used. Whilst the paper accompanying the table of LDCs states that the grid also includes variation in $\xi$, for values of 0, 1, 2, 4 and 8 km s$^{-1}$, this is only the case where [Fe/H] $= 0.0$, and $\xi = 2$ km s$^{-1}$ for models with non-solar metallicity. Since the variation in the form of the limb darkening induced by a change in [Fe/H] is stronger than that introduced by $\xi$, only models with $\xi = 2$ km s$^{-1}$ are included.

Values for the LDCs have been calculated for a range of photometric bandpasses. As the bandpass is not a continuous variable, and each lightcurve is for a single fixed bandpass, only the relevant LDC values are used to pre-condition the GP. The method has been tested for several common filters, and the efficacy is consistent for each. The technique has been designed for the four parameter non-linear limb darkening law (the form of the law is shown in Equation 3.7), but the same method could be used for any law that has a sufficiently well-sampled table of coefficients.

**Table 4.1:** Details of the hyper-parameters used in GP fit to tabulated LDCs for the V passband

| Parameter | Initial value | Boundaries | Optimised value |
|---|---|---|---|
| RBF 1 | | | |
| $\sigma_f$ | 1.0 | none | $0.103^2$ |
| $l_1$ | 10.0 | 1e-5, 1e5 | 0.0917 |
| $l_2$ | 30.0 | 1e-5, 1e5 | 1e-5 |
| $l_3$ | 10.0 | 1e-5, 1e5 | 1e-5 |
| RBF 2 | | | |
| $\sigma_f$ | 1.0 | none | $0.524^2$ |
| $l_1$ | 1.0 | 1e-5, 1e5 | 0.632 |
| $l_2$ | 3.0 | 1e-5, 1e5 | 0.327 |
| $l_3$ | 1.0 | 1e-5, 1e5 | 1.37 |
| White noise | | | |
| $\sigma_w$ | 1e-5 | 1e-10, 5e-5 | 4.98e-5 |

To implement the GP, the SciKit-Learn (Pedregosa et al., 2011) python machine learning package was utilised, which contains the `GaussianProcessRegressor` method. The method provides a mechanism to construct the covariance functions, invert the covariance matrix, and optimise the hyperparameters for the data given. Once the optimal hyperparameters are found, the method is also able to predict outputs from arbitrary inputs. The optimiser deployed is the Limited-memory BFGS algorithm (Byrd et al., 1995; Zhu et al., 1997) from the SciPy[4] library.

Unlike the example of the RBF covariance function given in Section 4.3.1, each sample in the training data from the tabulated LDCs has three inputs, and four outputs. The inputs are: $T_{\rm eff}$, $\log g$ and [Fe/H], where $T_{\rm eff}$ is expressed in kK so that the variation in the input parameters are on similar scales. The outputs are the four coefficients of the non-linear limb darkening law: $a_1$, $a_2$, $a_3$ and $a_4$. To account for the extra inputs, a length scale is applied for each input, so that $l_1$, $l_2$ and $l_3$ refer to $T_{\rm eff}$, $\log g$ and [Fe/H] respectively. The amplitude of the variation for the RBF is denoted $\sigma_f$. To fully map the variation in the data, a second RBF kernel is added to the first with the same form, and a white noise kernel is added to complete the definition of the covariance function. The kernel form and values of hyperparameters are shared between each of the four outputs.

The detail of the initial values, boundaries applied, and optimal values found are presented in Table 4.1. The residuals of the fit are shown in the lower plots in each of the windows in Figure 4.2, these are very small except for some larger residuals for particularly low values of

[4]The SciPy library is an open source software package of numerical tools for python https://www.scipy.org

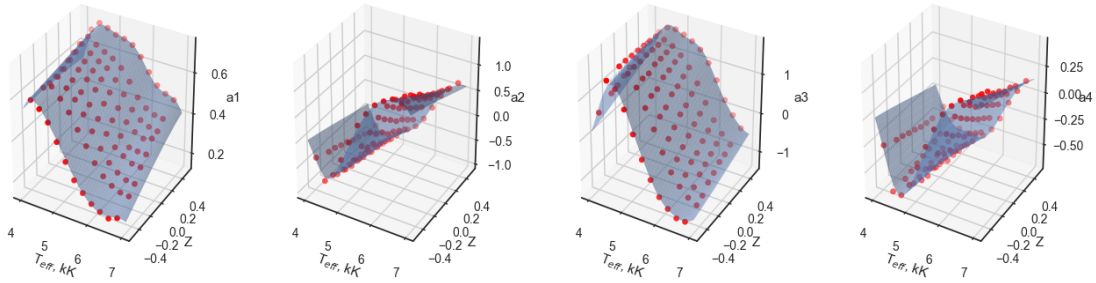



**Figure 4.3:** Figure demonstrating the GP fit (blue surface) for each of the four non-linear limb darkening law coefficients for the V band and $\log g = 4.5$ km s$^{-1}$ . The red points are the locations of the input data, which were used to condition and train the GP model. The plotted surface depicts the variation in each of the LDC coefficients induced by $T_{\rm eff}$ and metallicity (Z).

$\log g$ . An example of the fit for $\log g = 4.5$ km s$^{-1}$ is shown in Figure 4.3, where the smoothness of the output surface can clearly be seen. The variation induced by $T_{\rm eff}$ has the most complex shape – with a trend apparent across the full range of $T_{\rm eff}$ plotted, as well as variation structures over smaller scales. This structure in variation from $T_{\rm eff}$ justifies the use of two RBF kernels, and the two length scales determined for $T_{\rm eff}$ , 0.0917 and 0.632, reflect the scales of each of the variations seen in the plotted GP surface. The variation induced by [Fe/H] and $\log g$ does not appear to have the same complexity as for $T_{\rm eff}$ , and this is reflected in the relevant values for the hyperparameters detailed in Table 4.1 – a very small value for the length scales found for one of the RBF kernels indicates that the variation in the LDCs caused by these two parameters is well modelled by a single RBF kernel.

Now that the GP has been trained to optimise the hyperparameters, and conditioned for the training data, the GP can be used to predict the values of the LDCs for the non-linear law for a given exoplanet transit. The calculation of the predicted LDCs for a single transit model is very efficient, and the inverse of the covariance matrix has already been determined, so integrating this method into a transit lightcurve fit would not result in a significant computational overhead.

A GP is an appropriate method to apply in this case, over more algorithmically simpler approaches, as the functional form of the variation in the coefficients is not assumed a priori. The form of the variation in limb darkening coefficient values, is not known, so a GP can learn from the values derived from stellar atmosphere models. The other reason to apply a GP over other interpolation procedures is to take the full variation of the parameters into account, rather than just interpolating between neighbouring values, the GP calculates a multi-

**Table 4.2:** Details of the priors used and the result of each fit to the simulated transits

| Parameter | Notation | Prior[1] | Prior value | Result | | Input | Unit |
|---|---|---|---|---|---|---|---|
| | | | | Fit 1 | Fit 2 | | |
| Orbital period | $P$ | G | $\mu = 13.00000$, $\sigma = 0.00001$ | 13.00000±0.00001 | 13.00000±0.00001 | 13.0 | days |
| Transit centre | $t_C$ | G | $a = 13.0$, $b = 0.02$ | 12.9993±0.0015 | 13.00±0.02 | 13.0 | days |
| Transit depth | $d$ | U | $a = 0.001$, $b = 0.003$ | 0.0025±0.0002 | 0.0025±0.00002 | 0.0025 | – |
| Baseline | $c$ | U | $a = 0.999$, $b = 1.001$ | 0.99996±0.00007 | 1.000000±0.000007 | 1.0 | – |
| Impact parameter | $b$ | U | $a = 0.01$, $b = 0.85$ | 0.4±0.2 | 0.17±0.09 | 0.131 | – |
| Semi-major axis | $a$ | U | $a = 5.0$, $b = 25.0$ | 13.7±1.4 | 14.9±0.2 | 15.0 | AU |
| Effective temperature | $T_{\rm eff}$ | G | $\mu = 5500$, $\sigma = 100$ | 5500±100 | 5510±80 | 5500 | K |
| Surface gravity | $\log g$ | G | $\mu = 4.5$, $\sigma = 0.2$ | 4.5±0.2 | 4.4±0.2 | 4.5 | – |
| Metallicity | [Fe/H] | G | $\mu = 0.0$, $\sigma = 0.1$ | 0.0±0.1 | 0.0±0.1 | 0.0 | dex |

**Note:** $\mu$ and $\sigma$ are the mean and standard deviation of the Gaussian prior. $a$ and $b$ are the lower and upper bounds of the uniform prior respectively. [1] "G" denotes a Gaussian prior, and "U" denotes a uniform prior

dimensional surface across the entire space that training data is provided for and more readily takes correlations between the coefficients into account.

### 4.3.3 Example fits

To demonstrate this GP interpolation method as part of an exoplanet transit fitting procedure, two lightcurves have been generated. The parameters used to generate each from the BATMAN transit model (Kreidberg, 2015)[5] are detailed in the "input" column in Table 4.2. The model was used to generate two lightcurves with different noise levels – the first had random values drawn from a normal distribution with $\sigma = 0.001$, and the second with $\sigma = 0.0001$ added to the modelled flux values. The uncertainty on each data point was taken to be the $\sigma$ used to add noise to the model. The lightcurves were calculated for 500 evenly spaced times between 12.75 and 13.25 days. The non-linear limb darkening coefficients used in the simulation were [0.5742, -0.2175, 0.8311, -0.4144], which were the tabulated coefficients given by Claret & Bloemen (2011) for a star with $T_{\rm eff} = 5500$ K, $\log g = 4.5$, and [Fe/H] $= 0.0$ for the *Kepler* bandpass and determined from an ATLAS9 stellar atmosphere model.

The simulated lightcurves were modelled using an MCMC procedure and the BATMAN transit model[6]. The parameters for each fit, and the priors applied for each are described in Table 4.2. The GP model was trained as described in Section 4.3.2 for the *Kepler* bandpass. The MCMC implementation was the EMCEE package (Foreman-Mackey et al., 2013). Each of the 100 walkers were run for 5000 iterations, with the first 500 steps for each walker discarded as burn-in, and only every 50th step retained to avoid correlations between adjacent steps. The remaining trial steps are used to estimate the posterior distributions for each parameter

[5]More details about how the BATMAN model works in Section 6.4.1.
[6]More specifics about the transit fitting procedures used throughout this work is included in Section 6.4.2

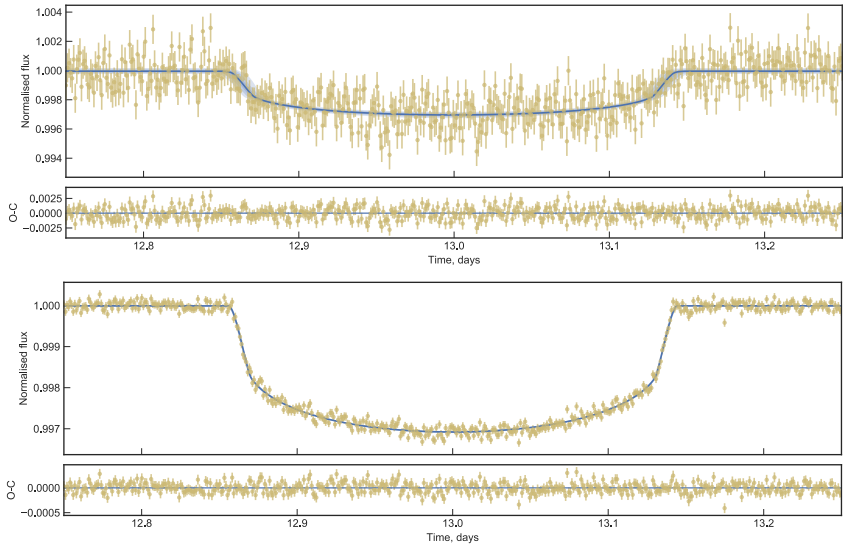



**Figure 4.4:** Output model (shown in blue) plotted for the simulated transit lightcurves (shown in yellow). The upper plot is for "fit 1" with a higher noise level, and the lower plot is for "fit 2". The models are plotted with 1-, 2- and 3-$\sigma$ uncertainty estimates from the posterior distributions. The residuals of the fit for each model are shown below the primary plot.

modelled. The results of the fit for each simulated lightcurve are shown in Table 4.2.

Figure 4.4 shows the results of the fits plotted with the simulated lightcurves, and the residuals of the fit. Both fits produce relatively flat residuals. Based on the differences between the input parameters and those found after fitting the simulated lightcurves, "fit 1" did not well determine the underlying parameters for the system. Whilst the results of "fit 1" did not well match the input parameters, the stellar parameters used to inform the limb darkening coefficients were not shifted away from the priors, which indicates that the high noise level in the simulated lightcurve was the primary driver behind the mismatch between input and determined parameters. The results of "fit 2" are very close to the input parameters, which demonstrates the efficacy of the GP interpolation procedure for modelling limb darkening in a transit lightcurve fit.

The GP interpolation method is seen to fit the simulated lightcurves very well, particularly for the lightcurve with lower noise levels. The limb darkening in the simulated lightcurves was modelled with coefficients for the non-linear limb darkening law. It is therefore unclear from these fits how effective the method is at modelling real limb darkening, to do this real rather than simulated data is required. More fits using the GP interpolation method are performed in Section 6.4, with the results discussed there.

### 4.3.4 Discussion and limitations

The method is very effective at producing values for the LDCs for stellar parameters that are not included in the tabulated values. The GP is able to map the variation surface in the LDCs, and thus allow a higher order limb darkening law to be included in the fit of an exoplanet transit constraining by existing knowledge of the stellar atmosphere parameters.

Whilst the precise surface of the LDC variation can be well mapped, the technique can only be applied for a single grid of tabulated LDCs at a time, and the accuracy of the modelled results are dependent on several other levels of modelling which occur before the GP is trained. The tabulated LDCs rely on the stellar atmosphere model accurately representing the CLIV of the stellar surface for the stated parameters, which is difficult to test when real stellar surfaces cannot be resolved. The method also depends on the assumption that the non-linear limb darkening law can provide a good description of the form of the CLIV, as well as the assumption that the fit of the LDCs to the model outputs is representative of the model atmosphere outputs.

To utilise the GP interpolation, a prior on the stellar atmosphere parameters must be provided. For the systems bright enough to observe and fit transit lightcurves for, some measure of these parameters will be known, so it is valuable to add this information and propagate the uncertainty through the transit fitting procedure.

# 5

# Artificial Neural Networks

## 5.1 Introduction

Neural networks, or artificial neural networks as they are more properly known, are based on how neurons in the brain work. The basic principle of the functioning of a biological neuron is that it consists of dendrites, a soma, an axon, and synapses, as shown in Figure 5.1. Each neuron is connected to thousands of others and electro-chemical signals are sent between them. If the sum of each of the signals received by the dendrites in the cell is large enough, the cell sends an electro-chemical signal along the axon (also known as the cell firing). The axon terminals pass the transmitted signal through synapses to the dendrites of connected neurons, and the process continues transmitting information around the brain.

Artificial neural networks (ANN) were first modelled as a mathematical representation of the neural processes of a brain. A formulation first proposed by McCulloch & Pitts (1943) was based on contemporary understanding of neuroscience, although the status of computer development at the time meant that the algorithm had little application. The advance which

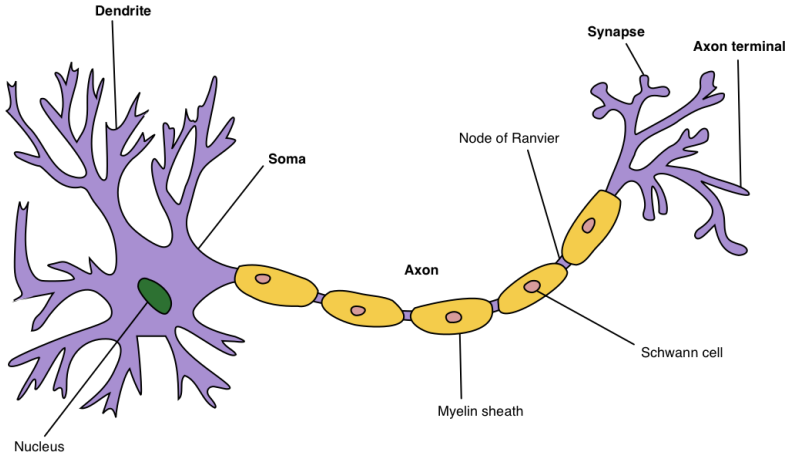



**Figure 5.1:** Schematic of the key features of a neuronal cell. Image from Quasar Jarosz at English Wikipedia, shared under Creative Commons.

led to the initial spike in interest in the field of neural computation was the introduction of the perceptron by Rosenblatt (1958). The perceptron was originally designed to function as a physical machine, where the inputs were photocells connected to the neurons, with the weights and biases controlled by potentiometers. The machine could then be trained with examples to predict which class the pattern matched, with the weights adjusted during training with electric motors. The algorithmic form of the perceptron can be applied in more modern computational environments as well, and its form is the basis for the neural networks utilised in this work.

## 5.2 Theory

### 5.2.1 Artificial neuron

The perceptron mentioned above is a specific type of neuron – the building blocks of any larger network. A schematic of the form of an artificial neuron is shown in Figure 5.2. The elements of an artificial neuron can be matched with a analogous element for a biological neuron – the inputs of the artificial neuron are comparable to the electro-chemical signals that reach the dendrites of the neuron; the weights are representative of the strengths of the synapses transferring the signal to the dendrites; the summing junction matches the behaviour of the whole soma of the neuron; the activation function is equivalent to the decision the neuron makes about whether to fire an electrochemical signal along the axon, and the outputs are

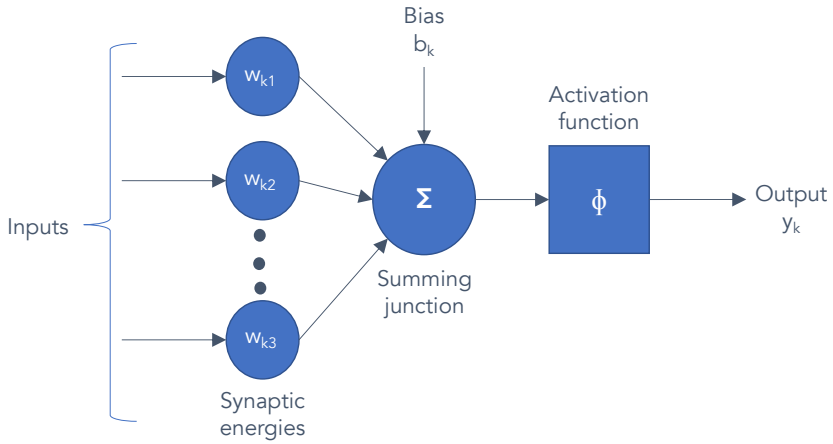



**Figure 5.2:** Schematic of an artificial neuron

like the axon terminals. The mathematical expression of the perceptron is simple:

$$y = \Phi(\mathbf{x}) = \begin{cases} 1 & \text{where } \mathbf{w} \cdot \mathbf{x} + b > 0 \\ 0 & \text{otherwise} \end{cases}, \tag{5.1}$$

where $\mathbf{x}$ is the input vector, $\mathbf{w}$ is the vector of the weights, which has the same shape as the input, $b$ is the bias term for that neuron, and $\Phi(\mathbf{x})$ is the activation function. The perceptron with well trained weights and an appropriate bias term acts as a binary classifier – it is very effective at learning to classify problems that are linearly separable. The perceptron uses a step function as the activation function, but a neuron can utilise a more specialised activation function to describe the output behaviour.

### 5.2.2 Activation functions

Several other activation functions are used in ANNs. The most common are sigmoid functions and rectified linear units (LeCun et al., 2015). Both of these functions differentiate themselves from the step function used in the perceptron as they are non-binary functions. The ability of these functions to take more subtle values of output allows the ANN to approximate more complex functional forms than a combination of step functions could.

The sigmoid function takes the form:

$$\Phi(z) = \frac{1}{1 + e^{-z}}, \tag{5.2}$$

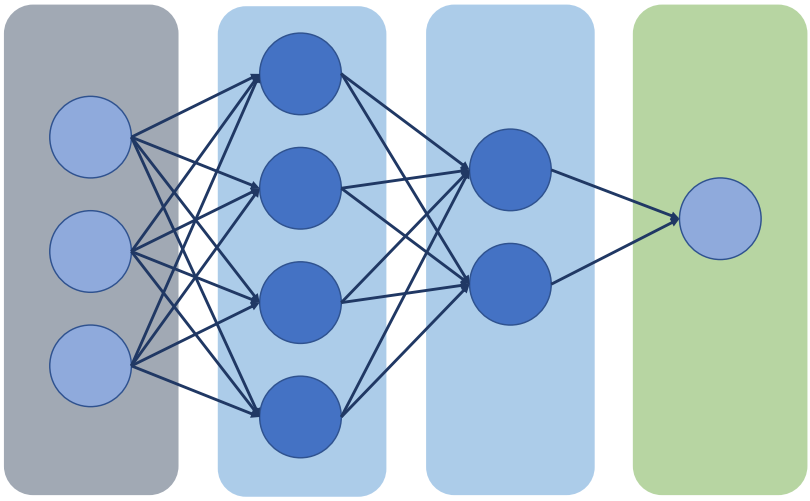


**Figure 5.3:** Schematic of a multi-layer perceptron, with the flow of information from left to right. The left-most column of circle represents the input into the network. Each input is connected to every neuron in the first hidden layer, and each output of that layer is connected to each input in the second hidden layer, and then to the final output. The circles represent neurons, and the connecting lines each carry a weight. There would also be a bias term for each neuron.

where $z = \mathbf{w} \cdot \mathbf{x} + b$. This function acts like a soft step that always has output in the range $[0,1]$. The rectified linear unit (ReLU) is:

$$\Phi(z) = \begin{cases} 0 & \text{for } z < 0 \\ z & \text{for } z \geqslant 0 \end{cases}. \tag{5.3}$$

The sigmoid function had been very commonly used as the standard activation function in multiple-layer perceptron networks, but as more layers were added, effective training of the ANN became challenging due to the vanishing gradient problem. The vanishing gradient problem arises because the gradient of the sigmoid activation function approaches zero as $z \to \pm\infty$, which ensures that any adjustment to weights is small, and thus training is slow. The introduction of the ReLU activation function prevented this problem, as the gradient does not tend to zero when $z \to +\infty$, and is therefore now prevalent in deep networks (Nair & Hinton, 2010).

### 5.2.3 Multi-layer perceptron

As alluded to above, neurons are not deployed on their own, but instead have more use when constructed as a network with multiple layers. An example of a multiple-layer perceptron

(MLP) network is shown in Figure 5.3, which has two hidden layers, each with multiple neurons. Whilst the term "perceptron" is in the name of this type of network, the neurons can have arbitrary activation functions.

The information in this type of network only flows in one direction, so it is termed a feed-forward network. Every neuron is connected with a weight to every other in neighbouring layers, so is referred to as fully connected. These are the form of networks used in this work, and the specific architectures are presented in Chapters 6 and 7.

MLPs can have an arbitrary number of layers, although the more layers there are, the more computationally heavy the training calculations are. Each layer can also have as many or as few neurons as function best for the problem, and the number of neurons does not have to be the same layer to layer, although the number of neurons usually decreases with increasing layer number. The exact classification of when a neural network becomes a deep network is undefined, but is usually reserved for applications with $> 3$ hidden layers.

The fully connected feed-forward network architecture described above is not the only form of useful ANN architecture. Convolutional neural networks (CNN) allow 2D spatial structure to be modelled by the network, rather than modelling each input independently, which is particularly useful for image classification problems because the identified features are then translationally independent. The first model of this form of network was proposed by Hubel & Wiesel (1959) based on the way that cells in animal visual cortexes responded to specific regions of their visual field.

Recurrent neural networks (RNN) are another different form of neural network, for which the flow of information does not follow a linear path from input to output. They were first developed by Hopfield (1982) and have found huge success in speech recognition applications.

### 5.2.4 Learning algorithm

The architecture of an ANN is described above, but in and of itself the form of the network does not provide much use without the weights and bias values being trained. Since this form of neural network is a supervised machine learning algorithm, a set of training inputs for which the real outputs are known is required to train the ANN.

For an MLP network, there are several steps to optimise the weights and biases. For the first iteration of learning, the values are usually initialised randomly to small values. Updating

the values can be performed in several modes – weights can be updated after a single training example (known as stochastic training), or after a batch of training examples. The stochastic method is useful to prevent getting stuck in local minima, but the process is slow particularly for large training sets as the calculation cannot be parallelised, and introduces noise in the training from individual examples. Batch training can be performed much faster, as the calculation can be hyper-parallelised, which allows more iterations of training through the entire training set, and reduces noise as averages are used. Batches that are particularly large can more easily get stuck in local minima, so a compromise is usually made to use mini-batches during training.

For each batch the output is predicted by the ANN, and compared with the known real output using a loss function. In this work the mean square error is used,

$$E(\mathbf{X}, \theta) = \frac{1}{T} \sum_{t=1}^{T} (y_t - \hat{y}_t(\theta))^2 \,, \tag{5.4}$$

where $\theta$ is the vector of all the trainable parameters of the system (weights and biases); $\mathbf{X}$ is the array of all the training example inputs in the batch; $T$ is the number of training examples in the batch; $y_t$ is the true output for that example, and $\hat{y}_t$ is the output predicted by the ANN. Once the loss function for the batch has been calculated, some adjustment to the weights needs to be applied to try and reduce the loss. It is not inherent to the loss value which parts of the network are performing well or badly, so the backpropagation algorithm is deployed to distribute the changes in weight values in the network.

Backpropagation for the training of multi-layer networks was introduced by Rumelhart et al. (1986). To aid the description of the method, nodes will denote inputs, outputs and neurons, which can all be connected. The algorithm relies on the following parameters from $\theta$: $w^k_{i,j}$, which is the weight between node $j$ in layer $l_k$ and node $i$ in layer $l_{k-1}$, and $b^k_i$, which is the bias for node $i$ in layer $l_k$.

The aim of backpropagation is to calculate the gradient of the error for each weight and bias term, so that the value can be updated. Since the intermediate layers do not have a target output, the algorithm works backwards from the final layer to calculate the partial derivative of the error with respect to each weight or bias. The error terms for the nodes in the final layer can be calculated from the predicted output and the derivative of the activation function, and then this can be propagated backwards through the layers until the input layer is reached.

There is then a partial derivative of the error for each node in the network for each training example in the batch, which can then be used to update the weights and biases:

$$\Delta w_{i,j}^{k} = -\alpha \frac{1}{T} \sum_{t=1}^{T} \frac{\partial E_t}{\partial w_{i,j}^{k}} \ , \tag{5.5}$$

where $E_t$ is the error for the individual training example, and $\alpha$ is the user selected learning rate. For large networks and large training sets, calculating the gradients is a significant computational expense, however the values for each training example can be computed in parallel.

A training epoch refers to how many times the entire training data set has been used to adjust the weights. This will be a lower value than the number of iterations of the gradient descent algorithm that have been called when batches of the training data are used rather than the entire data set.

A significant risk with training an ANN is over-fitting. The training data may well contain features or trends by chance that are not representative of the underlying features of the full data set. If the ANN model has been over-fit, then the ability of the ANN model to accurately produce predictions is compromised, as the predictions will be based on the features of the training data rather than the features of the full data set – the model does not generalise. There are several approaches that can be taken to prevent over-fitting, one common approach is to remove some of the training data, and use it as a validation data set once training is complete. Another proposition to prevent over-fitting is to add a regularisation parameter to the final loss function, this would take the form:

$$\delta \theta^T \theta \ , \tag{5.6}$$

where $\delta$ is the regularisation coefficient. $\delta$ would be a small number (eg. 0.05), and would act to penalise the fitting of localised spikes and encourage smoothness in the output.

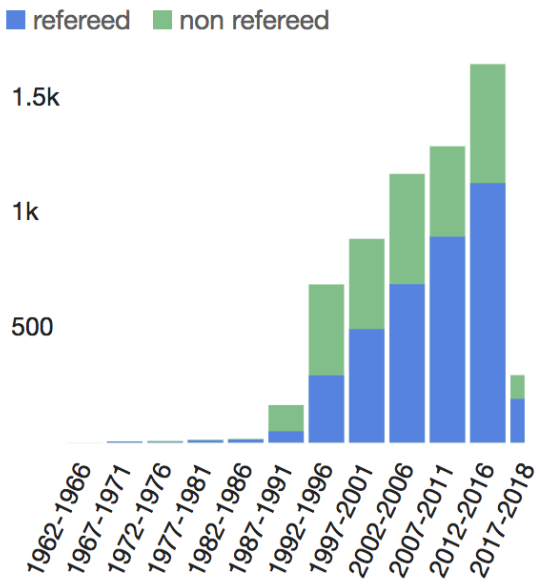



**Figure 5.4:** Plot of mentions of "neural networks" in astrophysics publications over time generated from the Astrophysical Data System [accessed in November 2017].

## 5.3 Examples in astronomy

ANNs have been used in astronomy for many years, which is demonstrated in Figure 5.4. The plot shows mentions of "neural networks" in astrophysical publications over time[1], which have been gradually increasing since the earliest applications. Alongside the increase in interest in neural networks in astrophysics research, has been an increase in interest in neural networks across most fields of academic research. This demand has led to the development of toolkits to train ANN architectures efficiently and easily, which has in turn made the use of ANNs simple in a broad range of research areas.

Early examples of ANNs in astrophysics used simple networks with some success, such as Odewahn et al. (1992) and Storrie-Lombardi et al. (1992) were able to train a network to classify galaxy images. Bailer-Jones (2000) was the first to construct an ANN to determine stellar parameters such as $T_{\rm eff}$ and $\log g$ directly from low resolution spectra.

More recently, the ability to classify galaxy images has improved (see eg. Schaefer et al. (2018)), and a broader variety of image classification aims have been realised. ANNs have also been deployed in different ways in exoplanet and stellar astrophysics, for example Kipping &

[1]The plot was generated using the Astrophysical Data System (ADS), which can be accessed here: https://ui.adsabs.harvard.edu/

Lam (2017) used ANNs to predict the most likely short period transiting planet systems to have companion planets on longer orbits. Hon et al. (2017) trained a deep convolutional neural network to perform astroseismic classification of red giants. Shallue & Vanderburg (2018) found a new multiple planet system in *Kepler* lightcurve data after training a deep convolutional network to identify transit candidates.

# 6

# Limb darkening interpolation with ANNs

## 6.1 Introduction

The method for modelling limb darkening presented in Chapter 4 provides a means to map the stellar atmosphere parameters to the coefficients of a parametric limb darkening law. This is effective at representing the variation in the coefficients with stellar atmospheric parameters, but through using tabulated grids of LDCs, the method is strongly dependent on the fit of the parametric law to the stellar atmosphere model. The fit may not have accurately matched the model output, Heyrovský (2007) showed that the values found for calculated LDCs are observably different when different fitting procedures are used. The form of the parametric law which has few enough parameters to be able to fit the lightcurve may not be able to fully represent the shape of the limb darkening effect.

Given the improvements to the ability to efficiently approximate complex multi-dimensional functions using artifical neural networks (ANN), which were introduced in Chapter 5, a new approach for modelling limb darkening is proposed in this chapter. Using the output specific

intensities calculated from stellar atmosphere models as training data, the ANN can be trained to produce a scaled intensity for a position on the surface of a star given $T_{\rm eff}$ , $\log g$ , [Fe/H] and $\mu$. This would remove the need to define a parametric law for limb darkening entirely. Stellar atmosphere models are known to well represent the limb darkening of the Sun (Asplund et al., 2009), and for other stars from high SNR transit fits (Müller et al., 2013).

The pre-trained ANN can then be sampled as a method within a transit fit, with stellar parameters as input, and the relative stellar intensity at a given position on the stellar surface as output, which can then be fed directly into a transit model. This chapter goes through the steps of how the ANN is built, trained and then deployed in some example transit fits.

## 6.2 Stellar atmosphere models

To provide the training data for the proposed ANN, a large grid of stellar atmosphere models is required. For this work, the Husser et al. (2013) grid of PHOENIX models is used. The physics of the PHOENIX atmosphere models is introduced in Section 3.3. The grid covers the following parameter space: 2300 K $\leq T_{\rm eff} \leq$ 7000 K in steps of 100 K, and 7000 K $\leq T_{\rm eff} \leq$ 12,000 K in steps of 200 K; 0.0 $\leq \log g \leq$ +6.0 in steps of 0.5, and -4.0 $\leq$ [Fe/H] $\leq$ -2.0 in steps of 1.0, and -2.0 $\leq$ [Fe/H] $\leq$ +1.0 in steps of 0.5. The published grid also includes alpha element enhancement, but any models with $[\alpha/\mathrm{Fe}] \neq 0.0$ are not included in this work.

The specific intensity spectra for the grid[1] consist of specific intensity values for a range of $\mu$ values for every 0.1 nm in the range 50 – 2600 nm. Each file has a specific intensity spectrum for approximately 80 $\mu$ values between 1 (the centre of the stellar disk) and a small value near 0 (the edge of the star). These angles are not evenly sampled in $\mu$, but rather more densely sampled near the stellar limb, so to provide more information where the largest change in the intensity is occurring.

Of the grids of stellar atmosphere models introduced in Section 3.3, the Husser et al. (2013) PHOENIX grid was selected as the basis for the training grid used in this application as the aim is to use as realistic and comprehensive an understanding of stellar atmospheres as possible to inform the ANN approach. The ATLAS models do not include as much physics, and the plane-parallel model geometry is particularly deficient at the edge of the star, which is where the ANN can be especially effective. A grid of fully 3D models with all of the known stellar

[1]The fits files containing the spectra were downloaded from ftp://phoenix.astro.physik.uni-goettingen.de/

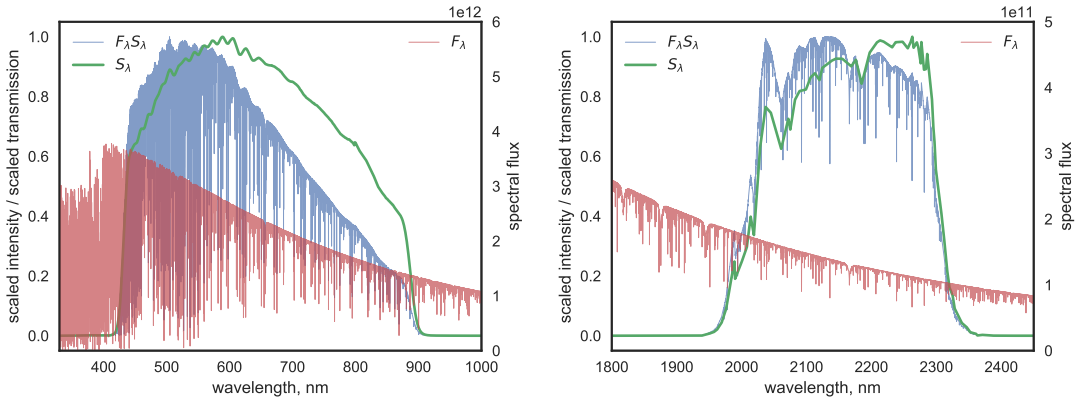



**Figure 6.1:** Example of the output of the model atmosphere models (shown in red) for a model at $\mu = 1$. The model shown has $T_{\rm eff} = 5500$ K; $\log g = 4.5$, and [Fe/H] $= 0.0$. The left plot is for the Kepler bandpass (transmission shown in green), and the right plot is for the 2MASS K band. The blue lines show the convolution of the model output with the bandpass, which is then integrated across wavelength to produce the value for specific intensity for that bandpass, model and $\mu$ value.

physics included would be the optimal grid to build the training data from, but no such grid of stellar atmosphere models exists. The STAGGER grid includes more physics than the PHOENIX models, but the grid is relatively small, so there is not enough data to train the ANN effectively (an example of an ANN limb darkening model trained with the STAGGER grid is demonstrated in Chapter 7). The PHOENIX grid combines being a large enough grid to be a useful training data set, and including enough physics to well represent the true limb darkening behaviour of the stars that the grid is modelling.

### 6.2.1 Model output pre-processing

The next step in creating the training data for the ANN is to pre-process the outputs of the stellar atmosphere models into the form of output required to feed into a transit model – the stellar intensity at a given position on the stellar surface (parameterised by $\mu$). An example of this process is described below, and shown for an example atmosphere in Figure 6.1.

The red spectra shown in each of the plots in Figure 6.1 demonstrate an example of the the output of the atmosphere models in the grid in the region of two photometric passbands – *Kepler* and 2MASS K. When fitting limb darkening exoplanet transit data, the lightcurve has been generated from observing with a specific filter, thus the model output must be convolved with the corresponding filter transmission curve to represent the form of limb darkening relevant to the observed lightcurve.

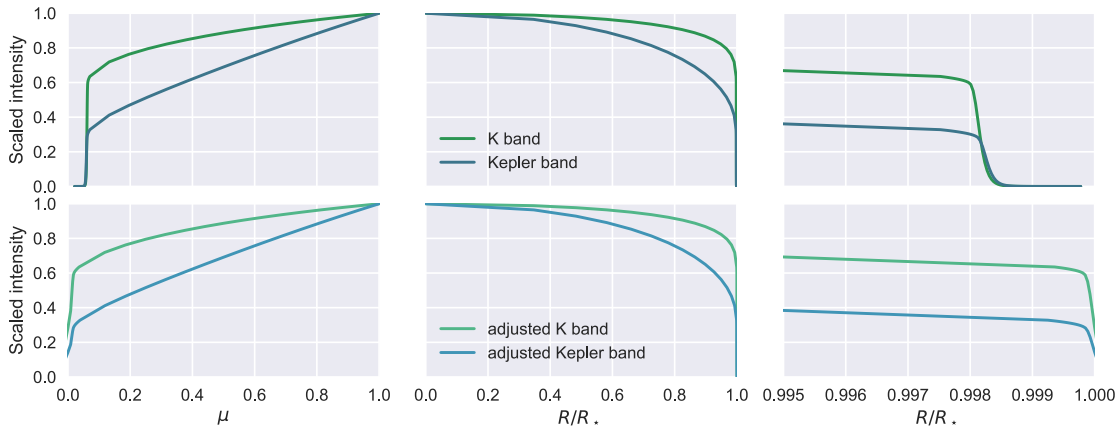



**Figure 6.2:** Example of limb profile from PHOENIX stellar atmosphere model for stellar properties $T_{\rm eff} = 5500$ K; $\log g = 4.5$, and [Fe/H] $= 0.0$. The upper row of plots show the limb profile as found directly from the convolution of the stellar atmosphere model output for the 2MASS K band, and the Kepler bandpass, plotted against $\mu$ on the left, and $R/R_\star$ in the middle and on the right, where the plot zooms into the limb region of the intensity profile. The lower plots are the same, but after the radius correction has been applied, so the values of $\mu$ and $R/R_\star$ are mapped from the effective radius defined by the correction, rather than that in the atmosphere model.

The thick green line in each plot in Figure 6.1 shows the transmission curves used, and the blue spectra show the result of convolving the spectra with the transmission curves. The total intensity for each model atmosphere for each $\mu$ value for each passband is obtained by integrating each of the convolution results with respect to wavelength. The resulting profile for each model and each bandpass is then normalised to equal unity at $\mu = 1$.

The convolution for calculating the intensity profiles from the PHOENIX models is represented by this integral:

$$I(\mu_i) = \int_{\lambda_1}^{\lambda_2} I(\lambda, \mu_i) S_\lambda(\lambda) \frac{\lambda}{hc} \, d\lambda \, , \tag{6.1}$$

where $I(\mu_i)$ is the stellar surface brightness at each $\mu_i$ from the model atmosphere; $\lambda_1$ and $\lambda_2$ are the lowest and highest wavelengths calculated for the atmosphere model respectively; $I(\lambda, \mu_i)$ is the model atmosphere intensity for each $\mu_i$ from the model; $S_\lambda(\lambda)$ is the wavelength dependent transmission curve for the selected photometric bandpass, and $\frac{\lambda}{hc}$ is the correction applied to adjust the model output from energy flux to photon number flux, which is what is measured by a CCD.

### 6.2.2 Plane-parallel versus spherical geometry

Unlike the plane-parallel ATLAS9 stellar atmosphere models, the PHOENIX models have a spherically symmetric geometry, which means that the way that the stellar radius is defined

in the model is different. In a plane-parallel model, the final layer of the model is the edge of the star, which means that a discontinuous drop in star light is seen at that edge, caused by the sharp transition from optical thickness to thinness regardless of wavelength in the model structure.

In an atmosphere model with spherically symmetric geometry, there is no clearly defined stellar edge, which better represents the realistic soft transition from optically thin to optically thick regions of the stellar atmosphere. Without a defined stellar edge, the modeller must choose some position to define the stellar radius as the reference radius for calculating $\mu$ values. This is usually a radius where the stellar atmosphere is known to be optically thin across all modelled wavelengths.

In an exoplanet transit fit, the radius of the star is understood to be the position at which the stellar atmosphere becomes optically thick, or where the decrease in observed flux begins at the start of a transit. The upper row of plots in Figure 6.2 demonstrates how the radius of the star in the PHOENIX models is defined beyond the region where the star is emitting in that wavelength.

A standard transit model using a parameterised limb darkening law will assume that beyond $R = R_\star$ ($\mu = 0$) there is no stellar emission, and that the sharp drop-off in intensity will occur at that radius. To be able to sensibly compare the transit parameters derived from the stellar radius for the method proposed here and for a parametric limb darkening law, the stellar radius must be defined in the same way, so an adjustment is applied to the PHOENIX atmosphere model outputs.

Wittkowski et al. (2004) proposed that the radial position of maximum gradient in intensity be denoted as the stellar radius. This method was also applied by Espinoza & Jordán (2015). For each of the limb profiles calculated, the position of maximum gradient in intensity with respect to radius was found, and the $\mu$ values inside of that position were rescaled to account for the new definition of $R_\star$. The results of the rescaling can be seen in the plots in the lower row of Figure 6.2 for the example limb profiles.

It can also be seen in Figure 6.2 (particularly in the upper right hand plot) that the position of maximum gradient is not the same for all passbands for the same stellar atmosphere. The wavelength dependence of the stellar radius position is the focus of Chapter 7. In this chapter however, the stellar radius is assumed to be constant for all wavelengths, which is one of the

results of applying the adjustment to the definition of $R_\star$.

### 6.2.3 Training data grid creation

The limb profiles were calculated for a range of photometric bandpasses – the details of the bandpasses used are collated in Appendix A. The ANN is trained for a single bandpass at a time, so the training grid consists of four input parameters – $T_{\rm eff}$, $\log g$, [Fe/H], and $\mu$, and one output parameter – $I$, which is the stellar surface intensity for the atmosphere model with that set of parameters at that $\mu$ value.

Two training grids are created, one with the $\mu$ values as provided in the model atmosphere outputs and then adjusted as described in Section 6.2.2, and the other grid with the $\mu$ values for each limb profile interpolated to 200 evenly sampled angles. Resampling tests how sensitive training the ANN is to a uniform grid in each parameter, as the initial $\mu$ values are tightly clustered around low values, which could bias the training. This form of resampling is performed before fitting LDCs to stellar atmosphere model outputs by Claret & Bloemen (2011); Parviainen & Aigrain (2015).

The final step in producing the grid of training data is to rescale the $T_{\rm eff}$ values to kK, rather than K, so that each input parameter is at a similar scale, which is significantly more efficient for training the network.

## 6.3 Artificial Neural Network architecture

To implement the ANN model, the structure was written into the framework offered by the KERAS python package[2]. The package offers a highly flexible API to build and train ANNs in python, with the use of THEANO[3] as a backend engine. THEANO is compiled to run on the GPU of a computer, which allows the ANN to be trained and run orders of magnitude faster than on a CPU.

Deciding on the exact form for the neural network architecture for this application is not a trivial problem. There are too many parameters for the structure that can be adjusted to perform a numerical optimization, as each instance of training can take several minutes. The

[2]KERAS is a high-level neural networks API (application programming interface) written in Python, and developed by Chollet et al. (2015).
[3]THEANO is a Python library that allows one to define, optimize, and evaluate mathematical expressions involving multi-dimensional arrays efficiently (Theano Development Team, 2016).

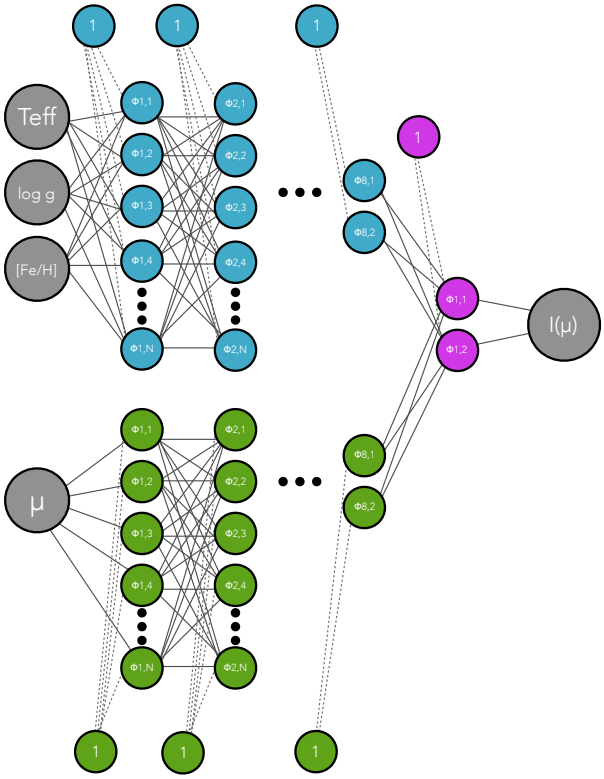



**Figure 6.3:** Schematic representation of the architecture used in the final ANN for this work, where the grey circles on the left represent the input neurons for an individual trial; the solid grey lines are the weighted connections between neurons; the dashed grey lines are the bias terms for each neuron; the coloured circles represent the neurons themselves organised in layers, and the grey circle on the right is the output neuron. The blue circles are the stellar parameter branch, and the green circles are the $\mu$ value branch.

key parameters to chose scales for are: the number of hidden layers used; the number of neurons in each hidden layer; the activation function applied for each layer; the learning rate, and how many training epochs to use. Each of these terms are defined and described in Chapter 5.

The form of the final model used is shown in schematic form in Figure 6.3, where input data (on the left) is transformed through the network to the intensity value (on the right). The split form of the model, where the stellar atmosphere parameters are modelled separately to the $\mu$ value, was decided upon for two reasons. The first justification for splitting the model was based on the astrophysics – two effects with very different forms are known in the training data: the overall reduction in intensity from the centre to limb for that passband, and the variations to that shape introduced by variation in the stellar parameters. The second justification for splitting the model was more pragmatic – it worked more effectively. This type of approach (for a much simpler ANN) was utilised in the final model of Kipping & Lam

(2017).

### 6.3.1 Model structure testing

Once the basic form of the ANN had been determined, the exact size and training parameters of the model had to be selected. The functional form of the variation induced by the stellar atmosphere parameters is much simpler than the variation in the $\mu$ value branch (shown with green circles in Figure 6.3) – particularly as the variation in the $\mu$ value branch has different behaviour on different scales. The variation has a shallow gradient from the centre of the star to near the edge, and then a steeper gradient in the remaining $\sim$10% of the $\mu$ values. The simpler output form of the stellar parameter branch (shown in blue in Figure 6.3) led to the expectation that a smaller (shallower and narrower) branch would be appropriate compared to that required for the $\mu$ value branch. Whilst training the different model architectures, any trials of architectures where the branches were not symmetric was not able to well represent the form of the CLIV, so the final form has symmetry between the two branches. Symmetry of the branches in this sense refers to each of the branches having the same number of neurons and layers, and well as the same activation functions for each neuron. The final number of layers deployed in the ANN was selected based on tests of models with different numbers of layers and finding a size that was deep enough to well represent the CLIV from the training data, but not so deep that each calculation of the model would be too slow to be useful for transit fitting.

In the first instance, the training loss (the mean square error for a regression fit, as defined in Section 5.2.4) is a useful metric to compare the efficacy of the fit for each form of the model to the training data. Figure 6.4 shows how the logarithm of the training loss reduces with increasing training epochs for four sizes of ANN models. The forms of each of the models and training parameters were kept the same, and only the number of neurons in each layer were adjusted relative to the final model detailed in Table 6.1. All of the example plots in this section used the Bessell V band training data.

Figure 6.5 shows an example limb profile from the training data with the same model fits as in Figure 6.4 overplotted. It is easy to see that the ANN approach in general is a good representation of the data used to train it, and it is positive to see that the output profile is smooth, as smoothness is not an inherent feature of this type of ANN model. The value of the loss function mentioned above as a good initial test of the goodness-of-fit of the model

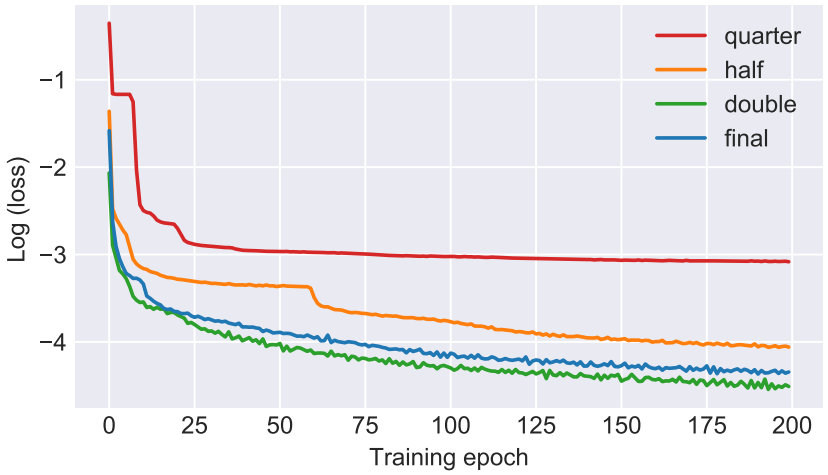



**Figure 6.4:** Plot demonstrating how the log of the loss function of the model training changes with the number of training epochs for different ANN model architectures. The models trained have the same form as the final ANN model detailed in Table 6.1, but for a quarter, half, and double as many neurons in each layer as reflected by the names.

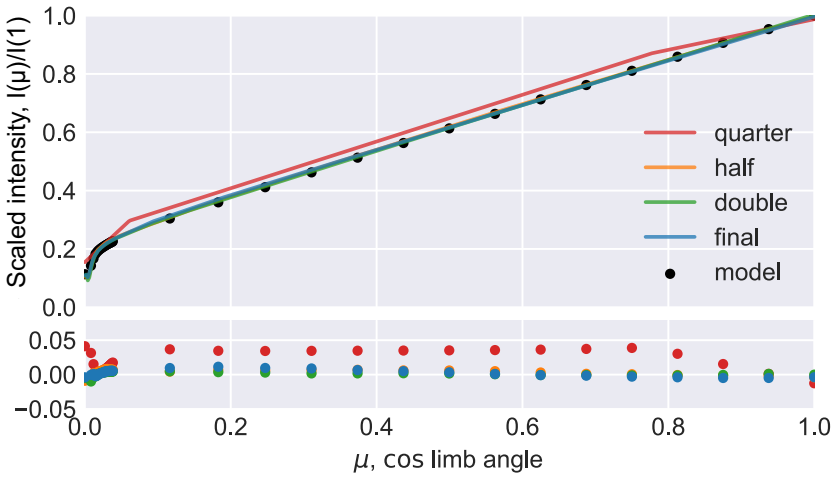



**Figure 6.5:** Plot is for an example stellar model atmosphere (used as training data points and shown as black circles) where $T_{\rm eff}$ = 4500 K, $\log g$ = 4.5 and [Fe/H] = -0.5. The outputs of the trained ANN models are shown as different coloured lines for varying numbers of neurons in each layer. The residuals of the fitted models to the example training data are shown below. The model architectures are the same as in Figure 6.4.

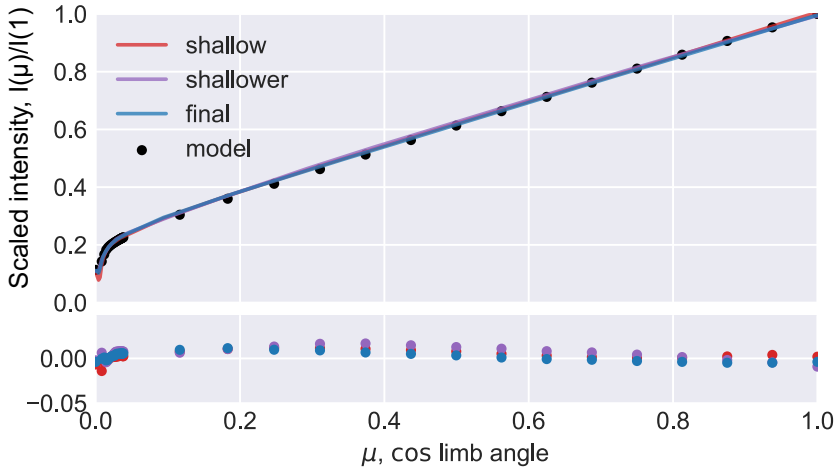



**Figure 6.6:** Plot is for the same example stellar atmosphere model output as in Figure 6.5, but this Figure shows the differences with changing the number of hidden layers in the model. The model architectures are based on the final ANN model described in Table 6.1, which is shown in dark blue line in this figure. The "shallow" model has only hidden layers 1, 2, 5 and 8 from the final ANN architecture, and the "shallower" model has only hidden layers 1 and 8.

translates well to how well each model fits the limb profile by eye in the Figure. As expected, the largest residuals in the fit are for low values of $\mu$, but the final model form does fit the limb profile well.

The models tested to construct Figure 6.5 were based on the architecture of the final model detailed in Table 6.1. The number of neurons in each hidden layer were determined by dividing the number of neurons in the final architecture by ten, and then multiplying the result by a given integer, $n$. For the "quarter" model, $n = 2$; for the "half" model, $n = 5$; for the "double" model, $n = 20$, and for the "final" model, $n = 10$. The models were trained with a batch size of 1000, and 200 training epochs, which is enough to illustrate the differences caused by the variation in the architecture, although different values are utilised for the final model training.

Varying the number of neurons in each layer has a strong effect on the ability of the model to fit the training data. With fewer neurons per layer, less detail of the limb profile is predicted, and the residuals of the fit are higher. The "double" model fits the training data well, although the differences between it and the "final" model are not large, so the significant increase in computation time for a prediction ($\sim 4\times$) for the larger model is not justified.

The number of hidden layers to use in the final model architecture was also challenging to decide, as there were too many variables to usefully search the parameter space to optimise them all. A MLP network with two hidden layers and a wide enough number of neurons in

each layer (and a non-linear activation function) should be able to represent the functional form of the output surface, however this proved difficult to train. Since deeper networks are tractable to train using GPU computational tools, and they are able to more efficiently train to represent complex functional forms than wide and shallow networks, the ANN architecture chosen had several layers in each branch. The deeper network was also more able to represent the smooth output in the CLIV that is expected physically.

To demonstrate the goodness-of-fit of models based on the number of hidden layers that are included, Figure 6.6 shows an example limb profile for three ANN architectures with differing numbers of hidden layers. The final model has 8 hidden layers in each branch, the "shallow" model has 4, and the "shallower" model has 2. Unlike for varying the number of neurons in each hidden layer (shown in Figure 6.5), reducing the number of layers does not have as pronounced an effect on the predicted output limb profile. The deeper final model does however have a smoother output, which is more physical, and the residuals of the fit are also lower for the deeper model.

To fully trust the value of the training loss, it would be ideal to remove a proportion of the training data from the training process, and use that to calculate the loss. When only testing the loss on the data that the model was trained with, there is a risk that the model is over-fitting the training data, rather than representing the underlying function that the training data samples. For this set of training data, there is only one example for each $\mu$ value for each model, so removing some of the training data degrades the amount of information that the ANN can represent. As a sanity check, the final model architecture was trained several times with random 20% portions of training data removed for testing the training, and the results were consistent. The final model was trained with the full data set to ensure the the CLIV surface was fully represented.

It was decided to use the ReLU activation function on all of the layers in the model, as the ANN is a deep network, and that activation function is viewed as most effective for deep networks (more detail on activation functions is given in Section 5.2.2). The activation functions for the final layers in each branch were chosen to be sigmoid functions, so that the output must lie in the interval [0,1]. The outputs of the training examples are all in the interval [0,1], but it is useful to force the network to that range of outputs to prevent unexpected results, since physically that is the range the output can be in.

**Table 6.1:** Detail of the structure of the final ANN model based on the schematic in Figure 6.3

| Layer | Number of neurons | Activation Function |
|---|---|---|
| *Stellar parameter branch* | | |
| Input | 3 – $T_{\rm eff}$ , $\log g$ , [Fe/H] | – |
| 1 | 100 | ReLU |
| 2 | 100 | ReLU |
| 3 | 100 | ReLU |
| 4 | 80 | ReLU |
| 5 | 60 | ReLU |
| 6 | 40 | ReLU |
| 7 | 40 | ReLU |
| 8 | 40 | Sigmoid |
| *$\mu$ value branch* | | |
| Input | 1 – $\mu$ | – |
| 1 | 100 | ReLU |
| 2 | 100 | ReLU |
| 3 | 100 | ReLU |
| 4 | 80 | ReLU |
| 5 | 60 | ReLU |
| 6 | 40 | ReLU |
| 7 | 40 | ReLU |
| 8 | 40 | Sigmoid |
| *Merger* | | |
| Merge | 1 | Multiply |
| Output | 1 | – |

The trial and error method for deciding on the model architecture was effective in this case, as the premise behind the use of the grid of stellar atmosphere models is that they are the ground truth of stellar atmosphere understanding, so the ANN model should fit the training data exactly.

### 6.3.2 Training final ANN model

The approach taken to set the form of the final model was to begin with a simple form and adjust the structure of the ANN until the improvements seen in the fit were not significant. Each trained model was assessed by eye, with sample limb profiles compared with the input training data to ensure that the model output was not only a good fit, but is also smooth across the valid range of $\mu$ – [0,1]. The final training epoch of the selected model had a loss value of $< 10^{-5}$, so given the by eye assessment and the low loss value, the trained ANN can be trusted as a good representation of the training data.

The final model architecture was trained with both the $\mu$ values given as output to the

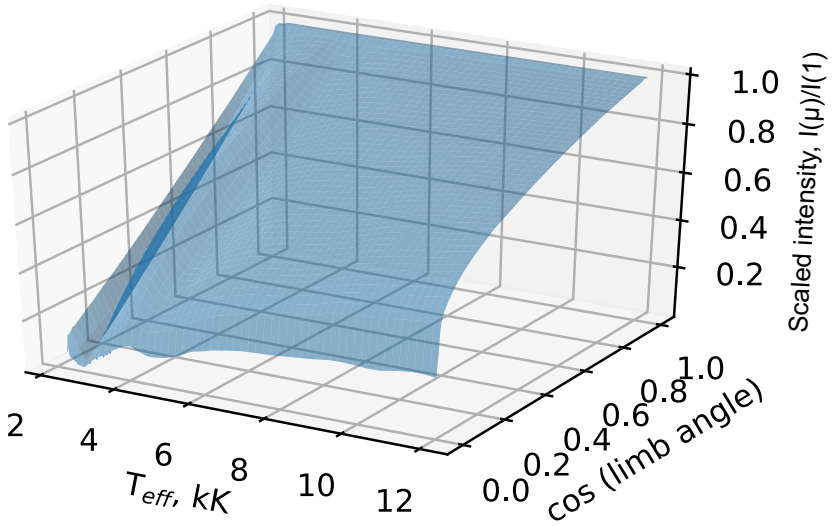



**Figure 6.7:** An example 3D surface of predicted outputs of the trained ANN model. The Figure shows the variation in the form of the limb profiles for a range of $T_{\rm eff}$ values for $\log g = 4.7$, and [Fe/H] = -0.2.

stellar atmosphere model, and the grid resampled to uniform intervals in $\mu$ (as described in Section 6.2.3). The outputs of each trained model were very similar, but because most of the information about the CLIV is found near the edge of the star, it was decided to use the original $\mu$ values as the final training grid. Unlike parameter fitting or algorithms like Gaussian processes, ANNs are not as sensitive to the evenness of the sampling of the parameter, so this approach is further justified. Adding another layer of modelling in order to resample each limb profile was also a negative, as that modelling process could easily introduce its own biases into the training data.

The detail of the final form of the ANN model is described in Table 6.1, where the number of neurons in each layer and the activation function for each are listed. The model was trained with a batch size of 50 training examples, which is approximately the number of $\mu$ samples for each limb profile in the training data. 1000 complete training epochs were used, to ensure the training data set was as well approximated by the ANN as possible. As can be seen in Figure 6.4, training beyond ∼100 epochs does not result in a significant improvement in the value of the loss for the model, but the model needs only be trained once for each bandpass, and the trained weights and biases can then be loaded into the model architecture when predictions of the CLIV are required in a transit lightcurve fit. The same model architecture and training parameters was also tested for several different bandpasses with comparable results.

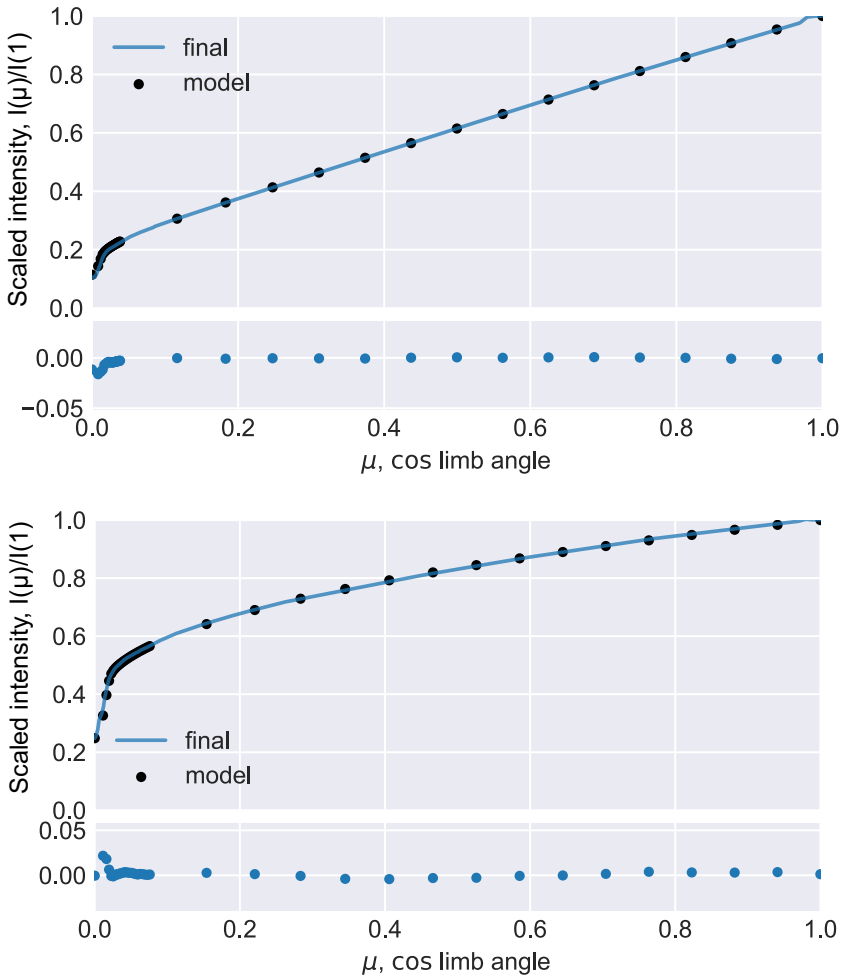



**Figure 6.8:** Figure showing two example limb profiles for different stellar parameters for the final ANN model. The upper plot is the predicted limb profile for $T_{\rm eff} = 4500$ K, $\log g = 4.5$, and [Fe/H] = -0.5. The lower plot has $T_{\rm eff} = 12{,}000$ K, $\log g = 4.5$, and [Fe/H] = 0.5

## 6.4 Example transit lightcurve fits

It is difficult to test this method with simulated transits, as some simulation must be made of the form of the limb darkening for the model host star. The stellar atmosphere models used to train the ANN are the most detailed view of our understanding of the form of limb darkening for stars other than the Sun. In order to test the model, real transit data is deployed. The rest of this section is dedicated to describing how the ANN limb darkening method was implemented into a transit fit, and then tested with real data.

### 6.4.1 Transit model

Since this work introduces a novel method of modelling limb darkening, the standard approaches to integrating limb darkening into an exoplanet transit model are not suitable, so this section introduces the transit model used to implement the limb darkening ANN into a transit model.

The most commonly used transit models in exoplanet studies are the models of Mandel & Agol (2002) and the BATMAN (Kreidberg, 2015) model. Each model is designed to operate with a parametric limb darkening law, which is called with stated coefficients and the stellar surface intensity is calculated for the given distance of the planet from the centre of the star for each time step to model a transit lightcurve.

The Mandel & Agol (2002) model utilises the small planet approximation to calculate the amount of star light occulted by the planet at any given time, which means that it is assumed that the model planet is small enough so that the stellar surface intensity is constant below the planet. As an implementation, this results in the limb darkening model being calculated once for each planet position, which is computationally efficient, but not necessarily justified, particularly for giant planets and near the edge of the star, where the gradient of the limb darkening effect is highest. Given the effort of this work is to accurately model the limb darkening, this approach is discounted so as not to degrade the detail of the limb darkening model used.

BATMAN does not make the same assumption about the stellar surface brightness variation below the planet. The algorithm slices the stellar surface into concentric circles of equal brightness, for which the surface areas occulted by a circular planet are analytically calculated.

To prevent unnecessary calculation, a scaling is used in the step size for the sector widths, so that more slices are used near the stellar limb and fewer near the stellar centre where the stellar brightness variation is lowest. For the first sector below the planet, the distance from the centre of the star to the centre of the relevant sector is used to calculate the expected stellar surface brightness for that position. The expected intensity is then multiplied by the surface area for that segment, the previous steps are repeated for more slices below the planet, and the intensity values for each segment are summed to provide the amount of light blocked by the whole planet at that time. The total occulted light is them normalised by the integrated entire stellar surface for the selected LDCs.

The BATMAN algorithm has been efficiently parallelised for a whole lightcurve calculation, since each sample time can be calculated independently. When using a parametric limb darkening law, each call to the law has a finite calculation time, so calculating the stellar intensity at each slice is efficient. In the case of the ANN model, each call to find an intensity value has a computational overhead associated with transferring the input values from the computer's CPU (where the majority of the transit model is calculated) to the GPU (where the ANN model is stored and run). This data transfer overhead is approximately the same amount of time regardless of how much data is transferred, so the current workflow of the BATMAN algorithm is not well optimised for using the ANN model.

The model developed in this study is based on the BATMAN method, but adjusted to require one call to the ANN model per lightcurve calculation. The same form of integration is deployed, with the order that each part of the calculation is performed altered. In this method, a constant number of stellar surface slices are used for each time step, whilst this will result in calculating more slices than are necessary close to the centre of the star, this allows the ANN to be called once per lightcurve, rather than once per time step (or once per slice as in the original BATMAN treatment). A 2D array is calculated of the surface areas of each slice, so each row of the array is for a different time step, and a column per slice. The positions of the centres of each slice relative to the centre of the star for the entire lightcurve are mapped to their $\mu$ values and fed as inputs to the ANN model. The output intensities from the ANN are shaped to match the 2D array of areas, and then the arrays are multiplied together. The rows of the resulting array are summed and normalised with the value of the integrated entire stellar surface, providing the model lightcurve.

### 6.4.2 Transit fitting procedure

To demonstrate the ANN limb darkening approach, different methods of implementing limb darkening are compared by fitting the same lightcurve with several limb darkening methods. The techniques compared are: freely fitting the quadratic limb darkening law; freely fitting the non-linear limb darkening law; the GP interpolation method from Chapter 4, and the ANN method presented here.

The core of the transit fitting procedure is in the use of Markov Chain Monte Carlo (MCMC) methods, which finds the location of the combination of parameters which correspond to maximum likelihood, and explores the parameter space around this region, which allows the posterior distributions of the parameters to be estimated. The MCMC implementation utilised for these transit fits was the EMCEE package (Foreman-Mackey et al., 2013), which is an Affine-Invariant Ensemble sampler. The ANN method for limb darkening is compared with other approaches, so two transit models are used – the transit model described in Section 6.4.1 for the ANN method, and BATMAN for the other approaches.

To parameterise the rest of the transit fit, the following jump parameters were used: orbital period ($P$), time of transit centre ($T_c$), transit depth ($\Delta F$), photometric baseline ($c$)[4], scaled impact parameter ($b$), and stellar density ($\rho_\star$). Wide uniform priors were applied to each of these jump parameters, but for $P$, which had a tight Gaussian prior using existing knowledge of the observed systems. Since the orbital period is not well constrained by a single transit observation, the Gaussian prior is justified to prevent biasing the fit. The limb darkening for the quadratic LD fits is parameterised by $q_1$ and $q_2$, as described in Kipping (2013), with a uniform prior in the interval [0,1]. The non-linear law is parameterised with $c_1$, $c_2$, $c_3$, and $c_4$, which are the coefficients for the law. The bounds of the free fit of the non-linear law were set by finding the maximum and minimum values for each coefficient from the Claret & Bloemen (2011) tabulated LDCs.

With the priors for each transit fit set, the MCMC was initialised with 100 walkers located randomly around the central values of the defined priors The random initial values were sampled from a Gaussian distribution with the mean as the centres of the priors, and $\sigma$ as 0.5 of the $\sigma$ for the Gaussian priors and 0.5 of the distance from the centre of the prior to the bound for the flat priors. The MCMC was then allowed to run for 10,000 iterations with the 100

[4]The photometric baseline ($c$) refers to the out of transit flux level of the lightcurve.

walkers, which resulted in 1,000,000 output samples of the parameter vector. The first 200 samples for each walker were discarded as burn-in, and then the other samples but for every 10th for each walker were also discarded to account for the correlation between neighbouring samples. The resulting 98,000 samples are then treated as the final output chains for each transit fit.

### 6.4.3 Transit of Venus observed with SDO/HMI

The lightcurve of the transit of a solar system planet across the surface of the Sun is the only set of real observations for which the underlying true values for the transit parameters are known to a high certainty. To make use of this opportunity, continuum intensity images from the HMI instrument on the *S*olar *D*ynamics *O*bservatory (*SDO*) during the transit of Venus in 2012 were used to create a photometric lightcurve of the transit.

The transit occurred over several hours on 5th/6th June 2012, so to ensure a long enough out of transit baseline, all of the images from the beginning of 4th June until the end of 7th June (UTC) were downloaded from JSOC[5]. The 7674 images were then used to create the "raw" lightcurve as shown in the upper row of Figure 6.9 – all the pixel values in each image were summed, but for the pixels with values $< 0$, or NaN. The lightcurve was initially detrended by dividing the summation for each image by the square of the apparent size of the solar disk in the image, which was found in the data header item `RSUN_OBS`. The lightcurve shown in Figure 6.9 is that data after being median normalised.

Unfortunately the SDO/HMI data products are subject to several time-dependent systematic signals introduced by the instrument itself and the data reduction techniques used. Because of these signals, the instrument is not optimal for use as a photometer as this work intends, but the timing of the transit is well known, so it is possible to model the systematic signal and remove it from the lightcurve. Schuck et al. (2016) note that the sources of the systematics in SDO/HMI data are not well understood, but the primary periodicity of the variation matches the orbital period of the spacecraft. I have chosen to model the signal using a Gaussian Process (GP), as it is not possible to build a parametric model to describe the systematic signal without an understanding of the source of the signal. GPs are introduced in Section 4.3.1.

[5]The Joint Science Operations Center (JSOC) host archival data from SDO: http://jsco.stanford.edu/

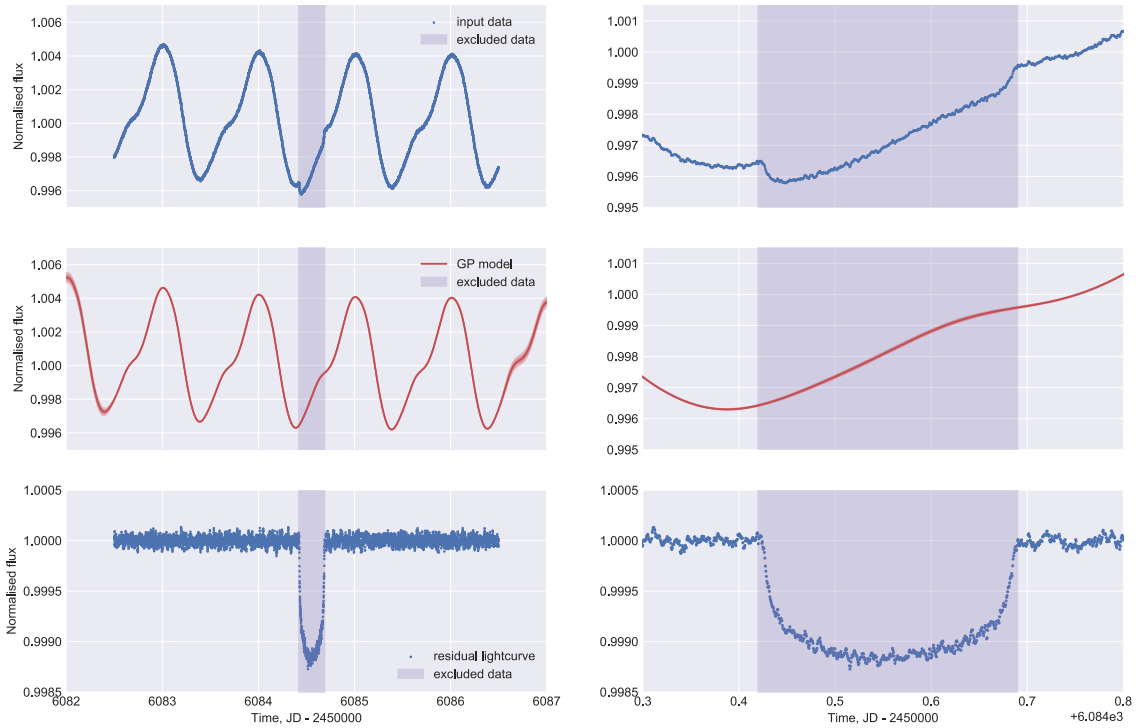



**Figure 6.9:** Upper plots show the initial lightcurve extracted from the SDO/HMI images with only a correction for the apparent size of the Sun on the CCD applied. The middle plots show the prediction made by the GP for the variation introduced by the spacecraft. The lower plots show the lightcurve residual of the GP fit, which is used for transit parameter fitting. In each plot, the purple shaded region depicts the time frame where the lightcurve data was excluded from the GP fit.

It would be ideal to fit the GP model for the systematic signal simultaneously with the transit fit, but each calculation of the GP likelihood requires the inversion of an $N \times N$ matrix, where $N$ is the number of data points in the lightcurve, which is 7674 in this case. The GP could be calculated on a subset (or binning) of the data, but since the systematics signal in the lightcurve is very smooth and at different amplitude and timescale to any solar activity seen in the data, it is justified to fit the GP to detrend the lightcurve before fitting the transit. The procedure for optimising the hyperparameters was slow because of the multiple inversions of the $7674 \times 7674$ required, but the procedure was only required to be performed once and the resulting residual lightcurve saved for fitting the transit profile.

The GP is implemented using the same tools as in Chapter 4, but the covariance kernel chosen is different. The kernel chosen is the quasi-periodic kernel, which takes the form:

$$k(x_i, x_j) = \sigma^2 \exp\left( -\frac{2\sin^2\left(\frac{\pi(x_i - x_j)}{P_1}\right)}{l_1^2} - \frac{1}{2}\left(\frac{x_i - x_j}{l_2}\right)^2 \right) , \tag{6.2}$$

where $\sigma$ is the amplitude of the signal; $x_i$ and $x_j$ are the inputs to which the covariance matrix element refers (in this case, the input variable is time); $P_1$ is the periodicity of the signal; $l_1$ is the length scale, which relates to how stable the periodicity remains in time, and $l_2$ is another length scale modulating how much harmonic content the signal has.

The choice of the quasi-periodic kernel was motivated by the large scale systematic signal in the data set – the signal is almost periodic, but the shape is varying between cycles, which is readily described as quasi-periodic. Since the signal is known to predominantly originate from the orbital motion of the spacecraft, the periodicity hyperparameter ($P_1$) was constrained to the value of the SDO orbital period. The exact prior on the value of $P_1$ is shown in Table 6.2 – the tight uniform prior was used rather than fixing the value to allow for the data to inform final value found.

A white noise kernel was also added to the fit, which has the form:

$$k(x_i, x_j) = \sigma_n^2 \delta_{i,j} , \tag{6.3}$$

where $\sigma_n$ is the variance of the noise signal, and $\delta_{ij}$ is a Kronecker delta, which is unity where $i = j$, and zero elsewhere (this is equivalent to the identity matrix). The white noise kernel

**Table 6.2:** Details of the hyperparameters used in GP fit of SDO/HMI Venus lightcurve

| Parameter | Initial value | Boundaries | Optimised value |
|---|---|---|---|
| $\sigma$ | 0.01 | none | $0.00316^2$ |
| $P_1$ | 0.997 | 0.996, 0.998 | 0.997 |
| $l_1$ | 0.5 | 1e-5, 1e5 | 1.07 |
| $l_2$ | 5.0 | 1e-5, 1e5 | 3.46 |
| $\sigma_n$ | 1e-5 | 1e-10, 5e-5 | 4.98e-5 |

accounts for the white noise in the data, and will represent the solar variation present in the lightcurve, such as granulation. Whilst the solar variation is not white noise, it is at a significantly different amplitude and timescale to the systematic signal, so the white noise kernel is enough to prevent the quasi-periodic kernel overfitting the solar activity as a systematic signal. The GP is designed to model only the systematic signal, as retaining the solar variation in the transit lightcurve allows the lightcurve to be more like that of an exoplanet transit.

As the GP is fit before the transit fit, the lightcurve data from during the transit are excluded from the fit – the data that that represents is highlighted in purple in all panels of Figure 6.9.

The GP fit was performed on the full remaining data set, using the same optimisation algorithm as in Chapter 4. The resulting hyperparameters are shown in Table 6.2, and those hyperparameters conditioned on the selected data were used to produce the GP model shown in the middle row of Figure 6.9. The expected fit is shown to be a good representation of the input data, including in the region of the data excluded from the modelling, where there was no data to apply a strict condition on the form of the expectation.

To produce the residual lightcurve, which should show solar effects and the transit of Venus without any instrumental signal, the original data was divided by the GP fit. The resulting lightcurve is shown in the lower panels of Figure 6.9, which is smooth and flat, and the transit of Venus appears symmetric.

Since the view SDO/HMI has of the transit of Venus is not exactly like the view of an exoplanet transit from Earth, some adjustment to the lightcurve or the transit model must be made to fit the lightcurve accurately. The primary aim of creating this data set is to test the transit model introduced in Section 6.4.1, so the lightcurve itself was adjusted so that the shape would represent that for an exoplanet transit.

A schematic of the differences in the geometries associated with an exoplanet transit, and

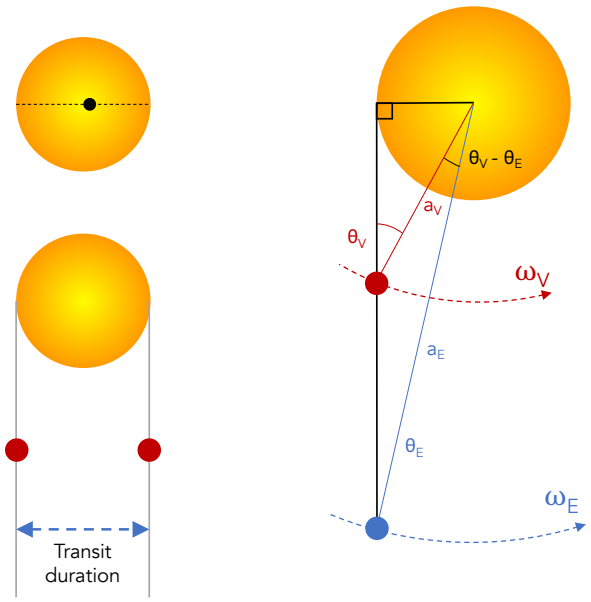



**Figure 6.10:** Schematic to demonstrate the differing orbital geometries for an exoplanet transit (left) and an observation of the transit of the Sun by a solar system body inside the Earth's orbit (right).

a solar system transit are shown in Figure 6.10. For an exoplanet transit, the transit depth is determined by the ratio of the radii of the transiting planet to the host star, for a solar system transit, the relative angular sizes of the bodies as seen by the observer determine the transit depth. To prevent degrading the information in the transit, I have assumed that the exoplanet transit measured is that of a larger planet than Venus, rather than altering the flux scale of the data.

In the solar system case the transit duration is shorter than would be expected for a transit of the same type of system when viewed as an extra-solar system. The source of this timing difference can be seen in Figure 6.10, where it is clear that the solar system transit duration represents a smaller proportion of the planetary orbit than the exoplanet case, and the Earth is also orbiting in the same direction as Venus. To adjust the lightcuve from a solar system transit to an extra-solar transit shape, the time scale is calibrated using the ratio of the expected transit durations of each type of system.

For the solar system transit, the duration is defined by twice the time taken for the angle $\theta_V - \theta_E$ (as shown in Figure 6.10) to reduce from its value at the beginning of the transit to

zero. Using this geometry, the solar system transit duration ($T'$) can be written as:

$$T' = 2\frac{R_\odot\left(\frac{1}{a_V} - \frac{1}{a_E}\right)}{\omega_V - \omega_E}\left(1 - b^2\right) , \tag{6.4}$$

where $R_\odot$ is the solar radius; $a_V$ is the orbital separation of the Sun and Venus; $a_E$ is the orbital separation of the Sun and Earth; $\omega_V$ is the orbital frequency of Venus; $\omega_E$ is the orbital frequency of Earth, and $b$ is the impact parameter of the orbit. The impact parameter can be measured from the images of the transit, and the other parameters are well known for the solar system, so a value for $T'$ can be readily calculated.

The duration of an exoplanet transit is defined by:

$$T = \frac{P}{\pi}\left(\frac{3\pi}{G\,P^2\rho_\star}\right)^{\frac{1}{3}}\left(1 - b^2\right) , \tag{6.5}$$

where $\rho_\star$ is the density of the host star. The orbital period of Venus is used, as is the density of the Sun, and as such a value for the expected transit duration of a exoplanet planet orbiting a solar type star with an orbital period matching that of Venus is known. The time values for original lightcurve are then scaled with the factor $T'/T$, producing an arbitrary timescale adjusting the transit to that of an exoplanet. Now that the data has been prepared, the lightcurve is ready to apply the transit model fitting procedure described above to.

The priors applied for each limb darkening method tested for the SDO/HMI data are detailed in Table 6.3.

**Results**

The results of each of the fits are described in Table 6.4, which are also listed with the known values for the lightcurve derived from measurements of the images, and knowledge of the Sun. The four parameter non-linear law fits were not possible for this system – the MCMC never found a stable solution for the limb darkening coefficients, as they were correlated with one another and not well constrained by the transit lightcurve. No values for the results of the non-linear law fit are given, as the posterior distributions are particularly wide and non-Gaussian. The full corner plot for the non-linear fit is included as Figure B.4 in Appendix B. The corner plots of the posteriors for the full set of parameters modelled are included for all of the fits in the same Appendix.

**Table 6.3:** Details of the priors used in each fit to the transit of Venus data

| Parameter | Notation | Prior type | Prior value | Unit | Fit used for |
|---|---|---|---|---|---|
| Orbital period | $P$ | Gaussian | $\mu = 224.701$, $\sigma = 0.001$ | days | All |
| Transit centre | $t_C$ | Uniform | $a = 2927.5$, $b = 2928.0$ | days | All |
| Transit depth | $d$ | Uniform | $a = 0.0005$, $b = 0.0020$ | – | All |
| Baseline | $c$ | Uniform | $a = 0.99990$, $b = 1.00001$ | – | All |
| Impact parameter | $b$ | Uniform | $a = 0.001$, $b = 0.99$ | – | All |
| Stellar density | $\rho_\star$ | Uniform | $a = 0.01$, $b = 5.00$ | g cm$^{-3}$ | All |
| Effective temperature | $T_{\rm eff}$ | Gaussian | $\mu = 5772$, $\sigma = 100$ | K | ANN and GP |
| Surface gravity | $\log g$ | Gaussian | $\mu = 4.45$, $\sigma = 0.10$ | – | ANN and GP |
| Metallicity | [Fe/H] | Gaussian | $\mu = 0.0$, $\sigma = 0.1$ | dex | ANN and GP |
| Quadratic coefficient | $q_1$ | Uniform | $a = 0.0$, $b = 1.0$ | – | Quadratic |
| Quadratic coefficient | $q_2$ | Uniform | $a = 0.0$, $b = 1.0$ | – | Quadratic |
| Non-linear coefficient | $c_1$ | Uniform | $a = 0.2099$, $b = 0.6935$ | – | Non-linear |
| Non-linear coefficient | $c_2$ | Uniform | $a =$ -0.6744, $b = 1.1355$ | – | Non-linear |
| Non-linear coefficient | $c_3$ | Uniform | $a =$ -1.0728, $b = 1.3609$ | – | Non-linear |
| Non-linear coefficient | $c_4$ | Uniform | $a =$ -0.5588, $b = 0.3687$ | – | Non-linear |

**Note:** $\mu$ and $\sigma$ are the mean and standard deviation of the Gaussian prior. $a$ and $b$ are the lower and upper bounds of the uniform prior respectively.

**Table 6.4:** The results of each fit to the SDO/HMI transit of Venus lightcurve

| Parameter | ANN fit | Quadratic fit | GP fit | Known values | Unit |
|---|---|---|---|---|---|
| $P$ | 224.7009 ± 0.0001 | 224.70010 ± 0.00010 | 224.70100 ± 0.00010 | 224.701 | days |
| $t_C$ | 2927.6990 ± 0.0004 | 2927.6989 ± 0.0004 | 2927.7001 ± 0.0004 | 2927.699 | days |
| $d$ | 0.00104 ± 0.00004 | 0.00094 ± 0.00006 | 0.0010 ± 0.0003 | 0.001035 | – |
| $c$ | 1.000000 ± 0.000002 | 0.999998 ± 0.000002 | 1.000002 ± 0.000002 | – | – |
| $b$ | 0.59 ± 0.09 | 0.37 ± 0.19 | 0.33 ± 0.14 | 0.61 | – |
| $\rho_\star$ | 1.5 ± 0.3 | 2.2 ± 0.5 | 2.4 ± 0.4 | 1.41 | g cm$^{-3}$ |
| $T_{\rm eff}$ | 5720 ± 100 | – | 5500 ± 90 | 5772 | K |
| $\log g$ | 4.43 ± 0.10 | – | 4.66 ± 0.08 | 4.45 | – |
| [Fe/H] | 0.04 ± 0.10 | – | 0.15 ± 0.13 | 0.0 | dex |
| $q_1$ | – | 0.66 ± 0.11 | – | – | – |
| $q_2$ | – | 0.32 ± 0.08 | – | – | – |

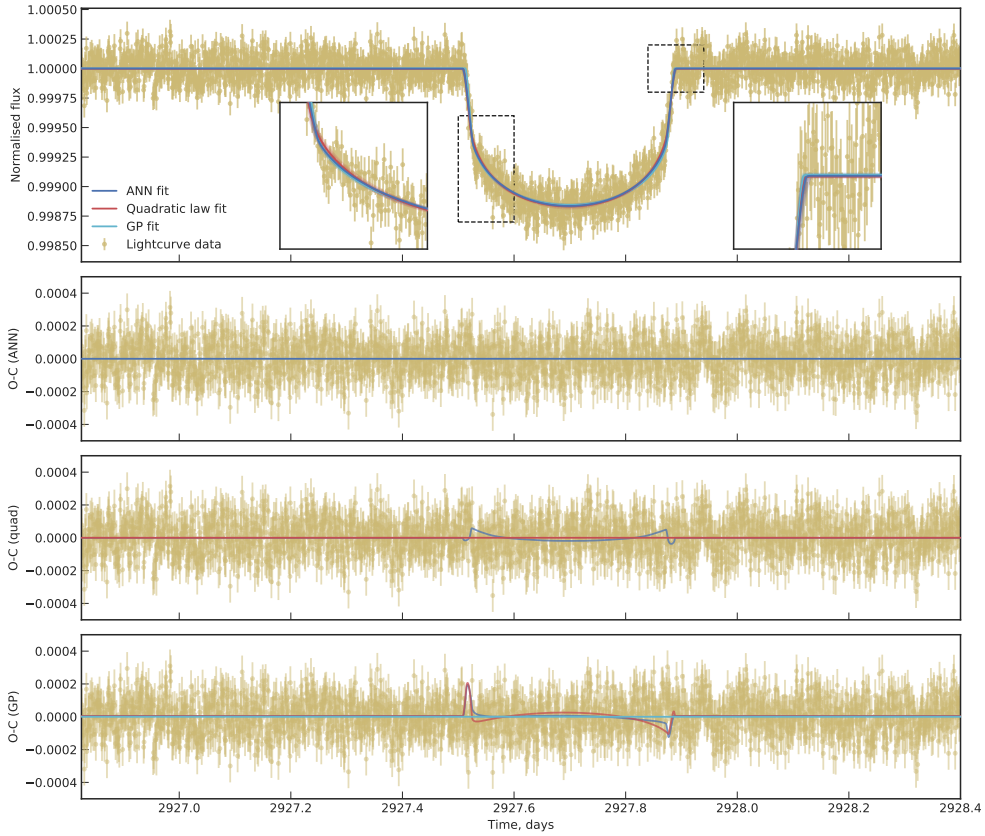



**Figure 6.11:** Upper plot shows the SDO/HMI lightcurve fitted with ANN model, quadratic law, and GP interpolation. The models are shown with 1, 2 and 3-$\sigma$ uncertainties derived from the posterior chains from each of the MCMC runs. The lower plots show the residuals of the fit to the data for the ANN model, quadratic law and GP interpolation. The lines over-plotted show the differences between the model fits with the colours referring to the comparison model.

The derived lightcurve model for each of the three fits detailed in Table 6.4 are shown in Figure 6.11. Each fit looks by eye to be a good representation of the lightcurve data. The residuals are also low, with the largest differences between the model fits in the transit ingress and egress, which is where the form of the limb profile has the largest effect. The high cadence of the lightcurve provides a lot of information about the shape of the ingress and egress, so the limb darkening can be well constrained.

There is evidence of some other time-dependent variation in the residuals of the lightcurve, which would indicate that a model for the stellar brightness variation should be modelled alongside the transit to fully represent the lightcurve. In the residual plots in Figure 6.11, there are differences seen between the fitting procedures that are asymmetric – the source of this asymmetry is not known. Besides the different treatments for limb darkening, the transit model for the ANN method is different to that used for the other limb darkening treatments, some part of which is likely to be the source of the asymmetry. I have investigated each step of the transit model built for the ANN method, and cannot find any algorithmic or geometric difference between it and the BATMAN transit model, so the underlying source of the observed imbalance remains unknown.

Figure 6.12 shows a subset of the posteriors for each of the fits, with the parameter values derived from the ANN fit shown with blue lines on each of the other corner plots, so that the values can be compared. Despite the tight correlations between transit depth, $b$ and $\rho_\star$, the posteriors for the ANN are relatively Gaussian, and centred around the expected values for the transit of Venus. The quadratic and non-linear law fit posteriors are less Gaussian, and the differences between the transit parameters found in these fits and the ANN fit are particularly significant.

The limb profiles predicted by each of the model fits are shown in Figure 6.13, with their associated 1-, 2- and 3-$\sigma$ uncertainty bounds derived from the MCMC fit. The limb profile for the quadratic law fit is the most uncertain, which is consistent with its coefficients being unconstrained by the prior for that fit, and with the form of the quadratic model not being a sufficient representation of the real solar limb darkening. The non-linear law fit appears to match very well to the ANN fit, although the profile has an uncertainty approximately an order of magnitude larger than that for the ANN fit. It is interesting to note that whilst the non-linear law fit appear very close to the ANN fit, the difference between them is enough that

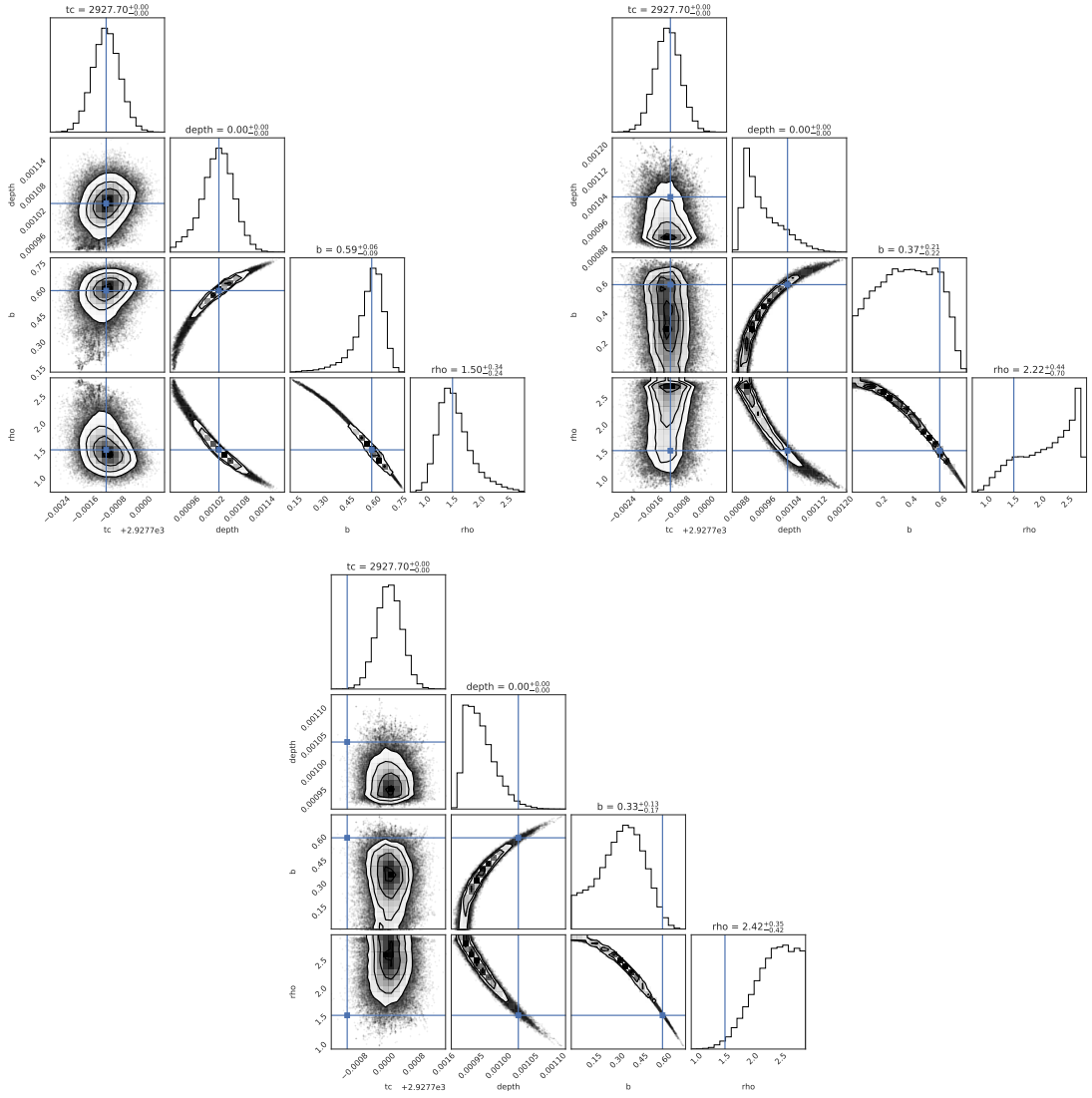



**Figure 6.12:** Plots of the MCMC posteriors for a subset of system parameters for each of the limb darkening fits for the SDO/HMI transit of Venus data. The blue lines in each plot represent the median parameter values found in the ANN fit. The upper left plot is for the ANN fit; the upper right plot is for the quadratic law fit, and the lower plot is for the GP interpolation fit.

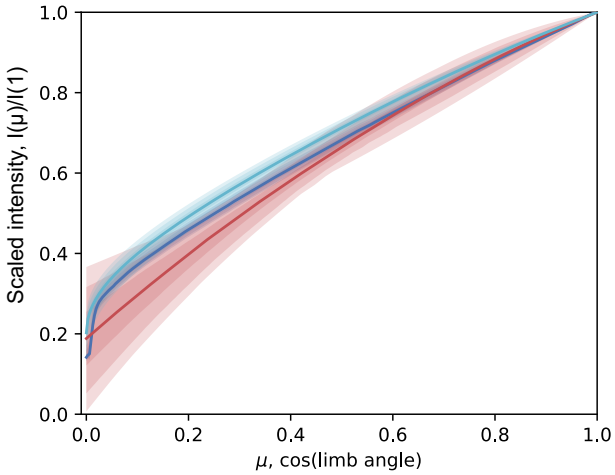



**Figure 6.13:** The output of the MCMC fit for each limb darkening model is plotted with reference to position on the stellar limb. The line for each colour shows the median limb variation profile, with the shaded regions referring to the 1, 2 and 3-$\sigma$ uncertainty ranges – the uncertainties in the profiles for the GP and ANN methods are plotted, but are much narrower than that for the quadratic fit. The colours used for each model fit are consistent with Figure 6.11.

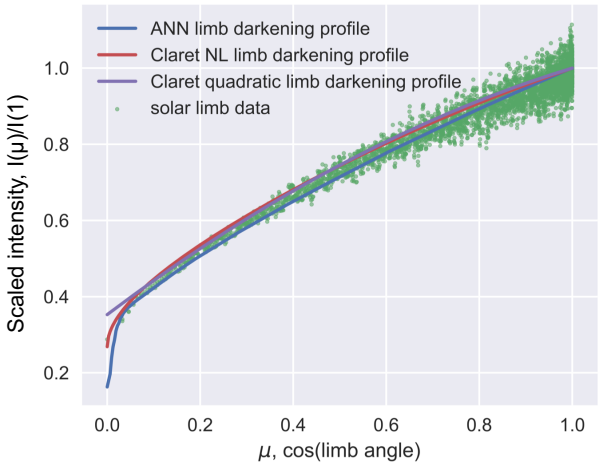



**Figure 6.14:** Plotted in blue is the limb profile predicted by the ANN for the stellar atmosphere parameters derived from the transit fit. Shown in green are the pixel values of a radial slice of the SDO/HMI continuum image from 4th June 2012 00:01:30 UTC. The red and purple lines show the limb profiles for the quadratic and non-linear coefficients for a solar type star from Claret & Bloemen (2011) respectively.

the prediction of the values of $b$ and $\rho_\star$ are significantly different for the two fits.

Figure 6.14 compares the prediction of the limb profile from the ANN lightcurve fit with an example radial limb profile extracted from one of the SDO/HMI continuum images. It is clear that the profiles match very well, despite the ANN only being trained on 1D stellar atmosphere model outputs.

**Additional tests of ANN method using SDO transit of Venus data**

To explore further the efficacy of the ANN lightcurve fit, two variations of the fitting approach were applied. The first variation was to add additional white noise to the SDO transit lightcurve, which allows for the investigation of the performance of the model for a lower SNR transit lightcurve where the underlying true values are known. The fitting procedure mirrored the set up for the original ANN fit, but for random noise added to the lightcurve (random noise sampled from a Gaussian with shape $\mu$=0, and $\sigma$=0.001), and the quoted uncertainties on each data point were also inflated to reflect the additional noise. The resulting lightcurve fit can be seen in the upper plot of Figure 6.15, with the posterior distributions for a subset of the fitting parameters shown in the left plot of Figure 6.16 (the full corner plot is shown in Appendix B in Figure B.5).

The second additional fit is to test the effectiveness of the modelling approach when fed with incorrect priors. The priors for each of the stellar atmosphere parameters were offset from their true values. The offset priors applied to the fit were as follows: $T_{\rm eff} = 6000 \pm 250$ K; $\log g = 4.45 \pm 0.10$, and [Fe/H] = -0.2 ± 0.1. The resulting lightcurve fit can be seen in the lower plot of Figure 6.15, with the posterior distributions for a subset of the fitting parameters shown in the left plot of Figure 6.16 (the full corner plot is shown in Appendix B in Figure B.6).

As can be seen in the plots of the posterior distributions from these additional fits (shown in Figure 6.16), the fits did not reproduce the parameters determined from the original ANN fit (that well maps to the true values for the system). In the case of the the lightcurve with increased white noise, it is difficult to see the shape of the transit, so it is unsurprising that the MCMC fit was unable to accurately find the true transit parameters, as the parameters relating to transit geometry were not particularly constrained by the priors applied and the transit shape is not well constrained by the lightcurve either.

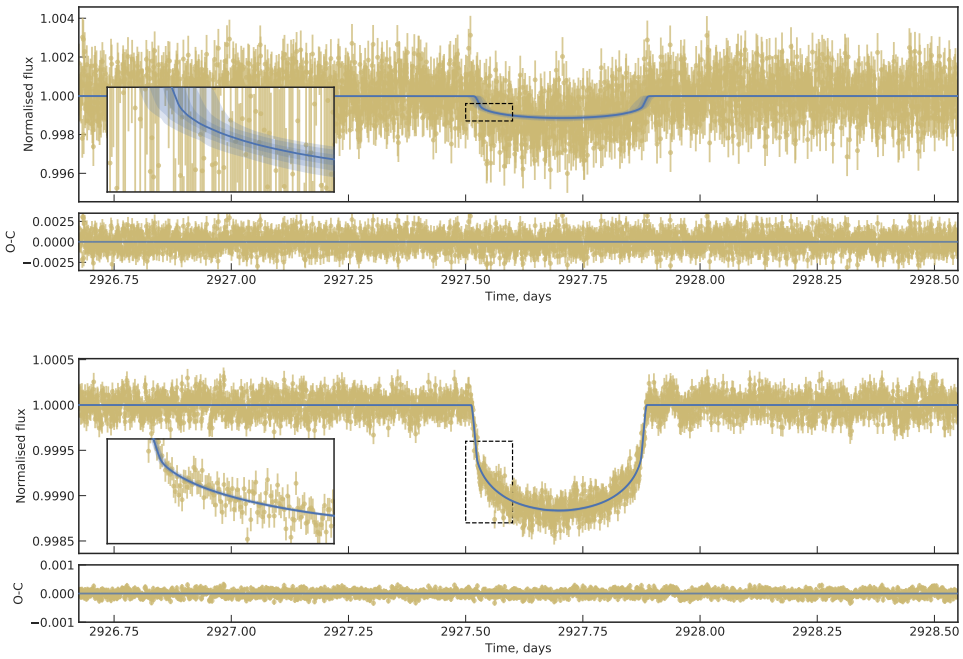



**Figure 6.15:** Upper plot shows the SDO/HMI lightcurve with additional noise fitted with ANN model. The model is shown with 1, 2 and 3-$\sigma$ uncertainties derived from the posterior chains from the MCMC run. The lower plot attributed to the upper portion of the Figure shows the residuals of the fit. The lower portion of the Figure is structured in the same way as above, but for the fit with stellar atmosphere parameter priors offset from their true values.

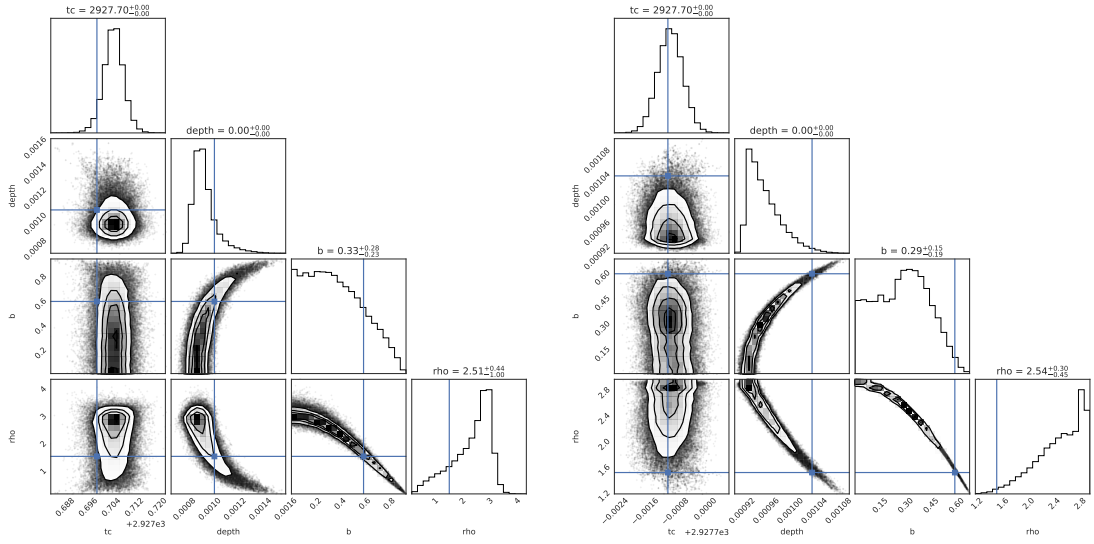



**Figure 6.16:** Plots of the MCMC posteriors for a subset of system parameters for each of the additional fits for the SDO/HMI transit of Venus data. The blue lines in each plot represent the median parameter values found in the original ANN fit. The left plot is for the fit with additional white noise added to the transit lightcurve, and the right plot is for the fit where the priors on the stellar parameters were offset from their true values.

For the initialisation of the model fit with the stellar atmosphere parameter prior positions offset to values divergent from the known values for the sun, the transit parameters found in the fit have been skewed relative to the known values, rather than adjusting the stellar parameters to better represent the limb darkening effects in the lightcurve. For the model posterior values for the stellar parameters to better reflect the true limb darkening in the lightcurve, there is a likelihood penalty to overcome, which is not the case for the model posterior values for the transit parameters with flat priors. This means that applying ANN method to a lightcurve fit should only be undertaken with appropriate priors on the stellar parameters, which fully reflect the existing knowledge of the stellar parameters for the system.

### 6.4.4 *Kepler* view of transit of TRES-2b

For a solar system transit the ground truth of the Sun's surface is known, which allows a direct test of the limb darkening model, but this only tests a small part of the parameter space that the ANN model covers. To broaden the testing of the model, we have used a transit lightcurve of TRES-2b, which was the first exoplanet to be discovered in the *Kepler* field of view (O'Donovan et al., 2006). Lightcurves of the system have been observed and analysed extensively (eg. (Southworth, 2011; Raetz et al., 2014)). The system was part of the short cadence (SC) *Kepler* data set, which means that there is a near continuous lightcurve for a period of five years, with data every one minute. The *Kepler* SC data was first analysed by Kipping & Bakos (2011), which looked at the first 90 days of the *Kepler* short cadence data and fitted the transit lightcurves. The analysis used the quadratic limb darkening law with the Pál (2008) reparameterisation to mitigate some of the correlation between the coefficients.

Given the high impact parameter of the orbit of the system, modelling the limb darkening well is challenging, since only a small region of the stellar surface is occulted by the planet during the transit. Recent work by Sandford & Kipping (2017) looked to model a range of *Kepler* lightcurves with well established measurements of the eccentricities of the orbit, so that the densities of the host stars could be precisely measured directly from the transit lightcurves without prior knowledge of the star applied in the fit. TRES-2b was one of the systems included in the study, which modelled the limb darkening by freely fitting the Sing (2010) 3-parameter limb darkening law utilising the Kipping (2016) reparameterisation of the coefficients to allow for uniform sampling of the parameters within a known interval. The TRES-2b system was anomalous in the study, as it was not possible to precisely determine a measurement of the

**Table 6.5:** Details of the priors used in each fit to the *Kepler* TRES-2b transit lightcurve

| Parameter | Notation | Prior type | Prior value | Unit | Fit used for |
|---|---|---|---|---|---|
| Orbital period | $P$ | Gaussian | $\mu = 2.47061317$, $\sigma = 0.00001$ | days | All |
| Transit centre | $t_C$ | Uniform | $a = 135.08$, $b = 135.15$ | days | All |
| Transit depth | $d$ | Uniform | $a = 0.01$, $b = 0.02$ | – | All |
| Baseline | $c$ | Uniform | $a = 0.99990$, $b = 1.00001$ | – | All |
| Impact parameter | $b$ | Uniform | $a = 0.001$, $b = 0.99$ | – | All |
| Stellar density | $\rho_\star$ | Uniform | $a = 0.01$, $b = 3.00$ | g cm$^{-3}$ | All |
| Effective temperature | $T_{\rm eff}$ | Gaussian | $\mu = 5850$, $\sigma = 50$ | K | ANN and GP |
| Surface gravity | $\log g$ | Gaussian | $\mu = 4.43$, $\sigma = 0.02$ | – | ANN and GP |
| Metallicity | [Fe/H] | Gaussian | $\mu$ = -0.15, $\sigma = 0.1$ | dex | ANN and GP |
| Quadratic coefficient | $q_1$ | Uniform | $a = 0.0$, $b = 1.0$ | – | Quadratic |
| Quadratic coefficient | $q_2$ | Uniform | $a = 0.0$, $b = 1.0$ | – | Quadratic |
| Non-linear coefficient | $c_1$ | Uniform | $a = 0.2099$, $b = 0.6935$ | – | Non-linear |
| Non-linear coefficient | $c_2$ | Uniform | $a$ = -0.6744, $b = 1.1355$ | – | Non-linear |
| Non-linear coefficient | $c_3$ | Uniform | $a$ = -1.0728, $b = 1.3609$ | – | Non-linear |
| Non-linear coefficient | $c_4$ | Uniform | $a$ = -0.5588, $b = 0.3687$ | – | Non-linear |

**Note:** $\mu$ and $\sigma$ are the mean and standard deviation of the Gaussian prior. $a$ and $b$ are the lower and upper bounds of the uniform prior respectively.

stellar density using their methodology, predominantly due to the flexibility of the limb darkening model, which settled to a physically unrealistic limb profile for that system more readily than the true profile, and in turn biased the values for the other parameters. Schröter et al. (2012) made use of the four parameter non-linear limb darkening law in their lightcurve fits, however the study fixed the coefficients from tabulated values, as there was not enough information about the stellar limb encoded in the transit lightcurve to model so many coefficients. Other previous studies of this system have all made use of two parameter limb darkening laws, such as O'Donovan et al. (2006); Kipping & Bakos (2011); Southworth (2011); Coughlin & López-Morales (2012); Raetz et al. (2014); Esteves et al. (2015).

The ANN method for limb darkening modelling is particularly useful for this system, where there is not a lot of information about the limb darkening in the transit. This study makes use of the *Kepler* SC data for the TRES-2 system. Since the transit is so deep, and the *Kepler* SC data is of such high quality (the low RMS scatter of the lightcurve can be seen in Figure 6.17), only a single transit is used. The data fitted here spans from *Kepler* JD 134.8 – 135.4 days.

The priors used for each of the fits are described in Table 6.5. It should be noted that all previous studies of the TRES-2 system found the orbit to be consistent with a circular orbit, so circularity is assumed in this analysis, and no parameterisation of eccentricity and longitude of periastron are used. The same four types of fit for limb darkening are presented here, as for the SDO/HMI transit of Venus data in Section 6.4.3.

**Table 6.6:** The results of each fit to the *Kepler* lightcurve of the transit of TRES-2b

| Parameter | ANN fit | Quadratic fit | GP fit | Literature values[1] | Unit |
|---|---|---|---|---|---|
| $P$ | 2.470613 ± 0.000010 | 2.470613 ± 0.000010 | 2.470613 ± 0.000010 | 2.470613 | days |
| $t_C$ | 135.11636 ± 0.00005 | 135.11636 ± 0.00005 | 135.11636 ± 0.00005 | – | days |
| $d$ | 0.01602 ± 0.00009 | 0.0160 ± 0.0005 | 0.01584 ± 0.00008 | 0.0157 | – |
| $c$ | 1.000018 ± 0.000009 | 1.000018 ± 0.000009 | 1.000018 ± 0.000009 | – | – |
| $b$ | 0.836 ± 0.002 | 0.840 ± 0.008 | 0.840 ± 0.002 | 0.844 | – |
| $\rho_\star$ | 1.59 ± 0.02 | 1.58 ± 0.06 | 1.56 ± 0.02 | 1.4 ± 0.1 | g cm$^{-3}$ |
| $T_{\rm eff}$ | 5850 ± 50 | – | 5850 ± 50 | 5850 ± 50 | K |
| $\log g$ | 4.43 ± 0.02 | – | 4.43 ± 0.02 | 4.43 ± 0.02 | – |
| [Fe/H] | -0.15 ± 0.10 | – | -0.15 ± 0.10 | -0.15 ± 0.10 | dex |
| $q_1$ | – | 0.34 ± 0.10 | – | – | – |
| $q_2$ | – | 0.4 ± 0.3 | – | – | – |

**Note:** [1] the literature values stated here are from Esteves et al. (2015), which is based on the full *Kepler* data set rather than a single transit.

### Results

The results of each of the fits are described in Table 6.6, which are detailed with the parameter values found by Esteves et al. (2015) with the analysis of the full *Kepler* data set. It is worth noting that the transit parameters derived in each of the fits are very similar, unlike for the much shallower transit in the SDO/HMI transit of Venus data.

The non-linear law fits were not possible for this system – the MCMC never found a stable solution, as there were too many parameters for limb darkening that were correlated with one another and not well constrained by the transit lightcurve. No values for the results of the non-linear law fit are given, as the posterior distributions are particularly wide and non-Gaussian. The full corner plot for the non-linear fit is included as Figure B.10 in Appendix B. The corner plots of the posteriors for the full set of parameters modelled are included for all of the fits in the same Appendix.

The final model fits for the TRES-2b transit lightcurve are shown in Figure 6.17, where each method for modelling limb darkening appears to fit the lightcurve very well. As is expected when using different modelling techniques for limb darkening, the largest differences between the model fits are in the transit ingress and egress. The asymmetries seen between the fitted transit lightcurves in the ingress and egress in Figure 6.17 originate from the same source discussed for Figure 6.11.

Given the similarity between the transit parameters in Table 6.6, it is expected that the subsets of the corner plots shown in Figure 6.18 would look very similar. It is interesting to observe that whilst the shape of the posteriors for the GP fit are very Gaussian, the median

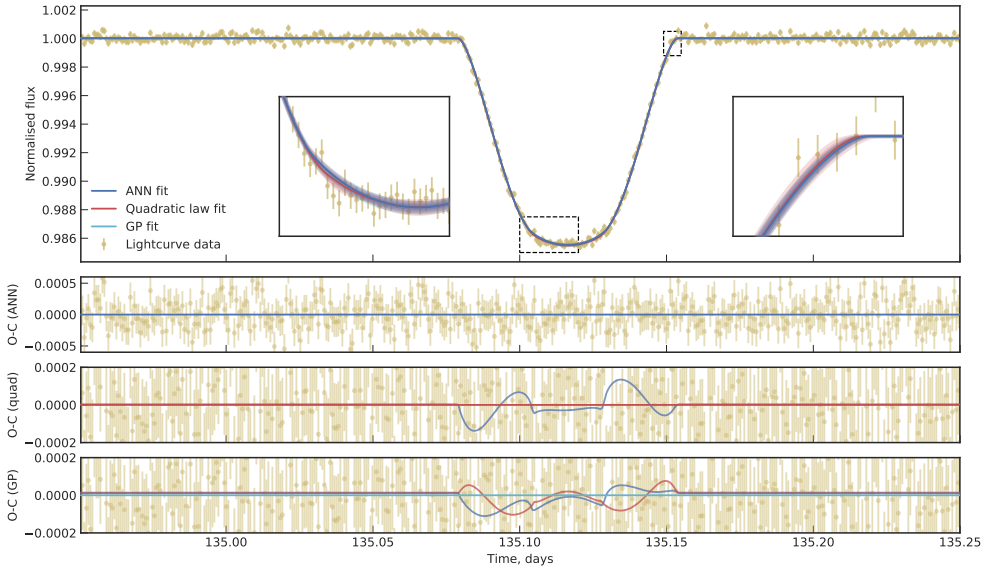



**Figure 6.17:** Upper plot shows the TRES-2b lightcurve fitted with ANN model, quadratic law, and GP interpolation. Each of the models are shown with 1, 2 and 3-$\sigma$ uncertainties derived from the posterior chains from each of the MCMC runs. The lower plots show the residuals of the fit to the data for the ANN model, quadratic law and GP interpolation. The lines over-plotted show the differences between the model fits with the colours referring to the comparison model.

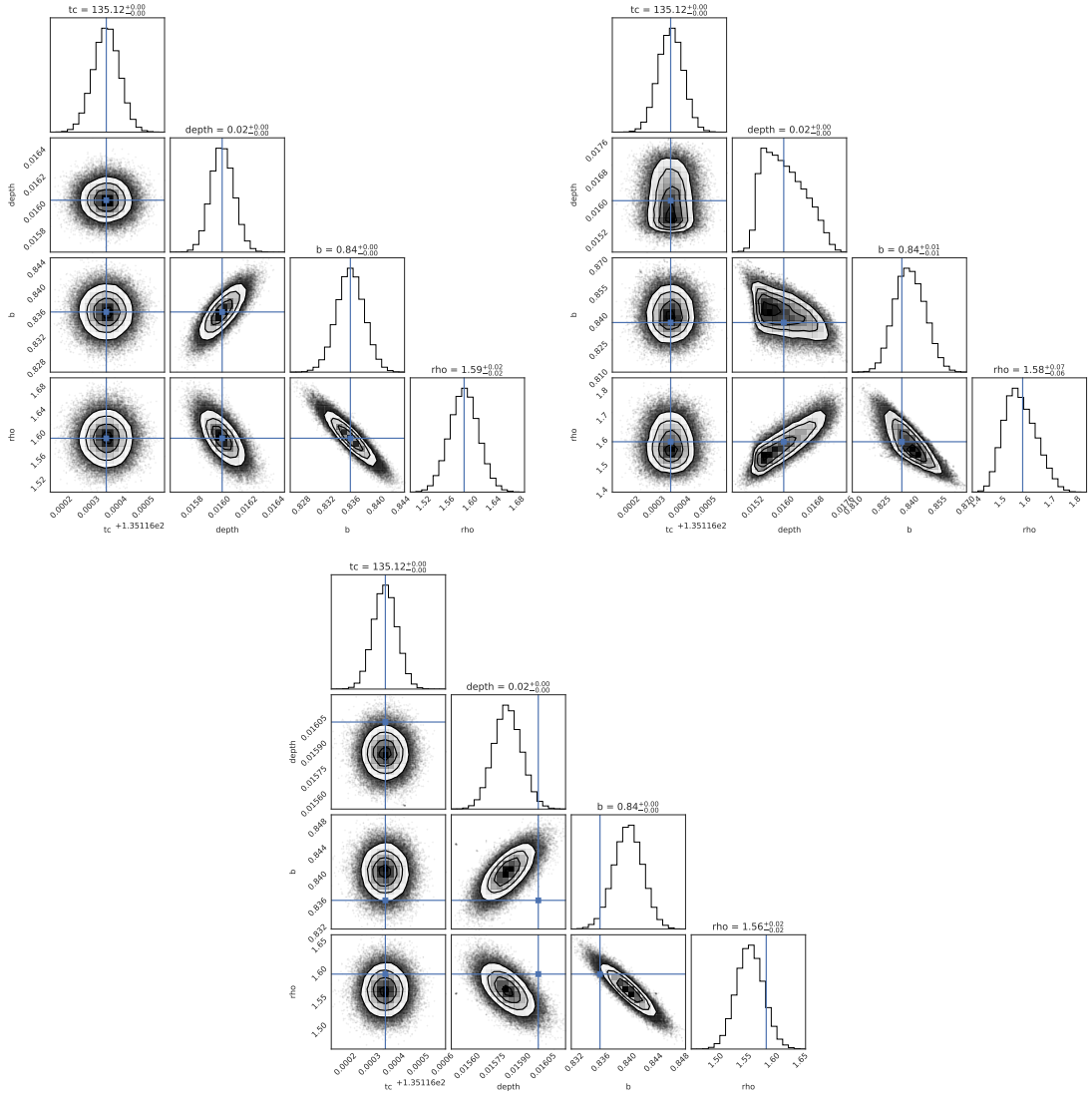



**Figure 6.18:** Plots of the MCMC posteriors for a subset of system parameters for each of the limb darkening fits. The blue lines in each plot represent the median parameter values found in the ANN fit. The upper left plot is for the ANN fit; the upper right plot is for the quadratic law fit, and the lower plot is for the GP interpolation fit.

values of the posterior are different from those found for the ANN fit. The GP method is used with a constraint on the stellar atmosphere parameters, which relatively tightly constrains the form of the limb darkening that the MCMC can explore based on the GP fit to the tabulated LDCs. The fact that the transit parameters for the GP fit are biased relative to the ANN values, would indicate that the GP model does not allow the form of the limb darkening profile to match the true limb profile, which has induced a bias in the transit parameters to increase the model likelihood.

The transit parameters found by the quadratic fit and the ANN fit are particularly similar, although the posteriors for the quadratic fit are less Gaussian. The quadratic limb darkening coefficients derived in the fit are however different from those in the Claret & Bloemen (2011) table of LDCs. Interpolation of the tabulated LDCs to predict the coefficients for TRES-2 performed by Schröter et al. (2012) resulted in values for $q_1$ and $q_2$ of 0.356 and 0.442 respectively, which are quite different to the $0.34 \pm 0.10$ and $0.4 \pm 0.3$ derived in the fit. The interpolated value for $q_1$ is close to the 1-$\sigma$ boundary for the value from the MCMC fit, and the value for $q_2$ is within the broad uncertainty bound found in the fit.

Figure 6.19 shows the limb profiles derived from the MCMC fits with the 1-, 2- and 3-$\sigma$ uncertainty bounds for each. The uncertainty of the quadratic fit is particularly large. It is expected that because the priors on the limb darkening coefficients for the quadratic fit were uninformative, the derived coefficients have been biased away from the expected coefficients from fitting the LDCs from the stellar atmosphere models, in order to better fit the other transit parameters. This bias in LDCs will also be induced by the two parameters of the quadratic law being insufficient to represent the true CLIV of TRES-2, particularly for low $\mu$ values.

The predicted limb profile from the GP fit shown in Figure 6.19 is significantly above that predicted by the ANN fit, which is the source of the biases in the transit parameters for the GP fit mentioned above.

## 6.5 Discussion

The ANN interpolation method has been shown to be very effective at predicting the limb profiles in the fits of the SDO/HMI transit of Venus data, and a single transit from the *Kepler* short cadence data. The predictions made for the solar limb profile match very well to the real profile. Where the stellar atmosphere parameters are less well known, the method is able to

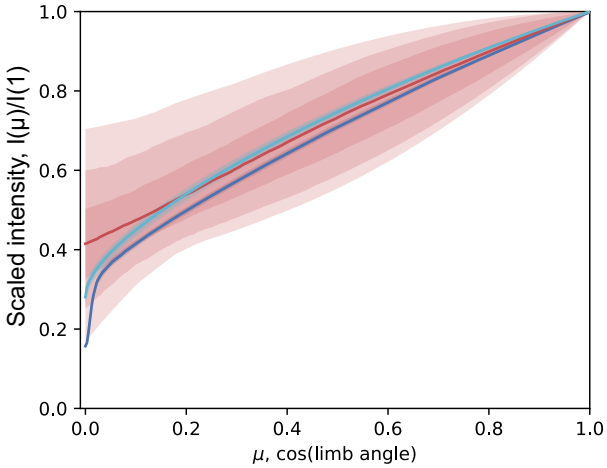



**Figure 6.19:** The output of the MCMC fit for each limb darkening model is plotted with reference to position on the stellar limb. The line for each colour shows the median limb variation profile, with the shaded regions referring to the 1, 2 and 3-$\sigma$ uncertainty ranges – the uncertainties in the profiles for the GP and ANN methods are plotted, but are much narrower than that for the quadratic fit. The colours used for each model fit are consistent with Figure 6.17.

explore the range of limb profiles that are physically expected, rather than available from the form of the parametric law chosen.

The accuracy of the technique is hinged on the accuracy of the stellar atmosphere models used to train the ANN, which is a limitation of this method. These atmosphere models are known to a better precision than is available from most exoplanet transit lightcurves, so there is less concern about the assumption of the atmosphere models providing an exact representation of the stellar limb profiles.

A significant limitation of using this ANN approach is the computational speed of the calculations associated with the ANN. The SDO/HMI lightcurve fitting took $\sim$ 4.5 hours using NVIDIA GTX 1070 GPU and AMD A8-7650K CPU. For a larger lightcurve data set, the timescale for performing an MCMC analysis could be restricting. It would be possible to speed up the processing by performing all of the modelling on the GPU, rather than just the predictions from the ANN – the data transfer from CPU to GPU is the most significant bottleneck in the computation time.

# 7

# Wavelength dependence of stellar radius

## 7.1 Introduction

In Chapter 6 a model of stellar limb darkening was constructed using ANNs, which was trained using the output of PHOENIX stellar atmosphere models. In Section 6.2.2 I explained why a correction is applied to account for the different positions of the stellar radius for different atmosphere models. Applying this correction makes sense when considering only transits in a single bandpass, but will introduce problems where lightcurves observed with multiple bandpasses are studied. Inconsistently defining the physical position where the planet is said to begin (or end) occulting the star in the limb darkening model for each bandpass will introduce biases in any parameters determined from that fit.

One of the primary uses of lightcurves obtained for several different wavelengths is to measure the atmosphere of the planet. The core principle behind this technique is that the amount of light from the star being blocked by the planet will be dependent on wavelength, as the planet's atmosphere will have a wavelength-dependent optical thickness at grazing incidence,

and thus a different apparent radius. The planetary radius will appear different because of overall temperature and pressure structure in the planetary atmosphere, and also for specific wavelengths because of absorption and emission lines for elements present in the atmosphere. This technique is known as transit spectrophotometry. The first measurement of the change in apparent planetary radius with wavelength for an exoplanet transit was made by Charbonneau et al. (2000), detecting sodium in the atmosphere of HD209458.

As the ability to measure the differences in transit depths for different wavelength ranges has improved, more planetary atmospheres have been characterised (eg. Sing et al. (2011); Wakeford et al. (2017); Parviainen et al. (2018)) from transit observations using space- and ground-based telescopes. All of these studies have assumed that the change in transit depth ($R_p^2/R_\star^2$) observed for different wavelengths is only introduced by the change in the apparent radius of the planet, and that the stellar radius remains constant. The core of the transit spectrophotometry method relies on the modelling of the dependence of observational wavelength on the planet's silhouette area. This dependence is also seen for the stellar atmosphere, but the effect is not normally included in the transit spectrophotometry modelling. This chapter explores the extent to which this assumption is reasonable, and introduces a new method to model the stellar limb darkening inclusive of the wavelength and stellar parameter dependence of the position of the stellar radius.

## 7.2 Stellar radius position

The correction applied to the radius position in stellar atmosphere model outputs described in Section 6.2.2 homogenises the radius position for all models and bandpasses to be at $\mu = 0$. This position is then defined to be where the gradient of the limb profile is steepest. If fitting multiple transit lightcurves for the same star, it is reasonable to assume that the overall stellar atmosphere parameters are shared for all of the light curves, so correcting for the differences in position of stellar radius between model atmosphere outputs is not going to bias the determination of the transit parameters. It is however not necessarily reasonable to apply the correction described when comparing the limb profiles for multiple bandpasses for the same star, since the radius defined by the correction is at a different position depending on the bandpass.

The difference in the radius positions found for the *Kepler* and 2MASS K bandpasses for the

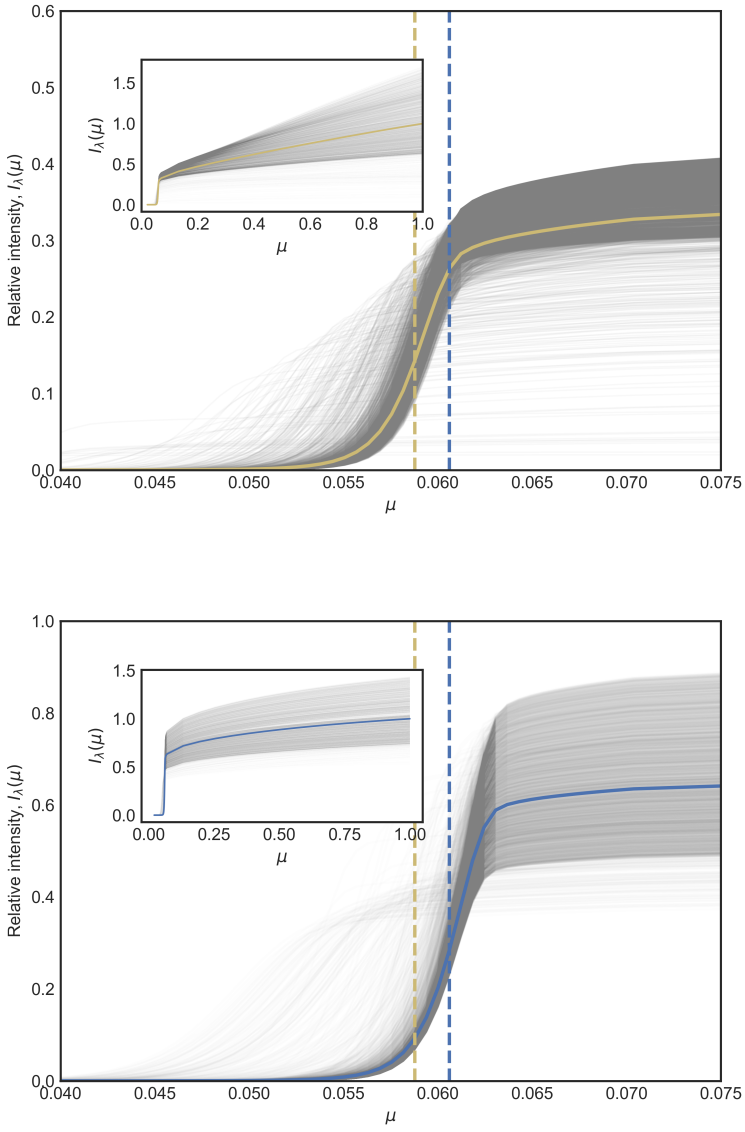



**Figure 7.1:** Demonstration of the difference in stellar radius when observing with different filters. Upper plot shows *Kepler* bandpass (yellow), and lower plot shows 2MASS K band (blue). The solid coloured line in each plot is the integrated flux for that bandpass dependent on $\mu$. The dashed vertical lines denote the positions identified as the stellar radius, where the colour indicated the passband referred to. Each thin grey line shows the limb profile for every 0.1 nm in the range of the FWHM of the passband. The example stellar atmosphere model used has $T_{\rm eff} = 5500$ K; $\log g = 4.5$, and [Fe/H] $= 0.0$.

same star are shown in Figure 6.2. The difference is small, but when looking for transit depth differences of order ∼50 ppm, small differences are significant. The detail of the difference between the apparent stellar radii in these bandpasses is explored in more detail in Figure 7.1. The Figure shows (in thin grey lines) the intensity variation with $\mu$ value for each individual wavelength of light included in the PHOENIX atmosphere model specific intensity spectrum (every 0.1 nm in the range of the FWHM of the two bandpasses) scaled to equal unity for the median intensity value at $\mu = 1$. The coloured solid lines in each plot show the intensities convolved with the bandpass and integrated over wavelength to produce the limb profile of the model atmosphere for that bandpass. The dashed vertical lines show the position that the stellar radius would be defined at for that bandpass.

The plotting of the intensities for the individual wavelengths in Figure 7.1 begins to demonstrate the source of the variation in apparent stellar limb position with wavelength of light. Most of the intensity profiles correspond to wavelengths in the stellar spectral continuum. They are formed deep in the atmosphere, and have very similar functional dependences on $\mu$. It is worth noting that the shapes of each of these profiles are not exactly the same, as even for a small shift in wavelength, the path the light takes through the stellar atmosphere because of the atmospheric temperature-pressure gradient is different enough to shift the apparent radius of the star a small amount. A particularly interesting feature in the plots is that some of the intensity profiles for individual wavelengths do not sit with the profiles for the continuum wavelengths – these are the wavelength values for absorption lines. The position of maximum gradient in the limb profile for the absorption lines is at a much higher $\mu$ value than for the continuum, as light at that wavelength is more readily absorbed by the stellar atmosphere. The position where the optical depth equals unity is therefore higher in the atmosphere, and thus the apparent radius of the star is bigger. This effect will be more prominent in bandpasses with large numbers of absorption lines or narrow bandpasses centred around a stellar absorption line, as the overall apparent stellar radius will be shifted to a higher value than suggested by just the continuum.

The characteristics of the variation in apparent stellar radius across the entire range of modelled wavelengths is demonstrated in Figure 7.2. The images have been created from the same example stellar model atmosphere as in Figure 7.1, where each row represents an individual wavelength from the shortest wavelength modelled (50 nm) at the top of the images, to the longest (2600 nm) at the bottom. The right-hand image shows the variation in inten-

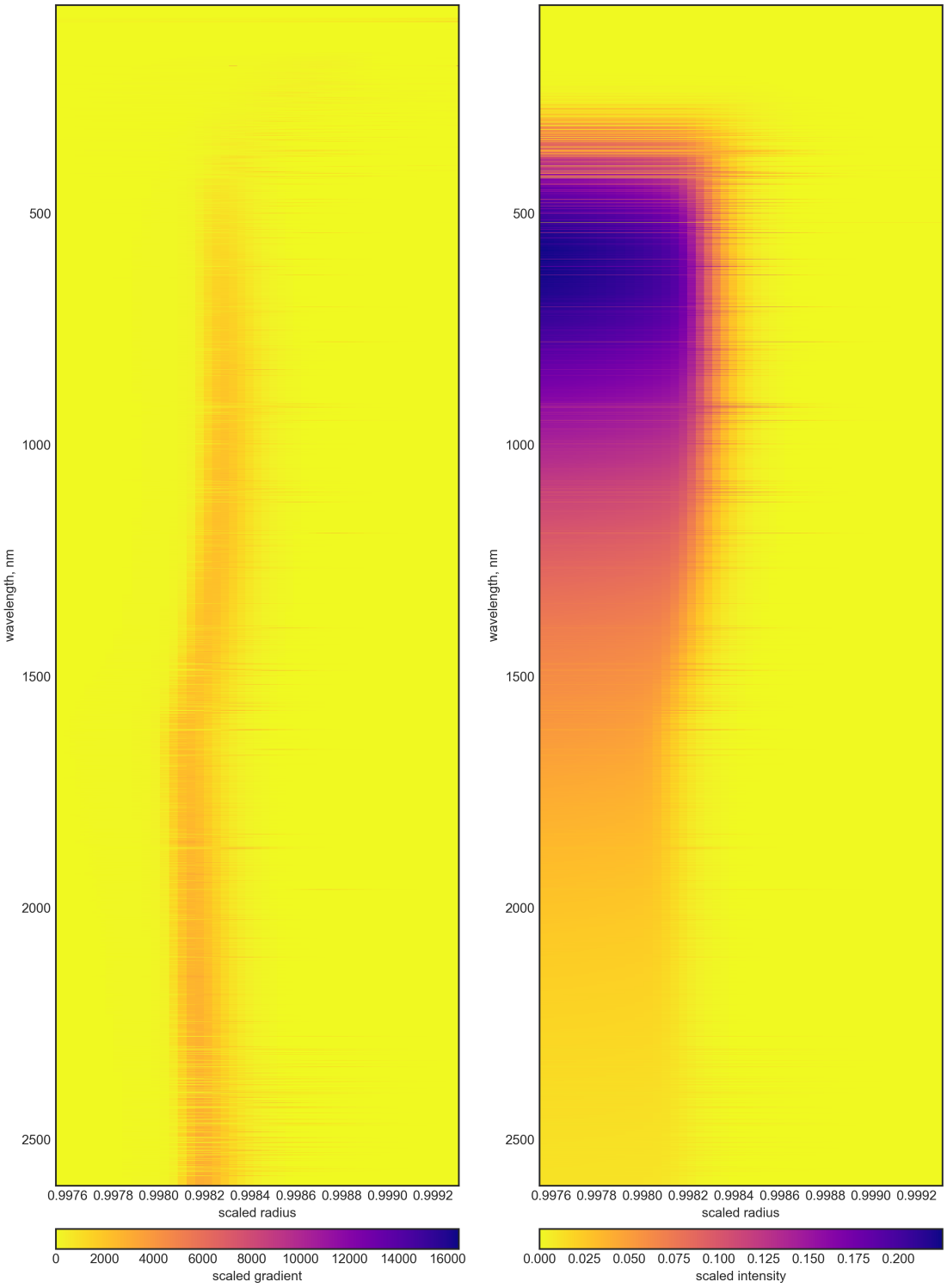



**Figure 7.2:** Demonstration of radius position variation with wavelength for the same example stellar atmosphere model as in Figure 7.1. Each row in the image is for a modelled wavelength, from 50 – 2600 nm. The left-hand image shows the variation in gradient of the limb profile for each individual wavelength modelled, and the right-hand image shows the intensity values scaled to the median intensity value at radius = 0 for all wavelengths.

sity across the narrow range of radii shown in the image, and the left-hand image shows the gradient in the limb profile at each radius plotted.

Morello et al. (2017) explored the scale of the difference between the radii determined for three stellar atmospheres – M5, M0 and F0 stars. It was found that the largest differences in transit depths induced by the differences in the stellar radius were 5 ppm for the M5 star, 4 ppm for the M0 star, and 10 ppm for the F0 star. There are a couple of reasons why the true values for the scale of this effect are slightly larger than the values quoted in Morello et al. (2017). The authors utilised the stellar reference radius used in the PHOENIX atmosphere models to define $\mu$ and $R_\star$ to compare differences in transit depth against. The radius in the PHOENIX models is selected so that it is beyond the point where the atmosphere is optically thick across all modelled wavelengths. This ensures that the full picture of emitted light is included in the model atmosphere outputs. Selecting this radius as the comparison reduces the scale of the difference between the apparent radii for the relevant observational bandpasses. Here it is proposed to set the comparison stellar radius as that found for a particular passband, so that the definition of the reference radius is equivalent to the apparent radius for a transit lightcurve. The second reason that the scale of the impact of the stellar radius variation is larger than suggested by Morello et al. (2017) (although not relevant to the broad bandpasses considered in their work) is that transmission spectrophotometry makes use of narrow filter bands, so the stellar absorption features can have strong effects on the apparent stellar radius, as demonstrated in Figures 7.1 and 7.2. With these considerations, the effect on an observed transit depth from stellar radius variation with wavelength is comparable to the precision of the transit spectrophotometric observations of transit depths obtained. For example Knutson et al. (2014) attain an average uncertainty in transit depth of 23 ppm.

## 7.3 ANN approach

In Chapter 6 a method to interpolate limb darkening profiles using ANNs was introduced. That method considered the stellar radius to be constant across all wavelengths, and modelled only light from within the determined stellar radius position. The limb profiles in Figures 7.1 and 7.2 clearly show stellar intensity beyond the point of maximum gradient in the limb profile. The approach introduced here proposes to model the entire profile, to account for the real soft edge of the star.

### 7.3.1 Adjusted training data

Since the definition of $\mu$ is dependent on the stellar radius, if the position of the stellar radius is defined at the point of maximum gradient in the limb profile, then any light emitted beyond that point would be attributed to negative values of $\mu$. Values of $\mu < 0$ are mathematically inconvenient when modelling the transit lightcurve, so instead the definition of the stellar radius directly from the PHOENIX models are not corrected as in Section 6.2.2, but all other steps described in Section 6.2.1 are the same for constructing the training data grid for this ANN model.

As the stellar radius used to construct the training data grid is no longer the same as that which would be observed for a transit in that bandpass, a second output is included alongside intensity, which models the stellar radius position. The second output included is the value of the stellar radius which would have been utilised to apply the correction described in Section 6.2.2 for that atmosphere model and bandpass. The radius correction factor is denoted $R_{cor}$ henceforth. The values of the radius factor are all very close to 1, and have low variation around that value, which is not trivial to model with an ANN, so the radius factor is remapped into a range [-1,1] with the maximum radius factor for that bandpass mapped to 1, and the minimum mapped to -1.

The limb profiles, when extended beyond the defined stellar radius, are particularly non-linear, so a reparameterisation is required to extend that region of the $\mu$ parameter space relative to higher values of $\mu$, which will allow the ANN to model the form of the limb profile in that region more readily. The reparameterisation applied is to use $\log_{10}(\mu)$ as the input to the ANN instead of $\mu$. For the example fits model training shown below, the V band data has been used as the training data set.

### 7.3.2 New ANN architecture

As in Chapter 6, a fully connected feed-forward ANN model is utilised to model the limb profiles. The form of the model found to be most successful for this application is shown in schematic form in Figure 7.3. The exact form of the architecture is described in Table 7.1. Since the value of the radius correction is only dependent on the the stellar parameters, and not the angle, the model is split into two part with separate output, rather than merging to produce the output as for the model in Chapter 6. The outputs are however modelled together, and

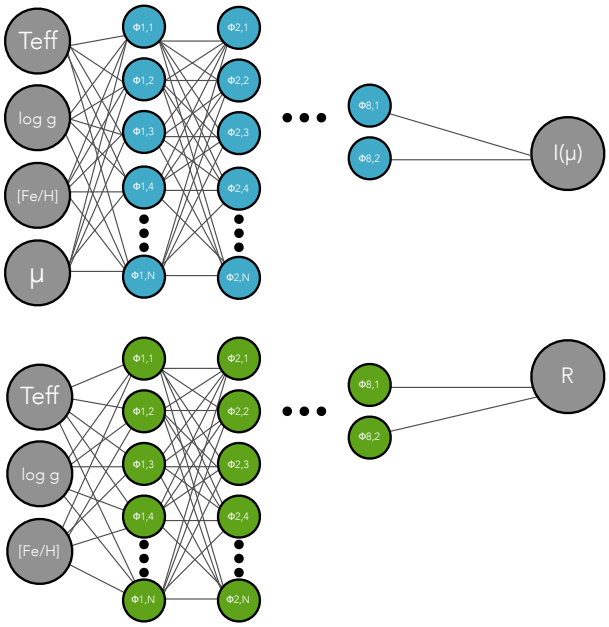



**Figure 7.3:** Schematic representation of the architecture used in the new ANN for the work in this Chapter, where the grey circles on the left represent the input neurons for an individual trial; the solid grey lines are the weighted connections between neurons; the dashed grey lines are the bias terms for each neuron; the coloured circles represent the neurons themselves organised in layers, and the grey circles on the right are the output neurons – the intensity, and the radius correction factor. The blue circles are the $\mu$ value branch, and the green circles are the radius correction branch.

**Table 7.1:** Detail of the structure of the new ANN model based on the schematic in Figure 7.3

| Layer | Number of neurons | Activation Function |
|---|---|---|
| *Radius correction branch* | | |
| Input | 3 – $T_{\text{eff}}$, $\log g$, [Fe/H] | – |
| 1 | 20 | ReLU |
| 2 | 20 | ReLU |
| 3 | 20 | ReLU |
| 4 | 10 | ReLU |
| 5 | 6 | ReLU |
| 6 | 4 | ReLU |
| Output | $1 - R_{cor}$ | – |
| *$\mu$ value branch* | | |
| Input | 4 – $T_{\text{eff}}$, $\log g$, [Fe/H], $\mu$ | – |
| 1 | 200 | ReLU |
| 2 | 100 | ReLU |
| 3 | 100 | ReLU |
| 4 | 100 | ReLU |
| 5 | 100 | ReLU |
| Output | $1 - I$ | Sigmoid |

the backpropagation algorithm is applied to both branches together throughout the training despite there being no connections between the two branches.

As can be seen in Table 7.1, every layer has the ReLU activation function – this was decided purely pragmatically as it offered the best training performance. The final (ouput) layer of the $\mu$ value branch made use of the sigmoid activation function as the output is physically required to be in the range [0,1]. Using the sigmoid activation function more generally resulted in difficulty training the model because of the vanishing gradient problem.

The ANN model was trained for 100 epochs to ensure the model was well trained to fit the data. The training batch size selected was 100, as the number of $\mu$ samples for each atmosphere model in this training set is closer to 100 than 50 as for the training data in Chapter 6.

### 7.3.3 Training ANN

The surface plotted in Figure 7.4 shows the variation in the limb profile output predicted by the new ANN model for different values of $T_{\text{eff}}$. The output is relatively smooth, although there are clearly a few regions of the parameter space where the model is not predicting a smooth profile. Smoothness is not assumed by the ANN model, so regions where discontinuities are

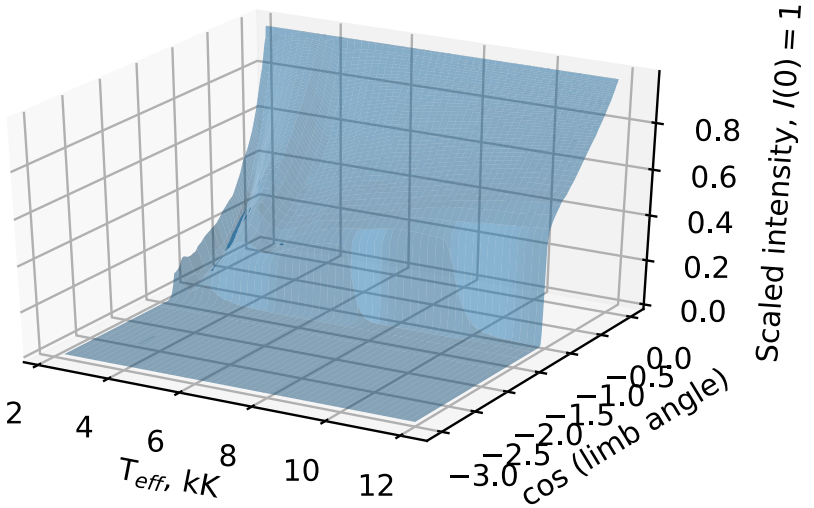



**Figure 7.4:** An example 3D surface of predicted outputs of the new trained ANN model. The Figure shows the variation in the form of the limb profiles for a range of $T_{\rm eff}$ values for $\log g = 4.7$, and [Fe/H] = -0.2. The x-axis is for $\log_{10}(\mu)$, so that the full modelled variation is visible.

observed are likely areas of the $log_{10}(\mu)$ space which are not well sampled in the training data.

A couple of example limb profiles modelled by the new ANN are shown in Figure 7.5 with the training data for those stellar parameters. The model predicts very well the form of these limb profiles, and it is particularly promising to see that the full shape of the behaviour at the limb is well represented by the model, despite the complex form of the profile. It would be very difficult to find a parametric law for limb darkening that could model that detail without an impractical number of coefficients.

## 7.4 Implementing into transit model

With the new ANN model trained, it can be implemented into a transit model in the same way as the original ANN method. The difference between the two methods is the way that the stellar radius is defined differently for each. This means that the values determined in the transit fit (as parameterised in this work) for the transit depth ($R_p/R_\star$) and the stellar density ($\rho_\star$) will require adjustment to reflect the definition of the stellar radius desired. The transit model using this ANN model will have included the starlight originating from outside the conventional definition of $R_\star$, which a real star has and will affect the shape of a real transit, but is usually not included in the model.

The ANN model will predict the factor required to transform the stellar radius definition

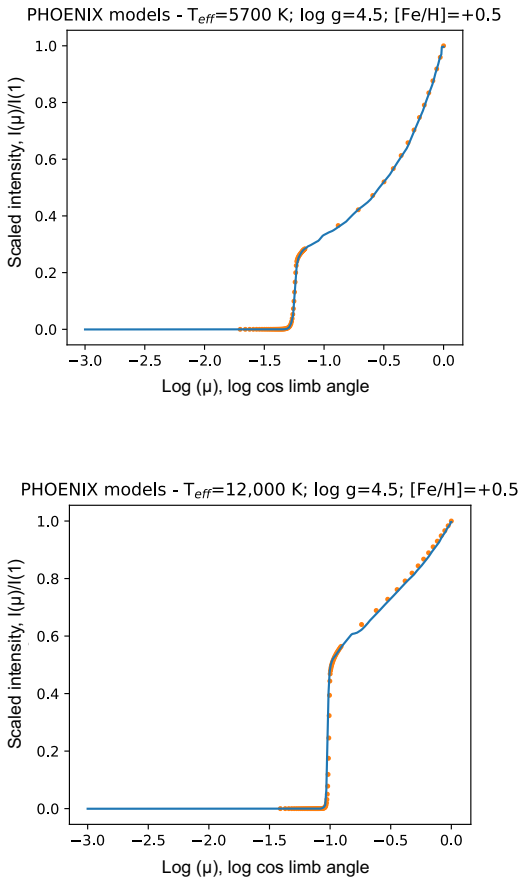



**Figure 7.5:** Figure showing two example limb profiles for different stellar parameters for the new ANN model. The upper plot is the predicted limb profile for $T_{\rm eff} = 5700$ K, $\log g = 4.5$, and [Fe/H] = +0.5. The lower plot has $T_{\rm eff} = 12{,}000$ K, $\log g = 4.5$, and [Fe/H] = +0.5.

from that defined in the atmosphere model to that expected for the definition described in Section 6.2.2. The factor predicted by the ANN model then allows the determined values of $R_p/R_\star$ and $\rho_\star$ to reflect the conventional definitions of the stellar radius for that bandpass.

## 7.5 Discussion

The new ANN approach is very effective at modelling the full shape of the limb profile across the stellar parameter space. The architecture and training of the model is not optimised in its current state, but despite this, the feasibility of the concept behind the method can be clearly seen.

To optimise the model, to ensure that the limb profiles across the stellar parameter space are well represented by the trained model, adjustments to the architecture of the model and potentially the form of the input parameterisation should be tested to find a good enough model for the required purpose. This process is time consuming, as each instance of training a model takes a significant amount of computation time ($\sim$ 1 hour for 100 training epochs and batch size 100).

Once an optimised model has been trained, the efficacy of the method should be validated through fitting example transit lightcurves. The optimal use for this method is in studies where transit lightcurves for multiple bandpasses for the same system are considered together. The current structure would requite training separate ANN models for each bandpass, or to use the bandpass as a categorical variable with the full training set used to inform the ANN model.

The approach described here is to use the radius defined in the atmosphere models as the comparison stellar radius, but this radius does not relate to any physical definition of the value of $R_\star$ – it is selected to be a position at which the star is known to be optically thin across the modelled wavelengths. A more practical comparison radius may be one for a selected wavelength where the position of the radius is not effected by elemental absorption lines.

# 8

# Conclusions and next steps

## 8.1 Conclusions

The field of exoplanet research is awash with exciting discoveries, and the work in this thesis provides some part in colouring in some part of that landscape. I present the discovery of some new planets, and I present work to improve the ability of future researchers to learn more about planets that are already detected.

In Chapter 2, I present the discovery and system parameter determination for three new hot Jupiter planets – WASP-92b, WASP-93b and WASP-118b. I modelled the transit lightcurves and RV measurements of the systems to determine the system parameters. For WASP-93b and WASP-118b there were time series spectral measurements during the transit, which were analysed for each to determine the potential spin-orbit alignment (or misalignment) of the system. I also investigated the tidal interactions of each of the planets with their host star, and was able to show the tidal migration that each planet will undergo depending on the spin-orbit alignment of the systems.

The remainder of the thesis contains studies of the modelling of stellar limb darkening. In Chapter 4, I describe a novel method for interpolating between tabulated coefficients for limb darkening, which utilises a Gaussian Process to train the interpolation surface in stellar atmosphere parameter space. The Gaussian Process is conditioned on an existing grid of tabulated coefficients, and with optimised hyperparameters can then be used to predict limb darkening coefficients for given stellar atmospheric parameters. The technique is demonstrated for fits of simulated transit lightcurves, and for real transit lightcurves in Chapter 6. The method is effective at constraining the values of the limb darkening coefficients based on the existing knowledge of the stellar atmosphere parameters, and prevents the introduction of biases from linearly interpolating between tabulated coefficients where the variation between them is not linear and is multi-dimensional.

Chapter 6 provides the description of another novel method for modelling limb darkening. The approach presented in this Chapter is prefaced on the knowledge that parametric models for limb darkening do not fully represent the CLIV for the stellar surface. The technique proposed the use of an Artificial Neural Network (ANN), which is trained directly from the outputs of stellar atmosphere models. The method removes the need for a parametric approximation of the form of the limb darkening, and is able to predict the limb darkening for any position on the stellar surface for the given stellar atmosphere parameters, even for stellar parameters not corresponding to an existing atmosphere model. The technique is validated through fitting a lightcurve of the transit of Venus, and for a transit of TRES-2b.

The final results Chapter, 7, explores the potential for the ANN method described in Chapter 6 to be extended to include the consideration of the wavelength-dependent nature of the stellar radius. I have developed a model architecture and training methodology that demonstrates the feasibility of using an ANN to model the full output shape from stellar atmosphere models with spherical geometry. The trained model matches well to the training data, and the potential to use this method to model limb darkening where very high precision is required has been demonstrated.

## 8.2 Next steps

To take the work in this thesis further, the primary goal would be to test training both ANN methods with 3D stellar atmosphere models. Despite the significantly smaller training sample

from the grids of 3D stellar atmosphere models, the inclusion of more realistic physics makes the comparison of the results from the 3D and 1D models an important step in defining the most accurate model of limb darkening to use in transit fitting applications.

A significant disadvantage of using an ANN model for limb darkening is that the required transit model for a non-parametric model is much slower to compute than for a parametric limb darkening model. The development of a transit model appropriate for the ANN approach that would perform all of the transit model computations on a GPU would offer a significant speed up in computation time.

The new ANN approach described in Chapter 7 is still under development, so the next steps for that technique would involve further work to optimise the architecture and training procedure. This would transform the approach from a modelling prospect to a viable alternative model for limb darkening.

# A

# Photometric bandpasses

The table below lists each of the transmission curves for which the limb profiles were calculated. The table contains:

- The name of the filter
- The letter code used to represent the filter
- Approximately what wavelength range the filter represents
- The source for the bandpass information

**Table A.1:** Table of the photometric bandpasses used in this study, and the source location for each

| Name | Code | Wavelength (nm) | Source |
|---|---|---|---|
| Kepler | Kp | 420–900 | http://keplergo.arc.nasa.gov/CalibrationResponse.shtml |
| Tycho V | vT | 460–670 | http://ulisse.pd.astro.it/Astro/ADPS/Systems/Sys_135/index_135.html |
| Tycho B | bT | 350–500 | http://ulisse.pd.astro.it/Astro/ADPS/Systems/Sys_135/index_135.html |
| 2MASS J | J | 1100–1350 | https://www.ipac.caltech.edu/2mass/releases/allsky/doc/sec6_4a.html |
| 2MASS H | H | 1500–1800 | https://www.ipac.caltech.edu/2mass/releases/allsky/doc/sec6_4a.html |
| 2MASS K | K | 2000–2300 | https://www.ipac.caltech.edu/2mass/releases/allsky/doc/sec6_4a.html |
| SDSS u | u_ | 310–490 | http://classic.sdss.org/dr7/instruments/imager/ |
| SDSS g | g_ | 420–560 | http://classic.sdss.org/dr7/instruments/imager/ |
| SDSS r | r_ | 550–700 | http://classic.sdss.org/dr7/instruments/imager/ |
| SDSS i | i_ | 700–870 | http://classic.sdss.org/dr7/instruments/imager/ |
| SDSS z | z_ | 820–1000 | http://classic.sdss.org/dr7/instruments/imager/ |
| Bessel U | U | 300–400 | http://spiff.rit.edu/classes/phys440/lectures/filters/filters.html |
| Bessel B | B | 380–500 | http://spiff.rit.edu/classes/phys440/lectures/filters/filters.html |
| Bessel V | V | 500–600 | http://spiff.rit.edu/classes/phys440/lectures/filters/filters.html |
| Bessel R | R | 560–750 | http://spiff.rit.edu/classes/phys440/lectures/filters/filters.html |
| Bessel I | I | 730–900 | http://spiff.rit.edu/classes/phys440/lectures/filters/filters.html |
| Stroemgren u | u | 340–370 | http://spiff.rit.edu/classes/phys440/lectures/filters/filters.html |
| Stroemgren v | v | 400–430 | http://spiff.rit.edu/classes/phys440/lectures/filters/filters.html |
| Stroemgren b | b | 460–480 | http://spiff.rit.edu/classes/phys440/lectures/filters/filters.html |
| Stroemgren y | y | 530–570 | http://spiff.rit.edu/classes/phys440/lectures/filters/filters.html |
| Spitzer IRAC band 1 | S1 | 3150–3900 | http://irsa.ipac.caltech.edu/data/SPITZER/docs/irac/calibrationfiles/spectralresponse/ |
| Spitzer IRAC band 2 | S2 | 3950–5050 | http://irsa.ipac.caltech.edu/data/SPITZER/docs/irac/calibrationfiles/spectralresponse/ |
| Spitzer IRAC band 3 | S3 | 5000–6450 | http://irsa.ipac.caltech.edu/data/SPITZER/docs/irac/calibrationfiles/spectralresponse/ |
| Spitzer IRAC band 4 | S4 | 6350–9500 | http://irsa.ipac.caltech.edu/data/SPITZER/docs/irac/calibrationfiles/spectralresponse/ |
| CHEOPS | Cp | 360–960 | *(private comm)* |

# B

# Full MCMC posterior figures

This appendix contains figures which show the full posterior plots for the MCMC transit fits in Chapter 6 for the SDO/HMI transit of Venus data, and the *Kepler* TRES-2b transit data. The definitions of each of the parameters plotted here are given in Table 6.3 for the SDO/HMI fit, and in Table 6.5 for the TRES-2b transit.

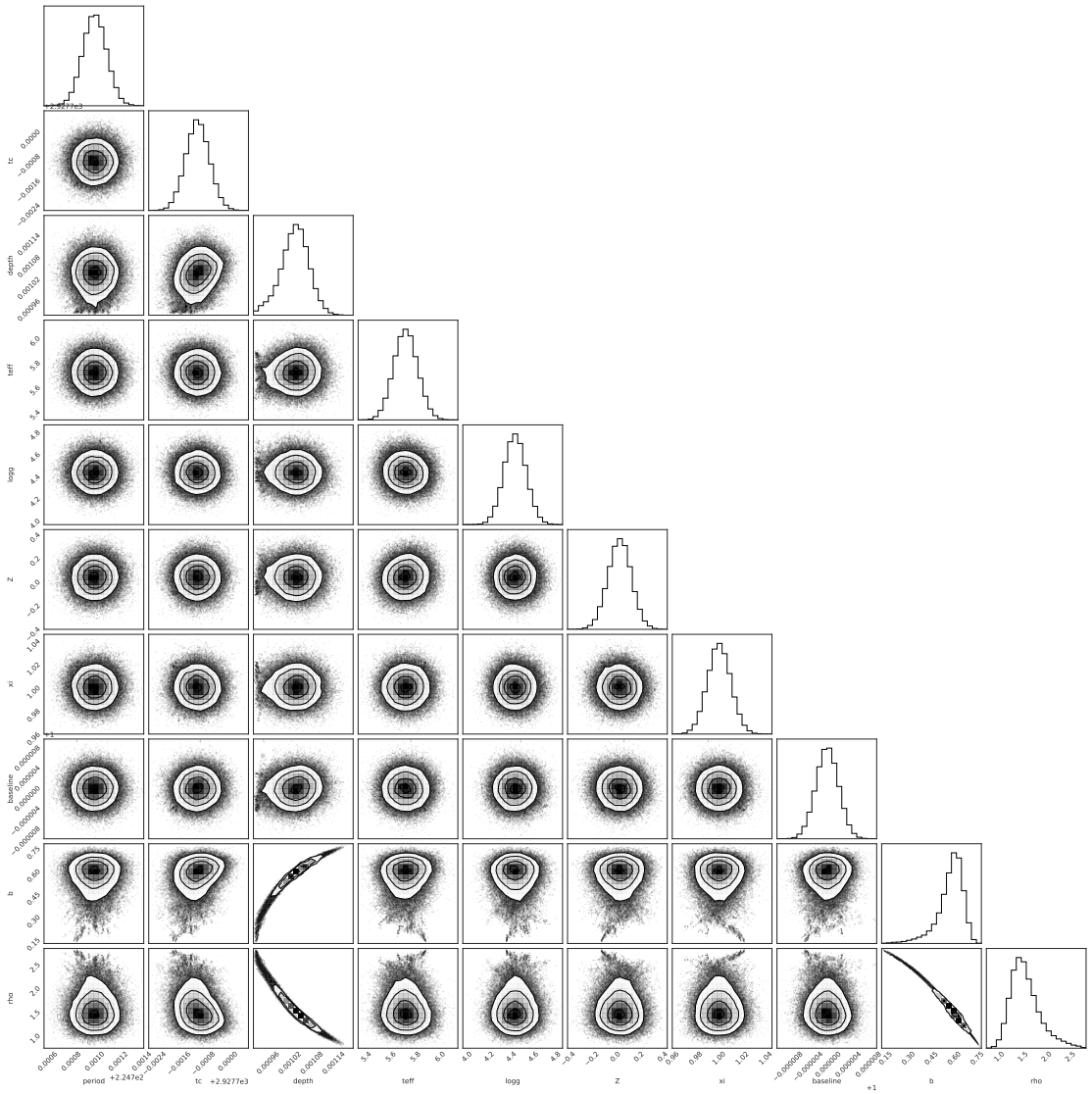



**Figure B.1:** Plot of the MCMC posteriors for the system parameters for the ANN limb darkening fit for the SDO/HMI transit of Venus data.

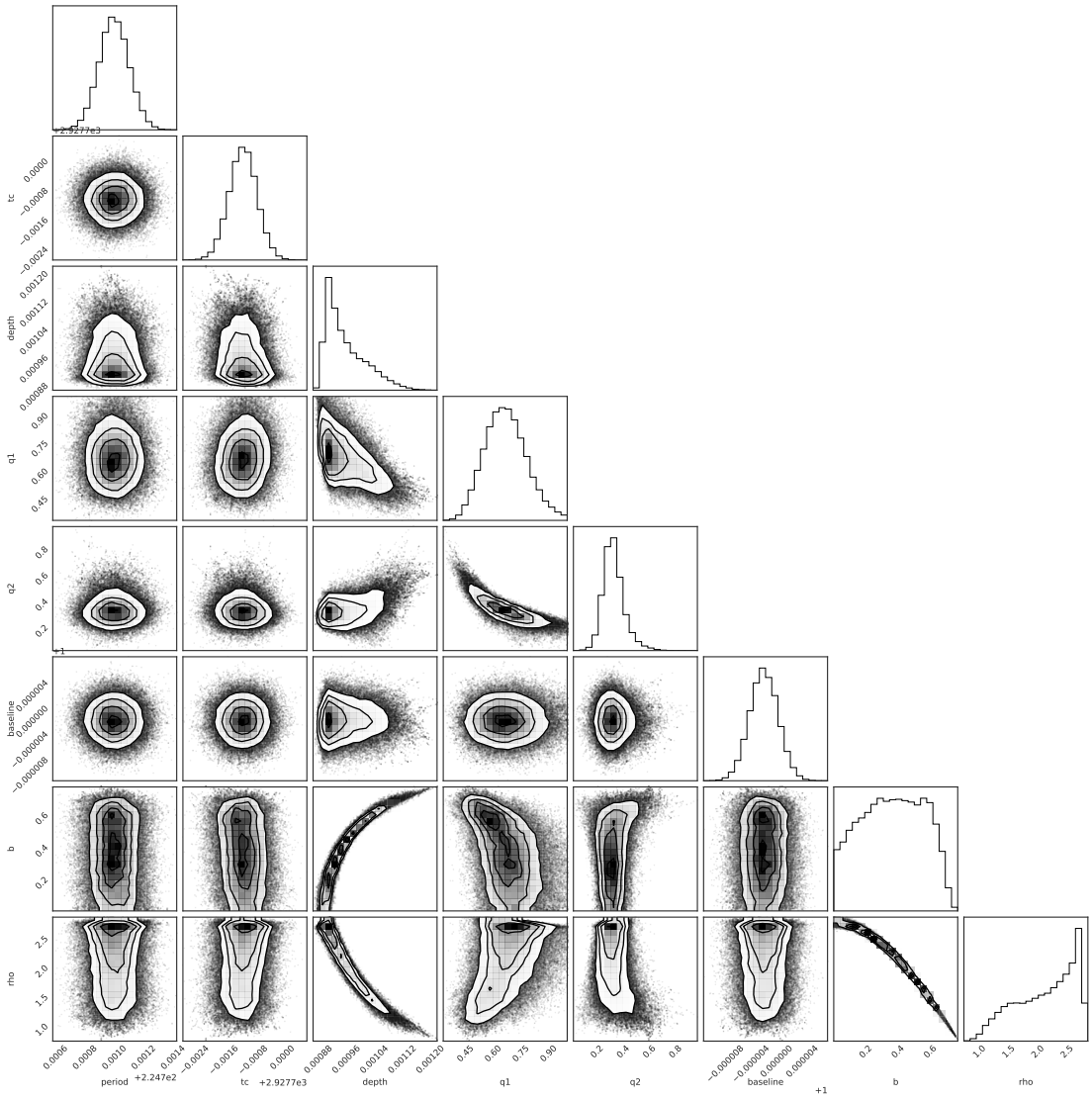



**Figure B.2:** Plot of the MCMC posteriors for the system parameters for the quadratic law limb darkening fit for the SDO/HMI transit of Venus data.

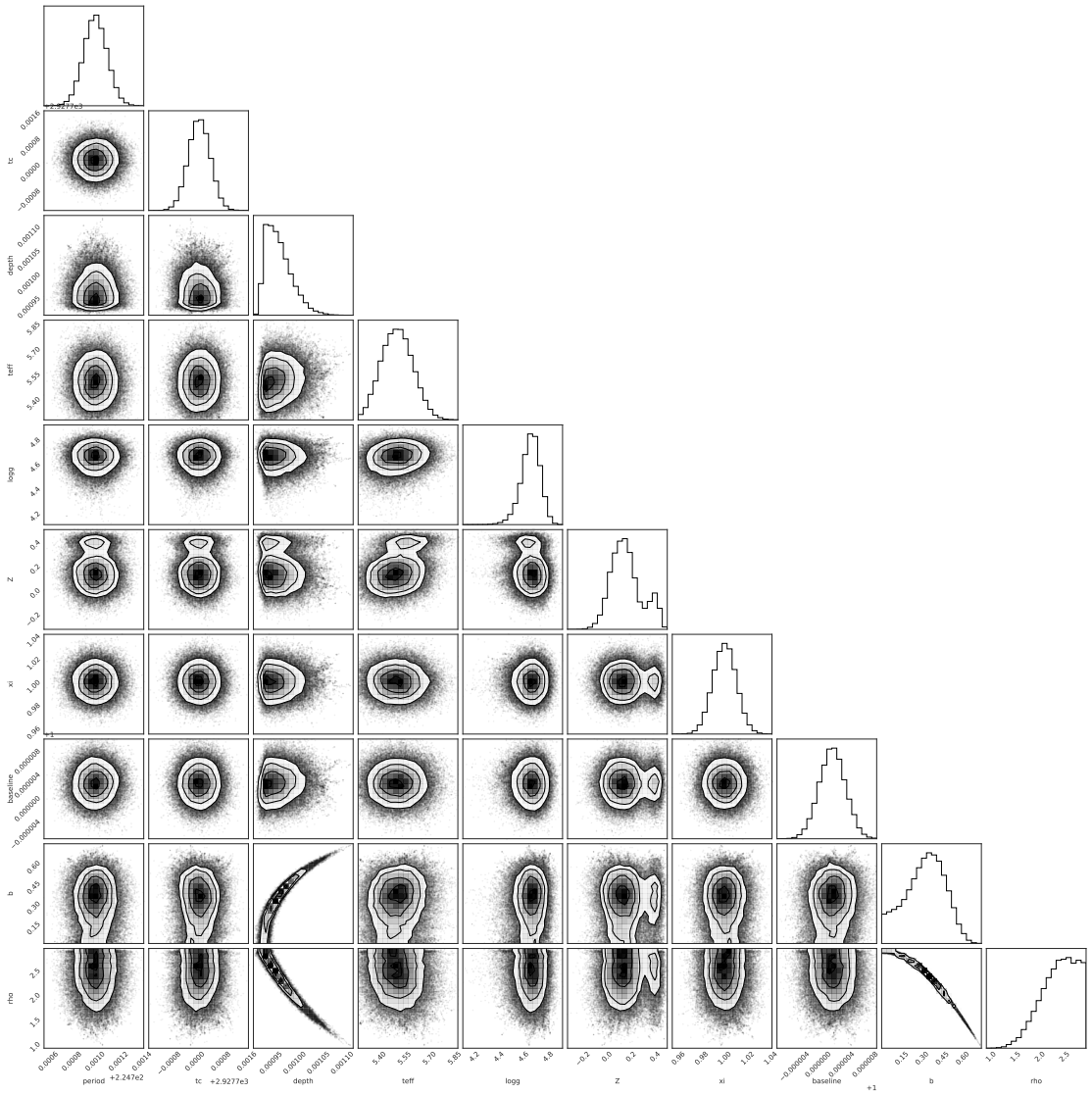



**Figure B.3:** Plot of the MCMC posteriors for the system parameters for the GP limb darkening fit for the SDO/HMI transit of Venus data.

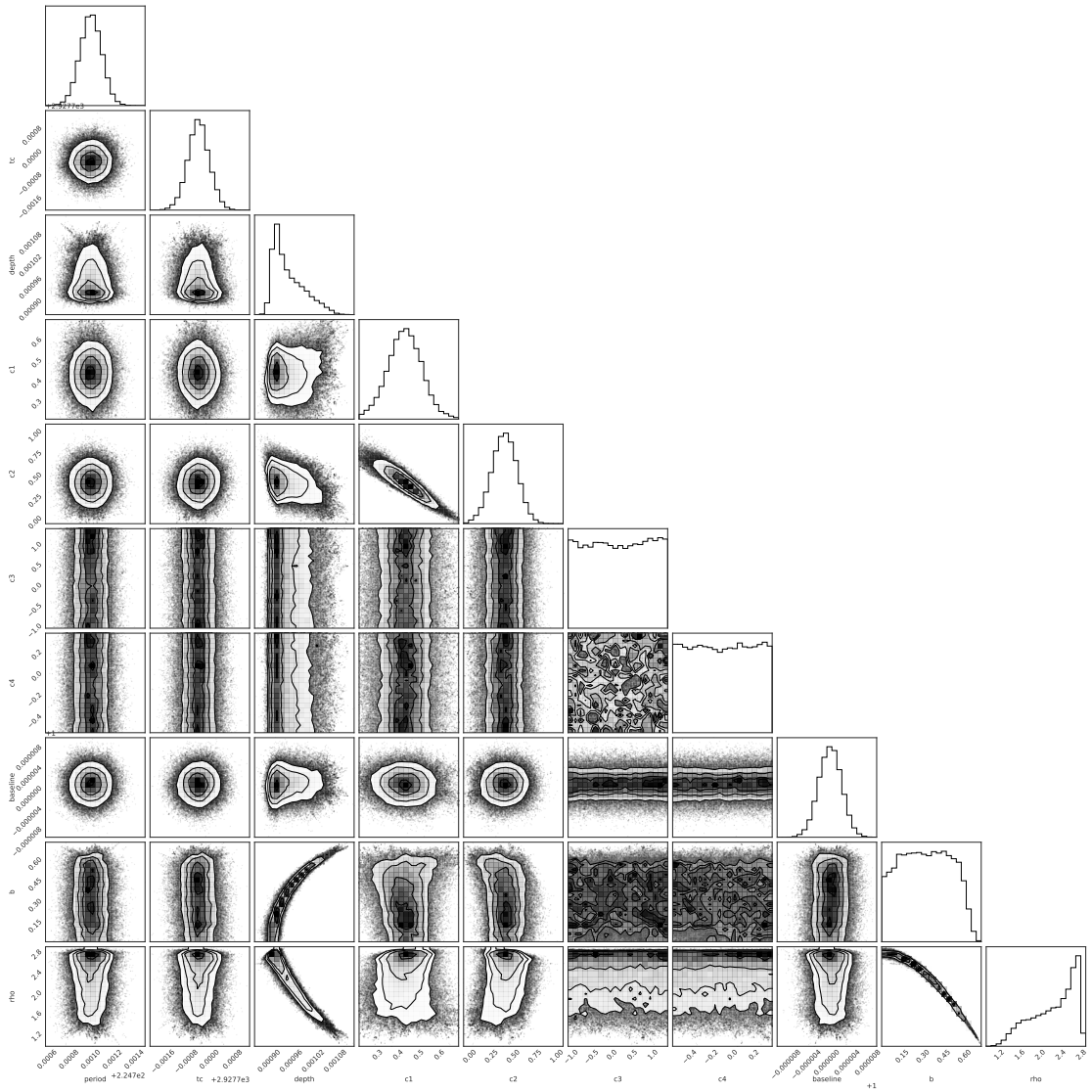



**Figure B.4:** Plot of the MCMC posteriors for the system parameters for the non-linear law limb darkening fit for the SDO/HMI transit of Venus data.

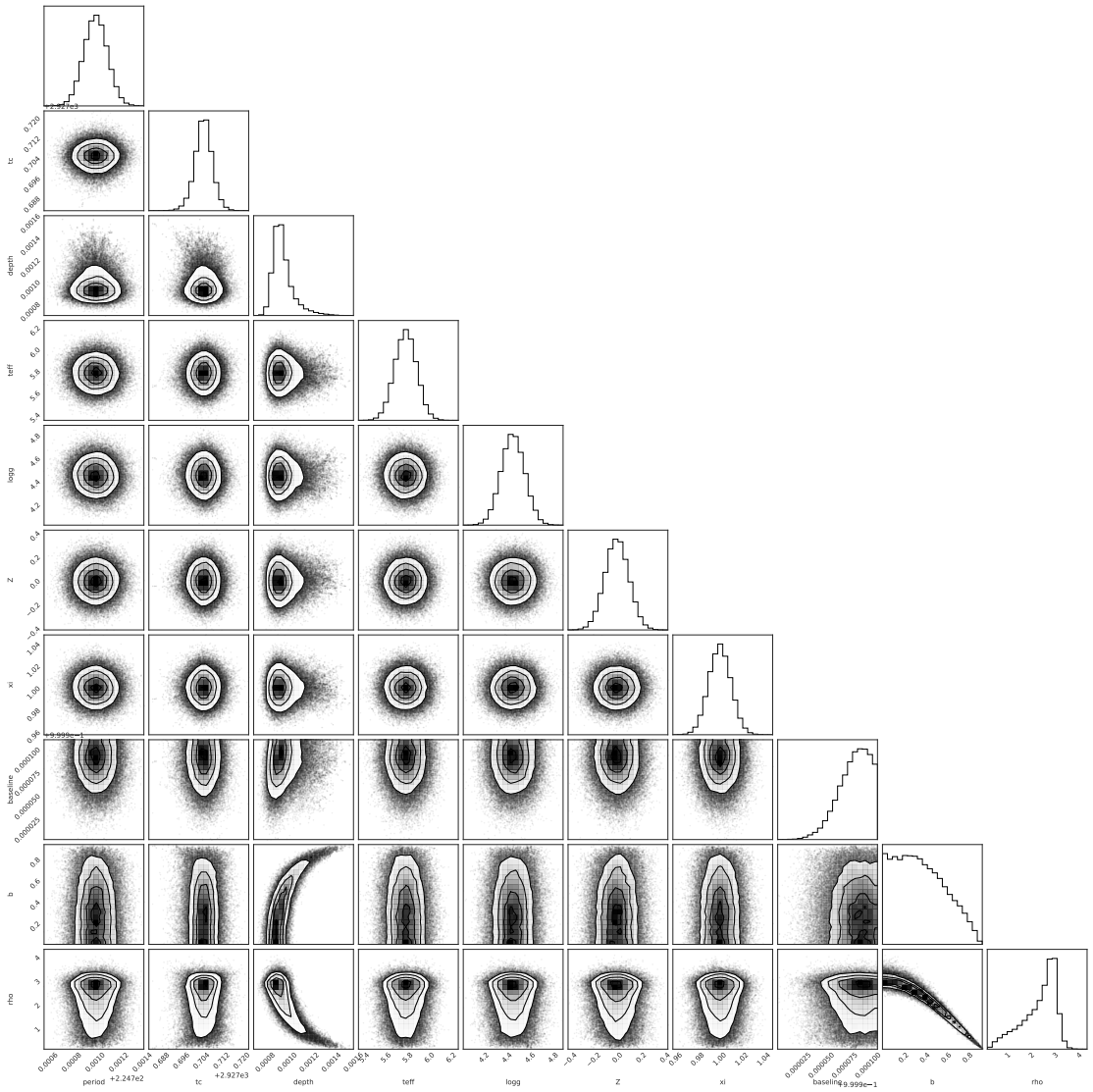



**Figure B.5:** Plot of the MCMC posteriors for the system parameters for the ANN method fit for the SDO/HMI transit of Venus data with significant random noise added to the transit lightcurve prior to fitting.

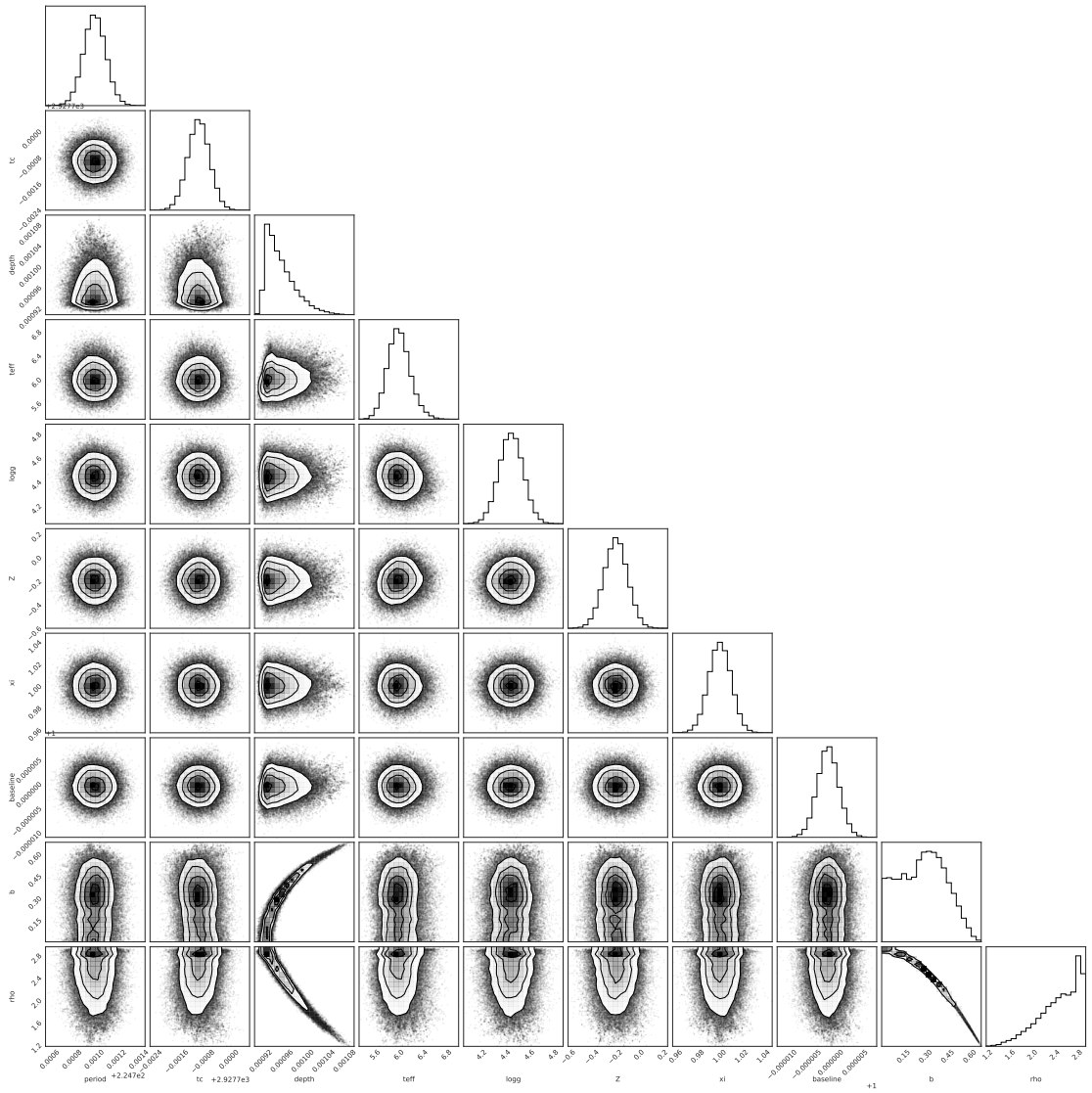



**Figure B.6:** Plot of the MCMC posteriors for the system parameters for the ANN method fit for the SDO/HMI transit of Venus data, where the priors on the stellar parameters were offset from their true Values.

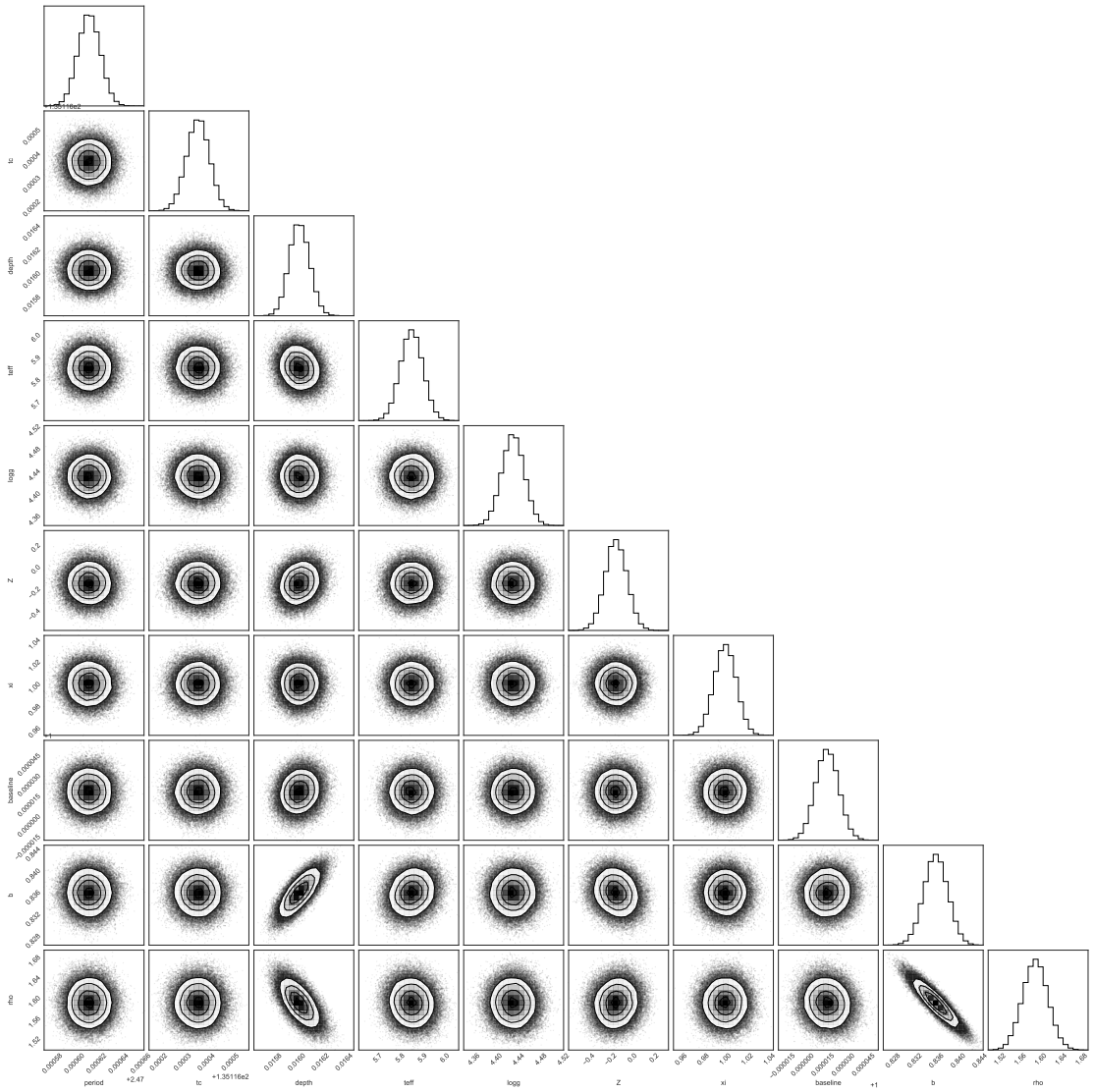



**Figure B.7:** Plot of the MCMC posteriors for the system parameters for the ANN limb darkening fit for the *Kepler* TRES-2b transit data.

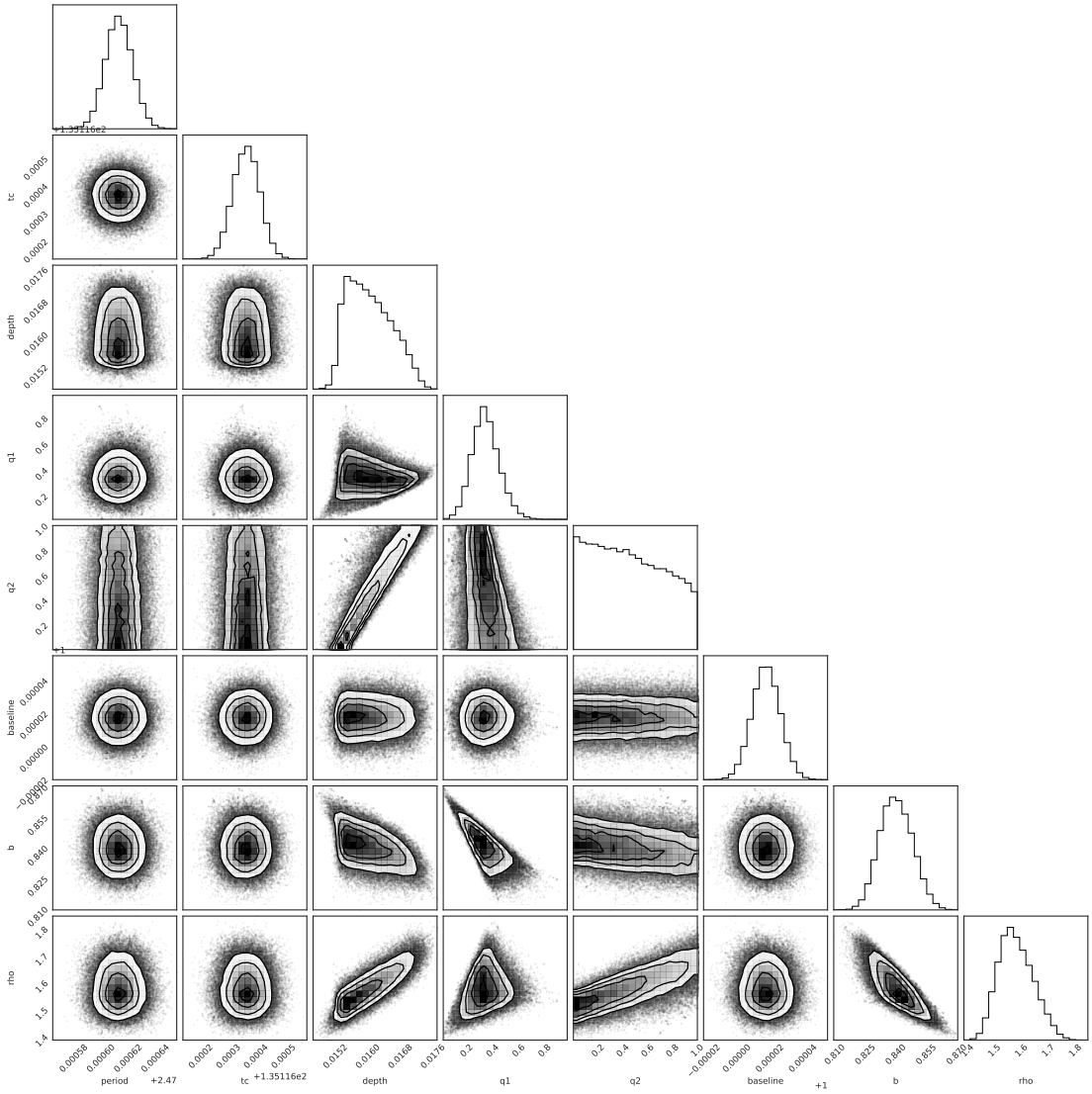



**Figure B.8:** Plot of the MCMC posteriors for the system parameters for the quadratic law limb darkening fit for the *Kepler* TRES-2b transit data.

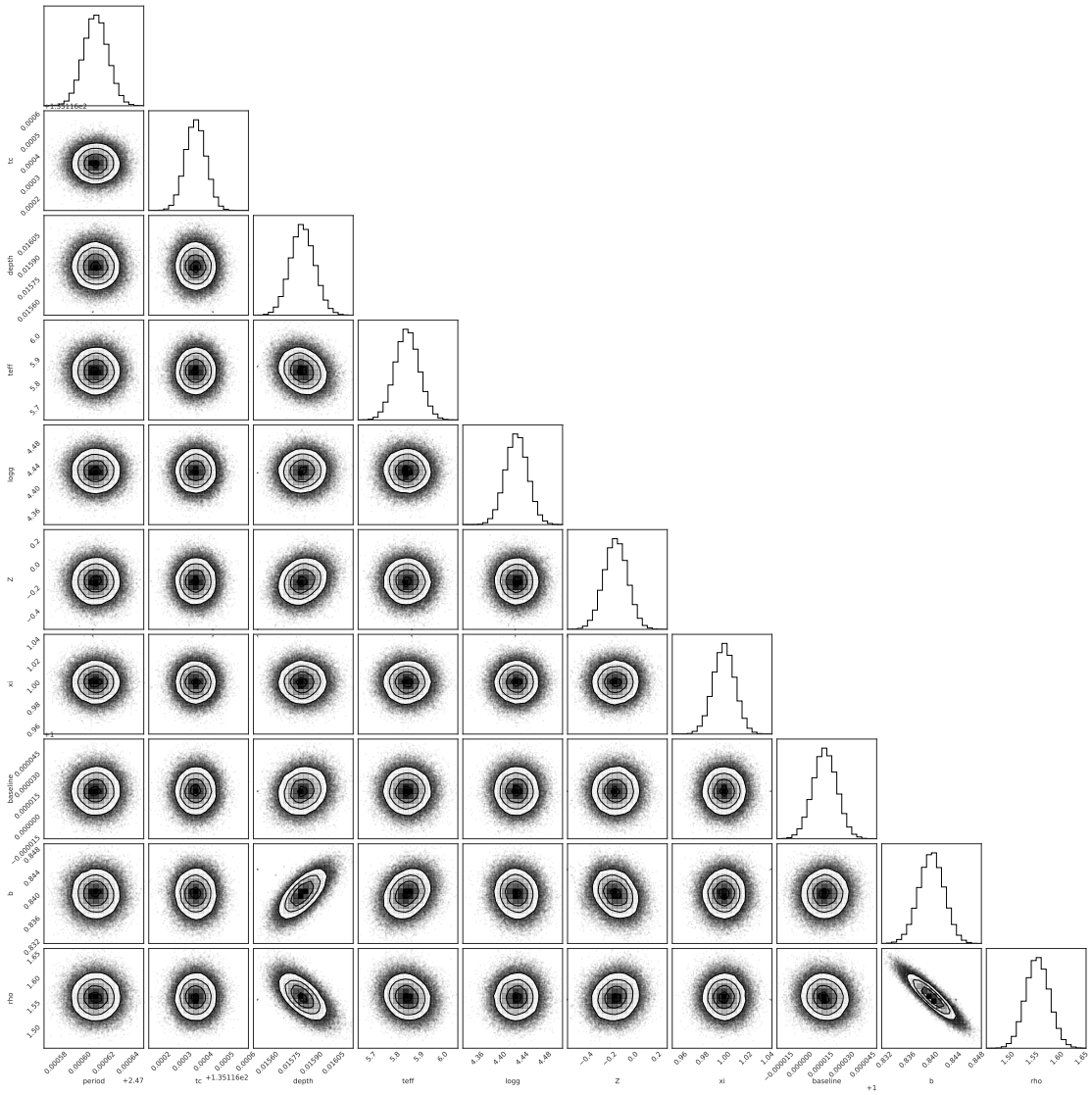



**Figure B.9:** Plot of the MCMC posteriors for the system parameters for the GP limb darkening fit for the *Kepler* TRES-2b transit data.

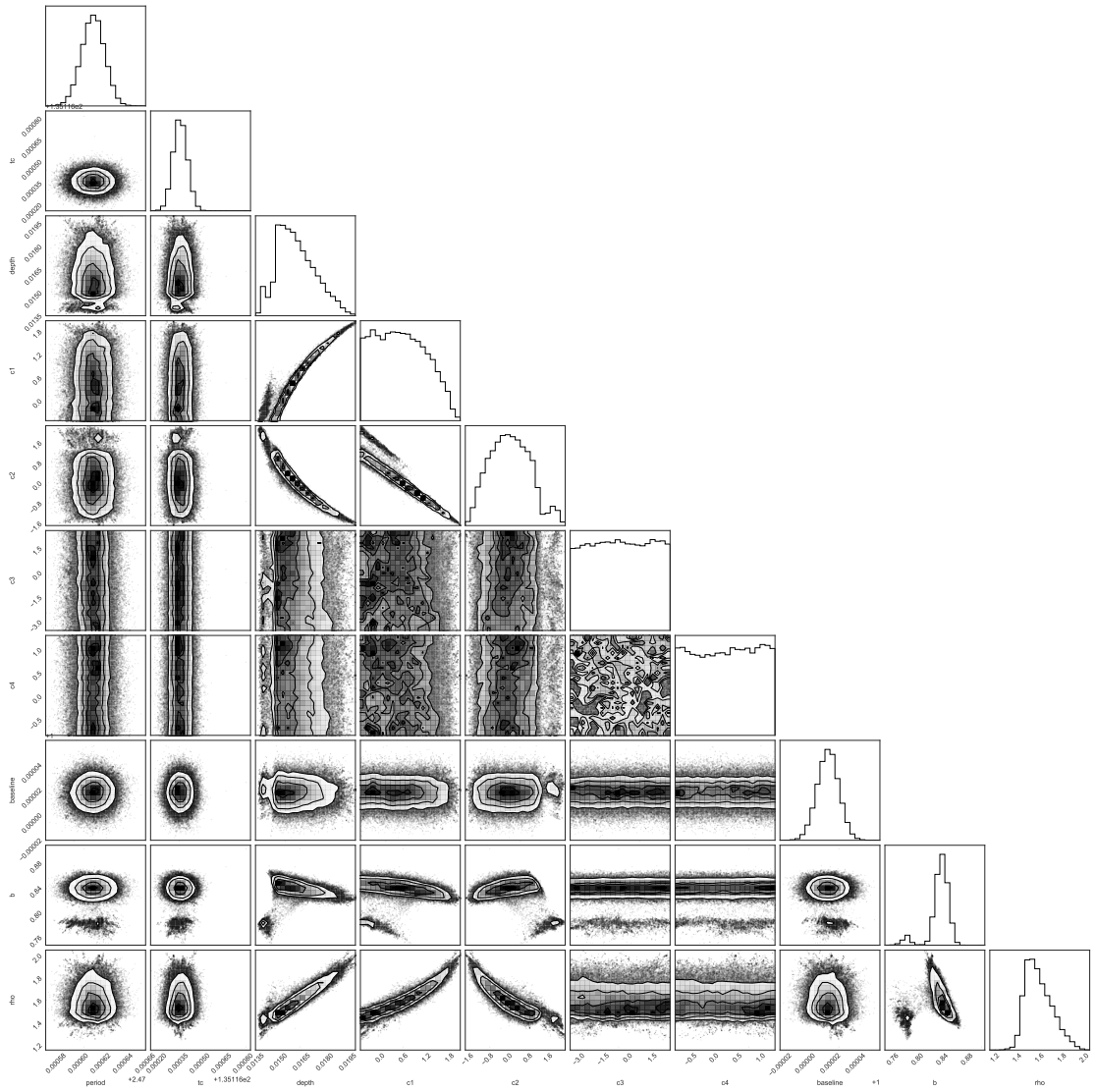



**Figure B.10:** Plot of the MCMC posteriors for the system parameters for the non-linear law limb darkening fit for the *Kepler* TRES-2b transit data.

# Bibliography


Abbot C., Fowle F., Aldrich L., 1913, Annals of the Astrophysical Observatory of the Smithsonian Institution, 3, 153

Abbot C. G., Fowle F. E., Aldrich L. B., 1923, Annals of the Astrophysical Observatory of the Smithsonian Institution, 4, 217

Adams E. R., et al., 2017, AJ, 153, 82

Albrecht S., et al., 2012, ApJ, 757, 18

Allard F., Hauschildt P. H., 1995, ApJ, 445, 433

Alonso R., et al., 2004, ApJ, 613, L153

Anderson D. R., et al., 2011, ApJ, 726, L19

Angus R., Morton T., Aigrain S., Foreman-Mackey D., Rajpaul V., 2018, MNRAS, 474, 2094

Arago F., 1858, Œuvres de François Arago, Mémoires scientifique Tome i. Paris, p. 235

Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481

Bailer-Jones C. A. L., 2000, A&A, 357, 197

Bakos G. Á., Lázár J., Papp I., Sári P., Green E. M., 2002, PASP, 114, 974

Bakos G. Á., et al., 2007, ApJ, 656, 552

Baranne A., et al., 1996, A&AS, 119, 373

Barnes S. A., 2007, ApJ, 669, 1167

Barragán O., et al., 2016, AJ, 152, 193

Bayliss D., et al., 2018, MNRAS, 475, 4467

Becker J. C., Vanderburg A., Adams F. C., Rappaport S. A., Schwengeler H. M., 2015, ApJ, 812, L18

Borucki W. J., et al., 2010, Science, 327, 977

Boss A. P., 1995, Science, 267, 360

Bouchy F., et al., 2009, A&A, 505, 853

Bouguer P., 1760, Traité d'optique sur la gradation de la lumière. H.L. Guerin & L.F. Delatour, Paris, pp 90–96

Brott I., Hauschildt P. H., 2005, in Turon C., O'Flaherty K. S., Perryman M. A. C., eds, ESA Special Publication Vol. 576, The Three-Dimensional Universe with Gaia. p. 565 (arXiv:astro-ph/0503395)

Brown T. M., Charbonneau D., Gilliland R. L., Noyes R. W., Burrows A., 2001, ApJ, 552, 699

Brown D. J. A., Collier Cameron A., Hall C., Hebb L., Smalley B., 2011, MNRAS, 415, 605

Byrd R. H., Lu P., Nocedal J., Zhu C., 1995, SIAM Journal on Scientific Computing, 16, 1190

Cameron A. C., et al., 2006, MNRAS, 373, 799

Cameron A. C., et al., 2007, MNRAS, 380, 1230

Castelli F., Kurucz R. L., 2003, in Piskunov N., Weiss W. W., Gray D. F., eds, IAU Symposium Vol. 210, Modelling of Stellar Atmospheres. p. A20

Charbonneau D., Brown T. M., Latham D. W., Mayor M., 2000, ApJ, 529, L45

Chatterjee S., Ford E. B., Matsumura S., Rasio F. A., 2008, ApJ, 686, 580

Chiavassa A., et al., 2017, A&A, 597, A94

Chollet F., et al., 2015, Keras, https://github.com/fchollet/keras

Christiansen J. L., et al., 2016, ApJ, 828, 99

Claret A., 2000, A&A, 363, 1081

Claret A., Bloemen S., 2011, A&A, 529, A75

Collier Cameron A., et al., 2007, MNRAS, 375, 951

Collier Cameron A., Bruce V. A., Miller G. R. M., Triaud A. H. M. J., Queloz D., 2010, MNRAS, 403, 151

Cosentino R., et al., 2012, in Ground-based and Airborne Instrumentation for Astronomy IV. p. 84461V, doi:10.1117/12.925738

Coughlin J. L., López-Morales M., 2012, AJ, 143, 39

Counselman III C. C., 1973, ApJ, 180, 307

Darwin G. H., 1879, Proceedings of the Royal Society of London, 29, 168

Dawson R. I., Murray-Clay R. A., Johnson J. A., 2015, ApJ, 798, 66

Diaz-Cordoves J., Gimenez A., 1992, A&A, 259, 227

Dittmann J. A., et al., 2017, Nature, 544, 333

Doyle A. P., et al., 2013, MNRAS, 428, 3164

Doyle A. P., Davies G. R., Smalley B., Chaplin W. J., Elsworth Y., 2014, MNRAS, 444, 3592

Enoch B., Collier Cameron A., Parley N. R., Hebb L., 2010, A&A, 516, A33

Espinoza N., Jordán A., 2015, MNRAS, 450, 1879

Espinoza N., Jordán A., 2016, MNRAS, 457, 3573

Esteves L. J., De Mooij E. J. W., Jayawardhana R., 2015, ApJ, 804, 150

Evans D. S., 1947, The Observatory, 67, 138

Ford E. B., Rasio F. A., 2006, ApJ, 638, L45

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Foreman-Mackey D., Morton T. D., Hogg D. W., Agol E., Schölkopf B., 2016, AJ, 152, 206

Gaia Collaboration et al., 2016, A&A, 595, A1

Gibson N. P., et al., 2008, A&A, 492, 603

Goldreich P., Tremaine S., 1980, ApJ, 241, 425

Gray D. F., 2008, The Observation and Analysis of Stellar Photospheres. Cambridge University Press, 2008

Grygar J., 1965, Bulletin of the Astronomical Institutes of Czechoslovakia, 16, 195

Hanbury Brown R., Davis J., Lake R. J. W., Thompson R. J., 1974, MNRAS, 167, 475

Hastings W. K., 1970, Biometrika, 57, 97

Hauschildt P. H., Baron E., 1999, Journal of Computational and Applied Mathematics, 109, 41

Hauschildt P. H., Allard F., Ferguson J., Baron E., Alexander D. R., 1999, ApJ, 525, 871

Hay K. L., et al., 2016, MNRAS, 463, 3276

Hayek W., Sing D., Pont F., Asplund M., 2012, A&A, 539, A102

Haywood R. D., et al., 2014, MNRAS, 443, 2517

Hestroffer D., 1997, A&A, 327, 199

Heyrovský D., 2007, ApJ, 656, 483

Holmes S., et al., 2011, PASP, 123, 1177

Hon M., Stello D., Yu J., 2017, MNRAS, 469, 4578

Hopfield J. J., 1982, Proceedings of the National Academy of Sciences, 79, 2554

Howard A. W., et al., 2012, ApJS, 201, 15

Howarth I. D., 2011, MNRAS, 418, 1165

Hubel D. H., Wiesel T. N., 1959, The Journal of physiology, 148, 574

Hubenet H., de Jager C., 1956, Bull. Astron. Inst. Netherlands, 13, 43

Husser T.-O., Wende-von Berg S., Dreizler S., Homeier D., Reiners A., Barman T., Hauschildt P. H., 2013, A&A, 553, A6

Hut P., 1980, A&A, 92, 167

Jehin E., et al., 2011, The Messenger, 145, 2

Julius W. H., 1906, ApJ, 23, 312

Julius W. H., 1913, ApJ, 37, 225

Kass R. E., Raftery A. E., 1995, Journal of the American Statistical Association, 90, 773

Kervella P., Bigot L., Gallenne A., Thévenin F., 2017, A&A, 597, A137

Kipping D. M., 2013, MNRAS, 435, 2152

Kipping D. M., 2016, MNRAS, 455, 1680

Kipping D., Bakos G., 2011, ApJ, 733, 36

Kipping D. M., Lam C., 2017, MNRAS, 465, 3495

Kipping D. M., Tinetti G., 2010, MNRAS, 407, 2589

Kirby E. N., 2011, PASP, 123, 531

Klinglesmith D. A., Sobieski S., 1970, AJ, 75, 175

Knutson H. A., et al., 2009, ApJ, 690, 822

Knutson H. A., et al., 2014, ApJ, 794, 155

Kolb U., 2014, in Revista Mexicana de Astronomia y Astrofisica Conference Series. pp 16–19

Konacki M., Torres G., Jha S., Sasselov D. D., 2003, Nature, 421, 507

Kopal Z., 1946, ApJ, 104, 60

Kopal Z., 1950, Harvard College Observatory Circular, 454, 1

Kovács G., Zucker S., Mazeh T., 2002, A&A, 391, 369

Kreidberg L., 2015, PASP, 127, 1161

Kron G. E., 1939, Lick Observatory Bulletin, 19, 59

Kurucz R. L., 1970, SAO Special Report, 309

Kurucz R. L., 1979, ApJS, 40, 1

Langley S. P., 1875a, American Journal of Science and Arts, 10, 489

Langley S. P., 1875b, in , Vol. 80, Comptes rendus hebdomadaires des séances de l'Académie des sciences. Gauthier-Villars, Paris

LeCun Y., Bengio Y., Hinton G., 2015, Nature, 521, 436

Lendl M., et al., 2012, A&A, 544, A72

Lester J. B., Neilson H. R., 2008, A&A, 491, 633

Levrard B., Winisdoerffer C., Chabrier G., 2009, ApJ, 692, L9

Liais E., 1866, in , Vol. XII, Mémoires de la Société impériale des sciences naturelles de Cherbourg. J.B. Baillière et fils, Paris, pp 277–342

Mädlow M., 1962, Astronomische Nachrichten, 286, 271

Magain P., 1984, A&A, 134, 189

Magic Z., Collet R., Asplund M., Trampedach R., Hayek W., Chiavassa A., Stein R. F., Nordlund Å., 2013, A&A, 557, A26

Magic Z., Chiavassa A., Collet R., Asplund M., 2015, A&A, 573, A90

Mandel K., Agol E., 2002, ApJ, 580, L171

Marsh T. R., 1989, PASP, 101, 1032

Matsumura S., Peale S. J., Rasio F. A., 2010, ApJ, 725, 1995

Maxted P. F. L., et al., 2011, PASP, 123, 547

Maxted P. F. L., Serenelli A. M., Southworth J., 2015, A&A, 575, A36

Mayor M., Queloz D., 1995, Nature, 378, 355

McCormac J., Skillen I., Pollacco D., Faedi F., Ramsay G., Dhillon V. S., Todd I., Gonzalez A., 2014, MNRAS, 438, 3383

McCulloch W. S., Pitts W., 1943, The Bulletin of Mathematical Biophysics, 5, 115

McLaughlin D. B., 1924, ApJ, 60

Mészáros S., et al., 2012, AJ, 144, 120

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

Milne E. A., 1921, MNRAS, 81, 361

Mizuno H., 1980, Progress of Theoretical Physics, 64, 544

Morello G., Tsiaras A., Howarth I. D., Homeier D., 2017, AJ, 154, 111

Mortier A., Sousa S. G., Adibekyan V. Z., Brandão I. M., Santos N. C., 2014, A&A, 572, A95

Mottram C. J., Steele I. A., Morales L., 2004, in Moorwood A. F. M., Iye M., eds, Proc. SPIEVol. 5492, Ground-based Instrumentation for Astronomy. pp 677–688, doi:10.1117/12.551337

Müller H. M., Huber K. F., Czesla S., Wolter U., Schmitt J. H. M. M., 2013, A&A, 560, A112

Munari U., Zwitter T., 1997, A&A, 318, 269

Nair V., Hinton G. E., 2010, in Proceedings of the 27th international conference on machine learning (ICML-10). pp 807–814

Neckel H., Labs D., 1994, Sol. Phys., 153, 91

Neilson H. R., Lester J. B., 2013, A&A, 556, A86

Neilson H. R., McNeil J. T., Ignace R., Lester J. B., 2017, ApJ, 845, 65

Nordlund A., 1982, A&A, 107, 1

Nordlund Å., Stein R. F., Asplund M., 2009, Living Reviews in Solar Physics, 6, 2

O'Donovan F. T., et al., 2006, ApJ, 651, L61

Odewahn S. C., Stockwell E. B., Pennington R. L., Humphreys R. M., Zumach W. A., 1992, AJ, 103, 318

Ohta Y., Taruya A., Suto Y., 2005, ApJ, 622, 1118

Pál A., 2008, MNRAS, 390, 281

Parviainen H., 2015, PyLDTk: Python toolkit for calculating stellar limb darkening profiles and model-specific coefficients for arbitrary filters, Astrophysics Source Code Library (ascl:1510.003)

Parviainen H., Aigrain S., 2015, MNRAS, 453, 3821

Parviainen H., Pallé E., Nortmann L., Nowak G., Iro N., Murgas F., Aigrain S., 2016, A&A, 585, A114

Parviainen H., et al., 2018, A&A, 609, A33

Pedregosa F., et al., 2011, Journal of Machine Learning Research, 12, 2825

Penev K., Sasselov D., 2011, ApJ, 731, 67

Pepe F., Mayor M., Galland F., Naef D., Queloz D., Santos N. C., Udry S., Burnet M., 2002, A&A, 388, 632

Pepper J., et al., 2007, PASP, 119, 923

Perruchot S., et al., 2008, in Ground-based and Airborne Instrumentation for Astronomy II. p. 70140J, doi:10.1117/12.787379

Perryman M., 2011, The Exoplanet Handbook

Pollacco D. L., et al., 2006, PASP, 118, 1407

Pollacco D., et al., 2008, MNRAS, 385, 1576

Pont F., Zucker S., Queloz D., 2006, MNRAS, 373, 231

Queloz D., et al., 2000a, A&A, 354, 99

Queloz D., Eggenberger A., Mayor M., Perrier C., Beuzit J. L., Naef D., Sivan J. P., Udry S., 2000b, A&A, 359, L13

Raetz S., et al., 2014, MNRAS, 444, 1351

Rasio F. A., Ford E. B., 1996, Science, 274, 954

Rasmussen C., Williams C., 2006, Gaussian Processes for Machine Learning. MIT Press

Ricker G. R., et al., 2014, in Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave. p. 914320 (arXiv:1406.0151), doi:10.1117/12.2063489

Rosenblatt F., 1958, Psychological review, 65, 386

Rossiter R. A., 1924, ApJ, 60

Rumelhart D. E., Hinton G. E., Williams R. J., 1986, Nature, 323, 533

Sandford E., Kipping D., 2017, AJ, 154, 228

Santerne A., et al., 2015, MNRAS, 451, 2337

Schaefer C., Geiger M., Kuntzer T., Kneib J.-P., 2018, A&A, 611, A2

Schröter S., Schmitt J. H. M. M., Müller H. M., 2012, A&A, 539, A97

Schuck P. W., Antiochos S. K., Leka K. D., Barnes G., 2016, ApJ, 823, 101

Schwarz G., et al., 1978, The annals of statistics, 6, 461

Schwarzschild K., 1906, Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen. Math.-phys. Klasse, 195, 41

Secchi A., 1870, Le Soleil, 1 edn. Gauthier-Villars, Paris

Secchi A., 1875, Le Soleil, 2 edn. Gauthier-Villars, Paris

Sestito P., Randich S., 2005, A&A, 442, 615

Shallue C. J., Vanderburg A., 2018, AJ, 155, 94

Sing D. K., 2010, A&A, 510, A21

Sing D. K., Désert J.-M., Lecavelier Des Etangs A., Ballester G. E., Vidal-Madjar A., Parmentier V., Hebrard G., Henry G. W., 2009, A&A, 505, 891

Sing D. K., et al., 2011, MNRAS, 416, 1443

Skrutskie M. F., et al., 2006, AJ, 131, 1163

Southworth J., 2008, MNRAS, 386, 1644

Southworth J., 2011, MNRAS, 417, 2166

Southworth J., 2015, JKTLD: Limb darkening coefficients, Astrophysics Source Code Library (ascl:1511.016)

Steele I. A., et al., 2004, in Oschmann Jr. J. M., ed., Proc. SPIEVol. 5489, Ground-based Telescopes. pp 679–692, doi:10.1117/12.551456

Steele I. A., Bates S. D., Gibson N., Keenan F., Meaburn J., Mottram C. J., Pollacco D., Todd I., 2008, in Ground-based and Airborne Instrumentation for Astronomy II. p. 70146J (arXiv:0809.3351), doi:10.1117/12.787889

Stetson P. B., 1987, PASP, 99, 191

Storrie-Lombardi M. C., Lahav O., Sodre Jr. L., Storrie-Lombardi L. J., 1992, MNRAS, 259, 8P

Theano Development Team 2016, arXiv e-prints, abs/1605.02688

Torres G., Andersen J., Giménez A., 2010, A&A Rev., 18, 67

Udalski A., Zebrun K., Szymanski M., Kubiak M., Soszynski I., Szewczyk O., Wyrzykowski L., Pietrzynski G., 2002, Acta Astron., 52, 115

Vogel H. C., 1872, in , Vol. 24, Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig. Mathematisch-Physische Classe. Weidmannsche Buchhandlung, Leipzig, pp 135–141

Vogel H. C., 1877, in , Vol. Jahre 1877 (1878), Monatsberichte der Königlichen Preussische Akademie des Wissenschaften zu Berlin.. Königliche Akademie der Wissenschaften, Berlin, pp 104–142

Wakeford H. R., et al., 2017, Science, 356, 628

Wang J., Fischer D. A., Horch E. P., Huang X., 2015, ApJ, 799, 229

Ward W. R., 1997, Icarus, 126, 261

Weiss A., Schlattl H., 2008, Ap&SS, 316, 99

Wesselink A. J., 1940, Bull. Astron. Inst. Netherlands, 9, 81

Wheatley P. J., et al., 2018, MNRAS, 475, 4476

Wiener N., Hopf E., 1931, Proc. Prussian Acad. Math.-Phys. Ser., page 696pp

Wilkins A. N., Delrez L., Barker A. J., Deming D., Hamilton D., Gillon M., Jehin E., 2017, ApJ, 836, L24

Winn J. N., 2009, in Pont F., Sasselov D., Holman M. J., eds, IAU Symposium Vol. 253, IAU Symposium. pp 99–109

Winn J. N., Fabrycky D., Albrecht S., Johnson J. A., 2010, ApJ, 718, L145

Wittkowski M., Aufdenberg J. P., Kervella P., 2004, A&A, 413, 711

Wright J. T., Marcy G. W., Howard A. W., Johnson J. A., Morton T. D., Fischer D. A., 2012, ApJ, 753, 160

Wu Y., Murray N., 2003, ApJ, 589, 605

Zhu C., Byrd R. H., Lu P., Nocedal J., 1997, ACM Transactions on Mathematical Software (TOMS), 23, 550

marquis de Laplace P. S., 1805, Traité de mécanique céleste. Vol. 4, Crapelet