Contrast Analysis of TTV-TDV Phase Difference Induced by Possible Exomoons

Team Exouls

Internal Guide: **Mr. Sundar**

Team Exoplanets

Internship and Projects Division

Society for Space Education Research and Development

October, 2020

The project report submitted in fulfilment of progress requirements for the 6-week research internship at SSERD.

To The Avengers

You know, for saving the world.

Exoulmates

1. §Aayushi Rajgor,1612035, Dept. of Electronics Engineering, K. J. Somaiya College of Engineering, Mumbai

2. †Ananya Shankar, 184702, Gargi College, New Delhi

3. †§Ayan Ahmad, 11191644, Dept. of Applied Physics, Delhi Technological University

4. §Binayyak B. Roy (Assistant team leader), Dept. of Physics, St. Stephen's College, University of Delhi

5. §Charmi Bhatt, Swarrnim Startup and Innovation University, Gandhinagar, Gujarat

6. †Junik Sengupta (Team leader), University of Calcutta , Calcutta , West Bengal

7. †Kedar Kulkarni, Priyadarshini College of Engineering, Nagpur

8. †§Pulkit Ojha, 1811118, SPS, NISER, Bhubaneswar, Odisha

9. Radha Mittal, Amity University Mumbai, Mumbai, Maharashtra

10. †Shreya Dey , 001820702010, Jadavpur University, Kolkata, West Bengal

11. Smit Macwan, St. Xavier's College, Ahmedabad, Gujarat

12. §Sneha C.A, 180393, Sri Darmasthala Manjunatheswara College, Dakshina Kannada

---

† Theoretical and Mathematical Analysis
§ Data Collection and Analysis

Acknowledgements

It is our earnest endeavour to express our gratitude on being assisted to corroborate our cause in the brevity of time. Working together as a team has been a wonderful and wholesome experience. Being benefited by the august and esteemed guidance of our mentor, M N Sundar, Jain University, Bangalore and coordinators, Prateek Boga and Anisha, has instilled the requisite vigour and dedication in us, to make our vision a reality. It has really been an honour to have our concerns answered by Chris Fox and Paul Wiegert, exoplanetary scientists from Department of Physics and Astronomy at The University of Western Ontario, Canada.

We would like to express our heartfelt gratitude towards Mahesh P for arranging the technical aids and technical talks by Pavan Kumar which en-compassed topics from training sessions on how to determine the authenticity of paper, writing papers in Latex to important cues while making career choices, has indeed helped us immensely in entirety. Concise and apt inputs from SSERD IPD B2 member Rutuja Attal has also been an important conduit for our work on exoplanets. Most importantly, we would like to thank SSERD, especially Nikhitha C and Sujay Sreedhar for providing us with this opportunity of cohorting with like-minded individuals.

Abstract

In our short duration research project, our team worked upon a target of understanding & using the theoretical model of the gravitational effects due to the presence of a satellite in the exo-planetary system using **variations in the transit signals (TTV and TDV) of the target exoplanet**, developed back in 2009. Using the model, we derived the expressions for the **instantaneous variations of the signals** and used them to confirm theoretically that the phase difference between the above mentioned signals for the possible exomoon candidates should be $\frac{\pi}{2}$. Further, we used this criteria along with some other parametric constraints like **SNR value** and the **Standard deviation** to short-list the possible candidates to check for the further possibility of the presence of an exomoon. Working with 2016 data set shows that only **1.9% of the candidates** have the quoted phase difference, we've tried to draw some theoretical conclusions which can be used to further analyse the reason behind the discrepancy between the prediction and the data-analysis. We're also eagerly waiting to try the data-analysis process with the latest data-set to gain better insights in the outcome.

Contents

List of Figures

List of Tables

Introduction

1.1 | Motivation

Ever since humanity has looked up into the night sky, we have wondered , are we alone in this universe? Or is there life beyond Earth? This question has bothered us since ages. It is the first time in human history, that now we have the technological means to find the answer of this question. Out of many approaches one is to find another Earth like planet; a planet with just the right temperature, density, atmospheric compositions and albedo which makes planet potentially habitable.

Since decades, scientists and researchers across the globe have conducted many missions like Hubble, Spitzer, TESS and mainly Kepler mission for Detection of Exoplanets. There are many upcoming missions lined up for next decade which will give us insights into the details about the exoplanets.

NASA's Kepler telescope was launched in the year of 2009 with an aim to facilitate the search for possible earth-like planets and since then the research has moved to a higher accelerated scale. All missions till date have taken the mark of confirmed exoplanets to more than 4000.

Apart from exoplanets there exists another possibility to find habitable conditions—Exomoons. Along with the fact that moon increases the habitability of planet , moons themselves can be habitable. For Example, the moon Europa has more ambient conditions that that of Jupiter.

Many inquisitive minds on this planet have contributed to the research on Exomoons. After the release of Kepler data, Exomoon research has gone up the hill. For this project, we tilt our curiosity towards Exomoons, its detection and theoretical model.

1.2 | Aims and Objectives

Till date there has been no concrete evidence for the existence of Exomoon candidates. There are multiple ways to look for them: the major one being photometric observations but they

produce extremely faint signal making it invisible to detectors. So we turn towards the signal produce by gravitational effects of Exomoon on Exoplanets.

The observations are mostly made via the NASA's Kepler space telescope, which was designed to detect a photometric signal corresponding to a **transit depth of 85 ppm—equal to the Earth-Sun analog**. Here, we target on possible Exomoons which are not detectable directly, but by these indirect signals.

The signals that we have focused on are the Transit Timing Variations and the Transit Duration variations. Our aim is to provide a better version of the theoretical model and sinusoidal fitting of TTV and TDV signals that can set well together, explaining the discrepancies which will be discussed in following sections 2.2.

<div align="right">**2**</div>

Background & Literature Overview

2.1 | Detection of Exo-moon

Given the large number of moons we have in our solar system, extrasolar moons are highly believed to exist. In pursuit of detecting them, many methods can be used like transit method Nesvornỳ et al. (2012), direct imaging method Peters and Turner (2013), and Gravitational microlensing method Ranc et al. (2019). No search to date has made a confirmed identification of an exomoon. With existing data from previous missions like Kepler Space Telescope, Exomoons candidates can be found by analyzing transits, specifically TTVs and TDVs, which are explained in the next section in detail. It is important to note that these satellites make a small impact in already faint signals of Exoplanets, which makes them even harder to detect Kipping (2009)

2.1.1 | Detection methods

The detection of Exo-moons depends on its size and photometric threshold of the telescope. If the Exo-moons are large enough, they can be detected by their own transit signals itself. There have been previous search missions for exo-moons in the Kepler data. The most sophisticated one is the **HEK (Hunt for Exo-moons with Kepler)**, Nesvornỳ et al. (2012) project, which uses a photodynamical approach to examine the signals of exo-planets closely.

On the other hand, If an exoplanet has a small moon around it, it will produce some perturbations in the orbit of its exoplanet which we can interpret from the data collected by telescopes. These perturbations can come in different forms, like Transit Timing Variations (TTV) and Transit Duration Variations (TDV), and Transit Photometric Variations (TPV).

Here, we explore the possibility of existence of exo-moons which are small enough to not produce their own Transit signal. These exo-moons are too small to create detectable photometric (transit) signals of their own but are large enough to create TTVs in their host planet's transit signal by displacing them with respect to their mutual center of mass Fox and Wiegert (2020)

2.1.2 | Exo-moon vs. Additional Planet Hypothesis

TTV Signals produced by Exo-moons (if present) are exactly reproducible by the additional non-transiting planet. To better understand the TTVs and TDVs in our data, we use **Bayesian inference**, where **chi-square fitting** helps to determine if the signal better fits the Exomoon hypothesis or Additional Planet.

2.2 | **Phase difference**

2.2.1 | Theoretical prediction

Transit Timing Variation signal produces a degeneracy in the measurement of the mass of the satellite and the orbital separation of the satellite, in the regard that one cannot be calculated without assuming the other. To solve this, a complementary method, using Transit Duration Variation signals was introduced. The TTV and the TDV waveforms are sinusoidal, with the TDV lagging behind the TTV by a phase difference of $\frac{\pi}{2}$. This arises from the fact that TTV is a **spatial effect** and TDV is a **velocity effect**.

If one assumes zero orbital eccentricity of the satellite orbit and that the orbital plane of the planet-moon and the planet-star system is the same as the plane of reference then the TTV signal is a relation between the projected distance between the planet and the planet-moon barycentre, and the component of the orbital velocity of this barycentre around the star in the same direction.

This projected distance, and consequently the **TTV, is a function of the true anomaly of the planet around the planet-moon barycentre**. The true anomaly is an angular parameter that defines the position of a body moving along a Keplerian orbit and varies from 0 to 2π. This relation is modified to include the effects of an **eccentric orbit of the satellite and planet**, and **longitude of the pericentre**—defined as the longitude of the point of closest approach to the planet, considering zero inclination of the orbit of satellite—of the satellite and planet.

Transit Duration Variation(TDV) is the change in the Transit Duration of the planet that is observed over multiple measurements. *This duration of transit is inversely proportional to the projected velocity of the planet on the star, hence making TDV a velocity derived effect.* The TTV and TDV waveforms have been plotted for the GJ436b system, assuming a hypothetical $1 M_{earth}$ exomoon, which indicates a $\frac{\pi}{2}$ phase difference. Kipping (2008)

In Heller, René et al. (2016), the TTV-TDV diagrams of an exoplanet-exomoon system are plotted. The TTV-TDV diagram for an exoplanetary system with one satellite and non zero eccentricity is described as an egg-shaped ellipsoid, wherein the orientation of the ellipse depends
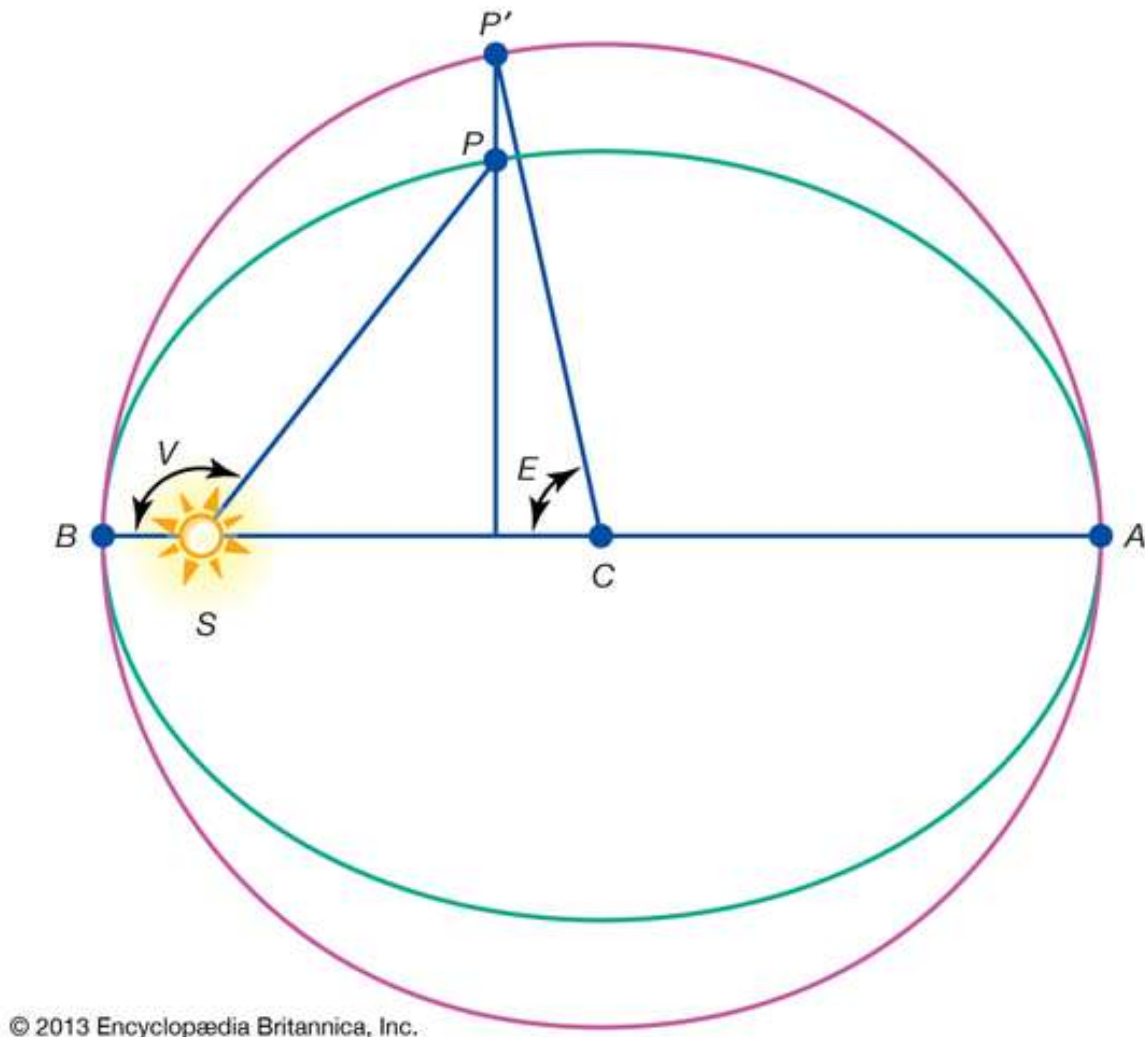
Figure 2.1: True Anomaly of the planet

on the value of the argument of the perihelion. The elliptical shape is derived from the fact that the amplitudes of the two signals differ significantly.

Experimentally, Fox and Wiegert (2020) shortlisted eight systems for testing their chi-square fitting for two hypotheses, an additional planet system, and an exomoon system, in accordance with a parameter space and found that for these systems, the TTV and TDV waveforms did not exhibit a shift in phase difference. This might be attributed to the **high error limit and the limited amount of data points**.

2.2.2 | Data outcome

For the data sources of exo-planetary systems our first go to was the NASA Kepler Exoplanet Archive. However, we started off with analysing the data-sets mentioned in the Holczer et al. (2016). The data-sets were **dated 03/2015 and 09/2016**. In order to get recent data-sets we did further digging in Vizier and Simbad sections of Exo-planetary archives. The recent data-sets had a large number of candidates but did not have all the parameters such as TDV and TDV error which was crucial for the analysis based upon these signals. Even after weeks of searching we could not find recent data-sets with TDV signals. Therefore we moved on with the data analysis of the 2016 data.

After looking at Fox and Weigert's paper we concluded that they could not have found any candidates with a phase difference of $\frac{\pi}{2}$ because they could have eliminated certain candidates via the transit depth being less than 85 ppm condition. That also significantly lowers the SNR because of which number of reliable candidates got lowered further. We wanted to confirm Kipping's theory and therefore **eliminated this transit depth condition** which also gave us significantly higher SNR values.

2.3 | **Mathematical Description**

In this section, let's have a closer look at the mathematical description of the theoretical model developed by David Kipping which have been subesquently followed by exoplanet researchers in the possible search of exomoons.

2.3.1 | Transit timing variation

As we learnt from studying Kipping's model, the RMS value of the TTV amplitude is for the case of circular orbit of satellite is

$$\delta_{TTV} = \frac{a_w}{\sqrt{2}v_{B\perp}} \tag{2.1}$$

where a_w is the radius of satellite around the barycentre of the planet and satellite system and the $v_{B\perp}$ is the velocity of barycentre around the star. Kipping (2009)

This value has been modified to include the eccentricity of the satellite's orbit e_s, position of pericentre of satellite ω_s and the corresponding quantities for the planet's orbit around the star e_p and ω_p.The modified version can be expressed as a function of above-mentioned parameters

in the following way.Kipping (2009)

$$\delta_{TTV} = \frac{1}{\sqrt{2}}\sqrt{a_p}a_s G_M \frac{\zeta_T(e_s\omega_s)}{\gamma(e_p,\omega_p)} \tag{2.2}$$

In the above expression, apart from the $\zeta(e_s\omega_s)$ and $\gamma(e_p,\omega_p)$ factors, all the factors are taken to be constants including G_M which is defined as $G_M = \frac{M_p M_{PRV}^{-1}}{\sqrt{g(M_{star}+M^{-1}PRV)}}$. Depending upon the ratio of periods of both the orbit of satellite and the planet, The values of a_p or a_s will change over a transit but here their RMS values have been taken as constants. *The expression also signifies that the signal is proportional to the product of M_s and M_s.*

$$\delta_{TTV} \propto M_s a_s \tag{2.3}$$

2.3.2 | Transit duration variation

Now, let's examine the expression of TDV signal as derived in the theoretical mode. The basic definition of this TDV starts from the expression $t_T = \frac{X}{v_{P\perp}}$, basically the time duration is the ratio of the transit distance covered by the planet and it's velocity $v_{P\perp} = v_{B\perp} + v_{w\perp}$ which is the velocity of the barycenter of the planet-moon system and the velocity of the planet around the barycentre.

Similar to TTV signals, this expression, when modified to encompass the eccentricity and the position of pericentres of the satellite, the rms expression becomes

$$\delta_{TDV} = \sqrt{\frac{a_p}{a_s}}G_M' \frac{t_T}{\sqrt{2}}\frac{\zeta_D(e_s\omega_s)}{\gamma(e_p,\omega_p)} \tag{2.4}$$

where G_M' stands for a constant factor of $G_M' = \frac{M_s^2}{\sqrt{M_{prv}(M_{prv}+M_{star})}}$. The above expression implies that

$$\delta_{TDV} \propto M_s a_s^{\frac{-1}{2}} \tag{2.5}$$

Now, these two signals can be used in order to figure out the mass and distance of satellite of the possible exomoon candidate.

It must be noted from the above formulas that the phase difference between the signals of TTV and TDV will only be generated due to the functions $\zeta(e_s\omega_s)$ and $\gamma(e_p,\omega_p)$. The fact that

these signals are at a phase difference of $\frac{\pi}{2}$ becomes a significant tool for short-listing candidates and checking their possibility.

3

Methodology

3.1 | Theoretical modelling

The aim of modifying the theoretically modeled transit timing and duration variations is to analyse the possible theoretical sources that give rise to the discrepancy down to have mainly three possible explanations:

- **Theoretical assumptions which hinder theoretical accuracy**
  There are constants in [Kipping (2008)](#) and our mathematical model ([3.1](#), [3.4](#)), denoted by C^T and C^D, can **change considerably** if given a **long enough epoch**. This can complicate mathematical modelling by a significant amount so considering the time constraint of the project, we will be taken them as constants, while mentioning the possibility of their modification in future scopes section.

- **Sensitive parameters not taken into account for theoretical modelling**
  There is always a possibility of sensitive and unnoticed parameters when we talk about far-away systems such as exoplanets. We will be analyzing each non-intrinsic parameter in order to see what role they play in varying transit timing variations and transit duration variations. If we do not see a parameter that can cause a change of phase difference from $\frac{\pi}{2}$, it would indicate the existence of an additional parameter/s.

- **Change of reference epoch to match limitations for data analysis**
  This will be the main focus of theoretical modelling for this report. Change of reference indicates not one but two modifications to previously known theoretical models. Let us discuss the two in brief below.

Epoch unit

The previous models **use transit number as their discrete epoch unit** as it is only at certain time intervals that we are able to notice the exoplanet transit in front of its host star. This is what we mean by limitations for data analysis. It is favorable that we have a **continuous epoch rather than a discrete one** so as to really be able to map the change in TTV and TDV over the

epoch.

The true anomaly of the wobble orbit is what causes the phase of TTV and TDV to shift (equally, therefore phase difference is maintained). It also describes the orientation of the exoplanet with respect to the planet-moon barycenter (our reference for the C component of $O - C$). This suits our conditions for epoch unit very well as it is related to the physical orientation of exoplanet and is also the main factor in changing the phase of transit variations.

Equations of variations

The previous models use RMS equations of variations given by

$$\delta TTV = \sqrt{\frac{1}{2\pi} \int_0^{2\pi} [TTV(f_W)]^2 df_W} \quad \delta TDV = \sqrt{\frac{1}{2\pi} \int_0^{2\pi} [TDV(f_W)]^2 df_W}$$

The integration of the square of $TTV(f_W)$ and $TDV(f_W)$ are non-trivial and require certain assumptions to be taken in order to get an approximate RMS value as done by Kipping (2008) and Kipping (2009). The above expressions also assume one wobble orbit in the transit duration as the integration is being done form $0 \to 2\pi$.

To overcome these assumptions we must derive a set of variation equations that do not require such assumptions and yet compliment the epoch unit as discussed above. In 3.1.1, we will derive our modified instantaneous variation equations which are a function of our selected epoch, true anomaly of exoplanet on wobble orbit.

3.1.1 | Instantaneous variations

The most generalized theoretical model of potentially exomoon-induced transit timing variations and transit duration variations up-to-date comes from Kipping (2008) and Kipping (2009). In order to analyze theoretically, the source of the aforementioned discrepancy in phase difference for the six exomoon candidates considered in Fox and Wiegert (2020), we must first **find an absolute epoch with respect to the planet-moon barycentre** as variations are expected to occur from the model of an exoplanet without an exomoon.

As the TTV and TDV phase is the main focus of our project, we have taken true anomaly, the orientation phase of the exoplanet w.r.t. planet-moon barycenter as the to be the absolute epoch for the theoretical model. Unlike the model used by Kipping (2020) and Fox and Wiegert (2020), we do not use the RMS value of TTV and TDV to check for phase as integrating the square of instantaneous variation equations we obtain in appendix A is non-trivial without assumptions that can hinder the theoretical accuracy.
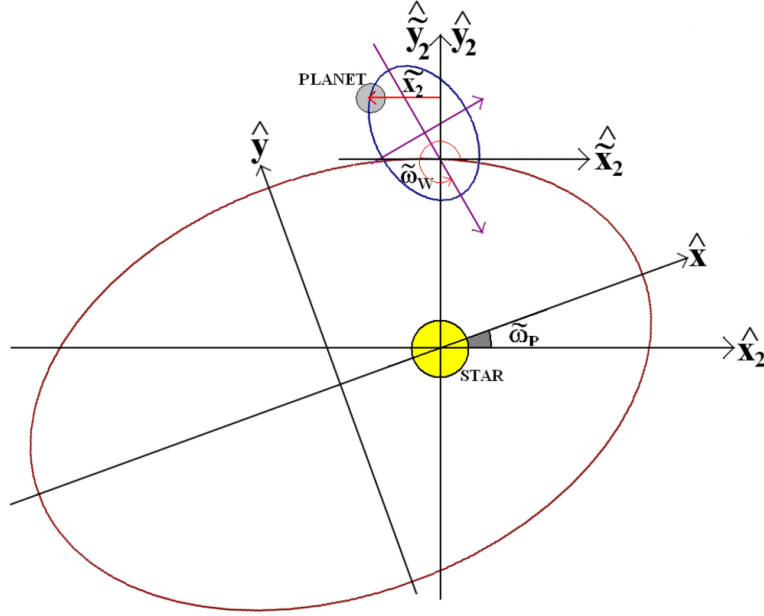
Figure 3.1: Reference diagram for our derivation - from Kipping (2008)

3.1.1.1 | Transit timing variations

The transit timing variation can be expressed as a function of true anomaly as follows

$$TTV(f_W) = \frac{C_1^T}{C_2^T + v_{W\perp}} \frac{cos(f_W + \bar{\omega}_W)}{1 + e_W cos(f_W)} \tag{3.1}$$

where

$$C_1^T = a_W(1 - e_W^2) \quad C_2^T = v_{B\perp} = Y(e_P, \bar{\omega}_P) \sqrt{\frac{G(M_{PRV} + M_*)}{a_P}} \tag{3.2}$$

$$Y(e_P, \bar{\omega}_P) = \cos\left[\arctan\left(\frac{-e_P\cos(\bar{\omega}_P)}{1 + e_P\sin(\bar{\omega}_P)}\right)\right] \sqrt{\frac{2(1 + e_P\sin(\bar{\omega}_P))}{(1 - e_P^2)} - 1} \tag{3.3}$$

where f_W represents the true anomaly of the planet in wobble orbit and $M_{ps} = M_p + M_s$.

The above expression gives the same result as that given by the expressions in Kipping (2008). Hence for TTV, we obtain the same phase and magnitude in our modelling as predicted and used in the paper.

3.1.1.2 | Transit duration variations

The transit duration variation can be expressed as a function of true anomaly as follows

$$TDV(f_W) = C^D\left(\frac{l_r}{1 + v_{r\perp}} - 1\right) \tag{3.4}$$

where

$$C^D = \frac{l_B}{v_{B\perp}} = \frac{\sqrt{(R_* + R_P)^2 - (bR_*)^2}}{v_{B\perp}} \tag{3.5}$$

$$l_r = \frac{l_W}{l_B} = \left(\frac{(R_* + R_P)^2 - (bR_*)^2(1 + \frac{1}{r_P})^2}{(R_* + R_P)^2 - (bR_*)^2} \right)^{\frac{1}{2}} \tag{3.6}$$

$$v_{r\perp} = \frac{v_{W\perp}}{v_{B\perp}} \tag{3.7}$$

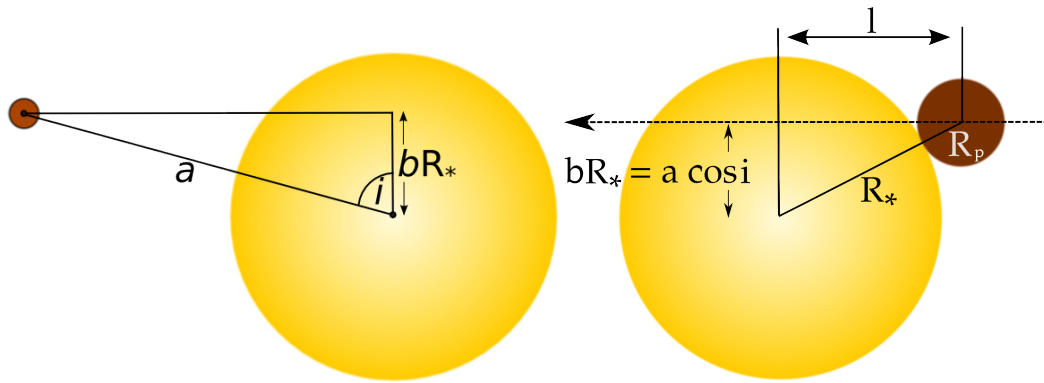where b represents the impact parameter as shown in 3.2.

Figure 3.2: Reference diagrams for impact parameter - from paulanthonywilson.com

The above expression deviates from the formula derived by Kipping (2009). Here, we have not taken transit duration as a parameter as it cannot be considered as a replacement for transit duration of a planet transmitting without an exomoon. So instead, we have replaced it with a theoretical expression for the transit duration of planet lacking moon.

Note that the phase of our TDV does not deviate from Kipping (2009)'s TDV expression, hence our theoretical model too predicts a phase difference of $\frac{\pi}{2}$. The change in magnitude can play a key role in RMS expressions for sinusoidal functions, hence further effecting the plots for TDV.

3.1.2 | Diagrammatic role of parameters

We will set a nomenclature for TDV graphs as we are obtaining a symmetric two point graphs which must be resolved into one per domain using theoretical understanding.

1. **TDV plots**: Separated by node, described by first letter of color, left to right

2. **Phase plots**: Separated by quadrant, described by first letter of color, chronological

Let us look at our reference/starting diagram 3.3 and explain the nomenclature using that.
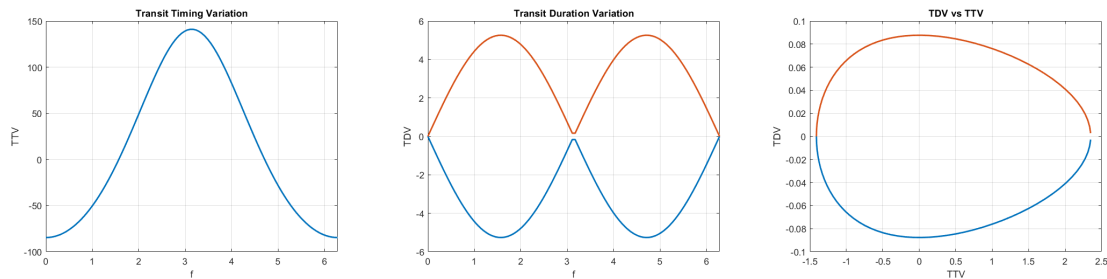

Figure 3.3: [Left: TTV plot] ; [Center: OB TDV plot] ; [Right: $O_2O_1B_4B_3$ Phase plot]

To be able to interpret different data plots and find a possible phase shift causing sensitive parameter, we must first identify the role of reach parameter diagrammatically. In this section we will individually look at the diagrammatic role of non-intrinsic parameters such as the following. The first three plots for each parameter represent a low value taken for it, the next three represent higher, both relative to the reference plot 3.3.

3.1.2.1 | Parameter 1: Barycenter eccentricity (e_P) - (4.1)

Result 1: The amplitude of TTV and TDV are directly proportional to e_P^n.
Result 2: The phase of TTV and TDV are independent of e_P^n.

3.1.2.2 | Parameter 2 and 3: Wobble eccentricity ($e_W = e_S$) - (4.2)

Result 1: The amplitude of TTV and TDV are directly proportional to e_W^n and e_S^n.
Result 2: The phase of TTV and TDV are independent of e_W^n and e_S^n.
Result 3: The rate of change of TTV changes with e_W^n and e_S^n.

3.1.2.3 | Parameter 4: Barycentre semi-major axis (a_P) - (4.3)

Result 1: The amplitude of TTV and TDV are directly proportional to a_P^n.
Result 2: The phase of TTV and TDV are independent of a_P^n.

3.1.2.4 | Parameter 5 and 6: Wobble semi-major axis ($a_S \propto a_W$) - (4.4)

Result 1: The amplitude of TTV is directly and TDV is inversely proportional to a_S^n and a_W^n.
Result 2: The phase of TTV and TDV are independent of a_S^n and a_W^n.
Result 3: The proportionality varifies the relations: $TTV \propto M_s a_S$ and $TDV \propto M_s a_S^{-\frac{1}{2}}$.

3.1.2.5 | Parameter 7: Position of barycenter pericenter ($\bar{\omega}_P$) - (3.4)

Dealing with proportionality of angles can be tricky as for their often symmetric nature. To represent the effect of $\bar{\omega}_P$ we refer to Kipping (2008) so as to see the change in value of Y^{-1}, the function that absorbs barycenter orbit properties, w.r.t. $\bar{\omega}_P$. Higher its value, higher the value for both TTV and TDV.
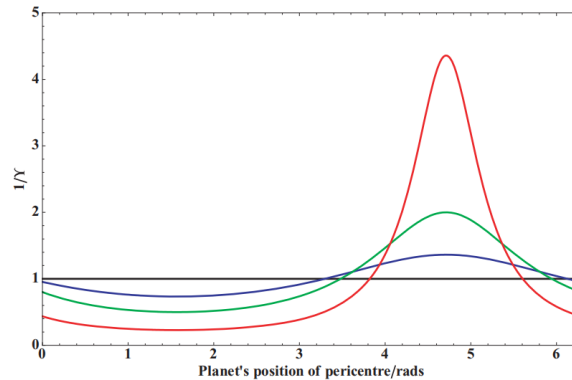

Figure 3.4: Y^{-1} dependency on $\bar{\omega}_P$ for different e_P - from Kipping (2008)

3.1.2.6 | Parameter 8: Position of wobble pericenter ($\bar{\omega}_S$) - (4.5, 4.6)

Result 1: The amplitude of TTV is independent of $\bar{\omega}_S$.
Result 2: The amplitude of TDV is altered per π cycle due to $\bar{\omega}_S$.
Result 3: The phase of TTV and TDV is shifted equally by $\bar{\omega}_S$, maintaining phase difference.

3.1.2.7 | Parameter 9 : Impact parameter (b) - (4.7)

Result 1: The amplitude of TTV is independent of b.
Result 2: The amplitude of TTV is inversely proportional to b^n.
Result 3: The phase of TTV and TDV are independent of b.

3.1.2.8 | Parameter 10 : True anomaly (f_W)

The true anomaly is the argument against which we plot TTV and TDV in our theoretical model. Hence it is trivial that
Result 1: The amplitude of TTV are independent of f_W.
Result 2: The phase of TTV and TDV is shifted equally by $\bar{\omega}_S$, maintaining phase difference.

3.2 | Data analysis

The data analysis done in Fox and Wiegert (2020) was very detailed and therefore we considered to work with that as our basis. Going by our objective, we would have modified the code a bit after having reconfirmed the results in their paper, to shortlist candidates with a $\frac{\pi}{2}$ phase difference. Since Kipping (2008) already suggests that TTV and TDV signals must have a phase difference of $\frac{\pi}{2}$ for exomoon systems. This would help us not only detect probable exomoon candidates but also inspect the fact as to why Fox and Weigert did not find the required phase difference.

3.2.1 | Systems shortlisting

For shortlisting candidates we went forward with the conditions mentioned in Fox and Wiegert's paper:

- Imposing a cut-off of standard deviation less than 40 minutes for the fitted TTV signals.

- Transit depth of 85 ppm or less.

- Finally selecting the ones with the best SNR values greater than 1.0

Using these points we **reconfirmed the results** as reported in paper. We got similar results as were reported in the paper. Next, we tried to add a condition which checks how many systems do show a phase difference of $\frac{\pi}{2}$.

We expected to see a majority of candidates with the required phase difference (excluding the SNR condition) as Kipping's theory doesn't assume anything else which could have been eliminated in our shortlisting conditions and therefore most of them (if exomoons) should be showing such a phase difference.

Next, we shortlisted candidates with just a phase difference of $\frac{\pi}{2}$ and SNR value greater than 1.0 since we did not want to look at systems with noise greater than signal. Even then we did not get a considerable number of candidates. We decided to look at candidates which showed a standard deviation of less than 40 minutes and showed $\frac{\pi}{2}$ phase difference with a tolerance of 0.1. Using these conditions we got 12 candidates 3.1 with which we would be using to find a common link as to why these are the ones which show the required phase difference.

3.2.2 | Plot fitting

For fitting the TTV and TDV data points we went with the least χ-square fit i.e., **global minimisation**. We used the SciPy *curve$_f$it* function module to fit the data points using a general

KepId	Kepoi Name	Kepler Name	Phase Difference	SNR
10187017	K00082.01	Kepler-102 e	-1.4825	2.38507
4914423	K00108.01	Kepler-103 b	1.5966	2.15896
8359498	K00127.01	Kepler-77 b	1.6313	3.2644
8644288	K00137.01	Kepler-18 c	1.5974	2.64098
5735762	K00148.02	Kepler-48 c	-1.57637	2.25042
2987027	K00197.01	Kepler-489 b	-1.52618	2.940218
9390653	K00249.01	Kepler-504 A b	1.6505	2.04999
8292840	K00260.02	Kepler-126 d	-1.65055	1.9133
5792202	K00841.02	Kepler-27 c	-1.505414	2.06893
10272640	K01074.01	Kepler-762 b	1.54297	2.003461
10864656	K01299.01	Kepler-432 b	1.6275	2.0652
12257999	K02577.01		1.63013	1.521138

Table 3.1: Shortlisted Candidates

function of the form $A\sin(Bx + C)$. We fit the TTV and TDV datapoints against the transit number. And calculate the difference in the value of C for both the fits to give the phase difference. Regarding the number data points we started out with just 5 and then increased it to 10 in order to **minimise random noise**. If the number of data points were to be higher than that then because of lack of observations some candidates get eliminated. In order to make sure the least χ-square fit was working properly we first matched the results with that in Fox and Weigert's paper. We got the same results for the KOI mentioned in their paper.

The module was providing certain erroneous results when the dataset provided had a shifted mean in different sections then the module failed to fit it properly because of an optimization error. For this we started considering more number of data points (as mentioned before). Next, we also started facing error with some datasets that could not be properly fitted using the default settings of the module such that maximum number of calls to the function being 800. This was changed to 5000 to get the best fit (most converged) result.

Results

4.1 | Theoretical Parameters Modelling

The results for checking the role of non-instrinsic parameters in TTV, TDV and phase plots are given below and hyperlinked by their title in the previous section.

4.2 | Data Analysis of Exoplanetary Systems

From our shortlisted candidates we fit the TTV and TDV signals upto 10 data points and their results are presented below. 4.8

The fitting in all of these cases seem reasonable as when other methods are used like fitting using higher order polynomials, different algebraic and geometric iterative methods as well, give fairly similar results. Since the fitted curves have different frequencies for TTV and TDV the phase keeps on changing, looking at the magnitude of this change and the speed would give us some information regarding the parameters in the theoretical model that controls it. 4.9

In order to look at a constant phase difference we tried to fit the observations of TTV using the least χ^2 method and use the frequency of that plot to fit TDV but that would be a biased fitting which would lead to wrong results. We also tried to do a **high frequency fitting** as the TTV would vary between transits, and going from the theory of the ratio $\frac{P_s}{P_p}$ the frequency should be greater than 9. Therefore we tried to fit using **the function** $A\sin\{(|B| + 9)x + C\}$. While this does lead to a **lesser variation in the phase difference** but it also brings up newer candidates. The best fitting function would be to use a composite form of the formulas provided by Kipping's theory but that wouldn't be for confirming his theory since we aren't considering any formula independent of his theory, it would just be helpful to find out exomoons given his theory works out fine.

There are a lot of other factors in the curve fitting that can be explored and have been mentioned in the future scope.

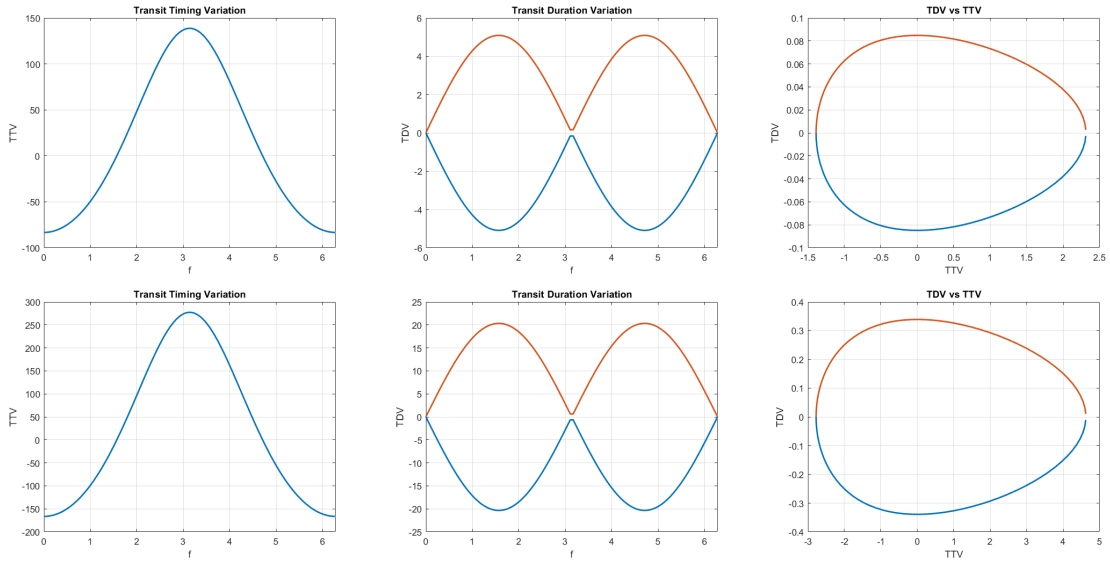[1st row: Low value for e_P] [2nd row: High value for e_P]

Figure 4.1: [Left: TTV plot] ; [Center: OB TDV plot] ; [Right: $O_2O_1B_4B_3$ Phase plot]

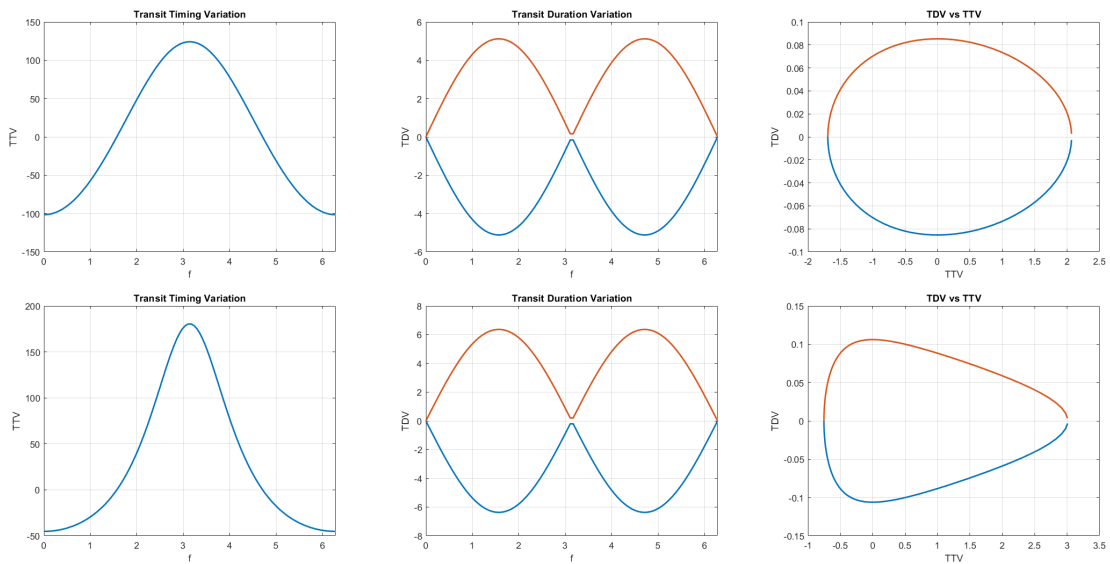[1st row: Low value for e_S] [2nd row: High value for e_S]

Figure 4.2: [Left: TTV plot] ; [Center: OB TDV plot] ; [Right: $O_2O_1B_4B_3$ Phase plot]

[1st row: Low value for a_P] [2nd row: High value for a_P]
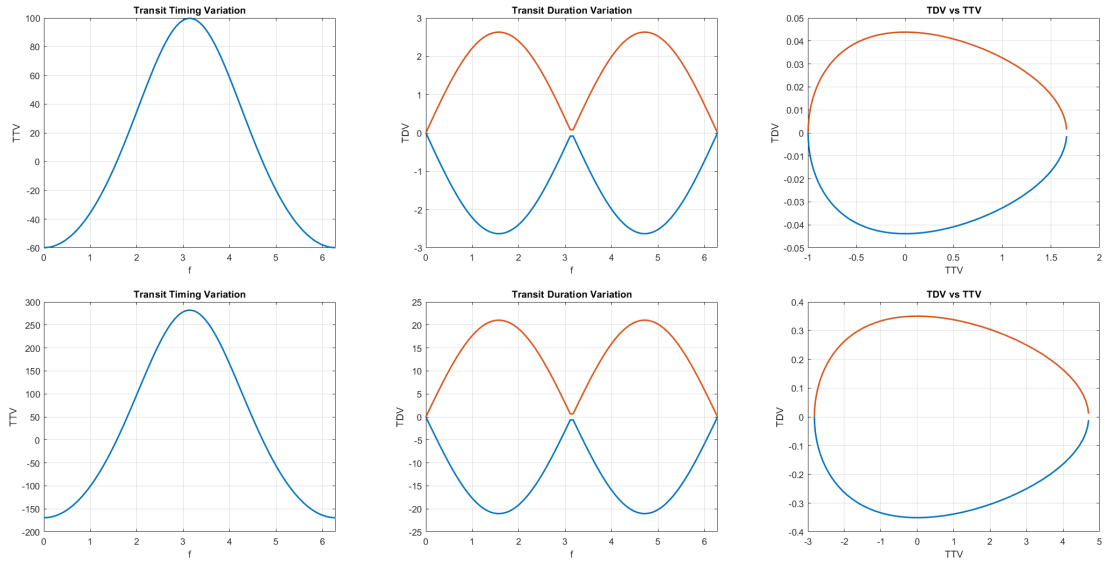

Figure 4.3: [Left: TTV plot] ; [Center: OB TDV plot] ; [Right: $O_2O_1B_4B_3$ Phase plot]

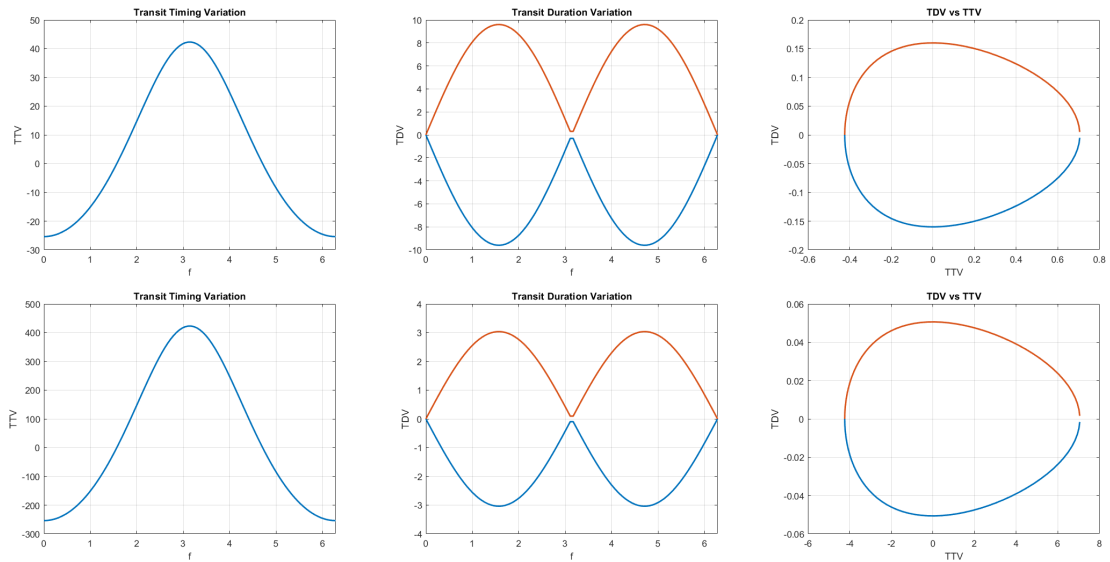[1st row: Low value for a_S] [2nd row: High value for a_S]

Figure 4.4: [Left: TTV plot] ; [Center: OB TDV plot] ; [Right: $O_2O_1B_4B_3$ Phase plot]

$[1^{st}$ row: $\bar{\omega}_S = \frac{\pi}{4}]$ $[2^{nd}$ row: $\bar{\omega}_S = \frac{3\pi}{4}]$
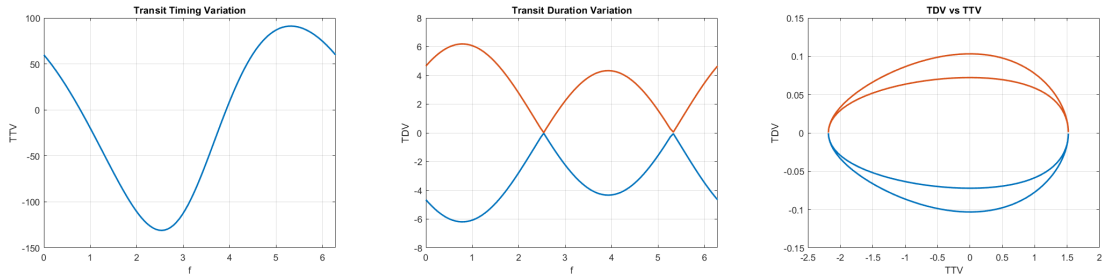

Figure 4.5: [Left: TTV plot] ; [Center: BOB TDV plot] ; [Right: $B_{4H}O_{1L}O_{2L}B_{3H}B_{4H}$ Phase plot]
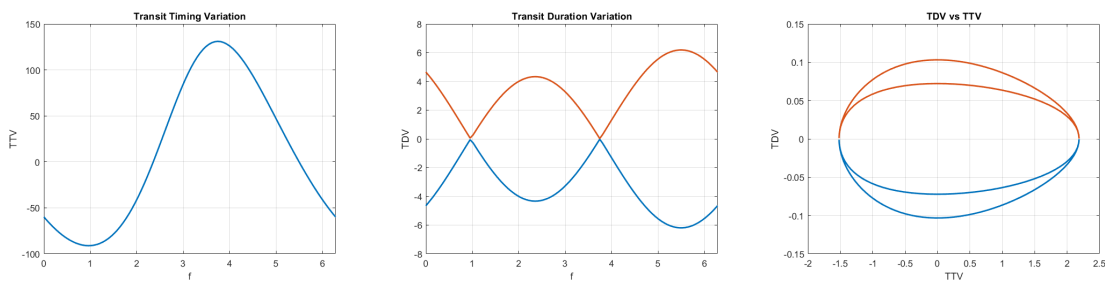

Figure 4.6: [Left: TTV plot] ; [Center: OBO TDV plot] ; [Right: $O_{2H}B_{3L}B_{4L}O_{1H}O_{2H}$ Phase plot]

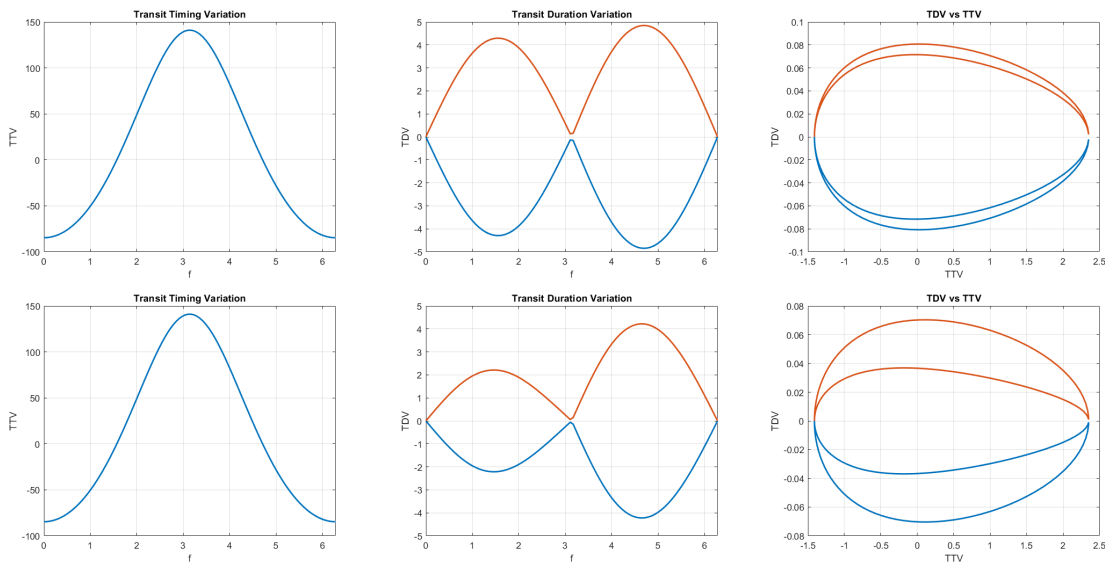$[1^{st}$ row: Low value for $b]$ $[2^{nd}$ row: High value for $b]$

Figure 4.7: [Left: TTV plot] ; [Center: OB TDV plot] ; [Right: $O_{2L}B_{3H}B_{4H}O_{1L}$ Phase plot]

(a)

(b)

(c)

(d)

(e)

(f)

(a)

(b)

(c)
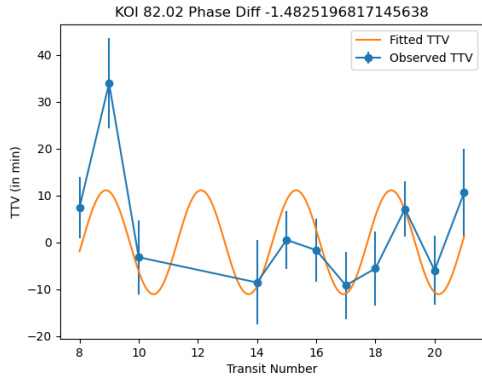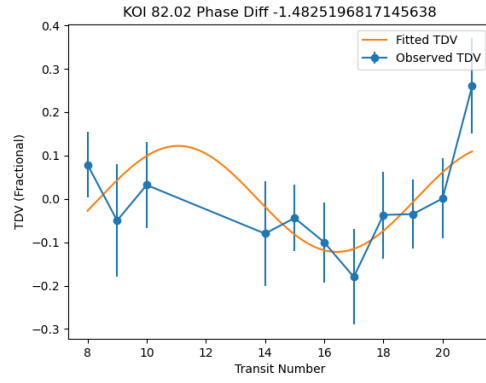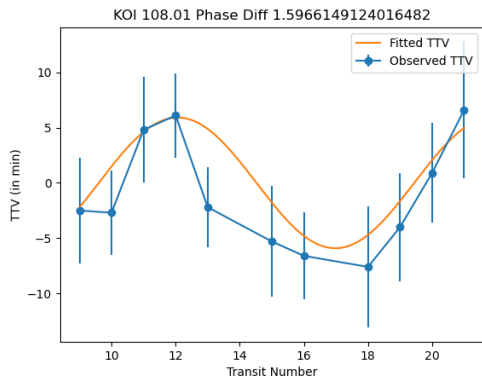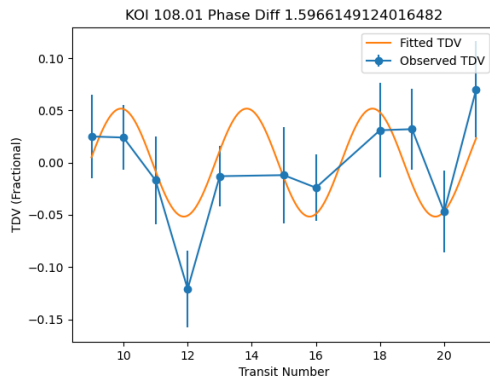
(d)
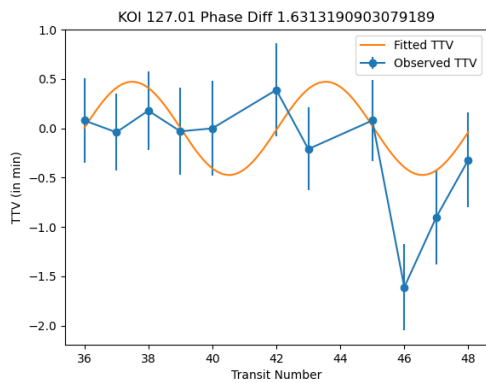
(e)
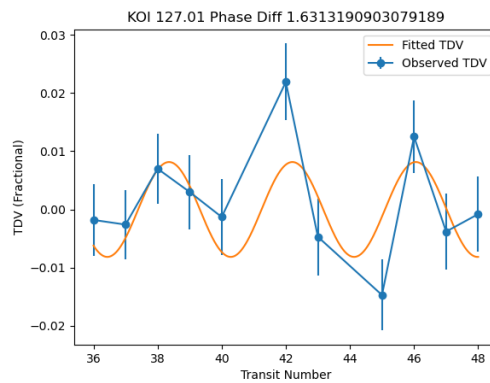
(f)
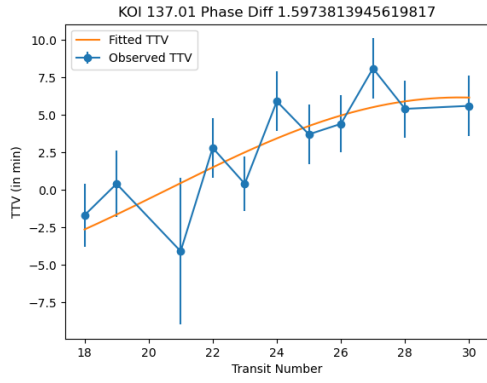
(a)
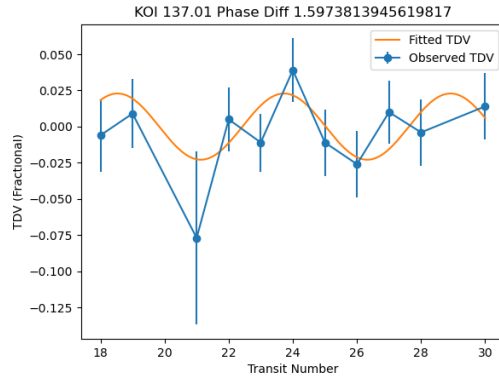
(b)
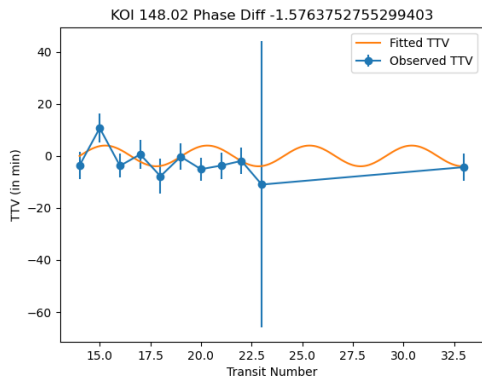
(c)
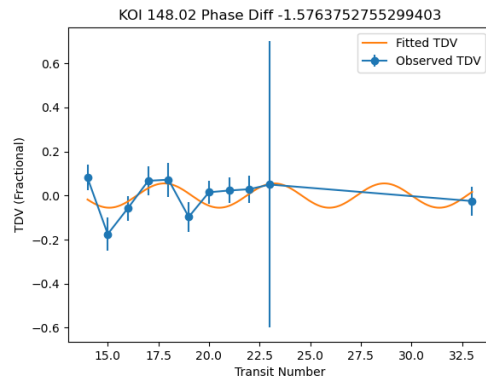
(d)
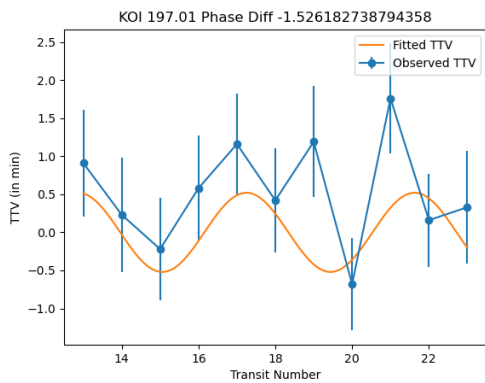
(e)
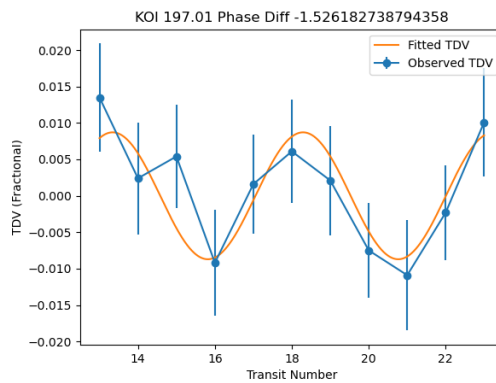
(f)
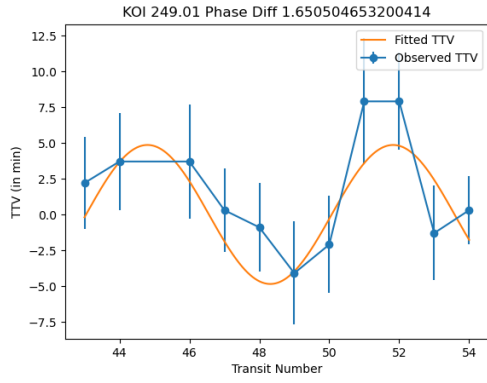
(a)
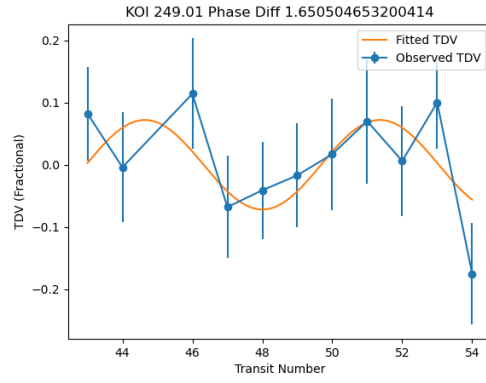
(b)
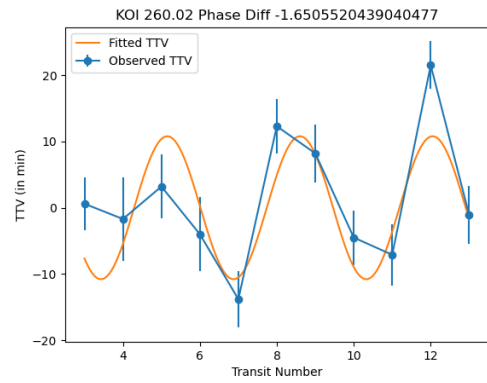
(c)
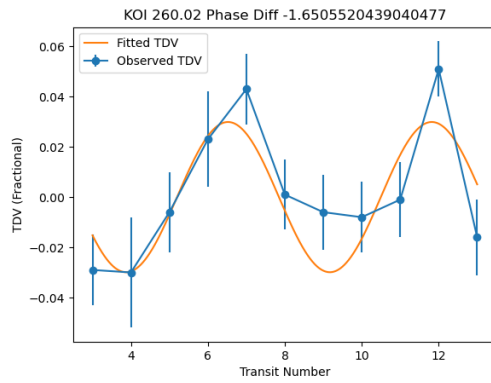
(d)
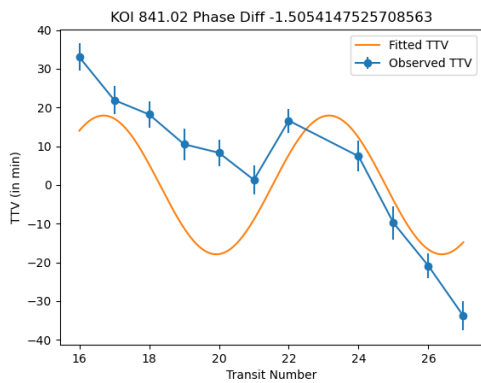
(e)
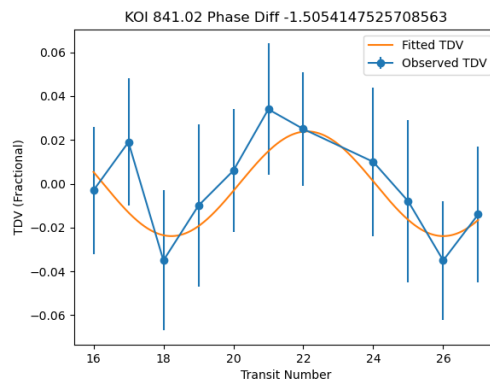
(f)
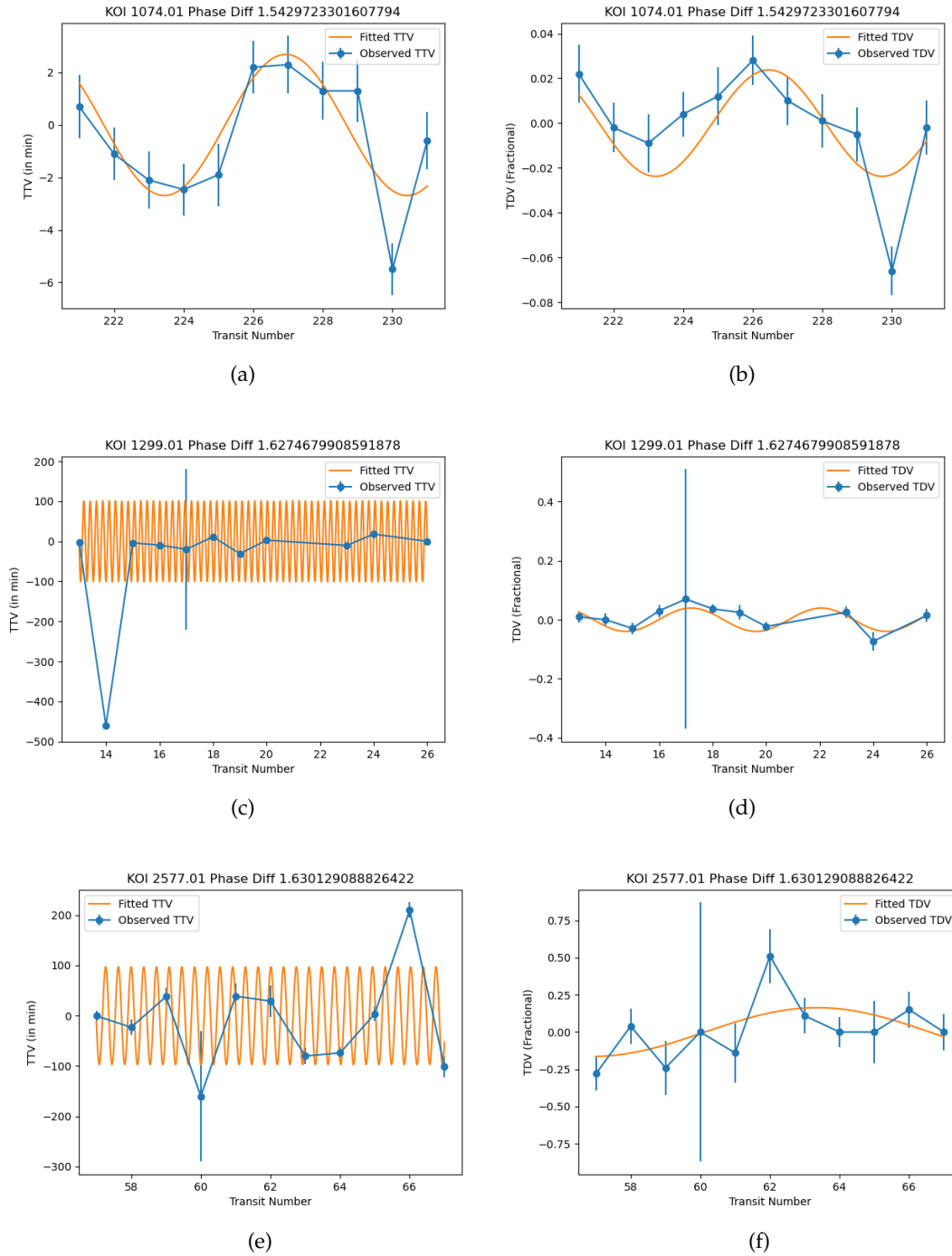
Figure 4.8: (a)Fitted TTV signals (Y) against Transit number (X) (orange) and data points along with errorbars (blue) (b)Fitted TDV signals (Y) against Transit number (X) (orange) and data points along with errorbars (blue)

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

(j)

(k)

(l)
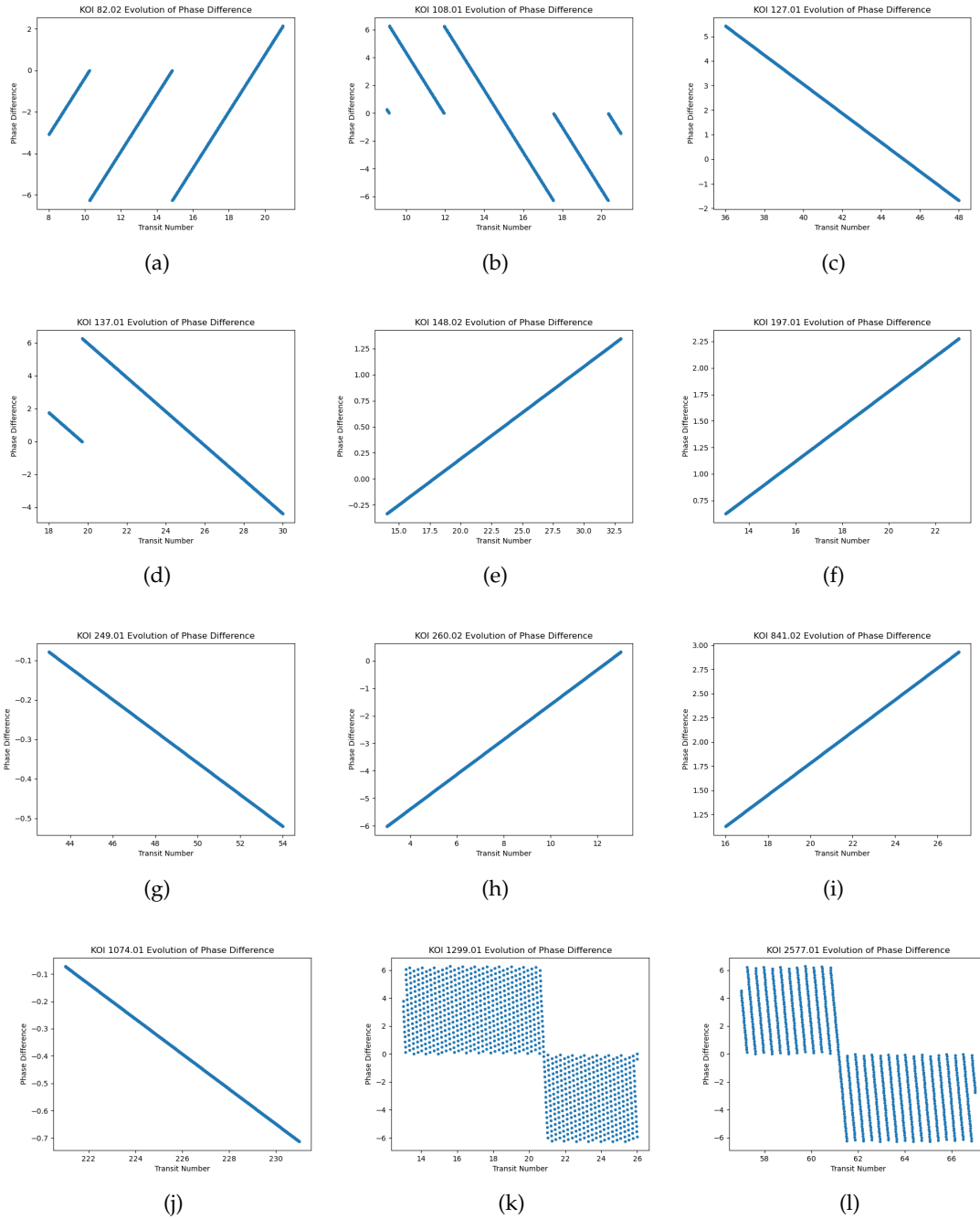
Figure 4.9: Variation of Phase Difference between TTV and TDV over Transit Number

Discussion & Conclusions

5.1 | Theoretical

There were mainly two theoretical modifications made in order to compensate for the limitations that hinder data analysis (such as being able to observe the state of TTV and TDV only on transit rather than having continuous observations so as to be able to measure the true changes in TTV and TDV):

- **Use of instantaneous variation equations rather than RMS variation** so as to remove the assumptions made for integration.

- **True anomaly of exoplanet** on its wobble orbit being the **measure of epoch** unit rather than transit numbers which can be discrete.

These modifications show four merits over previously used theoretical methods:

- The method make sure the TTV and TDV have the same frequency so as to avoid change of phase difference by assuming a common, shorter epoch - true anomaly of exoplanet on its wobble orbit.

- On analysis, it gives out more candidates with a phase difference of $\frac{\pi}{2}$ than found using previous methods of analysis.

- It gets us a step closer to removal of discrepancy between the theoretical model for exomoon-induced TTV-TDV and data analysis methods.

- The period of exomoon orbit P_s is taken as a parameter, the factor of which can be obtained after fitting the instantaneous variations on modified epoch.

5.2 | Observational

On having modified the shortlisting conditions used by Fox and Weigert we got 53 candidates with an epoch of $\frac{\pi}{2}$ (excluding the standard deviation and SNR condition). The following table provides the list of KOIs that showed an epoch of approximately $\frac{\pi}{2}$. (Now the precision is to

0.5 because the SD condition has been removed and we also need to account for the noise in the low SNR candidates and that's why the tolerance level has been increased) 5.1

All of these candidates can be looked into further as they do not vary significantly from Kipping's theory and might stand out to be probable exomoon candidates. Better curve fitting methods maybe used to do the analysis other wise these can also be used as a part of the neural network method explained in future scope.

The other conclusions based on just the data analysis are:

- Through the datasets till 2016 there were only 1.9% of the candidates which did show a phase difference comparable to $\frac{\pi}{2}$.

- Initially for the curve fitting module for less number of data points the graph had **concentrated data points leading to a shifted mean** and therefore the least χ^2 method ended up giving logically erroneous results or the module for curve fit in SciPy was not converging even when the maximum calls to the function was increased to 9000. Therefore there must be a **pattern in the variation of TTV which must be periodic** (although the order of variation maybe small but overtime accumulates).

- A high frequency plotting was also tried as described earlier while that did considerably decrease the phase difference variation for our shortlisted candidates it also **gave rise to newer candidates**. This could have been possible because of the fact that some candidates start out with a phase difference not close to $\frac{\pi}{2}$ and then the high frequency plot actually accounts for this change.

- In order to comment or confirm Kipping's theory a more independent and rigorous curve fitting analysis has to be undertaken. Using any sort of pre-defined functions to fit the data points would be a bit biased and therefore using convoluted neural networks would be the best way to go forward.

- The change in phase difference shown in the plots actually hint towards the fact that the ω_p and ω_s variables which were assumed to be constant by Kipping's theory aren't actually constant, they do vary somewhat. A better way to plot the theoretical models would have been to just do a numerical integration and for that the form of the function describing the previous 2 variables have to be known, which as of yet we do not.

In the end the fact stands that using any module with pre-defined curve fitting methods might not work here as we do not want to bias our fitting. The plots of the TTV and TDV might give better results with more results (as that might reduce the error because of random noise).

Serial No.	KOI	Phase Difference
1	2.01	-2.0156
2	12.01	1.27826
3	20.01	1.42405
4	46.01	-2.0634
5	70.01	1.4444
6	82.01	-1.4825
7	82.02	-1.4335
8	103.01	1.5966
9	108.01	-1.4314
10	110.01	1.2680
11	111.01	1.14898
12	116.01	-1.9258
13	116.02	1.6313
14	127.01	1.9293
15	128.01	1.5973
16	137.01	1.13160
17	142.01	-1.5763
18	149.01	1.7468
19	152.02	1.4184
20	162.01	-1.0878
21	165.01	1.1309
22	197.01	-1.526
23	203.01	1.8189
24	227.01	1.8297
25	249.01	1.6505
26	252.01	-1.94436
27	254.01	2.0503
28	255.01	-1.1239
29	257.01	1.1589
30	260.02	-1.6505
31	275.01	-2.035
32	282.01	-1.373
33	288.02	1.7280
34	304.01	-1.419
35	306.01	-1.8629
36	314.02	-1.3976
37	315.01	1.2489
38	354.01	1.6941
39	398.01	1.94631
40	456.01	-1.1332
41	657.02	-1.3819
42	723.02	1.8721
43	841.02	-1.5054
44	921.02	-1.9889
45	984.01	-1.6896
46	1074.01	1.5429
47	1299.01	1.62746
48	1426.01	-1.26108
49	1573.01	-1.8433
50	1781.01	-1.25435
51	1805.01	2.03585
52	2577.01	1.630129
53	2672.02	-1.9235

Table 5.1: KOIs showing a phase difference of $\frac{\pi}{2}$ within a 30% agreement

6

Future Work

Towards the ending of our project we realised that a lot more can be done and we have just touched the tip of the iceberg. Given as follows are some verticals which can be worked upon to give a better analysis and more reliable conclusions on the observations.

6.1 | Modifications of theoretical constants

For a long epoch interval, the parameters considered constant in both our and Kipping (2008) can change considerably enough so as to play a role in changing the magnitude of TTV and TDV. **Parameters such as $\bar{\omega}_S$, $\bar{\omega}_S$, a_P and a_S can change slightly over the course of a single planetary orbit**. For a confirmed hypothesis, we need observations from multiple transits usually in the order of 10^1. So it is pretty evident that in this order of transit numbers, the constants mentioned above are bound to change by some amount. If not taken into account, they can result in discrepancies between the theoretical models and data analysis. Figure 6.1 is an exaggerated description of how the evolution of the aforementioned constants can have an effect on the magnitude of TTV and TDV over long epoch interval.
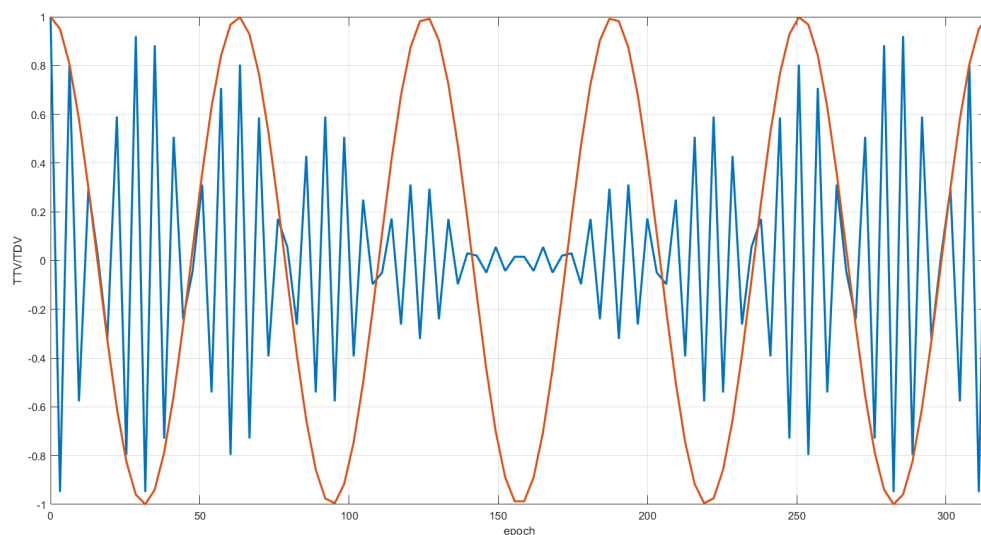
Figure 6.1: Exaggerated evolution of TTV/TDV (blue line) over long transit epoch interval

Such a function (figure 6.1) will require further mathematical modelling. Fourier transform can be a promising method to find close-to analytical solution as our epoch - true anomaly - changes with time depending on other parameters and it would be useful to express transit timing and duration variations as sinosoidal frequency domain function.

$$TV(\omega) = \int_{-\infty}^{+\infty} TV(f_W(t))e^{j\omega t}dt \tag{6.1}$$

where TV (transit variation) can be TTV and TDV in compact form

$$TTV(f_W) = \frac{C_1^T}{C_2^T + v_{W\perp}} \frac{cos(f_W + \bar{\omega}_W)}{1 + e_W cos(f_W)} \tag{6.2}$$

$$TDV(f_W) = C^D \left(\frac{l_r}{1 + v_{r\perp}} - 1 \right) \tag{6.3}$$

Although the current models are very general, they can be be further generalized if this change in constants is taken into account so as to be able to expand the range of applicability of our model to any given epoch interval. It is almost certain that for the above generalization, analytical solutions can be quiet difficult to obtain and for this reason, we will need computational analysis to play an important role alongside theoretical and mathematical analysis.

6.2 | Modelling using Convoluted Neural Networks

A completely independent and unbiased way of confirming Kipping (2008) theory would be to train a convoluted Neural Netwrok using the latest dataset possible. There are already well developed models for analyzing lightcurves of different TTV and TDV signals from possible exomoon candidates. One easy model Alshehhi et al. (2020) to set up would be to consider a binary classification where a false positive would be 0 and an exomoon candidate would mean 1. Considering a dataset going from 0<n<N. The loss function can be defined as follows,

$$L_err(\Phi) = \frac{1}{N}\Sigma_0^{N-1}|C(X_m; \Phi) - c_m| \tag{6.4}$$

where each term can only take values 1 or 0. For getting the required convoluted neural network (CNN) one is required to minimise this function (The number of data should be more than the number of Φ parameters). In fitting the data points there might be noise generated by more (unnecessary) parameters and even more the weights associated with them. In order to smoothen out the last convoluted layer i.e., to eliminate noise one would have to take up a convolution with a centered Gaussian Kernel G of the form,

$$S_m(p) = \Sigma_{j=0}^l Y(p - j)G(j) \tag{6.5}$$

where Y is the final layer and l denotes the last layer number.

Using these modifications one can easily define a preliminary neural network to analyse the dataset and form one with well-defined variables and weights associated with each. Once we have a neural network one can work with all different sorts of activation functions to check how the fitting changes. Once we have converging results we can then do a **missing variable analysis** by comparing the variables of the CNN with that of Kipping's theory. This will help us exactly confirm his theory and also have a well-defined code to check for exomoons.

6.3 | Wider dataset analysis

The analysis done by our code was for datasets till the year 2016. Which contained sufficient number of observations (more than 5) for 2599 candidates. When the NASA exoplanetary archive was examined they displayed more than 26000 candidates and therefore more observations for the TTV and TDV of these systems would help in terms of statistical analysis as by examining a bigger dataset we can form more inferences or better might even get to see some interesting signals.

Also from a statistical viewpoint it would help us set better shortlisting conditions and more number of data points will also cancel out on the random TTV/TDV signal error i.e., increase the average SNR value. Because of the low SNR in case of some candidates with transit depth less than 85 ppm leads to false positives which can only be verified with more data points and by eliminating this one would be working with a smaller dataset and therefore it would be easier to do the contrast analysis for them.

6.4 | Testing exomoon hypothesis

The aforementioned theoretical and data analysis models are made from scratch, purely for exomoon-induced TTV and TDVs. To confirm an exomoon, we collect data and try fitting it with our theoretical model, and we do so for multiple transits as mentioned in 6.1. Considering the assumptions made in theoretical model for TTV and TDV instantaneous equations, such as periodic values for TTV and TDV, and so on, it is very unlikely for a system to follow the model accurately for multiple transits. With the fitted values of $\frac{M_S}{M_P}$, $\frac{a_S}{a_P}$, $\frac{P_S}{P_P}$ and other satellite-planet ratio parameters taken into account as well. If we obtain a fitting for the ratio parameters that describe a satellite-planet ratio and do so consistently for multiple transits, it is likely to be an exomoon-candidate.

Derivation of generalized instantaneous variations

Frame switch operators

A general linear translation of magnitude x_0, from α to β frame can be expressed as

$$x_\beta = x_\alpha - x_0 \tag{A.1}$$

A general 2-D rotational transform of magnitude θ, from α to β frame can be expressed as

$$\begin{pmatrix} x_\beta \\ y_\beta \end{pmatrix} = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix} \begin{pmatrix} x_\alpha \\ y_\alpha \end{pmatrix} \tag{A.2}$$

$\hat{\tilde{x}}_1 - \hat{\tilde{y}}_1$ to $\hat{\tilde{x}} - \hat{\tilde{y}}$ frame

This is a purely rotational frame switch and hence can be expressed in rotation matrix form.

$$\begin{pmatrix} \tilde{x}_1 \\ \tilde{y}_1 \end{pmatrix} = \begin{pmatrix} cos(\bar{\omega}_W) & -sin(\bar{\omega}_W) \\ sin(\bar{\omega}_W) & cos(\bar{\omega}_W) \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} \tag{A.3}$$

$\hat{\tilde{x}}_2 - \hat{\tilde{y}}_2$ to $\hat{\tilde{x}}_1 - \hat{\tilde{y}}_1$ frame

This is a purely translational frame switch and hence can be expressed as seperate linear equations.

$$\tilde{x}_2 = \tilde{x}_1 - a_W e_W cos(\bar{\omega}_W) \tag{A.4}$$

$$\tilde{y}_2 = \tilde{y}_1 - a_W e_W sin(\bar{\omega}_W) \tag{A.5}$$

$\hat{x} - \hat{y}$ to $\hat{x}_2 - \hat{y}_2$ frame

This is a combination of rotational and translational frame switch and hence can be expressed as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} cos(-\bar{\omega}_W) & -sin(-\bar{\omega}_W) \\ sin(-\bar{\omega}_W) & cos(-\bar{\omega}_W) \end{pmatrix} \begin{pmatrix} x_2^0 \\ y_2 \end{pmatrix} \tag{A.6}$$

$$x_2 = x_2^0 - a_P e_P \tag{A.7}$$

$\hat{\tilde{x}} - \hat{\tilde{y}}$ to $\hat{\tilde{x}}_2 - \hat{\tilde{y}}_2$ frame

This is a combination of rotational and translational frame switch and hence can be expressed as

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} cos(-\bar{\omega}_W) & -sin(-\bar{\omega}_W) \\ sin(-\bar{\omega}_W) & cos(-\bar{\omega}_W) \end{pmatrix} \begin{pmatrix} \tilde{x}_2^0 \\ \tilde{y}_2 \end{pmatrix} \tag{A.8}$$

$$\tilde{x}_2 = \tilde{x}_2^0 - a_W e_W \tag{A.9}$$

Barycenter terms

The distance of any point on the ellipse from the foci is given by

$$r(f) = \frac{a(1 - e^2)}{1 + e cos(f)} \tag{A.10}$$

The above equation when applied to orbital motion of barycenter around the star, takes the following form

$$r_P(f_P) = \frac{a_P(1 - e_P^2)}{1 + e_P cos(f_P)} \tag{A.11}$$

where r_P is roughly the distance between star

and exoplanet.

Assuming $\lim_{\Delta t \to \tau_0}(\Delta f_P) \to 0$, we obtain the expression $f_P = \frac{\pi}{2} - \bar{omega}_P$, a constant as the barycenter must roughly lie on the line joining star and observer for an observable transit. That gives us the constant expression for position of barycenter during transit.

$$r_P = \frac{a_P\left(1 - e_P^2\right)}{1 + e_P \sin\left(\bar{\omega}_P\right)} \qquad \text{(A.12)}$$

As we have seen in Kipping (2008), the apparent velocity of barycenter around the star is given by

$$v_{B\perp} = Y\left(e_P, \bar{\omega}_P\right)\sqrt{\frac{G\left(M_{\text{PRV}} + M_*\right)}{a_P}} \qquad \text{(A.13)}$$

where

$$Y\left(e_P, \bar{\omega}_P\right) = \cos\left[\arctan\left(\frac{-e_P\cos\left(\bar{\omega}_P\right)}{1 + e_P\sin\left(\bar{\omega}_P\right)}\right)\right]$$
$$\sqrt{\frac{2\left(1 + e_P\sin\left(\bar{\omega}_P\right)\right)}{\left(1 - e_P^2\right)} - 1} \qquad \text{(A.14)}$$

Wobble terms

Equation A.10 when applied to wobble motion of exoplanet around the planet-moon barycenter, takes the following form

$$r_W\left(f_W\right) = \frac{a_W\left(1 - e_W^2\right)}{1 + e_W\cos\left(f_W\right)} \qquad \text{(A.15)}$$

where r_W is the distance between barycenter and exoplanet. The perpendicular distance of the exoplanet from line joining star, barycenter and observer will be given by

$$\tilde{x}_2 = r_W cos(f_W + \bar{\omega}_W) \qquad \text{(A.16)}$$

$$\tilde{x}_2 = a_W(1 - e_W^2)\frac{cos(f_W + \bar{\omega}_W)}{1 + e_W cos(f_W)} \qquad \text{(A.17)}$$

Once again, we will consider Kipping (2008) for the wobble velocity of the exoplanet on its wobble orbit, given by

$$v_W\left(f_W\right) = \mu_W^{\frac{1}{2}}\left(\frac{2}{r_W\left(f_W\right)} - \frac{1}{a_W}\right)^{\frac{1}{2}} \qquad \text{(A.18)}$$

where

$$\mu_W = \frac{\text{GM}_S^3}{M_{\text{PRV}}^2} \qquad \text{(A.19)}$$

This is not the final expression for the apparent wobble velocity. We must now take the component perpendicular to line joining star and observer

$$v_{W\perp}\left(f_W\right) = v_W\left(f_W\right)\cos\left[\tilde{\theta}\left(f_W\right)\right] \qquad \text{(A.20)}$$

$$\tilde{\theta}\left(f_W\right) = \arctan\left(\frac{d\tilde{y}_2}{d\tilde{x}_2}\right) \qquad \text{(A.21)}$$

The expression for $\tilde{\theta}(f_W)$ is obtained by differentiating (w.r.t. $\tilde{x}_2$) equation of ellipse in the $\tilde{x}_2 - \tilde{y}_2$ frame after replacing them with $\tilde{x} - \tilde{y}$ by performing the inverse frame switch operator A.8. Doing so, we get

$$\tilde{\theta}\left(f_W\right) = \arctan\left(\frac{\tilde{y}\sin\left(\bar{\omega}_W\right) - \left(1 - e_W^2\right)\tilde{x}\cos\left(\bar{\omega}_W\right)}{\tilde{y}\cos\left(\bar{\omega}_W\right) + \left(1 - e_W^2\right)\tilde{x}\sin\left(\bar{\omega}_W\right)}\right)$$
$$\text{(A.22)}$$

Transit timing variation

Transit timing variations by definition is the magnitude of lag or lead in the observed transit time of the exoplanet from the calculated one. Our reference for calculations is the barycenter which represents the motion of exoplanet without an exomoon. So it is trivial now, that the lag

or lead in transit time will depend on the apparent distance between exoplanet and barycenter, and also will depend on the total apparent velocity of the exoplanet on its orbit.

$$TTV(f_W) = TTV_O - TTV_C = \frac{\tilde{x}_2}{v_{B\perp} + v_{W\perp}} - 0 \tag{A.23}$$

$$TTV(f_W) = \frac{\tilde{x}_2}{v_{B\perp} + v_{W\perp}} \tag{A.24}$$

Substituting in the above equation from equations A.13, A.17 and A.20, we obtain A.35, the expression for instantaneous TTV, as a function of true wobble anomaly.

Writing the purely constant TTV terms as C^T, we get

$$TTV(f_W) = \frac{C_1^T}{C_2^T + v_{W\perp}} \frac{cos(f_W + \bar{\omega}_W)}{1 + e_W cos(f_W)} \tag{A.25}$$

Impact parameter

The impact parameter plays a huge role in determining the magnitude of the theoretical TDV value. The larger the impact parameter, the shorter will be the distance the exoplanet has to travel in order to fully transit the star.

In this section we will be deriving expressions for relative lengths and velocities, in the case of any general impact parameter (zero or nonzero). For this purpose, we use 3.2 as our reference. We can notice the following relations from the diagram-

$$bR_* = r_P cos(i) \tag{A.26}$$

where r_P is position of exoplanet w.r.t. star,

given by A.12 and b is the impact parameter.

$$l_B = \sqrt{(R_* + R_P)^2 - (bR_*)^2} \tag{A.27}$$

In the case of exoplanet containing exomoon, the apparent length is altered due to perturbation by an amount Δl, given by Kipping (2009)

$$\Delta l = cos(i) r_W sin(f_W + \bar{\omega}_W) \tag{A.28}$$

We now substitute our relation A.26 in the above formula to get

$$\Delta l = \frac{bR_*}{r_P} r_W sin(f_W + \bar{\omega}_W) \tag{A.29}$$

Therefore the apparent length in presence of an exomoon will be given by

$$l_W = \sqrt{(R_* + R_P)^2 - (bR_*)^2(1 + \frac{1}{r_P})^2} \tag{A.30}$$

Transit duration variation

Transit duration variations by definition is the magnitude of lag or lead in the observed transit duration (time taken for the exoplanet to complete its transit). Our reference for calculations is again, a barycenter which represents motion of exoplanet without an exomoon. So it is trivial now, that the lag or lead in transit duration will depend upon both the length of transit distance and on the total apparent velocity.

$$TDV(f_W) = TDV_O - TDV_C = \frac{l_w}{v_{B\perp} + v_{W\perp}} - \frac{l_b}{v_{B\perp}} \tag{A.31}$$

$$TDV(f_W) = \left(\frac{\frac{l_W}{l_b}}{1 + \frac{v_{W\perp}}{v_{B\perp}}} - 1 \right) \frac{l_B}{v_{B\perp}} \tag{A.32}$$

We define the ratios as $\frac{l_W}{l_b} = l_r$ and $\frac{v_{W\perp}}{v_{B\perp}} = v_{r\perp}$

which denote relative apparent length and velocity.

$$TDV(f_W) = \left(\frac{l_r}{1 + v_r} - 1 \right) \tau_0 \qquad \text{(A.33)}$$

Here, τ_0 represents the transit duration if the exomoon were to be absent, or in other words our reference transit duration.

Writing the purely constant TDV terms as C^D, we get

$$TDV(f_W) = C^D \left(\frac{l_r}{1 + v_{r\perp}} - 1 \right) \qquad \text{(A.34)}$$

$$TTV(f_W) = \frac{a_W(1 - e_W^2)}{Y(e_P, \bar{\omega}_P) \sqrt{\frac{G(M_{\mathrm{PRV}} + M_*)}{a_P}} + v_W(f_W) \cos\left[\tilde{\theta}(f_W)\right]} \frac{cos(f_W + \bar{\omega}_W)}{1 + e_W cos(f_W)} \qquad \text{(A.35)}$$

MATLAB code for plotting derived instantaneous variations

Listing B.1: Earth-Moon System

```matlab
% MATLAB CODE FOR PLOTTING DERIVED INSTANTANEOUS VARIATIONS
% EARTH-MOON SYSTEM

% PARAMETERS
% UNIVERSAL CONSTANT
G=6.67408*10^-11;

% MASS
Mstar=1.989*10^30;
Mp=5.972*10^24 ;
Ms=7.34767309*10^22;
Mprv=Mp+Ms;

% DENSITY
rhostar=1.41*10^3;
rhop=5.51*10^3;
rhos=3.34*10^3;

% RADIUS
Rstar=696340*10^3;
Rp=6371*10^3;
Rs=1737.1*10^3;

% ORBITAL PERIOD
Pp=365*(86400);
Ps=27*(86400);

% ORBITAL ECCENTRICITY
ep=0.0167086;
```

```matlab
30  es=0.25;
31  ew=es;
32
33  % SEMI-MAJOR AXIS
34  ap=0.5*1.496*10^11;
35  as=3.84748*10^8;
36  aw=as*(Ms/Mprv);
37
38  % ORBITAL ORIENTATION
39  b=0;
40  fp=pi/2-wp;
41  wp=1.5*pi;
42  ww=0;
43  ws=-ww;
44
45  % TRANSIT TIME
46  tT=13*60*60;
47
48  % CHANGING QUANTITY
49  f=linspace(0,2*pi);
50  %t=;
51
52  % ORBITAL POSITION OF EXOPLANET
53  rw=(aw*(1-ew^2)./(1+ew.*cos(f)));
54  rp=(ap*(1-ep^2)./(1+ep.*cos(pi/2-wp)));
55
56  % APPARENT X-Y POSITION
57  xtil2=rw.*cos(ww+f);
58  ytil2=rw.*sin(ww+f);
59
60  % REFERENCE FRAME SHIFTS
61  xtil=-xtil2.*cos(ww)-ytil2.*sin(ww)-aw*ew;
62  ytil=xtil2.*sin(ww)-ytil2.*cos(ww);
63
64  % TANGENTIAL ANGLE
65  dytil2=ytil.*sin(ww)-(1-ew^2).*xtil.*cos(ww);
66  dxtil2=ytil.*cos(ww)+(1-ew^2).*xtil.*sin(ww);
67  thtil=atan((dytil2)./(dxtil2));
```

```matlab
68
69 % APPARENT BARYCENTER VELOCITY
70 up=cos(atan((-ep*cos(wp))/(1+ep*sin(wp))))*sqrt((2+2*ep*sin(wp))/(1-
      ep^2)-1);
71 vbper=up*sqrt((G*(Mprv+Mstar))/ap);
72
73 % APPARENT WOBBLE VELOCITY
74 muw=(G*Ms^3)/(Mprv^2);
75 vw=(muw^(1/2)).*((2./rw)-(1/aw)).^(1/2);
76 vwper=vw.*cos(thtil);
77
78 % TTV EXPRESSION
79 ttv=xtil2./(vbper+vwper);
80
81 % TDV EXPRESSIONS
82 q=b*Rstar;
83 dq=rw.*sin(ww+f).*((b*Rstar)/(rp));
84
85 % KIPPING (2009)'S TDV
86 eb=sqrt(((Rstar+Rp)^2-(q+dq).^2)./((Rstar+Rp)^2-q^2));
87 tdvk=((eb.*vbper)./(vbper+vwper)-1).*tT;
88
89 % MODIFIED TDV
90 lr=sqrt(((Rstar+Rp)^2-(q+dq).^2)./((Rstar+Rp)^2-q^2));
91 t0=(sqrt((Rstar+Rp)^2-(b*Rstar)^2))./(vbper);
92 tdv=((lr.*vbper)./(vbper+vwper)-1).*t0;
93
94 % TDV DIVERGENCE FROM KIPPING'S MODEL
95 plot(f,tdv,f,-tdv,f,tdvk,f,-tdvk,'linewidth',1.7);
96 xlim([0 2*pi])
97
98 % TDV PLOT
99 plot(f,tdv,f,-tdv,'linewidth',1.7);
100 title('Transit Duration Variation')
101 xlabel('f')
102 ylabel('TDV')
103 xlim([0 2*pi])
104 grid on
```

```matlab
105
106 % TTV PLOT
107 plot(f,ttv,'linewidth',1.7)
108 title('Transit Timing Variation')
109 xlabel('f')
110 ylabel('TTV')
111 xlim([0 2*pi])
112 grid on
113
114 % TTV-TDV MAGNITUDE COMPARISON
115 TTV and TDV
116 plot(f,ttv,f,tdv,f,-tdv,'linewidth',1.7)
117 title('Transit Variation')
118 xlabel('f')
119 ylabel('TTV vs TDV')
120 xlim([0 2*pi])
121 grid on
122
123 % TTV-TDV PHASE DIAGRAM (IN MINUTES)
124 plot(ttv/60,tdv/60,ttv/60,-tdv/60,'linewidth',1.7)
125 title('TDV vs TTV')
126 xlabel('TTV')
127 ylabel('TDV')
128 grid on
```

C

Python Code for Data Analysis

```python
# -*- coding: utf-8 -*-
"""
Created on Thu Sep 24 09:25:15 2020

@author: binay
"""
import numpy as np
import pandas as pd
import shlex
from scipy.optimize import curve_fit
import matplotlib.pyplot as plt


arr = []
Tnum = []
KTTV = []
i = 0

def f(x,A,B,C):
    #M = max(KTTV[i])
    #(((M-abs(A))*np.sin(((np.abs(B)+90)*x)+C)))
    return ((A*np.sin((B*x)+C)))

def f2(x,A,C):
    params,coparam = curve_fit(f, Tnum[i], KTTV[i],maxfev=5000)
    B = params[1]
    return ((A*np.sin((B*x)+C)))

def find(ele):
    for i in range(len(arr)):
        if(ele==arr[i]):
```

```python
                return False
        return True

def CheckString(var):
    array = list(var)
    sum1 = ""
    for i in range(len(array)):
        if(array[i]=='*'):
            continue
        else:
            sum1 += array[i]
    return(float(sum1))
temp2 = np.array(pd.read_csv("C:/Users/binay/Downloads/kepler_O-C.txt",\
skiprows=27))

#declaring variable
KOI3 = []
KOII = []
KTTV_s = []
KTDV_s = []
KTTV_err = []
KTDV_err = []
KTTV_err_s = []
KTDV_err_s = []
Tnumt = []
#storing the text file in an array format
for i in range(len(temp2)):
    KOI3.append(np.array(shlex.split(temp2[i][0])))
temp = 1.01
i = 0
z = 0
KTDV = []
TDV = []
TDV_s = []
print(len(KOI3))
while(i<len(KOI3)):
    a = float(KOI3[i][0])
    TTV = []
```

42

```python
TTV_s = []
TDV = []
TDV_s = []
TTV_err = []
TDV_err = []
TTV_err_s = []
TDV_err_s = []
Tnumt = []
#print(i)
h = 0
#Storing TTV and transit number data in the arrays
while(a==temp)and(h<11):
    if((KOI3[i][3]!= "NaN")and(KOI3[i][5]!="Nan"))and\
    ((KOI3[i][11]!= "NaN")and(KOI3[i][13]!="Nan")):
        if((KOI3[i][4]!= "NaN")and(KOI3[i][6]!="Nan"))and\
        ((KOI3[i][12]!= "NaN")and(KOI3[i][14]!="Nan")):
            TTV.append(float(KOI3[i][3]))#TTV
            Tnumt.append(int(KOI3[i][1]))#no. of transit
            TDV.append(float(KOI3[i][5]))
            TTV_s.append(float(KOI3[i][11]))#TTV
            TDV_s.append(float(KOI3[i][13]))
            TTV_err.append(CheckString(KOI3[i][4]))#TTV
            TDV_err.append(CheckString(KOI3[i][6]))
            TTV_err_s.append(CheckString(KOI3[i][12]))#TTV
            TDV_err_s.append(CheckString(KOI3[i][14]))
            h += 1
    if(i<len(KOI3)-1):
        i += 1#to move on to the next row
        temp = float(KOI3[i][0])#KOI
    if(i>=len(KOI3)-1):
        break
if(i==len(KOI3)-1):
    break

#storing all the TTV and Transit number data in a bigger
#array in the form of a DDA
if(len(TTV)>10)and(len(Tnumt)>10):
    KOII.append(a)
```

43

```python
            KTTV.append(np.array(TTV))
            KTDV.append(np.array(TDV))
            Tnum.append(np.array(Tnumt))
            KTTV_s.append(np.array(TTV_s))
            KTDV_s.append(np.array(TDV_s))
            KTTV_err.append(np.array(TTV_err))
            KTDV_err.append(np.array(TDV_err))
            KTTV_err_s.append(np.array(TTV_err_s))
            KTDV_err_s.append(np.array(TDV_err_s))
            while(a==temp)and(i<len(KOI3)-1):
                i += 1
                temp = float(KOI3[i][0])
    z = 0
    c = []
fit = []
params,coparam = [],[]
params1,coparam1 = [],[]
count = 0
phase = 0.0
Phi = []
count1 = 0
for i in range(len(KTTV)):
    if(find(KOII[i])):
        #print(KOII[i])
        params,coparam = curve_fit(f, Tnum[i], KTTV[i],maxfev=5000)
        params1,coparam1 = curve_fit(f, Tnum[i], KTDV[i],maxfev=5000)
        phase = params[2]-params1[2]
        while(phase<-2.0*np.pi)or(phase>2.0*np.pi):
            if(phase<-2.0*np.pi):
                phase+=2.0*np.pi
            elif(phase>2.0*np.pi):
                phase-=2.0*np.pi
        #print(phase)
        #plotting
        count+=1
        #if(params[0]*1.4142<5000.0):
        if(abs(abs(phase)-abs(np.pi/2.0))<0.5):
            count1+=1
```

44

```python
    print(KOII[i],phase)
    #plt.clf()
    #plt.title("KOI "+str(KOII[i])+" Phase Diff "+str(phase))
    #plt.ylabel("TTV (in min)")
    #plt.xlabel("Transit Number")
    #plt.errorbar(Tnum[i],KTTV[i],yerr = KTTV_err[i],fmt='-o',\
    label = "Observed TTV")
    #plt.errorbar(Tnum[i],KTDV[i],yerr = KTDV_err[i],fmt='-o',\
    label = "Observed TDV")
    #X = np.linspace(Tnum[i][0],Tnum[i][10],1000)
    #plt.plot(X,params[0]*np.sin((params[1]*X) + params[2]),\
    label = "Fitted TTV")
    #plt.plot(X,params1[0]*np.sin((params1[1]*X) + params1[2]),\
    label = "Fitted TDV")
    #plt.legend()
    #plt.savefig("Diff KOI "+str(KOII[i])+" TTV.png")
    '''
    for j in range(len(X)):
        phase = ((X[j]*params[1]) +\
        params[2])-((X[j]*params1[1])+params1[2])
        while(phase<-2.0*np.pi)or(phase>2.0*np.pi):
            if(phase<-2.0*np.pi):
                phase+=2.0*np.pi
            elif(phase>2.0*np.pi):
                phase-=2.0*np.pi
        Phi.append(phase)
    '''
    #plt.title("KOI "+str(KOII[i])+" Evolution of Phase Difference")
    #plt.ylabel("Phase Difference")
    #plt.xlabel("Transit Number")
    #plt.plot(X,Phi,'.')
    #plt.savefig("Phase_KOI "+str(KOII[i])+" TTV.png")
    #Phi = []
```

References

Rasha Alshehhi, Kai Rodenbeck, Laurent Gizon, and Katepalli R Sreenivasan. Detection of exomoons in simulated light curves with a regularized convolutional neural network. *arXiv preprint arXiv:2005.13035*, 2020.

Chris Fox and Paul Wiegert. Exomoon candidates from transit timing variations: Six kepler systems with ttvs explainable by photometrically unseen exomoons. *arXiv preprint arXiv:2006.12997*, 2020.

Heller, René, Hippke, Michael, Placek, Ben, Angerhausen, Daniel, and Agol, Eric. Predictable patterns in planetary transit timing variations and transit duration variations due to exomoons. *A&A*, 591:A67, 2016. doi: 10.1051/0004-6361/201628573.

Tomer Holczer, Tsevi Mazeh, Gil Nachmani, Daniel Jontof-Hutter, Eric B Ford, Daniel Fabrycky, Darin Ragozzine, Mackenzie Kane, and Jason H Steffen. Transit timing observations from kepler. ix. catalog of the full long-cadence data set. *The Astrophysical Journal Supplement Series*, 225(1):9, 2016.

Daniel Huber, William J Chaplin, Jørgen Christensen-Dalsgaard, Ronald L Gilliland, Hans Kjeldsen, Lars A Buchhave, Debra A Fischer, Jack J Lissauer, Jason F Rowe, Roberto Sanchis-Ojeda, et al. Fundamental properties of kepler planet-candidate host stars using asteroseismology. *The Astrophysical Journal*, 767(2):127, 2013.

Daniel Jontof-Hutter, Eric B Ford, Jason F Rowe, Jack J Lissauer, Daniel C Fabrycky, Christa Van Laerhoven, Eric Agol, Katherine M Deck, Tomer Holczer, and Tsevi Mazeh. Secure mass measurements from transit timing: 10 kepler exoplanets between 3 and 8 m with diverse densities and incident fluxes. *The Astrophysical Journal*, 820(1):39, 2016.

David Kipping. An independent analysis of the six recently claimed exomoon candidates, 2020.

David Kipping and Alex Teachey. Impossible moons–transit timing effects that cannot be due to an exomoon. *arXiv preprint arXiv:2004.04230*, 2020.

David M. Kipping. Transit timing effects due to an exomoon. *Monthly Notices of the Royal Astronomical Society*, 392(1): 181–189, 12 2008. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2008.13999.x.

David M. Kipping. Transit timing effects due to an exomoon – II. *Monthly Notices of the Royal Astronomical Society*, 396 (3):1797–1804, 06 2009. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2009.14869.x.

David M Kipping, Allan R Schmitt, X Huang, Guillermo Torres, D Nesvornỳ, Lars A Buchhave, Joel Hartman, and Gáspár Á Bakos. The hunt for exomoons with kepler (hek). v. a survey of 41 planetary candidates for exomoons. *The Astrophysical Journal*, 813(1):14, 2015.

Tsevi Mazeh, Tomer Holczer, and Simchon Faigler. Dearth of short-period neptunian exoplanets: A desert in period-mass and period-radius planes. *Astronomy & Astrophysics*, 589:A75, 2016.

Sean M Mills and Tsevi Mazeh. The planetary mass–radius relation and its dependence on orbital period as measured by transit timing variations and radial velocities. *The Astrophysical Journal Letters*, 839(1):L8, 2017.

David Nesvornỳ and Alessandro Morbidelli. Mass and orbit determination from transit timing variations of exoplanets. *The Astrophysical Journal*, 688(1):636, 2008.

David Nesvornỳ, David M Kipping, Lars A Buchhave, Gáspár Á Bakos, Joel Hartman, and Allan R Schmitt. The detection and characterization of a nontransiting planet by transit timing variations. *Science*, 336(6085):1133–1136, 2012.

András Pál and Bence Kocsis. Periastron precession measurements in transiting extrasolar planetary systems at the level of general relativity. *Monthly Notices of the Royal Astronomical Society*, 389(1):191–198, 2008.

Mary Anne Peters and Edwin L Turner. On the direct imaging of tidally heated exomoons. *The Astrophysical Journal*, 769(2):98, 2013.

Clément Ranc, David P Bennett, Yuki Hirao, Andrzej Udalski, Cheongho Han, Ian A Bond, Jennifer C Yee, Michael D Albrow, Sun-Ju Chung, Andrew Gould, et al. Ogle-2015-blg-1670lb: A cold neptune beyond the snow line in the provisional wfirst microlensing survey field. *The Astronomical Journal*, 157(6):232, 2019.

Kai Rodenbeck, René Heller, and Laurent Gizon. Exomoon indicators in high-precision transit light curves. *Astronomy & Astrophysics*, 638:A43, 2020.

Sujan Sengupta and Mark S Marley. Detecting exomoons around self-luminous giant exoplanets through polarization. *The Astrophysical Journal*, 824(2):76, 2016.

AE Simon, Gy M Szabó, and K Szatmáry. Exomoon simulations. *Earth, Moon, and Planets*, 105(2-4):385–389, 2009.

Mario Sucerquia, Jaime A Alvarado-Montes, Jorge I Zuluaga, Nicolás Cuello, and Cristian Giuppone. Ploonets: formation, evolution, and detectability of tidally detached exomoons. *Monthly Notices of the Royal Astronomical Society*, 489(2):2313–2322, 2019.

Mario Sucerquia, Vanesa Ramírez, Jaime A Alvarado-Montes, and Jorge I Zuluaga. Can close-in giant exoplanets preserve detectable moons? *Monthly Notices of the Royal Astronomical Society*, 492(3):3499–3508, 2020.

Alex Teachey, David Kipping, Christopher J Burke, Ruth Angus, and Andrew W Howard. Loose ends for the exomoon candidate host kepler-1625b. *The Astronomical Journal*, 159(4):142, 2020.