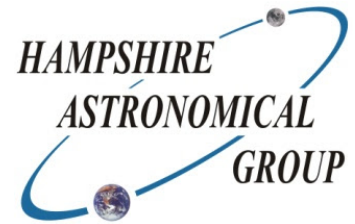




UNIVERSITY OF
PORTSMOUTH



School of Mathematics and Physics

Title: Determining The Parameters of Exoplanetary Candidates
From Transit Timing Variations

Name: Jack Lloyd-Walters FRAS

Course: MPhys (Hons) Physics, Astronomy, and Cosmology

Student No: UP881788

Year: 2022

Word Count: 6531 words

School of Mathematics and Physics

Disclaimer

The enclosed project is entirely the work of the undergraduate student, and any opinions expressed are not necessarily those of any member of the staff of the School of Mathematics and Physics, University of Portsmouth. The text does not show any corrections of fact or calculations, and significant errors and omissions are possible. Any external reader or user of this dissertation does so entirely at their own risk and responsibility, and neither the School of Mathematics and Physics nor the University of Portsmouth can be held responsible for anything contained within this student project report.

Determining The Parameters of Exoplanetary Candidates From Transit Timing Variations

Jack Lloyd-Walters FRAS^{1,2}, Supervised by Stephen Fitcher FCA FRAS^{1,3}, Dr. Hooshyar Assadullahi², and Prof. Daniel Thomas²

¹Royal Astronomical Society, Burlington House, Piccadilly, London W1J 0BQ, UK

²School of Mathematics and Physics, Portsmouth University, Lion Terrace, Portsmouth PO1 3HF, UK

³Hampshire Astronomical Group, Portsmouth Reservoir, Hinton Manor Ln, Clanfield PO8 0QP, UK

Master's Thesis, May 2022

ABSTRACT

Transit Timing Variation (TTV) provides a powerful tool to probe the dynamical configuration of exoplanetary systems from historical transit data (Holman & Murray 2005; Agol et al. 2004). TTV analysis has allowed both verification of planetary parameters (Wang et al. 2017) and the discovery of new planetary bodies (Ballard et al. 2011) from transit observation alone. As part of this work, additional transit light curves have been collected with the 24" Ritchey-Chrétien telescope at Clanfield observatory and combined with the ExoClock database (Kokori et al. 2021, 2022), Exoplanet transit database (Poddaný et al. 2010), and TESS light curves (Ricker 2014) to create a set of historical TTV data for analysis. A set of extensible TTV models have been developed to analytically approximate the chaotic n-body nature of real planetary systems. A computational pipeline to automate model fitting using various parameter optimisation (Storn & Price 1995; Xiang et al. 1997) and model comparison (Akaike 1992, 1974; Hurvich & Tsai 1993; Schwarz 1978) techniques has been developed in-situ, allowing verification of model validity and analysis of TTV candidates using a combination of simulation and historical TTV data. The models developed were found to accurately describe TTV, and could determine the initial system parameters of simulated TTV systems to reasonable accuracy. Future work will allow extensions to these models, providing a more powerful suite of analytical tools for exoplanetary science, and the application of these methods to real exoplanetary systems with the possibility of new planetary discoveries.

Key words: astronomical databases: miscellaneous – software: development – software: simulations – techniques: photometric

CONTENTS

1	Introduction	3.3	Simulation pipeline
1.1	Transit photometry	3.3.1	Initialisation
1.2	Transit timing	3.3.2	Simulated TTV
1.2.1	Barycentre motion	3.4	Analytical Models
1.2.2	Orbital perturbation	3.4.1	System and TTV
1.2.3	mean motion libration	3.4.2	Model comparison
1.3	Project objectives	3.4.3	Parameter optimisation
2	Methodology	3.4.4	Model selection
2.1	TTV Models	3.4.5	Most likely system
2.1.1	Interior perturbation	3.5	Discussion
2.1.2	Exterior perturbation	3.5.1	Evaluation
2.2	Parameter search	3.5.2	Limitations
2.2.1	Optimisation	3.5.3	Errors
2.2.2	Model comparison	4	Conclusion
2.2.3	Markov Chain, Monte Carlo		Acknowledgements
3	Results / Discussion	A	List of Figures
3.1	Observations	B	List of Tables
3.2	Light curve analysis	C	Extra material
3.2.1	De-trending	C1	Transit duration derivation
3.2.2	Transit fitting		
3.2.3	TTV Residuals		

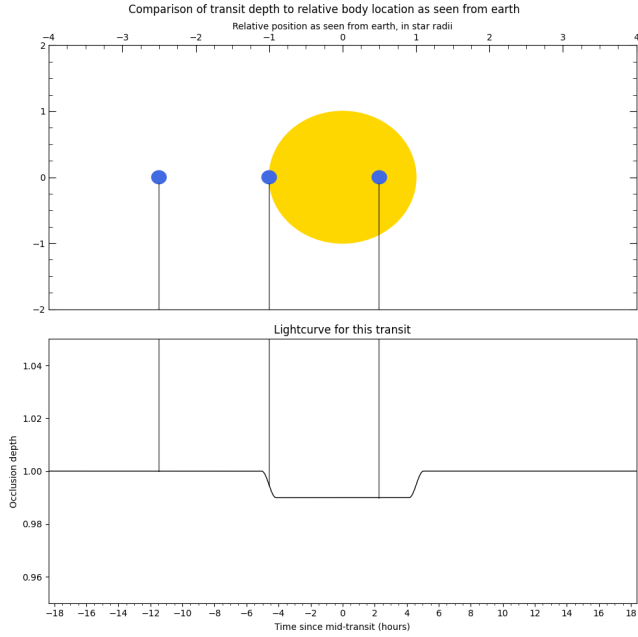


Figure 1. How the location of an exoplanet corresponds to the observed light curve. Stellar limb darkening was ignored for simplicity.

1 INTRODUCTION

In the last quarter-century, exoplanet detections have seen astronomical success, in no small part due to the launch of several space borne telescopes. Of the 5000 planets found within 3800 planetary systems, around three quarters have been discovered through transit photometry (Akeson et al. 2013). To this day, over 130 million light curves have been observed, providing a large base for historical analysis.

Of particular interest for this project is a specific subset of transit photometry that makes heavy use of this historical data. “Transit Timing Variation”, or “TTV”, reanalyses historical transit observations for deviation from 2-body Keplerian motion to provide an insight into the dynamics of an exoplanetary system.

1.1 Transit photometry

In the case that an exoplanetary system is oriented ‘edge on’ from our vantage point on earth, then planets within the system will periodically occlude the central star. By observing the occlusion, it is possible to determine many of the physical and orbital parameters of an exoplanet.

$$\Delta_L = \frac{\theta_{planet}}{\theta_{star}} = \frac{R_{planet}^2}{R_{star}^2} \approx \frac{r_{planet}^2}{r_{star}^2} \quad (1)$$

The reduction in stellar light observed during each transit, “Occlusion depth”, is proportional to the angular areas of both the planet and star. An expression for this is given in equation 1, where θ is the angular area, and R the angular radius of each object. In the small angle approximation, the ratio of R becomes equal to the ratio of r , the true radii of each object.

Under the approximation that an exoplanetary transit is small compared to the size of its orbit, the velocity of the exoplanet, v_{planet} ,

will remain constant. By approximating the transit as a line segment, the time duration of this transit will be proportional to the combined diameters of the body and star, as demonstrated by figure 1, and expressed in equation 2.

$$T_{transit} = \frac{2(R_{star} + R_{planet})}{v_{planet}} \quad (2)$$

By instead considering the transit as a circular arc (and assuming the orbital eccentricity is low), the expression for transit duration is given in equation 3, where a is the semi-major axis of the orbit.

$$T_{transit} = \frac{2a}{v_{planet}} \arcsin \frac{R_{star} + R_{planet}}{a} \quad (3)$$

Introducing the effect of orbital inclination further complicates this approximation, as the exoplanet will not transit across the equator of the stellar disc, as demonstrated in figure 2. This introduces the quantity known as the “Impact parameter”, and is given in equation 4, where i is the orbital inclination and Ω the longitude of the ascending node.

$$b = a \cos i \sin \Omega \quad (4)$$

$$T_{transit} = \frac{2a}{v_{planet}} \arcsin \frac{\sqrt{(R_{star} + R_{planet})^2 - b^2}}{a} \quad (5)$$

Introducing the impact parameter, b , to equation 3 gives the expression given in equation 5. This allows information about both the inclination and radius of the orbit to be deduced from a single transit. A full derivation of equations 2, 3, and 5 is given in section C1.

1.2 Transit timing

TTV, by definition, is any departure from the predicted transit timing for an exoplanet as given by its linear ephemerides. While this can be due to a multitude of reasons, of primary interest for this paper are those transit timing variations caused by additional planets in the system (Holman & Murray 2005; Agol et al. 2004).

In the case where a system contains only a central star and orbiting planet, the equations of motion for the system are given by the closed-form Keplerian equations. Introducing an additional body to the system, however, causes the equations of motion to become chaotic, with no closed-form solution possible. This departure from Keplerian motion due to an additional gravitating body results in any transiting planets exhibiting TTV, whose parameters are related in some way to the orbital and physical properties of the perturbing body.

In the case where gravitational interactions between orbiting bodies are small compared to the gravity of the central star, the motion of each body can be described as some small perturbation atop the closed-form Keplerian equations. Measuring TTV allows the dynamical configuration of the system to be probed, as these are entirely caused by gravitational perturbation. This was demonstrated for the TRAPPIST-1 system, where the masses and orbital configurations of the planets were deduced from transit timing variations (Wang et al. 2017), and for the Kepler-19 system, where an additional planet was

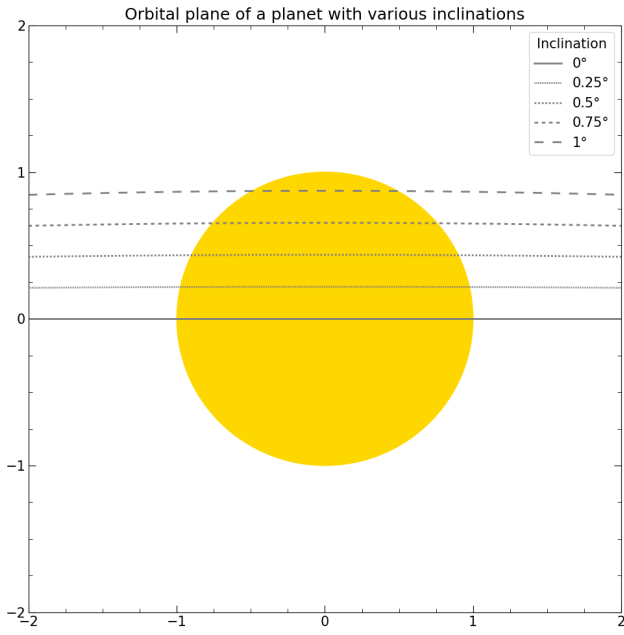


Figure 2. How the orbital inclination affects the length of a transit. The axis scale is in stellar radii, and it can be seen in this configuration that a transit would only be visible if the relative inclination were in the region of ± 1 degree.

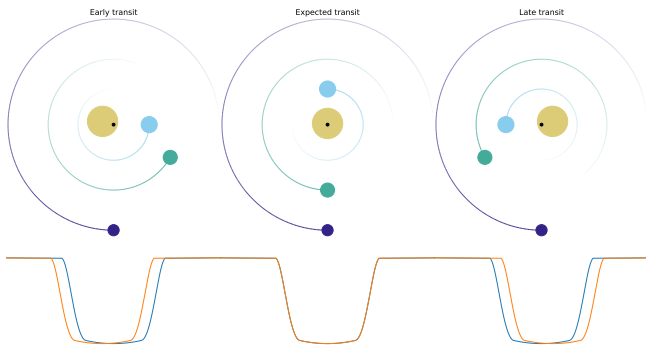


Figure 3. Transit timing variations due to interior planetary motion shifting the barycentre. From left to right are the early-time, expected time, and late-time transits. Note how, as the graphics are co-rotating with the transiting planet-barycentre reference frame, it is the apparent motion of the star that causes TTV.

discovered through TTV analysis (Ballard et al. 2011), to name but two examples.

Of particular interest for this project are the three following causes for transit timing variations: Barycentre motion; Orbital perturbation; and mean motion libration. While other causes are evident, these are a useful starting ground for intuition.

1.2.1 Barycentre motion

As all bodies in an exoplanetary system orbit about the common barycentre, rather than the centre of the parent star, the position of

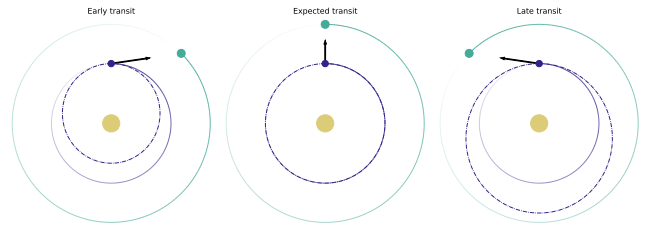


Figure 4. Transit timing variations due to exterior planet(s) perturbing the orbit of the transiting planet. The solid line is the unperturbed orbit, the dashed is the instantaneous perturbed orbit, and the arrow marks the instantaneous perturbing force vector. From left to right are the early-time, expected time, and late-time transits.

the star will appear to shift over time. This motion underpins both Doppler spectroscopic and astrometric methods for exoplanetary detection, and also results in variations in transit times.

As a transit occurs when the star and planet are aligned, the motion of the star relative to the barycentre will affect the timing of a transit. As the star moves, the planet must also move along its orbit for a transit to occur. The difference in position of the exoplanet as compared to if the star had not moved causes transits to occur earlier or later than predicted. This variation is a function of the position of the barycentre, as shown in figure 3, and is especially sensitive to massive planets that orbit interior to the transiting planet.

1.2.2 Orbital perturbation

As discussed, the equations of motion have no closed-form solution for systems with three or more bodies, but can be expressible as a set of perturbed 2-body solutions, as seen in figure 4.

As perturbing planets orbit about the system, their gravitational interactions pull the transiting planet further ahead or behind its unperturbed position. This primarily affects the semi-major axis of the transiting planet, and thus its period, causing the transit to appear earlier or later as the planet is in a lower or higher orbit than predicted from simple Keplerian motion. This motion is especially sensitive to massive perturbing planets exterior to the transiting planet on elliptical orbits.

1.2.3 mean motion libration

In the case where perturbing planets are in or near mean motion resonance with the transiting planet, the effects of orbital perturbation will be greatly exaggerated. As the planets are in mean motion resonance, where their orbital periods are expressed as some small integer ratio, they will only experience conjunctions at discrete locations along their orbits. This causes orbital perturbations to accumulate, rather than average out over many orbits.

As the planets orbit about the system, perturbation causes their respective periods to drift, as seen in figure 4. This in turn causes the conjunction positions to drift along the orbit, causing the perturbations to act opposite once the conjunction locations have rotated 180° relative to the initial position. This causes cyclic changes in orbital elements with larger magnitudes over longer timescales than any of the previously mentioned causes. These changes are especially

sensitive to perturbing planets of much lower mass than seen in other TTV causes, particularly planets in first order resonance with the transiting planet, such as is seen in the kepler-19 system (Ballard et al. 2011).

1.3 Project objectives

As the detection and analysis of transit timing variations rely on historical data, additional observations have been made throughout this project with the 24" Ritchey-Chrétien telescope at Clanfield observatory. These observations were analysed with Holomon Photometric Software (HOPS) (Tsiaras 2021), and uploaded to the ExoClock database, where other transit observations are combined and verified (Kokori et al. 2021, 2022).

Transit data from both ExoClock and the exoplanet transit database (ETD) (Poddaný et al. 2010) have been combined with light curves from the *TESS* spacecraft (Ricker 2014). These light curves have been analysed with the Juliet python package to obtain mid-transit times (Espinoza et al. 2019). A set of linear ephemerides were fit with linear regression to the observed transit data and compared with those published in both the ExoClock database and the NASA Exoplanet Archive (Akeson et al. 2013). Computing predicted transit times from these linear ephemerides have allowed a set of TTV to be computed for each planetary target.

A set of analytical TTV Models have been developed to fit to this TTV data, providing a method for determining the configuration of each exoplanetary system from transit observation. These models have had parameters determined through various minimisation techniques using the SciPy package (Virtanen et al. 2020), such as least squares regression (Branch et al. 1999), bounded limited memory Broyden–Fletcher–Goldfarb–Shanno (Byrd et al. 1995; Zhu et al. 1997), differential evolution (Storn & Price 1995), and dual annealing (Xiang et al. 1997). The parameter distribution was determined with Markov Chain Monte Carlo (MCMC) analysis using the emcee package (Foreman-Mackey et al. 2013). This was used to provide uncertainty bounds on reported values by taking the 16th, 50th, and 84th quantiles of the MCMC samples.

2 METHODOLOGY

2.1 TTV Models

As a first step in transit timing variation analysis, a model, or set of models, is required against which data can be fitted. As these models will be used for any given simulation or real-world system setup, they must have the following properties:

- Computationally inexpensive, models will need to be executed thousands of times for parameter fitting and should be vectorisable to take advantage of optimisation.
- Extensible, TTV models should be formulated with standard equations of motion, such that they can be further re-written to take advantage of more complex processes.
- System independent, many TTV models in the literature (Agol et al. 2004; Agol & Deck 2016; Agol & Fabrycky 2018) are presented in only the three body case. These models should be provided in a form that allows arbitrarily many planets to be accounted for.

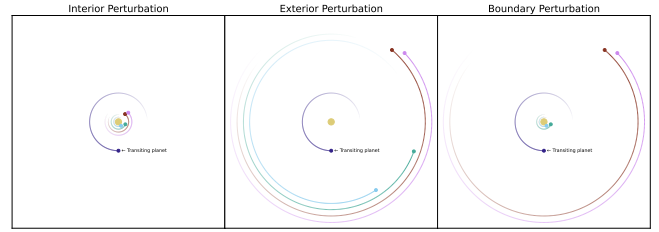


Figure 5. Three classes of TTV cause, from left to right, are interior perturbation, exterior perturbation, and combined (or boundary) perturbation. The classes each represent a subtly different way of describing TTV, and the form of each model reflects the different approximations and approaches made.

Strictly speaking, there are three categories into which TTV models can be subdivided. Those are a transiting planet exterior to perturbing planets, a transiting planet interior to perturbing planets, and a transiting planet both interior and exterior to perturbing planets. These three model classes, demonstrated visually in figure 5, represent an increasing level of complexity for derivation. Of those, only the interior and exterior perturbations require derivation, as the boundary model is formed from a combination of the two.

2.1.1 Interior perturbation

In the case where a transiting exoplanet orbits exterior to the perturbing planets, the transit timing variations seen will be dominated by the motion of the barycentre, as mentioned in section 1.2.1. This setup is the most intuitive to understand, and makes a useful starting point.

2.1.1.1 Initial derivation The distance between the barycentre and primary body is given in equation 6, where r is the distance between the primary and secondary, m_0 and m_1 are the masses of the primary and secondary respectively, and μ is the reduced mass of the secondary.

$$r_b = r \frac{m_1}{m_1 + m_0} = r\mu \quad (6)$$

If we take the assumption that gravitational influences are dominated by that of the central star, any effects between planets will be negligible. Thus, the position of the global barycentre can be described as a linear combination of each planet in the system, as given in equation 7.

$$r_b = \sum_i^n r_i \mu_i \quad (7)$$

The position of an exoplanet along its orbit is given in equation 8, where P_i is the orbital period of the planet, $t_{0,i}$ is the initial transit time of the planet, and ϕ is the angle between the planet, and the closest point of its orbit as seen from earth.

$$\phi = \frac{2\pi(t - t_{0,i})}{P_i} \quad (8)$$

This expression of the angle of exoplanet position is equivalent to the mean anomaly of the planet, and can be combined with equation

7 to obtain an expression for the position of the barycentre of the system at any given time, t , as given in equation 9. The introduction of the sine term is due to the nature of TTV. Only the motion of the barycentre perpendicular to the vantage point of the earth is relevant.

$$r_b = \sum_i^n \left[r_i \mu_i \sin \frac{2\pi(t - t_{0,i})}{P_i} \right] \quad (9)$$

As this is the distance of the barycentre from the star, we can negate the expression to obtain the distance of the star from the barycentre. This also defines the additional distance the transiting planet must cover for the transit to occur.

$$\delta_x = - \sum_i^n \left[r_i \mu_i \sin \frac{2\pi(t - t_{0,i})}{P_i} \right] \quad (10)$$

If we assume the distance associated with a transit is small compared to the size of the orbit of the transiting planet, the velocity of all bodies in the system will remain constant during the transit. Thus, to convert the distance, δ_x , to a transit timing variation, we can apply standard equations of motion, as given in equation 11. As the stellar mass will be many orders of magnitude larger than the planet, it's velocity contribution to the TTV will be negligible.

$$t = \frac{\delta_x}{v_{planet} - v_{star}} \approx \frac{\delta_x}{v_{planet}} \quad (11)$$

$$v_{planet} = \sqrt{\frac{G(m_{star} + m_{planet})}{a_{planet}}} = \frac{2\pi a}{P_{planet}} \quad (12)$$

If we assume the orbit of the transiting planet to be circular, the orbital velocity is that given in equation 12. This can be combined with equation 11 and equation 10 to provide an expression for the transit timing variation in this case, as given in equation 13.

$$\delta_T = - \frac{P_T}{2\pi a_T} \sum_i^n \left[a_i \mu_i \sin \frac{2\pi(t - t_{0,i})}{P_i} \right] \quad (13)$$

2.1.1.2 Extension to eccentric orbits While many of the approximations and assumptions made in the derivation of equation 13 are valid for real planetary systems, the exclusion of orbital eccentricity will cause measurable deviation.

As the orbits are no longer circular, we cannot use a and r interchangeably, instead, we must use the planets' true anomaly, f , and eccentricity, e , as given in equation 14. Additionally, as orbital velocity varies over an elliptical orbit, we need to use the vis-viva equation as given in equation 15.

$$r = \frac{a(1 - e^2)}{1 + e \cos f} \quad (14)$$

$$v = \sqrt{G(m_{star} + m_{planet}) \left(\frac{2}{r} - \frac{1}{a} \right)} \quad (15)$$

Introducing equation 14 to equation 13 gives the following expression for TTV, given in equation 16. As the orbital distance equation

requires introduction of the true anomaly, the term, $\sin \frac{2\pi(t - t_{0,i})}{P_i}$, has also been replaced with the equivalent but more accurate expression using true anomaly and argument of periapsis, $\sin(f_i + \omega_i)$.

$$\delta_T = - \frac{P_T}{2\pi a_T} \sum_i^n \left[a_i \mu_i \frac{1 - e_i^2}{1 + e_i \cos f_i} \sin(f_i + \omega_i) \right] \quad (16)$$

Combining the vis-viva equation with the equation for orbital period, $P = 2\pi \sqrt{a^3/G(m_{star} + m_{planet})}$, gives equation 17, where the gravitational constant, G , is abstracted away.

$$v = \frac{2\pi a}{P} \sqrt{\left(\frac{2a - r}{r} \right)} \quad (17)$$

Introducing equation 14 to equation 17 gives equation 18, an expression for orbital velocity that does not require knowledge of distance from the star.

$$v = \frac{2\pi a}{P} \sqrt{\frac{1 + 2e \cos f + e^2}{1 - e^2}} \quad (18)$$

This can be further combined with equation 16, replacing our circular orbital velocity with the new eccentric orbital velocity, to give equation 19, which is our model for TTV extended to include orbital eccentricity.

$$\delta_T = - \frac{P_T}{2\pi a_T} \left(\frac{1 - e_T^2}{1 + 2e_T \cos f_T + e_T^2} \right)^{\frac{1}{2}} \cdot \sum_i^n \left[a_i \mu_i \frac{1 - e_i^2}{1 + e_i \cos f_i} \sin(f_i + \omega_i) \right] \quad (19)$$

2.1.2 Exterior perturbation

To derive the effects of orbital perturbation, we follow a derivation for a two-planet case (Agol & Fabrycky 2018), and extend this to n planets. We assume the transiting planet to have zero eccentricity and all planets are on coplanar orbits. We give the equation of motion for a body acting under gravity in equation 20.

$$\ddot{\mathbf{R}} = \sum_{j \neq i} \left[G m_j \frac{\mathbf{R}_j - \mathbf{R}_i}{|\mathbf{R}_j - \mathbf{R}_i|^3} \right] \quad (20)$$

Where the bold indicates that the position of the planet, \mathbf{R} , is a vector. This can trivially be shown to satisfy 21

$$\sum_i m_i \ddot{\mathbf{R}}_i = 0 \quad (21)$$

Which is a demonstration that the centre of mass of the system, $\mathbf{R}_{C.o.M}$ is fixed, and no external forces are at play. This set of equations are most commonly used in numerical approaches; for an analytical approach to perturbation, it is more convenient to deal with the Jacobi coordinates of the system (Malhotra 1993; Murray & Dermott 1999).

This gives a set of new coordinates, r_i , describing the position

of the i^{th} body relative to the mass interior to its orbit, as given in equation 22.

$$\begin{aligned} \mathbf{r}_0 &= \mathbf{R}_{C.o.M} = 0 \\ \mathbf{r}_1 &= \mathbf{R}_1 - \frac{m_0 \mathbf{R}_0}{m_0} = \mathbf{R}_1 - \mathbf{R}_0 \\ \mathbf{r}_2 &= \mathbf{R}_2 - \frac{m_0 \mathbf{R}_0 + m_1 \mathbf{R}_1}{m_0 + m_1} \\ &\vdots \\ \mathbf{r}_{n+1} &= \mathbf{R}_{n+1} - \frac{\sum_{j=0}^n m_j \mathbf{R}_j}{\sum_{j=0}^n m_j} \end{aligned} \quad (22)$$

We can reformulate the equations of motion in Jacobi coordinates, given in equation 23.

$$\ddot{\mathbf{r}}_{n+1} = \ddot{\mathbf{R}}_{n+1} - \frac{\sum_{j=0}^n m_j \ddot{\mathbf{R}}_j}{\sum_{j=0}^n m_j} \quad (23)$$

As we are investigating the case of an exterior perturbation, we only consider the equation for the innermost body, $\ddot{\mathbf{r}}_1$, given in equation 24.

$$\begin{aligned} \ddot{\mathbf{r}}_1 &= \ddot{\mathbf{R}}_1 - \ddot{\mathbf{R}}_0 \\ \ddot{\mathbf{r}}_1 &= \sum_{j \neq 1} \left[Gm_j \frac{\mathbf{R}_j - \mathbf{R}_1}{|\mathbf{R}_j - \mathbf{R}_1|^3} \right] - \sum_{j \neq 0} \left[Gm_j \frac{\mathbf{R}_j - \mathbf{R}_0}{|\mathbf{R}_j - \mathbf{R}_0|^3} \right] \end{aligned} \quad (24)$$

The double summation terms can be collected, to give equation 25,

$$\begin{aligned} \ddot{\mathbf{r}}_1 &= Gm_0 \frac{\mathbf{R}_0 - \mathbf{R}_1}{|\mathbf{R}_0 - \mathbf{R}_1|^3} - Gm_1 \frac{\mathbf{R}_1 - \mathbf{R}_0}{|\mathbf{R}_1 - \mathbf{R}_0|^3} \\ &+ \sum_{j=2} \left[Gm_j \frac{\mathbf{R}_j - \mathbf{R}_1}{|\mathbf{R}_j - \mathbf{R}_1|^3} - Gm_j \frac{\mathbf{R}_j - \mathbf{R}_0}{|\mathbf{R}_j - \mathbf{R}_0|^3} \right] \end{aligned} \quad (25)$$

For the Jacobian coordinates, We introduce the notation $|\mathbf{r}_i| \equiv r_i$. That is to say, coordinates in bold is the vector position, while non-bold is the magnitude of that vector. We also introduce the reduced mass, μ , given as $\mu_i = m_i/M$, where M is the total mass of the system. As the central star typically dominates the mass of the system, this can also be written $\mu_i \approx m_i/m_0$.

$$\begin{aligned} \ddot{\mathbf{r}}_1 &= -Gm_0 \frac{\mathbf{r}_1}{r_1^3} - Gm_1 \frac{\mathbf{r}_1}{r_1^3} \\ &+ \sum_{j=2} \left[Gm_j \frac{\mathbf{R}_j - \mathbf{R}_1}{|\mathbf{R}_j - \mathbf{R}_1|^3} - Gm_j \frac{\mathbf{R}_j - \mathbf{R}_0}{|\mathbf{R}_j - \mathbf{R}_0|^3} \right] \end{aligned} \quad (26)$$

This can be further simplified to equation 27,

$$\begin{aligned} \ddot{\mathbf{r}}_1 &= -G(m_0 + m_1) \frac{\mathbf{r}_1}{r_1^3} \\ &+ \sum_{j=2} \left[Gm_j \frac{\mathbf{R}_j - \mathbf{R}_1}{|\mathbf{R}_j - \mathbf{R}_1|^3} - Gm_j \frac{\mathbf{R}_j - \mathbf{R}_0}{|\mathbf{R}_j - \mathbf{R}_0|^3} \right] \end{aligned} \quad (27)$$

In the 2-body case, where there is only a single body and star, this equation is the Keplerian equation of motion, as given in equation 28. As the central star dominates the mass of the system, we can use $m_0 + m_i \approx m_0$ to simplify.

$$\begin{aligned} \ddot{\mathbf{r}}_i &= -G(m_0 + m_i) \frac{\mathbf{r}_i}{r_i^3} \\ &\approx -Gm_0 \frac{\mathbf{r}_i}{r_i^3} \end{aligned} \quad (28)$$

From this, we can see that the acceleration in Jacobian coordinates for an n-body system as given by equation 27 is the standard Keplerian with some perturbative acceleration applied,

$$\ddot{\mathbf{r}}_i = -G(m_0 + m_1) \frac{\mathbf{r}_1}{r_1^3} + \delta \ddot{\mathbf{r}}_1 \quad (29)$$

With the perturbing acceleration given in equation 30,

$$\delta \ddot{\mathbf{r}}_1 = \sum_{j=2} \left[Gm_j \frac{\mathbf{R}_j - \mathbf{R}_1}{|\mathbf{R}_j - \mathbf{R}_1|^3} - Gm_j \frac{\mathbf{R}_j - \mathbf{R}_0}{|\mathbf{R}_j - \mathbf{R}_0|^3} \right] \quad (30)$$

If we introduce terms to the first fractional part, we have equation 31.

$$\begin{aligned} \delta \ddot{\mathbf{r}}_1 &= \sum_{j=2} \left[Gm_j \frac{\mathbf{R}_j - \mathbf{R}_1 + \mathbf{R}_0 - \mathbf{R}_0}{|\mathbf{R}_j - \mathbf{R}_1 + \mathbf{R}_0 - \mathbf{R}_0|^3} - Gm_j \frac{\mathbf{R}_j - \mathbf{R}_0}{|\mathbf{R}_j - \mathbf{R}_0|^3} \right] \\ &= \sum_{j=2} \left[Gm_j \frac{\mathbf{R}_j - \mathbf{R}_0 - \mathbf{r}_1}{|\mathbf{R}_j - \mathbf{R}_0 - \mathbf{r}_1|^3} - Gm_j \frac{\mathbf{R}_j - \mathbf{R}_0}{|\mathbf{R}_j - \mathbf{R}_0|^3} \right] \end{aligned} \quad (31)$$

We can consider the value of $\mathbf{R}_j - \mathbf{R}_0$ by using the general case in equation 22, to give the values for \mathbf{R}_j ,

$$\begin{aligned} \mathbf{r}_{n+1} &= \mathbf{R}_{n+1} - \frac{\sum_{j=0}^n m_j \mathbf{R}_j}{\sum_{j=0}^n m_j} \\ \mathbf{R}_{n+1} &= \mathbf{r}_{n+1} + \frac{\sum_{j=0}^n m_j \mathbf{R}_j}{\sum_{j=0}^n m_j} \end{aligned} \quad (32)$$

Which can be evaluated to give equation 33. Note the expression obtained in line 3, the value for $\mathbf{R}_j - \mathbf{R}_0$ is some combination of $\mathbf{R}_k - \mathbf{R}_0$ where $0 < k < j$.

$$\begin{aligned} \mathbf{R}_j - \mathbf{R}_0 &= \mathbf{r}_j + \frac{\sum_{k=0}^{j-1} m_k \mathbf{R}_k}{\sum_{k=0}^{j-1} m_k} - \mathbf{R}_0 \\ &= \mathbf{r}_j + \frac{\sum_{k=1}^{j-1} m_k \mathbf{R}_k}{\sum_{k=0}^{j-1} m_k} \\ &= \mathbf{r}_j + \sum_{k=1}^{j-1} \mu_k (\mathbf{R}_k - \mathbf{R}_0) \end{aligned} \quad (33)$$

We define a new value \mathbf{r}_{j0} , as a shorthand for $\mathbf{R}_j - \mathbf{R}_0$, given in equation 34,

$$\mathbf{r}_{j0} \equiv \mathbf{R}_j - \mathbf{R}_0 = \mathbf{r}_j + \sum_{k=1}^{j-1} \mu_k (\mathbf{R}_k - \mathbf{R}_0) \quad (34)$$

By combining equations 31 and 33, we obtain equation 35,

$$\delta \ddot{\mathbf{r}}_1 = \sum_{j=2} \left[Gm_j \frac{\mathbf{r}_{j0} - \mathbf{r}_1}{|\mathbf{r}_{j0} - \mathbf{r}_1|^3} - Gm_j \frac{\mathbf{r}_{j0}}{r_{j,0}^3} \right] \quad (35)$$

Which can be simplified slightly as equation 36, which gives the perturbing acceleration on the transiting planet. Additionally, this can be expanded in a Legendre series to first order, giving the second line in equation 36.

$$\begin{aligned} \delta \ddot{\mathbf{r}}_1 &= \sum_{j=2} \left[Gm_j \left(-\frac{\mathbf{r}_1 - \mathbf{r}_{j0}}{|\mathbf{r}_1 - \mathbf{r}_{j0}|^3} - \frac{\mathbf{r}_{j0}}{r_{j0}^3} \right) \right] \\ &= \sum_{j=2} \left[-\frac{Gm_j}{r_j^3} \left(\mathbf{r}_1 - 3\frac{\mathbf{r}_1 \cdot \mathbf{r}_j}{r_j^2} \mathbf{r}_j \right) + \mathcal{O}(r_1/r_j)^2 \right] \end{aligned} \quad (36)$$

To find the perturbed period of the transiting planet, we compute how this acceleration changes when averaged over the orbital period. The angular position of the transiting planet from the vantage point of earth, θ_1 , is given by the sum of its true anomaly, f_1 , and argument of periapsis, ω_1 . This is given by equation 37. As the eccentricity of this planet is assumed zero, the true and mean anomaly are equal. We can also introduce the transit number, n_1 , and epoch τ_1 , to this expression.

$$\begin{aligned} \theta_1 &= f_1 + \omega_1 \\ &= n_1 (t - \tau_1) \omega_1 \end{aligned} \quad (37)$$

Differentiating equation 37 with respect to time gives equation 38.

$$\dot{\theta}_1 = \dot{n}_1 (t - \tau_1) + n_1 - n_1 \dot{t}_{0,1} \quad (38)$$

Following (Murray & Dermott 1999, section. 2.9), we express \dot{n} as a function of the semimajor axis: $\dot{n}_1 = -3n_1/2a_1 \dot{a}_1$,

$$\begin{aligned} \dot{\theta}_1 &= \frac{3n_1 \dot{a}_1}{2a_1} (t - \tau_1) + n_1 - n_1 \dot{\tau} \\ &= n_1 \left(\frac{3\dot{a}_1}{2a_1} (t - \tau_1) + 1 - \dot{\tau} \right) \end{aligned} \quad (39)$$

Expressing the time derivatives \dot{a}_1 , $\dot{\tau}$, and $\dot{\omega}$ in terms of $\ddot{\mathbf{r}}_1$ gives equation 40,

$$\dot{\theta}_1 = n_1 \left(1 - \frac{2a_1^2}{G(m_0 + m_1)} \sum_{j=2} \left[\frac{1}{2} \frac{Gm_j a_1}{r_j^3} \right] \right) \quad (40)$$

This demonstrates why the orbital period of the transiting planet increases, the addition of planets in the system causes an increase in the effective mass interior to its orbit by $\frac{1}{2}m_j (a_1/r_j)^3$.

We can finally obtain the timing of the (N+1)th transit, given in equation 41 (see Agol & Fabrycky 2018, Section 4).

$$\begin{aligned} t - t_0 &= \int_{f_0}^{f_0+2\pi N} df_1 \dot{\theta}_1^{-1} \\ &= \int_{f_0}^{f_0+2\pi N} df_1 n_1^{-1} \left[1 + \frac{1}{m_0 + m_1} \sum_{j=2} \left[m_j \left(\frac{a_1}{r_j} \right)^3 \right] \right] \end{aligned} \quad (41)$$

Following (Borkovits et al. 2003), we can express the true anomaly of the transiting planet in terms of the true anomalies of the perturbing planets,

$$df_1 = \sum_{j=2} \left[\frac{P_j}{P_1} \frac{r_j^2}{a_j^2 (1 - e_j^2)^{1/2}} df_j \right] \quad (42)$$

As r depends on f , we can introduce equation 8 (also given below as equation 43) to equation 42, giving equation 44.

$$r_i = \frac{a_i (1 - e_i^2)}{1 + e_i \cos f_i} \quad (43)$$

$$\begin{aligned} df_1 &= \sum_{j=2} \left[\frac{P_j}{P_1} \frac{1}{a_j^2 (1 - e_j^2)^{1/2}} \left(\frac{a_j (1 - e_j^2)}{1 + e_j \cos f_j} \right)^2 df_j \right] \\ &= \sum_{j=2} \left[\frac{P_j}{P_1} \frac{(1 - e_j^2)^{3/2}}{(1 + e_j \cos f_j)^2} df_j \right] \end{aligned} \quad (44)$$

As the original variable of integration in equation 41, f_1 , changes due to the perturbation, we rewrite the integral in terms of the unperturbed f_j . As an approximation, we consider only the gravitational forces acting on the transiting planet, and treat the perturbing planets as following Keplerian orbits. The substitution for df_1 is given in equation 44, and gives equation 45.

$$\begin{aligned} t - t_0 &= \sum_{j=2} \left[\int_{f_0}^{f_0+2\pi N} \left(df_j n_1^{-1} \frac{P_j}{P_1} \frac{(1 - e_j^2)^{3/2}}{(1 + e_j \cos f_j)^2} \right. \right. \\ &\quad \left. \left. \cdot \left(1 + \frac{m_j}{m_0 + m_1} \left(\frac{a_1}{r_j} \right)^3 \right) \right) \right] \end{aligned} \quad (45)$$

Which can be evaluated to give the timing of the (N+1)th transit, as shown in equation 46.

$$\begin{aligned} t - t_0 &= NP_1 + \frac{P_1^2}{2\pi (m_0 + m_1)} \\ &\quad \cdot \sum_{j=2} \left[\frac{m_j (f_j + e_j \sin f_j) (1 - e_j^2)^{-3/2}}{P_j} \right] \end{aligned} \quad (46)$$

To find the departure from linear ephemerides, and thus obtain the TTV, we subtract the mean transit time NP_1 from equation 46. As the f_j terms also includes the mean motion, $n(t - \tau_j)$, of the perturbing planets, we subtract that too, giving equation 47. We have now found a model of TTV caused by the perturbation due to outer planets.

$$\begin{aligned} \delta t_1 &= \frac{P_1^2}{2\pi (m_0 + m_1)} \\ &\quad \cdot \sum_{j=2} \left[\frac{m_j (f_j - n_j (t - \tau_j) + e_j \sin f_j) (1 - e_j^2)^{-3/2}}{P_j} \right] \end{aligned} \quad (47)$$

2.2 Parameter search

With a suite of analytical TTV models at our disposal, we need to determine methods for fitting to simulated or physical TTV signals,

and determining the marginalization and uncertainty associated with the fit.

2.2.1 Optimisation

To determine this set of best fit parameters, we leverage the field of computational optimisation and the large quantity of implementations that have been written for it through the SciPy package (Virtanen et al. 2020). There are various metrics and methods for optimisation, the most notable of which are discussed below.

2.2.1.1 Least squares regression An appealing yet naïve approach to parameter determination, where the squared sum of the residuals to a fit is minimised. While this has benefits in the form of being expressible as linear algebra, and efficient implementation in many programming packages, it suffers on two accounts:

Least squares is a method of local minimisation, there is no way of determining if a solution is the global optimum without searching all possible solutions. Searching through all solutions is described by the brute-forcing algorithm, the complexity of which grows exponentially with the number of model parameters. This exponential growth leads to computationally expensive searches.

This method will also often under- or over-estimate the fitting parameters, as the residuals to the fit is not necessarily a good metric from initial parameters. This is most susceptible to random error in the data, and is less an issue on larger datasets.

2.2.1.2 Maximum likelihood Maximum likelihood estimation, or MLE, provides a numerical optimisation method for models that more accurately describe real world data, as opposed to the more idealised case within which linear regression was originally formulated. The logarithm of the likelihood function for Bayesian analysis is given in equation 48

$$\ln p(y|x, \theta, M) = -\frac{1}{2} \sum_n \left[\frac{(y_n - M(x, \theta))^2}{s_n^2} + \ln(2\pi s_n^2) \right] \quad (48)$$

Where x and y is the set of data, M is the model to be fit, θ the parameters of the model, and s is given by:

$$s_n^2 = \sigma_n^2 + f^2 [M(x, \theta)]^2 \quad (49)$$

Where the additional σ is the error on each data point. The logarithm of the likelihood is used due to its concavity, and that a given log likelihood is the sum of individual log likelihoods, providing an intuitive base for usage.

2.2.1.3 Minimisation In computational parameter optimisation, it is standard to try to minimise, rather than maximise, the values of a function within a given parameter space. As we have highlighted the usefulness of maximum likelihood estimation, it is useful to quickly adapt this to the computational workflow. This is very easily achieved by simply attempting to minimise the *negative* log likelihood of a model.

As mentioned in least squares regression, 2.2.1.1, it is preferable to find, or approximately find, the global optimal parameters, two methods for which are differential evolution (Storn & Price 1995)

and dual annealing (Xiang et al. 1997). These global optimisation methods are accessed through the `scipy.optimize` python package (Virtanen et al. 2020).

Differential evolution takes inspiration from evolutionary science, and attempts to optimise a problem through iterative improvements to some metaheuristic, such as the log likelihood, of the solution. This is specifically achieved by initialising some large set of “agents”, each with random starting parameters, and allowing them to converge to the global solution.

Dual annealing implements a different metaheuristic to differential evolution, taking its name from the controlled cooling of material (annealing) used in metallurgy. The method begins by considering the neighbouring parameters s^* to its current solution s , and probabilistically deciding to move between those states. This probabilistic method is dependent upon the “Temperature” of each state, which is defined as some function of the likelihood of each state. The free energy of the system defines how much each parameter may vary, and is slowly reduced, eliminating solutions with lower likelihoods. This continues until the free energy is zero, and the system has found the optimum with maximum likelihood.

As a typical final step in global optimisation, we polish the found parameters with some local optimisation method. The specific tool used in this project is the bounded limited memory Broyden–Fletcher–Goldfarb–Shanno algorithm (Byrd et al. 1995; Zhu et al. 1997), accessed through the `scipy.optimize.minimize` method call. This method, like many other quasi-newton algorithms, iteratively converges to an optimal solution through gradient descent. This is specifically achieved using the hessian matrix of the function, the square matrix of second-order partial derivatives with respect to the free parameters of the function.

2.2.2 Model comparison

In the case where models have an equivalent set of parameters, simply comparing the value of the maximum likelihood is sufficient to draw conclusions about their accuracy. If the number of parameters between models is inconsistent, this approach will naturally result in overfitting.

This is especially relevant for our transit models. For example, the initial TTV model seen in equation 13 has two fixed parameters and an additional four free parameters per body in the system, while the extension to this, as seen in equation 19 has four fixed parameters and an additional five free parameters per body in the system. As the number of bodies in the system is an additional free parameter, some method of comparison is required.

2.2.2.1 Akaike information criterion To determine the quality for each model, some weighting dependent on their free parameters is introduced to the likelihood value (Akaike 1992, 1974).

$$\text{AIC} = 2k - 2 \ln L \quad (50)$$

In the case where a model has k free parameters, with a maximum likelihood L , the AIC value is as given in equation 50. This provides a middle ground between goodness of fit, and model simplicity, handling both under- and over-fitting simultaneously.

2.2.2.2 Corrected Akaike information criterion Where the sample size is small, the standard AIC method can lead to overfitting, and a correction is required to better describe the fit (Hurvich & Tsai 1993). This extension to AIC attempts to fix the overfitting by introducing a further penalty from the number of data points, n , as seen in equation 51.

$$\text{AICc} = \text{AIC} + \frac{2k^2 + 2k}{n - k - 1} = \frac{2kn}{n - k - 1} - 2 \ln L \quad (51)$$

The exact form of the expression for the corrected Akaike information criterion (AICc) is not fixed, but is determined for the individual models to be scored. The equation given in 51 is the one used in this project, and is not representative of some general AICc expression. The commonality between all AICc expressions is the dependence on k^2 , as all AICc are second-order estimates.

2.2.2.3 Bayesian information criterion Another model selection metric closely related to the corrected Akaike information criterion, BIC introduces a larger penalty for free parameters by weighting them with the number of observed data points (Schwarz 1978).

$$\text{BIC} = k \ln n - 2 \ln L \quad (52)$$

In the case where a model has k free parameters, n data points, and a maximum likelihood L , the BIC value is as given in equation 52. Models that have lower BIC values are generally preferred, though this does not always yield the optimal fit.

2.2.3 Markov Chain, Monte Carlo

To determine parameter bounds, a Markov Chain Monte Carlo (MCMC) method is used through the *Emcee* package (Foreman-Mackey et al. 2013). This class of algorithms use the probability distribution of a function to construct a sample chain that converges to the desired distribution, typically one that is normally distributed around the true parameters. This begins very similarly to the differential evolution optimisation technique described in section 2.2.1.3, where a set of “walkers” are initialised with arbitrary starting parameters, and allowed to walk around the parameter space according to some ensemble method.

The ensemble method used for this project is the stretch move (Goodman & Weare 2010), a method that significantly outperforms the more traditional Metropolis-Hastings algorithm (Hastings 1970). The position of a walker, X_k , is determined by randomly selecting another walker, X_j , and proposing a new position by interpolating between the two with some variable randomly drawn from the distribution, as given in equation 53, (see Foreman-Mackey et al. 2013, eq. 7).

$$X_k(t) \rightarrow Y = X_j + Z [X_k(t) - X_j] \quad (53)$$

3 RESULTS / DISCUSSION

3.1 Observations

Six transit observations were planned at the beginning of this project to observe four different exoplanetary systems: HAT-P-13b, HAT-

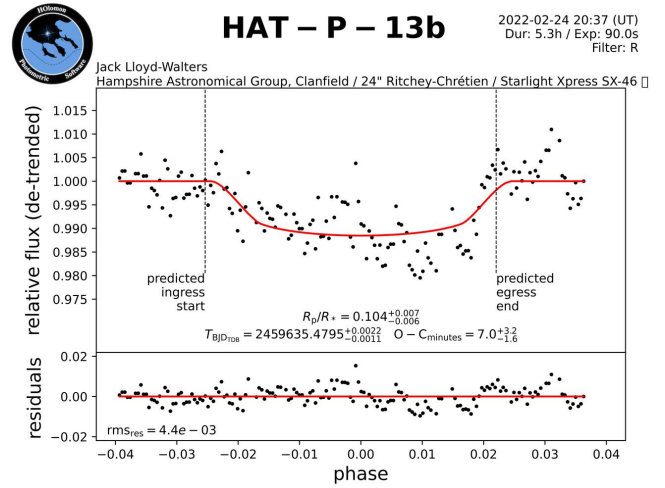


Figure 6. De-trended transit light curve for an observation of HAT-P-13b taken 2022-02-24 and analysed with HOPS (Tsiaras 2021). Note the anomaly near the mid-transit time due to light cloud cover. Reported R_p/R_* is 0.0844 ± 0.0013 (Bakos et al. 2009).

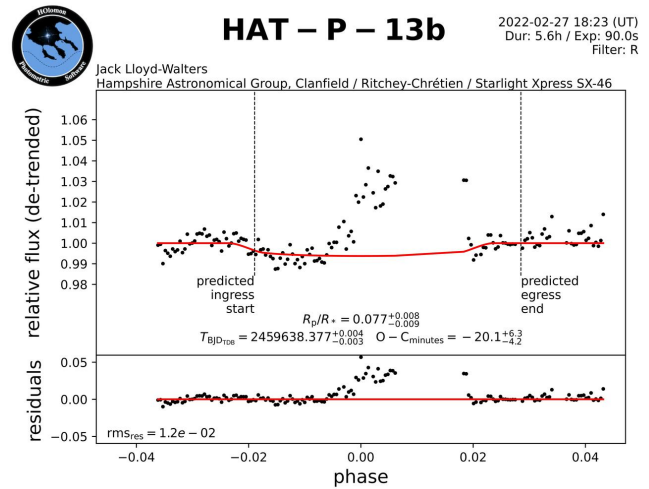


Figure 7. De-trended transit light curve for an observation of HAT-P-13b taken 2022-02-27 and analysed with HOPS (Tsiaras 2021). Note the large residuals and missing data in the second half of the transit, caused by clouds completely obscuring the star for several hours.

P-44b, and K2-19b. Of those, only two observations were made, both of HAT-P-13b, and are given in figures 6 and 7. Significant cloud cover prevented further observations, and severely hampered the observation made on 2022-02-27 (figure 7).

The light-curves have been fit with HOPS (Tsiaras 2021), and have provided values for R_p/R_* close to the reported literature (Bakos et al. 2009) when considering the large variance in flux caused by suboptimal weather conditions.

The observation on 2022-02-27, given in figure 7, shows large residuals and has had data points after the mid-transit time removed. This was due to large cloud cover that reduced sky visibility to 0% for several hours during the middle of the transit. Re-introducing these data points to the HOPS fitting, the result in figure 8 is obtained.

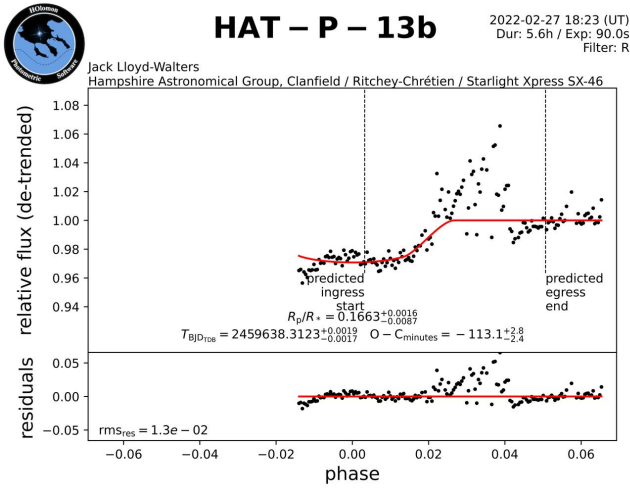


Figure 8. De-trended transit light curve for the observation of HAT-P-13b taken 2022-02-27 without the central data points removed. Note how the relative change in flux due to the cloud-cover is larger than the predicted occlusion depth.

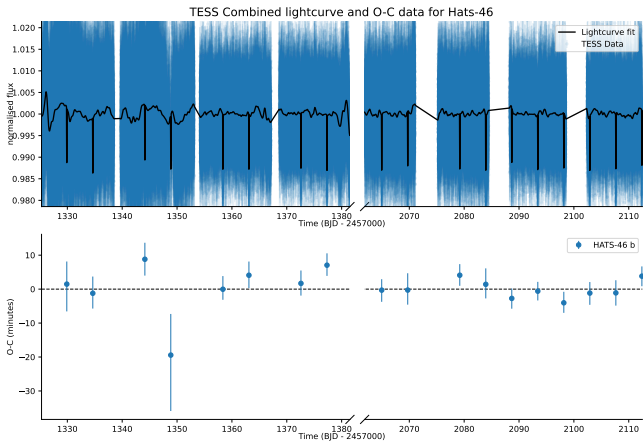


Figure 9. Combined light curve (top) and TTV data (bottom) for HATS-46, as observed by the *TESS* spacecraft. Only HATS-46b transits the star, and a large disparity is noted in the middle of the dataset due to observations over multiple *TESS* sectors.

The relative change in flux due to the cloud cover is larger than the occlusion depth of the transit by a significant factor, causing hops to fit transit egress to this position.

3.2 Light curve analysis

With *TESS* light curves collected from the Mikulski archive for space telescopes, additional mid-transit times could be computed to complement those collected from the exoplanet transit database and ExoClock database, as was briefly touched upon in section 1.3.

3.2.1 De-trending

Many of the *TESS* light curves show strong long-term trends that can make transit detection difficult. To de-trend the data, a Gaussian

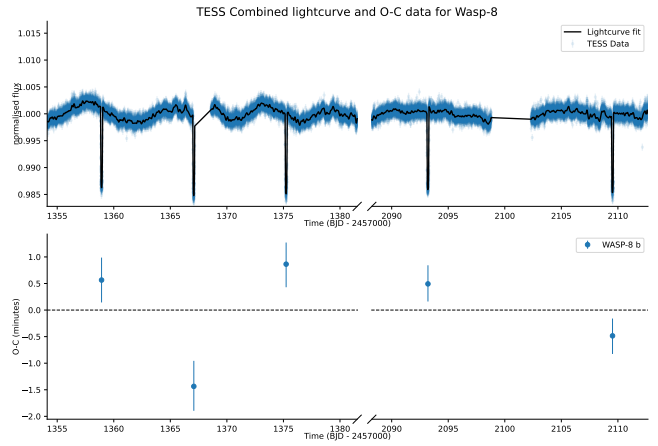


Figure 10. Combined light curve (top) and TTV data (bottom) for Wasp-8, as observed by the *TESS* spacecraft. Note the large transit depths as compared to HATS-46, and the correspondingly small uncertainties in the TTV residuals that result from that.

Juliet-Literature comparison

Source	(Brahm et al. 2018)	Juliet posteriors
a/R_*	$13.55^{+0.45}_{-0.65}$	$14.699^{+1.541}_{-1.830}$
b	$0.63^{+0.042}_{-0.034}$	$0.480^{+0.266}_{-0.171}$
i	$87.32^{+0.22}_{-0.31}$	$88.126^{+1.119}_{-1.014}$
R_p/R_*	0.1088 ± 0.0027	$0.10369^{+0.00496}_{-0.00444}$
P	4.7423729 ± 0.0000049	$4.7423836070^{+0.0000139110}_{-0.0000114410}$

Table 1. Comparison between literature parameters and Juliet parameters from *TESS* for HATS-46b. Many of the values satisfy the ‘Good enough’ criterion despite not having the complementary radial velocity measurements used in the detection paper to refine the parameters.

process is fit to the out-of transit data, using the linear ephemerides for the planetary system to deduce the locations of transits. This was performed using an approximate Matern kernel using the Juliet wrapper (Espinoza et al. 2019) to the Celerite package (Foreman-Mackey et al. 2017). This gives the black line seen in figures 9, and 10, which very closely matches the overall light curve trend.

3.2.2 Transit fitting

Transit fits were then performed on the de-trended data, using the Juliet wrapper to both the Batman (Kreidberg 2015) and Dynesty (Speagle 2020) packages. The parameters for the transit models are initialised by randomly selecting from the priors, and iteratively walked through parameter space. The general parameters for the exoplanet are returned to the posteriors of the fit, which adequately match the confirmed literature despite having few transits and no initial parameters to work from.

Taking the posteriors found for HATS-46b, as compared to the detection paper (Brahm et al. 2018), we have the results as given in table 1

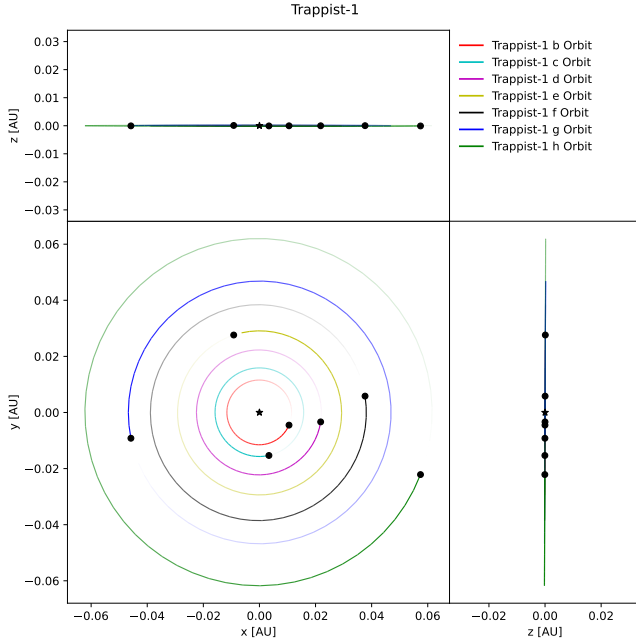


Figure 11. system layout as generated by the simulation pipeline for TRAPPIST-1.

3.2.3 TTV Residuals

By subtracting the found transit times from computed ephemerides, TTV residuals are computed for each transit. The transit uncertainty is read directly from the Juliet posteriors, and can be seen in the bottom half of figures 9, and 10.

The uncertainty in the residuals are proportional to the depth of the transit divided by the variance in the data, as seen in the two provided figures. From this, we would expect the magnitude and standard deviation of TTV to be approximately the same order of magnitude. Of particular note in that regard are the Wasp-8 TTV residuals: all of them lie more than one sigma from the centre of the O-C diagram.

3.3 Simulation pipeline

3.3.1 Initialisation

The simulation pipeline makes use of the REBOUND (Rein & Liu 2012) and REBOUNDx (Tamayo et al. 2020) packages for numerical integration. The pipeline has been created to accept either a given planetary system, as seen in figure 11, or a .csv file holding orbital elements.

The system is set up from provided parameters, and iterated forward using IAS15 (Rein & Spiegel 2015). Transits are evaluated where the position of the star, \hat{R}_* , and position of the target, \hat{R}_T , satisfy the following,

$$\hat{R}_{*.y} < \hat{R}_{T.y} \quad (54)$$

$$\hat{R}_{*.x} \approx \hat{R}_{T.x} \quad (55)$$

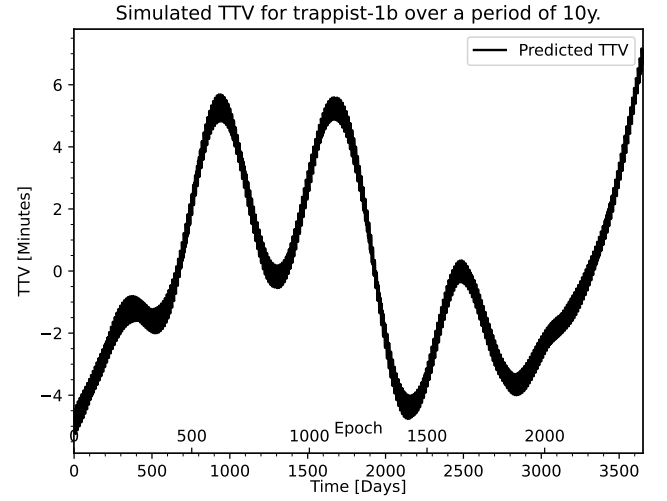


Figure 12. TTV Residuals for TRAPPIST-1b according to simulation. Note the complex sinusoidal motion and large magnitudes due to the highly-resonant nature of the system. Additionally, note the thickness of the TTV curve, caused by low period variations due to non-resonant interactions.

3.3.2 Simulated TTV

The transit time is found by iteratively decreasing simulation step time, and integrating until the desired precision is met, which is 1 millisecond by default. This provides a set of simulated transit times which are fit to TTV using least squares regression as mentioned in section 2.2.1.1. In the case of TRAPPIST-1b, the simulation output is as seen in figure 12.

The simulation pipeline integrates the effect of General Relativity, but could be further extended to include effects such as tidal deformation, stellar evolution, to even those as subtle as the Yarkovsky effect.

3.4 Analytical Models

Following the derivations in section 2.1, analytical models were converted to vectorisable code to be executed by the computation pipeline. The translation from equation to code is mostly intuitive, and can be seen in this project's GitHub repo, <https://github.com/SK1Y101/TransitProject>.

3.4.1 System and TTV

To demonstrate the capabilities of the model, we selected exoplanetary systems with known TTV signals and generated synthetic systems whose properties were representative of the real systems, to allow precise control of TTV magnitude during testing phases. An example system layout is given in figure 13.

Running the TTV simulation over a timescale of 200 years gives the residuals seen in figure 14. A 200-year integration is not representative of any known transiting planet, the earliest transiting data dates to 1999 (Charbonneau et al. 2000), and was simply chosen to better demonstrate the long-term evolution of the transit timing variation. The variance of data returned by the simulation is the standard

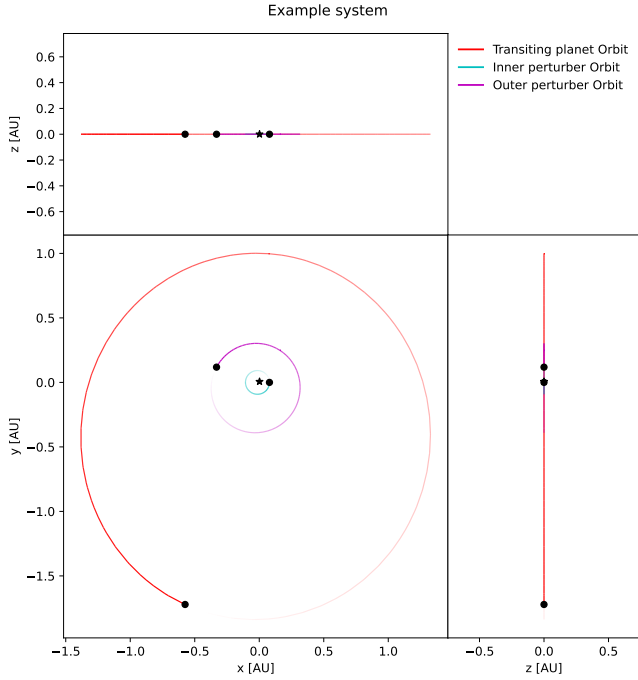


Figure 13. The system layout for one of the synthetic test systems. The exoplanets in this example were initialised in low-eccentricity non-resonant orbits, with masses of 0.8, 0.5, and 5 Jupiter mass from inner to outer.

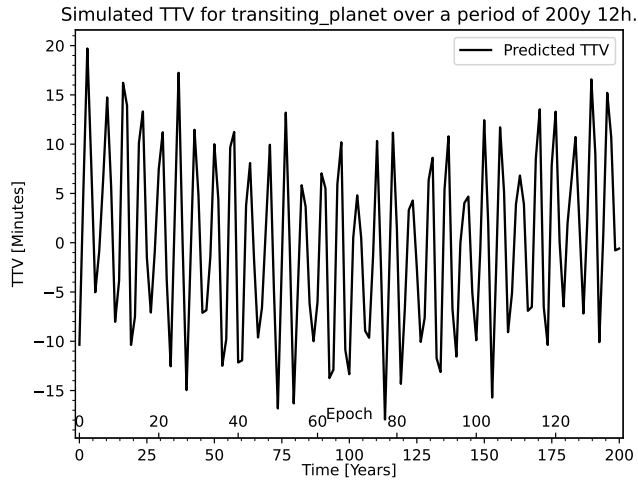


Figure 14. Simulated TTV residuals for the synthetic system used in model testing. The TTV simulated are for the outermost planet in the system, as seen in figure 13.

deviation of measured TTV. For real systems, however, this variance is computed from the uncertainties of known parameters.

The initial parameters used in this synthetic system are given in table 2.

3.4.2 Model comparison

To determine the validity of the analytical models, they were executed with the initial parameters used to set up the system. As the transiting

Initialisation parameters

Body	m	p (Days)	e	ω (Radians)
Star	$1.32M_{\odot}$			
Perturber 1	$0.8M_J$	9	0.15	0
Perturber 2	$0.5M_J$	65	0.15	1
Target	$5M_J$	537	0.30	$\pi/2$

Table 2. Initial parameters for the synthetic system: mass, orbital period, eccentricity, and argument of periapsis. Other parameters not given here are initialised to zero, or (in the case of a , computed in-situ).

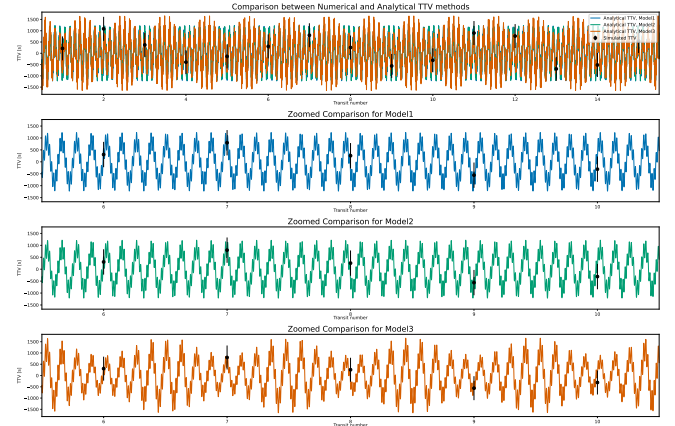


Figure 15. Analytical TTV curves for the synthetic system, using the initial parameters. Model1, Model2, and Model3 are the models described in equations 13, 16, and 19 respectively. Each graph shows a 5 transit snippet from the TTV curve. Simulated TTV times are overlaid in black.

planet is the outermost in the system, the first class of TTV models 2.1.1 is used to analytically approximate the system. The models as derived in this paper are given in equations 13, 16, and 19, which are re-summarised below,

$$\delta_T = -\frac{P_T}{2\pi a_T} \sum_i^n \left[a_i \mu_i \sin \frac{2\pi (t - t_{0,i})}{P_i} \right] \quad (56)$$

$$\delta_T = -\frac{P_T}{2\pi a_T} \sum_i^n \left[a_i \mu_i \frac{1 - e_i^2}{1 + e_i \cos f_i} \sin (f_i + \omega_i) \right] \quad (57)$$

$$\delta_T = -\frac{P_T}{2\pi a_T} \left(\frac{1 - e_T^2}{1 + 2e_T \cos f_T + e_T^2} \right)^{\frac{1}{2}} \cdot \sum_i^n \left[a_i \mu_i \frac{1 - e_i^2}{1 + e_i \cos f_i} \sin (f_i + \omega_i) \right] \quad (58)$$

The TTV curves generated by these models are shown in figure 15. Despite the various approximations used in the derivation for each model, and the subsequent applicability space for which they are valid, they match the simulated TTV to within the 1-sigma bounds.

Additionally, note how the overall shape of the analytical curves change as a result of first introducing perturbed eccentricity, and then transiting planet eccentricity. Strictly speaking, the sections of the analytical curves between each transit do not correspond to anything

One perturbing planet					
Model	Method	μ (10^{-6})	a (10^{-6} AU)	e	ω
Model 1	diff. evo.	348.996	380017.037	0.245	6.083
Model 1	dual an.	612.062	216632.299	0.033	5.589
Model 2	diff. evo.	504.325	185229.581	0.344	2.018
Model 2	dual an.	321.476	419718.646	0.116	-0.019
Model 3	diff. evo.	428.884	225677.205	0.043	-4.821
Model 3	dual an.	625.847	158625.666	0.114	-3.129

Table 3. Best fit parameters for the synthetic system TTV with only a single perturbing planet.

Two perturbing planets					
Model	Method	μ (10^{-6})	a (10^{-6} AU)	e	ω
Model 1	diff. evo.	75.484	1242571.08	0.239	-0.864
		41.269	135176.665	0.195	-1.165
Model 1	dual an.	206.855	226389.639	0.463	-6.241
		227.427	585132.122	0.483	-2.24
Model 2	diff. evo.	398.235	118829.517	0.019	-0.285
		146.886	833966.926	0.135	1.518
Model 2	dual an.	268.507	500367.032	0.101	-2.655
		138.984	335523.36	0.142	1.404
Model 3	diff. evo.	427.237	225677.039	0.0	6.283
		0.001	58070.523	0.085	4.119
Model 3	dual an.	35.051	995449.059	0.113	3.73
		102.098	966847.114	0.082	2.8

Table 4. Best fit parameters for the synthetic system TTV with two perturbing planets.

physically within the system. They describe the predicted TTV, were the transiting planet to have a slightly offset time of periapsis passage.

3.4.3 Parameter optimisation

As we have now demonstrated the accuracy of the models, to some set of initial parameters, the next stage is in the reverse: finding some parameters for which the model fits the data set.

To do this, the three models were each initialised with the fixed parameters, those being the properties of the host star and transiting planet, and three sets of arbitrary initial parameters corresponding to a system with one, two, and three planets additional to the transiting target, giving an effective set of 9 models to optimise. Each of these were optimised with both dual annealing (Xiang et al. 1997) and differential evolution (Storn & Price 1995), for which the best fit of the two methods was selected as that model’s maximum likelihood solution.

The optimal TTV curves found for each model are given in figure 16, where the table of the optimal solution values for each model is given in tables 3, 4, and 5.

3.4.4 Model selection

How a model is selected as most likely is two-fold. To begin with, we sample the solutions for each model to determine marginalisation and uncertainties. The parameter space is sampled using MCMC (Foreman-Mackey et al. 2013), and a normal parameter distribution

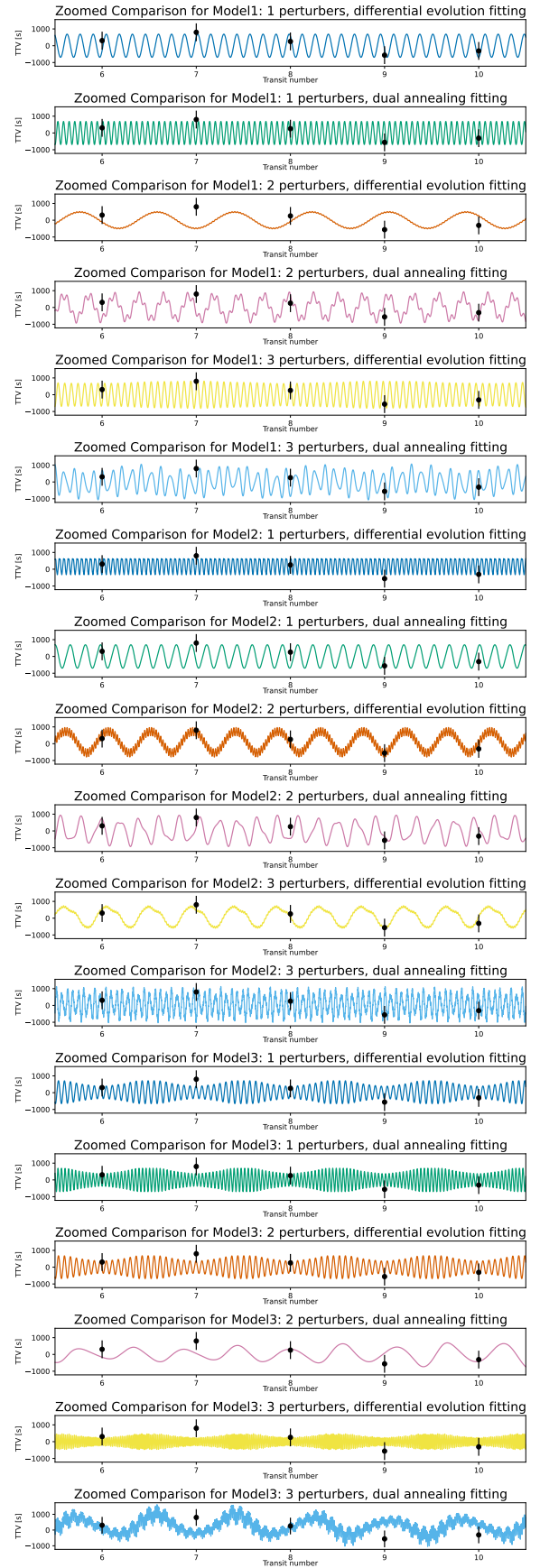


Figure 16. Fit TTV curves from the set of 18 models after being optimised with the synthetic system.

Three perturbing planets					
Model	Method	μ (10^{-6})	a (10^{-6} AU)	e	ω
Model 1	diff. evo.	589.909	241886.321	0.386	-2.961
		7.508	263814.017	0.449	-5.18
		53.287	236535.741	0.009	0.337
Model 1	dual an.	19.684	1411636.58	0.205	3.55
		196.18	240995.442	0.354	0.185
		428.518	310462.859	0.059	-6.026
Model 2	diff. evo.	145.253	833960.172	0.262	2.613
		0.001	134746.642	0.288	-3.71
		103.699	109160.755	0.142	-1.761
Model 2	dual an.	286.502	167374.87	0.14	2.521
		625.306	212912.077	0.001	-4.735
		471.528	67248.984	0.434	0.301
Model 3	diff. evo.	5.06	216218.273	0.311	1.451
		18.73	240838.21	0.309	4.963
		616.988	100349.363	0.021	-4.403
Model 3	dual an.	78.76	1242718.32	0.012	2.37
		123.85	281959.982	0.192	-4.953
		602.58	95865.617	0.5	1.574

Table 5. Best fit parameters for the synthetic system TTV with three perturbing planets.

Information criterion					
Model	Method	k	AIC	AIC_c	BIC
Model 1	diff. evo.	4	1430.588159	1430.961991	1441.462154
Model 1	dual an.	4	1430.953796	1431.327627	1441.827791
Model 1	diff. evo.	7	1642.175578	1643.252501	1661.20507
Model 1	dual an.	7	1606.839154	1607.916077	1625.868646
Model 1	diff. evo.	10	1460.812691	1462.990909	1487.99768
Model 1	dual an.	10	1512.057709	1514.235926	1539.242697
Model 2	diff. evo.	5	1465.179736	1465.745774	1478.772231
Model 2	dual an.	5	1464.585283	1465.151321	1478.177777
Model 2	diff. evo.	9	1452.463533	1454.228239	1476.930023
Model 2	dual an.	9	1454.402606	1456.167312	1478.869096
Model 2	diff. evo.	13	1472.298081	1476.012367	1507.638567
Model 2	dual an.	13	1727.322675	1731.03696	1762.66316
Model 3	diff. evo.	5	1477.916183	1478.48222	1491.508677
Model 3	dual an.	5	1454.979605	1455.545643	1468.5721
Model 3	diff. evo.	9	1620.473796	1622.238502	1644.940286
Model 3	dual an.	9	1564.258843	1566.023549	1588.725333
Model 3	diff. evo.	13	1592.098363	1595.812649	1627.438849
Model 3	dual an.	13	1697.010862	1700.725148	1732.351348

Table 6. Free parameter count, maximum log likelihood, and information criterion scores for each of the 18 models. The number of data-points used was 112, corresponding to a simulation time of around 200 years.

constructed from the sampler. This is then converted to a set of quantiles, specifically 16%, 50%, and 84%, representing the reportable parameters and one sigma variance. The entire normal distribution is reported in a corner plot using the corner package (Foreman-Mackey 2016).

With the set of best fit parameters and error bounds for each model, the accuracy of their fit is determined by scoring with the three information criterion discussed earlier: AIC (Akaike 1992, 1974), AICc (Hurvich & Tsai 1993), and BIC (Schwarz 1978). The information criterion values for each model for this data is given in table 6

Initial-‘best fit’ parameters			
Source	Initial	best-fit	second-contender
perturbers	2	1	2
μ (10^{-6})	763.834	347.104 ^{+42.043} _{-36.969}	307.455 ^{+266.399} _{-179.331}
	477.396		142.286 ^{+66.645} _{-39.806}
a (10^{-6})	902905.881	380015.561 ^{+22.941} _{-21.390}	118936.225 ^{+6823.023} _{-65818.750}
	347103.538		833957.621 ^{+144.770} _{-192.663}
p Days	65.002	74.471 ^{+0.007} _{-0.006}	13.040 ^{+1.002} _{-9.002}
	8.999		242.119 ^{+0.058} _{-0.081}
e	0.15	0.245 ^{+0.159} _{-0.170}	0.019 ^{+0.900} _{-0.220}
	0.15		0.135 ^{+0.080} _{-0.002}
ω	1	1.082 ^{+5.132} _{-4.661}	-0.285 ^{+0.826} _{-0.0368}
	0		1.518 ^{+0.593} _{-0.013}

Table 7. Comparison of the found parameters, with uncertainties, and the initial parameters used for system setup. The periods are computed from the semimajor axes, and each planet is given its own row. Model 1 with a single perturber produced the best fit parameter, while Model 2 with two perturbers produced the second-contender

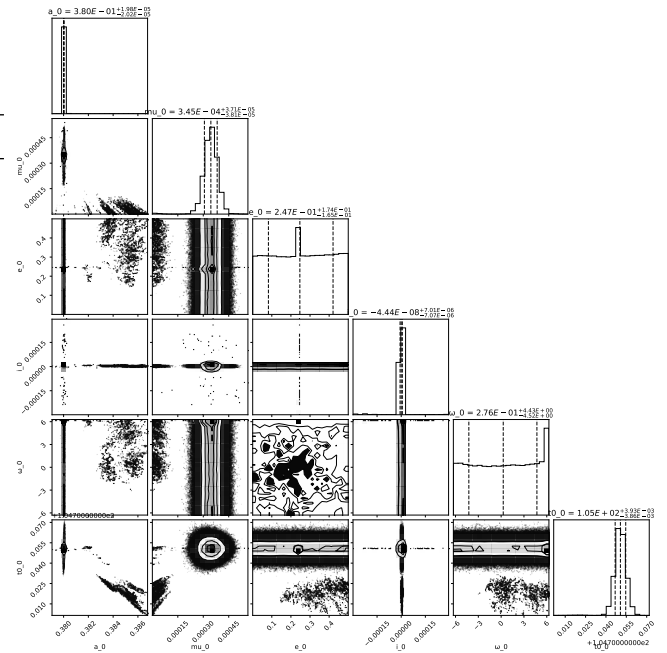


Figure 17. Sample corner plot for the best-fit model output. Along the diagonal is the normal distribution for each parameter, with the remainder filled out with two-dimensional slices of the parameter space. The 16th, 50th, and 84th quantiles are marked with the dashed lines.

3.4.5 Most likely system

We can determine the most likely model by selecting that with the lowest AIC, AICc, and BIC from the information criterion table 6. The parameters for the best fit model are given in table 7, with the initial system parameters also listed. Additionally, the corner plot for the best-fit model is given in figure 17, and the system layout is given in figure 18. As a comparison, the second-best fit model parameters are also listed in figure 7, with the system layout given in 19.

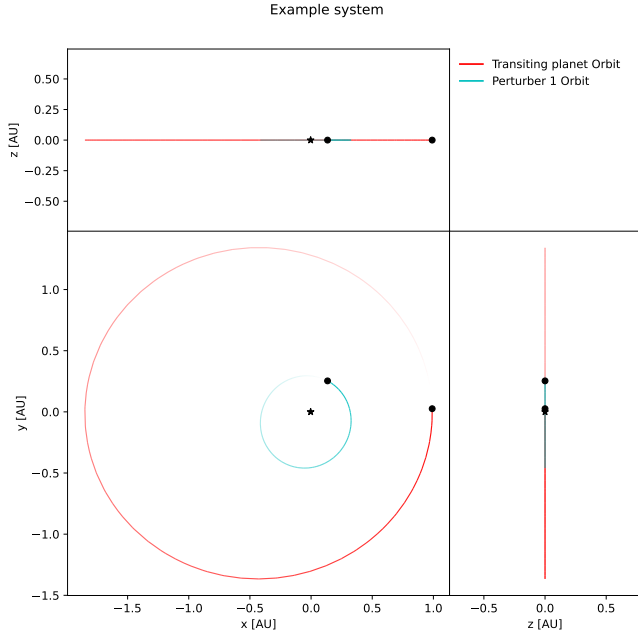


Figure 18. System layout for the best-fit model parameters for the synthetic system. Note how the found planet roughly correlates to the second transiting planet in the system.

3.5 Discussion

3.5.1 Evaluation

We have successfully demonstrated the accuracy of analytical transit timing variation models as compared to TTV signals, seen in figure 15, as well as the potential for this methodology to be applied to determining the parameters of TTV systems, seen in figure 18. We have also demonstrated the capabilities of our computational pipeline, and the possibilities for further work that could be implemented using it.

Therefore, although it has not yet been demonstrated for a known exoplanetary system, the method and computational pipeline outlined in this paper should be capable of producing similar results when using real transit timing variation data. Further study in this field would allow this to be proven, and further refinements should reduce the computational time required to perform each analysis.

3.5.2 Limitations

One of the primary goals was the observation of exoplanetary transits, with six observation windows planned at the beginning of the project. One of those observations was severely hampered by cloud cover, and four others had to be cancelled.

While part of this project focused on finding the globally optimal solution, this does not necessarily always exist. Due to the highly chaotic nature of gravitational interactions, multiple systems layouts can provide identical or near identical TTV signals. This is best demonstrated in the detection paper for kepler-19c, (see Ballard et al. 2011, fig. 14), where a single TTV signal had no fewer than eight possible configurations for the orbit of an extra planet.

This demonstrates a possible drawback of this method, and would require additional constraints based on complementary observations

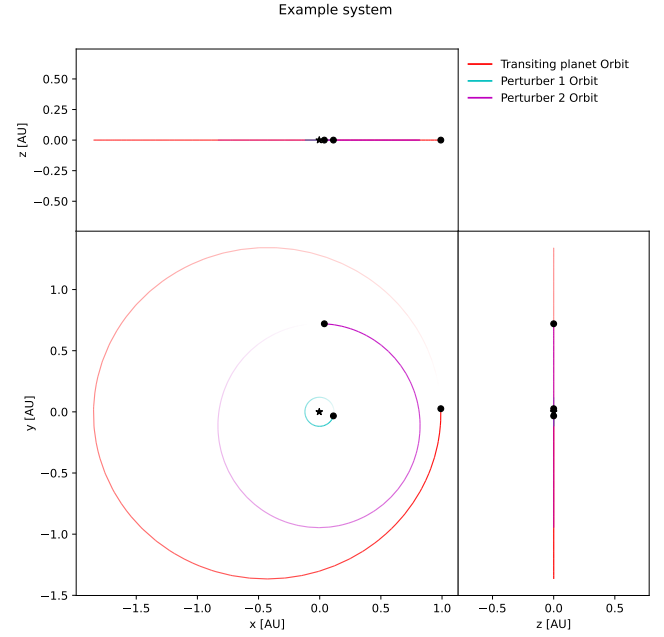


Figure 19. System layout for the second-best-fit model parameters for the synthetic system. Note how both of the found planets have slightly larger orbits than the initial system layout in figure 13, due to the lower μ found by the optimiser, as seen in table 7.

for any tentative detections made using transit timing variations to be considered.

The best-fit for the synthetic system, given in table 7, gave only a single perturbing planet as the maximum likelihood, as opposed to the two that were used to initialise the simulation. This is due to the outer perturbing planet dominating the TTV signal, as seen in figure 14, due to its larger semi-major axis. This demonstrates the second drawback of this methodology, and ties into the statement about non-unique TTV solutions: the additional effect on TTV due to more than one perturber can become negligible.

This could be circumvented by fitting for a TTV cause to find the most dominant planetary effect, and then subtracting that from the known TTV signal. This would leave only the effects of other planets in the system, which could be fed back into the fitting pipeline to determine the parameters of the second perturber, and so on. This would, however, require a longer integration time for the computational pipeline.

As an additional point, we see the second-best fit given in table 7 does pick out the second planetary signal. The higher Akaike information criterion for this model is almost singularly decided by the model having 9 free parameters, as opposed to 4. Further research may benefit from re-evaluating the information criterion used. A disparity in our implementation may come from the assumptions made when determining the number of free parameters for each model: The original AIC assumed each free parameter contributed an additional polynomial power, while our paper assumed each orbital element contributed an additional free parameter.

One perturbing planet with broken pipeline					
Model	Method	μ (10^{-6})	a (10^{-6} AU)	e	ω
Model 1	diff. evo.	1.035	96.079	0.328	5.544
Model 1	dual an.	6.705	492.195	0.247	4.653
Model 2	diff. evo.	1.87	16.48	0.05	-1.199
Model 2	dual an.	0.0	1000000.0	0.0	-6.283
Model 3	diff. evo.	7.159	25.489	0.721	0.331
Model 3	dual an.	230.674	32.877	0.95	-1.076

Table 8. Best fit parameters for the synthetic system TTV with only a single perturbing planet, generated with a version of the fitting pipeline that was not working correctly.

Two perturbing planets with broken pipeline					
Model	Method	μ (10^{-6})	a (10^{-6} AU)	e	ω
Model 1	diff. evo.	4.25	6.047	0.689	-0.429
		5.191	13.68	0.682	3.073
Model 1	dual an.	10000.0	1000000.0	0.089	3.107
		0.0	1000000.0	0.168	3.556
Model 2	diff. evo.	8.378	5.079	0.334	-4.329
		3.081	3.842	0.315	3.798
Model 2	dual an.	142.003	1470.238	0.904	-6.245
		135.908	363.211	0.446	-2.719
Model 3	diff. evo.	1.242	91.763	0.345	3.337
		69.118	1.51	0.774	-0.539
Model 3	dual an.	9566.976	1749.583	0.946	-6.156
		44.774	39782.621	0.021	-1.303

Table 9. Best fit parameters for the synthetic system TTV with two perturbing planets, generated with a version of the fitting pipeline that was not working correctly.

3.5.3 Errors

A sample output was generated while writing this paper to show some problems associated with the TTV fitting pipeline during its development. For this broken state, Figure 20 shows the best-fit TTV curves generated by model 1, and tables 8, 9, and 10, show the output parameters found.

We can see that some magnitudes predicted are on the order of 10^9 seconds, and some did not generate a TTV at all. This is clearly incorrect, and there are a few reasons as to why.

First, while the distances and times in the initial data *should* have been in astronomical units and years when loaded, the simulation pipeline used metres and days. A smaller issue was seen in the parameter optimisation output, where many values for μ are zero and many values for e are close to 1. To align the planets with values seen in real exoplanets, we modified the eccentricity selection to prefer lower values and defined a lower bound on μ at 10^{-9} . As a reference, the lowest planet to star mass ratio of any discovered planet is PSR B1257+12 b at $4.092 * 10^{-9}$.

Additionally, many of the ‘best fit’ solutions gave planets whose orbits are near identical. This solution, while technically satisfying the minimisation criterion, does not aptly describe physical planetary systems, and so an additional factor must be introduced when determining maximum likelihood.

Three perturbing planets with broken pipeline						
Model	Method	μ (10^{-6})	a (10^{-6} AU)	e	ω	
Model 1	diff. evo.	14.363	5.521	0.713	3.778	
		26.087	1.68	0.07	-0.992	
		54.047	1.908	0.472	0.803	
Model 1	dual an.	373.562	1019.304	0.09	4.129	
		460.821	3787.52	0.63	-0.409	
		429.981	4730.766	0.856	5.576	
Model 2	diff. evo.	0.031	651.633	0.438	1.067	
		3.066	8.581	0.609	1.124	
		667.252	0.164	0.076	-1.809	
Model 2	dual an.	10876.298	19380.767	0.948	0.079	
		2263.278	116193.389	0.91	-0.201	
		52.942	199344.612	0.38	-3.02	
Model 3	diff. evo.	0.933	5.807	0.346	1.041	
		13.972	0.649	0.606	5.087	
		0.0	25.101	0.786	-0.541	
Model 3	dual an.	1963.766	49260.785	0.95	-3.42	
		2002.719	33767.03	0.948	-0.245	
		1268.371	5562.579	0.935	-3.162	

Table 10. Best fit parameters for the synthetic system TTV with three perturbing planets, generated with a version of the fitting pipeline that was not working correctly.

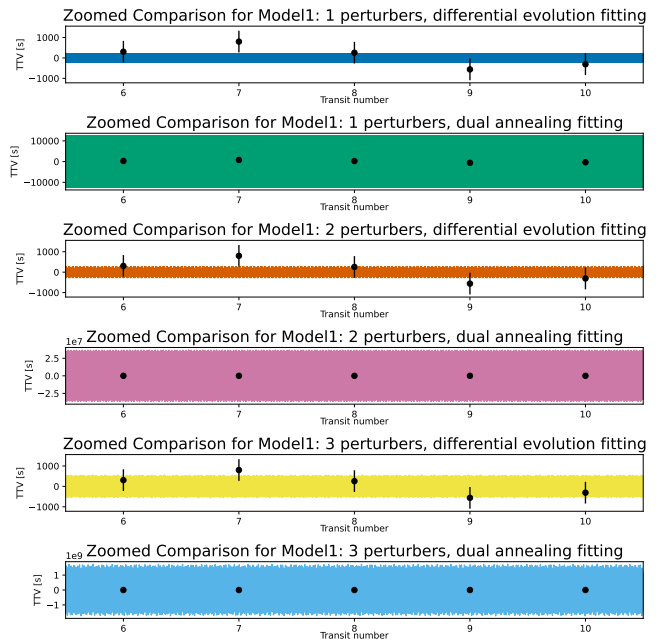


Figure 20. Fitted TTV Curves for the six fitting sets for model 1, generated with a version of the fitting pipeline that was not working correctly, where the number of perturbing bodies and the fitting method are given. Simulated TTV times are overlaid in black.

$$r_{H,min} = a(1 - e) \sqrt[3]{\frac{m}{3M}} \quad (59)$$

$$r_{H,max} = a(1 + e) \sqrt[3]{\frac{m}{3M}}$$

The Hill sphere of a celestial body is the region over which it dominates the attraction of satellites, as compared to its orbital parent.

The radii of the hill sphere are given in equation 59, where M is the mass of the primary body, and m , a , and e are the mass, semi-major axis, and eccentricity of the secondary body respectively. As the extent of the hill sphere is dependent upon the distance between the body and its parent, we require two expressions for the maximum and minimum radii. We can use this to define the boundaries of a region of instability that other planets in the system cannot cross, given in equation 60, (see Jones et al. 2006, eq. 7), where n is some additional term that increases the effective size of the hill sphere, as orbital perturbations in the immediate vicinity can cause bodies initially outside the boundary region to intersect with it. While this is not a strict requirement for long-term stability, as the orbits of Pluto and Neptune show, we will not consider those exoplanets with intersecting orbits.

$$\begin{aligned} \text{Boundary}_{inner} &= a(1-e) \left[1 - n \sqrt[3]{\frac{m}{3M}} \right] \\ \text{Boundary}_{outer} &= a(1+e) \left[1 + n \sqrt[3]{\frac{m}{3M}} \right] \end{aligned} \quad (60)$$

4 CONCLUSION

The aim of this project was to increase historical transit data by observing exoplanet transits, to develop a set of analytical TTV models to approximate transit timing variation, and to create a computational pipeline for TTV analysis that should be capable of determining the best fit parameters for those models. On all three accounts, this project has been successful, with observations logged to the ExoClock database to be included in their updated ephemerides paper later this year, the models as derived in the methodology section of this paper, and a well documented codebase available in this project's GitHub repository. As mentioned in sections 3.1 and 3.5.2, only 2 transits could be observed, less than originally intended.

We have demonstrated that the analytical models are a valid descriptor for TTV signals, provided they are used within the bounds of their approximation, and that the computational pipeline is adequate for fitting those models to data. As the main programming objectives have been met, further development with the computational pipeline would focus on code optimisation and refactoring.

Although real TTV data was used to test the computational pipeline and models, no analysis was performed on them due to time constraints, especially the length of time required to fit light curves to *TESS* data when using the Gaussian process. Further research would build on this project and perform analysis of real planetary systems, as we have demonstrated the methodology to be sound.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Steve Futcher, for his excellent and continued guidance throughout this project. From steering me towards the ExoClock project, to providing excellent conversation through long observations at the observatory, and guidance with the work I've produced. I would also like to thank Prof. Daniel Thomas and Dr. Hooshyar Assadullahi, for their own invaluable support and guidance throughout this project, which most certainly wouldn't have

been possible without the combined input of all three of my supervisors.

I would also thank the Hampshire Astronomical Group for the use of their equipment (primarily their telescopes) that not only made a large portion of this project possible, but allowed this project to be offered at all. I would also mention my thanks to the ExoClock community. Through them, I have been able to both log my own observation for the scientific community at large, and have been able to source high quality data for my own usage.

Simulations in this paper made use of the REBOUND N-body code (Rein & Liu 2012). The REBOUNDx package was used to incorporate additional physics (Tamayo et al. 2020). The simulations were integrated using IAS15, a 15th order Gauss-Radau integrator (Rein & Spiegel 2015).

Light curve analysis in this paper made use of the Juliet package (Espinoza et al. 2019), which provided an interface to the following packages and methods: Transit fits were performed using Batman (Kreidberg 2015); Gaussian processes were performed using Celerite (Foreman-Mackey et al. 2017); MCMC sampling was performed using Dynesty (Speagle 2020). Additionally, sample limb-darkening coefficients used the method outlined by Kipping et al. (Kipping 2013), and uninformative samples for radii and impact parameters used those outlined by Espinoza et al. (Espinoza 2018).

This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. This paper also includes data collected with the *TESS* mission, obtained from the MAST data archive at the Space Telescope Science Institute (STScI). Funding for the *TESS* mission is provided by the NASA Explorer Program. STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–26555.

DATA AVAILABILITY

The datasets were derived from sources in the public domain:

- ExoClock – <https://www.exoclock.space/>, (Kokori et al. 2021, 2022)
- Exoplanet Archive – <https://exoplanetarchive.ipac.caltech.edu/> (Akeson et al. 2013).
- Exoplanet Transit Database – <https://var2.astro.cz/ETD/> (Poddany et al. 2010)
- *TESS* MAST – <https://archive.stsci.edu/missions-and-data/tess> (Ricker 2014)

Data for this paper were sourced from the NASA Exoplanet Archive and *TESS* MAST using the astroquery package (Ginsburg et al. 2019).

Additionally, the data and software underlying this project are available in GitHub at <https://github.com/SK1Y101/TransitProject>.

REFERENCES

- Agol E., Deck K., 2016, Transit Timing to First Order in Eccentricity, *ApJ*, **818**, 177
- Agol E., Fabrycky D. C., 2018, Transit-Timing and Duration Variations for the Discovery and Characterization of Exoplanets, in Deeg H. J., Belmonte J. A., eds., *Handbook of Exoplanets*. Springer International Publishing AG, part of Springer Nature, p. 7, doi:10.1007/978-3-319-55333-7_7
- Agol E., Steffen J., Sari R., Clarkson W., 2004, On detecting terrestrial planets with timing of giant planet transits, arXiv e-prints, pp astro-ph/0412032
- Akaike H., 1974, A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, **19**, 716
- Akaike H., 1992, *Information Theory and an Extension of the Maximum Likelihood Principle*. Springer New York, New York, NY, pp 610–624, doi:10.1007/978-1-4612-0919-5_38, https://doi.org/10.1007/978-1-4612-0919-5_38
- Akeson R. L., et al., 2013, The NASA Exoplanet Archive: Data and Tools for Exoplanet Research, *PASP*, **125**, 989
- Bakos G. Á., et al., 2009, HAT-P-13b,c: A Transiting Hot Jupiter with a Massive Outer Companion on an Eccentric Orbit, *ApJ*, **707**, 446
- Ballard S., et al., 2011, The Kepler-19 System: A Transiting 2.2 R_{\oplus} Planet and a Second Planet Detected via Transit Timing Variations, *ApJ*, **743**, 200
- Borkovits T., Érdi B., Forgács-Dajka E., Kovács T., 2003, On the detectability of long period perturbations in close hierarchical triple stellar systems, *A&A*, **398**, 1091
- Brahm R., et al., 2018, HATS-43b, HATS-44b, HATS-45b, and HATS-46b: Four Short-period Transiting Giant Planets in the Neptune-Jupiter Mass Range, *AJ*, **155**, 112
- Branch M., Coleman T., Li Y., 1999, A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems, *SIAM Journal on Scientific Computing*, **21**
- Byrd R. H., Lu P., Nocedal J., Zhu C., 1995, A Limited Memory Algorithm for Bound Constrained Optimization, *SIAM Journal on Scientific Computing*, **16**, 1190
- Charbonneau D., Brown T. M., Latham D. W., Mayor M., 2000, Detection of Planetary Transits Across a Sun-like Star, *ApJ*, **529**, L45
- Espinoza N., 2018, Efficient Joint Sampling of Impact Parameters and Transit Depths in Transiting Exoplanet Light Curves, *Research Notes of the American Astronomical Society*, **2**, 209
- Espinoza N., Kossakowski D., Brahm R., 2019, juliet: a versatile modelling tool for transiting and non-transiting exoplanetary systems, *MNRAS*, **490**, 2262
- Foreman-Mackey D., 2016, corner.py: Scatterplot matrices in Python, *The Journal of Open Source Software*, **1**, 24
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, emcee: The MCMC Hammer, *PASP*, **125**, 306
- Foreman-Mackey D., Agol E., Angus R., Ambikasaran S., 2017, Fast and scalable Gaussian process modeling with applications to astronomical time series, *AJ*, **154**, 220
- Ginsburg A., et al., 2019, astroquery: An Astronomical Web-querying Package in Python, *AJ*, **157**, 98
- Goodman J., Weare J., 2010, Ensemble samplers with affine invariance, *Communications in Applied Mathematics and Computational Science*, **5**, 65
- Hastings W. K., 1970, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, **57**, 97
- Holman M. J., Murray N. W., 2005, The Use of Transit Timing to Detect Terrestrial-Mass Extrasolar Planets, *Science*, **307**, 1288
- Hurvich C. M., Tsai C.-L., 1993, A CORRECTED AKAIKE INFORMATION CRITERION FOR VECTOR AUTOREGRESSIVE MODEL SELECTION, *Journal of Time Series Analysis*, **14**, 271
- Jones B. W., Sleep P. N., Underwood D. R., 2006, "Habitability of Known Exoplanetary Systems Based on Measured Stellar Properties", *ApJ*, **649**, 1010
- Kipping D. M., 2013, Efficient, uninformative sampling of limb darkening coefficients for two-parameter laws, *MNRAS*, **435**, 2152
- Kokori A., et al., 2021, ExoClock project: an open platform for monitoring the ephemerides of Ariel targets with contributions from the public, *Experimental Astronomy*,
- Kokori A., et al., 2022, ExoClock Project. II. A Large-scale Integrated Study with 180 Updated Exoplanet Ephemerides, *ApJS*, **258**, 40
- Kreidberg L., 2015, batman: BASIC Transit Model cALculation in Python, *Publications of the Astronomical Society of the Pacific*, **127**, 1161
- Malhotra R., 1993, Three-Body Effects in the PSR 1257+12 Planetary System, *ApJ*, **407**, 266
- Murray C. D., Dermott S. F., 1999, *Solar system dynamics*
- Poddaný S., Brát L., Pejcha O., 2010, "Exoplanet Transit Database. Reduction and processing of the photometric data of exoplanet transits", *New Astron.*, **15**, 297
- Rein H., Liu S. F., 2012, REBOUND: an open-source multi-purpose N-body code for collisional dynamics, *A&A*, **537**, A128
- Rein H., Spiegel D. S., 2015, IAS15: a fast, adaptive, high-order integrator for gravitational dynamics, accurate to machine precision over a billion orbits, *MNRAS*, **446**, 1424
- Ricker G. R., 2014, "The Transiting Exoplanet Survey Satellite Mission", *Joint American Association of Variable Star Observers*, **42**, 234
- Schwarz G., 1978, Estimating the Dimension of a Model, *The Annals of Statistics*, **6**, 461
- Speagle J. S., 2020, DYNesty: a dynamic nested sampling package for estimating Bayesian posteriors and evidences, *MNRAS*, **493**, 3132
- Storn R., Price K., 1995, Differential Evolution: A Simple and Efficient Adaptive Scheme for Global Optimization Over Continuous Spaces, *Journal of Global Optimization*, **23**
- Tamayo D., Rein H., Shi P., Hernandez D. M., 2020, REBOUNDx: a library for adding conservative and dissipative forces to otherwise symplectic N-body integrations, *MNRAS*, **491**, 2885
- Tsiaras A., 2021, HOPS: a user-friendly data analysis software to open exoplanet research, in *European Planetary Science Congress*. pp EPSC2021–602, doi:10.5194/esp2021-602
- Virtanen P., et al., 2020, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, *Nature Methods*, **17**, 261
- Wang S., Wu D.-H., Barclay T., Laughlin G. P., 2017, Updated Masses for the TRAPPIST-1 Planets, arXiv e-prints, p. arXiv:1704.04290
- Xiang Y., Sun D., Fan W., Gong X., 1997, Generalized simulated annealing algorithm and its application to the Thomson model, *Physics Letters A*, **233**, 216
- Zhu C., Byrd R. H., Lu P., Nocedal J., 1997, Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization, *ACM Trans. Math. Softw.*, **23**, 550–560

APPENDIX A: LIST OF FIGURES

- 1 Exoplanet location to observed light curve
- 2 Exoplanet inclination to transit length
- 3 TTV due to interior perturbation
- 4 TTV due to exterior perturbation
- 5 Three classes of TTV
- 6 De-trended transit curve for HAT-P-13b, observed 2022-02-24
- 7 De-trended transit curve for HAT-P-13b, observed 2022-02-27
- 8 De-trended transit curve for HAT-P-13b, observed 2022-02-27, without removed data points
- 9 Combined light curve and TTV data for HATS-46
- 10 Combined light curve and TTV data for Wasp-8
- 11 TRAPPIST-1 System Layout
- 12 TRAPPIST-1b simulated TTV
- 13 System layout for one of the synthetic systems
- 14 TTV Residuals for the synthetic system
- 15 Analytical TTV given synthetic system initial parameters

- 16 Best fit curves for the 18 models with the synthetic data set
- 17 Corner-plot for best-fit model output
- 18 Best fit system layout
- 19 Second Best fit system layout
- 20 Fitted curves for the six fitting sets for model 1 for the synthetic system, generated with broken fitting pipeline.

APPENDIX B: LIST OF TABLES

- 1 Comparison between literature and computed transit data for HATS-46b
- 2 Synthetic system parameters
- 3 Best fit parameters for the synthetic system TTV with only a single perturbing planet.
- 4 Best fit parameters for the synthetic system TTV with two perturbing planets.
- 5 Best fit parameters for models for the synthetic system
- 6 Information criterion for fitting models for the synthetic system
- 7 Comparison between best fit and initial parameters for the synthetic system
- 8 Best fit parameters for the synthetic system TTV with only a single perturbing planet, generated with a version of the fitting pipeline that was not working correctly.
- 9 Best fit parameters for the synthetic system TTV with two perturbing planets, generated with a version of the fitting pipeline that was not working correctly.
- 10 Best fit parameters for the synthetic system TTV with three perturbing planets, generated with a version of the fitting pipeline that was not working correctly.

APPENDIX C: EXTRA MATERIAL

C1 Transit duration derivation

As demonstrated in figure 1, a transit begins the moment the planetary disk overlaps the stellar disk, reaches maximum occlusion when the centres of the star and planet are aligned, and ends once the planetary disk no longer overlaps the stellar disk. The centre-to-centre distance between the exoplanet and star at transit ingress and egress is thus the sum of their angular radii,

$$s_{ingress} = s_{egress} = R_{star} + R_{planet} \quad (C1)$$

As a transit occurs between ingress and egress, the total angular distance the exoplanet must cover is simply twice this,

$$s_{transit} = 2R_{star} + 2R_{planet} \quad (C2)$$

If we assume the exoplanets orbit is large compared to the distance travelled during a transit, this linear distance will approximately equal the distance travelled by the exoplanet during the transit. If we additionally assume the orbit of the exoplanet is circular, then the tangential velocity will remain constant,

$$v = \text{const.} = \sqrt{\frac{\mu}{a}} = \sqrt{\frac{G(m_{star} + m_{planet})}{a}} \quad (C3)$$

The total duration of the transit in this case is given from the standard equations of motion,

$$T_{transit} = \frac{s}{v} = \frac{2R_{star} + 2R_{planet}}{v_{planet}} \quad (C4)$$

Which is the expression for transit duration given in equation 2.

As a useful stepping stone, we instead consider the transit distance as that of a circle arc. The radius of this circle is the semi-major axis, a , and the angle swept by the transit, β , in radians. Thus, the transit time is given,

$$T_{transit} = \frac{a\beta}{v} \quad (C5)$$

The swept angle, β , is also the angle swept by an isosceles triangle of base $2R_{star} + 2R_{planet}$ and sides a , as demonstrated by observing the transit from above. The angle, β , can therefore be given by the law of cosines,

$$c^2 = a^2 + b^2 - 2ab \cos \gamma \quad (C6)$$

Which, if written using our orbital parameters, becomes

$$(2R_{star} + 2R_{planet})^2 = 2a^2(1 - \cos \beta) \quad (C7)$$

Rearranging gives,

$$1 - \cos \beta = \frac{(2R_{star} + 2R_{planet})^2}{2a^2} \quad (C8)$$

This can be reformulated from trigonometric identity, $1 - \cos x = 2 \sin^2 x/2$,

$$2 \sin^2 \frac{\beta}{2} = \frac{(2R_{star} + 2R_{planet})^2}{2a^2} \quad (C9)$$

Which simplifies as,

$$\sin \frac{\beta}{2} = \frac{R_{star} + R_{planet}}{a} \quad (C10)$$

obtaining the expression for β ,

$$\beta = 2 \arcsin \frac{R_{star} + R_{planet}}{a} \quad (C11)$$

Which gives the transit duration,

$$T_{transit} = \frac{2a}{v} \arcsin \frac{R_{star} + R_{planet}}{a} \quad (C12)$$

This of course exactly matches the expression given in equation 3. As any given exoplanet may not transit exactly aligned with the equator of the stellar disk due to its relative inclination, the distance covered during a transit must also account for this.

There are two orbital elements that describe the orientation of a

circular orbit, the inclination, i , and longitude of the ascending node, Ω . As we do not have a defined “zero angle” for an arbitrary exoplanet, these will be described as their relative inclination, defined as zero when the planet transits the stellar equator, and relative longitude of the ascending node, defined as zero when perpendicular to our line of sight.

The offset from the stellar equator from our point of view is thus,

$$b = a \cos i \sin \Omega \quad (\text{C13})$$

The transverse distance of the transit is then described by a right-angled triangle of hypotenuse $R_{star} + R_{planet}$ and height b , which can be recovered from Pythagorean theorem,

$$s_{transverse}^2 + b^2 = (R_{star} + R_{planet})^2 \quad (\text{C14})$$

which gives the transverse distance as,

$$s_{transverse} = \frac{s_{transit}}{2} = \sqrt{(R_{star} + R_{planet})^2 - b^2} \quad (\text{C15})$$

following the same derivation between equations C7 and C11 gives the modified swept angle of the transit as,

$$\beta_b = 2 \arcsin \frac{\sqrt{(R_{star} + R_{planet})^2 - b^2}}{a} \quad (\text{C16})$$

Which gives a transit duration of,

$$T_{transit} = \frac{2a}{v} \arcsin \frac{\sqrt{(R_{star} + R_{planet})^2 - b^2}}{a} \quad (\text{C17})$$

which is of course our expression given in equation 5.

This paper has been typeset from a \TeX/L\AA\TeX file prepared by the author.