MNRAS **000**, 1–8 (2022) Preprint 6 December 2022 Compiled using MNRAS LaTeX style file v3.0

Constraints on Secular Perturbations in Impact Parameter Amongst Kepler Planets

6 December 2022

ABSTRACT

We analyze the usage of impact parameters to get accurate data on exoplanets. We first calculate the signal-to-noise ratio based on the change of impact parameter and the planet-star ratio, which could help scientists choose the exoplanets that have the clearest signal. We then model the changing of the impact parameter of the system that contains different numbers of exoplanets and derives a polynomial best regression function between the number of clearly changed impact parameters and the number of exoplanets in the system, which could help future scientist to discover the unnoticed planets. In the same section, we also derive the relationship between the theoretical $\dot{b}$ with respect to the mass and period of the planets that could exemplify the ideal characteristic of the exoplanets that be suitable for approximate mass based on $\dot{b}$. We finally search for various Kepler planets and applied the linear function of the impact parameters to get accurate parameters for multiple Kepler planets and analysis the corresponding accuracy statistically. For the occurrence of long-term linear increases in impact parameter values, we provide the following possibilities that exhibit strong signal-to-noise ratios, clear light curves, and significant z-scores. We can optimize the characteristics of verified Kepler planets by controlling orbital precession, and we can also hypothesize the existence of non-transit planets in the Kepler system.

Key words: Impact Parameters – Mutual Inclination – Transit Duration Variation

1 INTRODUCTION

Since the start of civilization, humans never extinguish their curiosity about outer space. With the launching of the Kepler telescope, enormous breakthrough took place in the understanding of exoplanets by identifying thousands of more confirmed planets, optimizing the parameters Batalha et al. (2013), and revealing the general statistic of exoplanets along with their system Fang & Margot (2012).

The majority of the exoplanets are detected by observing the changes in the flux of the star in the system Borucki et al. (2010). All of the planets in the system will transit from the star in our telescopes requires those planets' approximate co-planar (relatively low mutual inclination angle). Even though the majority of the Kepler system follows this pattern, a portion of the system still obtained a larger inclination angle Zhu & Dong (2021), which will cause the transit of some exoplanets in the system not visible by earth telescope. Due to the existence of non - transiting planets, there is a significant distinction between the detected and expected number of Kepler planets, which is known as the Kepler dichotomy. To resolve this, we have attempted to detect the non-transit planets by observing their gravitational influence on transit planets via the constraints of mutual inclination.

Mutual Inclination is more significant in planets with a short period that is placed in a close pack system composed of other large mass planets, which commonly exist in the Kepler population as revealed by Batalha et al. (2013), Borucki et al. (2011), Borucki et al. (2010), and Fabrycky et al. (2014). Because of their large distribution, various previous research has combined mutual inclination with Transit Time Variation Hadden & Lithwick (2014) Judkovsky et al. (2020), Transit Duration Variation from orbital precession Hamann et al. (2019), and Radial Velocity survey Tremaine & Dong (2012), to discover the non-transit planet.

All previous Kepler models assumed the impact parameter is constant beside a very recent publication Millholland et al. (2021). Impact parameters can change because of various effects: the uneven distribution of gravity due to the inclination angle and eccentricity, the changing of the period due to orbital resonance, and the variation of the semi-major axis that resulted from transit time variation. Because of those, the study of impact parameters combined with transit time variation can help us understand the variation of planets' period, semi-major axis, and inclination thus deriving more accurate parameters for exoplanets.

We model this effect and search for candidates that obtained planet star ratio bigger than 0.03, but smaller than 0.1, mass bigger than four earth mass, period less than, but bigger than, and the number of confirmed planets in that system less or equal to four. We also model the expected signal-to-noise(SNR) and the number of viable candidates and found SNR obtains a proportional relationship with the mass and radius of the planets relative to its star, and an inverse relationship with the period of the planets and the number of planets in the system. We cross-validate our search results to place constraints on Kepler planets into the planets with higher mass, shorter period, larger planet-star radius ratio, and the system with a lower number of planets.

⋆ E-mail: houxiong@ucsb.edu (HX)

Table 1. Notation Table.

Notation	Description
t	time
a	semi-major axis
b	Impact Parameters
R_*	radius of the star
M_*	mass of the star
m	mass of the planet
W	window for savigol
t_0	Transit Epoch
$\dot{b}$	The changing of impact parameters respect to time
T	Period
p	Planet-star ratio
u_1	Limb Darkening Parameter 1
u_2	Limb Darkening Parameter 2
ρ_*	$\frac{M_*}{r_*^3}$, the density of the star
ω	orbital frequency
$\overrightarrow{\Theta}$	Parameter Vector for Light curve deternding
$\overrightarrow{\Theta}_b$	$\overrightarrow{\Theta}$when b is changing
$\overrightarrow{\Theta}^{(n)}$	$\overrightarrow{\Theta}$ after nth optimization
n	number of data points
N	number of planets in that system
N_T	number of transits
z	normalized separation between centers
$\bar{F}$	raw flux from Kepler telescope
$\hat{F}$	Model flux generated from $\overrightarrow{\Theta}$
F	Actual flux from Kepler telescope
σ	Standard deviation for the error of the F and $\hat{F}$
σ_b	Standard deviation for the error of $\hat{F}(\overrightarrow{\Theta}_b)$ and$\hat{F}(\overrightarrow{\Theta})$
σ_I	Standard deviation for the inclination of planets

Similar computation has also been done by Daniel Fabrycky from the Department of Astronomy and Astrophysics at the University of Chicago, Sahar Shahaf and Tsevi Mazeh from the School of Physics and Astronomy, Shay Zucker from Porter School of the Environment and Earth Sciences, at Tel Aviv University. In their paper, "Systematic search for long-term transit duration changes in Kepler transiting planets" from 2020 and "Light-Curve Evolution due to Secular Dynamics and the Vanishing Transits of KOI 120.01" from 2021, both Transit Time Variation and Transit Depth Variation were applied and deduced the possible existence of a nontransit planet in KOI 120 system among with 16 other system. However, our approach is different as we mainly focus on the effect of the changing of impact parameters.

Section 2 of the paper presented our theoretical calculation of the signal-to-noise ratio of planets with secular changing impact parameters, Section 3 shows the derivation of impact parameter changes from their mutual inclination and the perturbation of two planet systems and estimated the relative frequency of those planets in the Kepler population from two different models, and demonstrates our process in light curve modeling, and Section 4 presents our diffraction model.

6. The notation that appears in the paper is presented in table 1.

2 METHOD

This chapter presented our sample selection method. 2.1 reveals the assumption of this research. 2.2 present the relationship between the signal-to-noise ratio and the planet's impact parameters and planet-star ratio. 2.3 demonstrates our model of the planet perturbations and presents the relationship between the changing of impact parameters and the planet's mass and the correlation of the number of planets in a system with the probability of detecting planets with $\dot{b} > 0.1$ per year. 2.4 presented how the candidates in the research are selected based on models 2.2 and 2.3.

2.1 Assumption

There are three assumptions we have made to conduct this research. First of all, we assume the impact parameter changes linearly. Secondly, we assume quadratic limb darkening with Agol Mandel light curve models and the planet is small enough that limb darkening will not take effect inside the planets. Thirdly, we correct for Transit Time Variation and assume no stellar variability and no Transit Depth Variation due to other causes, like inclination or eccentricity.

2.2 SNR Calculation

Generate a plot of the difference between constant b and db/dt model, and another plot of SNR as a function of b, R/R*, etc... If equations are too long - we put them in the appendix

The flux from a star during a planet transit is given by the equation below from Mandel & Agol (2002).

In the equation below, d is the center-to-center distance between the star and the planet, r_p is the radius of the planet, r_* is the stellar radius, $z = d/r_*$ is the normalized separation of the centers, and $p = r_p/r_*$ is the size ratio

$$F = 1 - \left[p^2\kappa_0 + p^2\kappa_1 - \sqrt{\frac{4z^2 - (1+z^2-p^2)^2}{4}} \right] \frac{1}{\pi} \tag{1}$$

$\kappa_1 = \cos^{-1}\left[\frac{1-p^2+z^2}{2z}\right]$ $\kappa_0 = \cos^{-1}\left[\frac{p^2+z^2-1}{2pz}\right]$

We then expand z into its horizontal components x and vertical components b, which are also the impact parameters of the planets, based on the Pythagorean theorem.

$$z = \sqrt{x^2 + b^2} \tag{2}$$

After that, we plug z into the flux function and calculate the numerical value of $\frac{dF}{db}$ and then calculate δF^2 based on the equation below, where $\frac{db}{dt}$ is set equal to 0.01 per year.

$$\left\langle \delta F^2 \right\rangle \approx \frac{1}{2(1+p)} \int_{-1-p}^{1+p} (\frac{\mathrm{d}F_0(x)}{\mathrm{d}b} \frac{\mathrm{d}b}{\mathrm{d}t} T N_T)^2 dx \tag{3}$$

We finally calculated SNR^2based on equation below

$$SNR^2 = \left\langle \delta F^2 \right\rangle \cdot (\frac{\tau}{cad}) \cdot \frac{1}{2\sigma^2} \tag{4}$$

Here $\left\langle \delta F^2 \right\rangle$ is the mean-square difference of two models, $\tau$is the transit duration, and cad is the cadence of Kepler Telescopes. We applied the short cadence which is equal to 2 minutes, and the corresponding σ equal to 8e-4. N_T is the number of transits in the data set, which is given by

$$N_T = \frac{1}{T} \tag{5}$$

in which T is the period of the planets equals to 30 days.

τ is given by

$$\tau = \frac{R*}{a} T \tag{6}$$

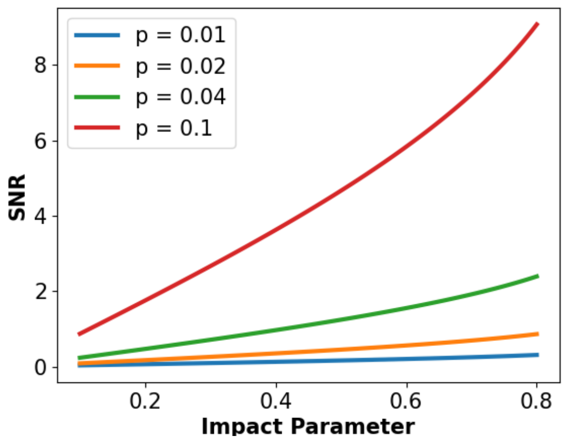

Figure 1. The signal-to-noise ratio (SNR) over the impact parameters of the planets for four different planet-star ratios. x-axis is the impact parameter and y-axis is the SNR. The colorful lines represent the results of different planet-star ratios respectively.

Table 2. R Squared Value of Four Impact Parameter to SNR lines.

Planet-star Ratio	R-Squared
0.01	0.99848
0.02	0.99849
0.04	0.99846
0.1	0.99827

which R* means the radius of the star. In this case, we set $\frac{R*}{a} = 50$

We then plot the signal-to-Noise Ratio as the function of the impact parameter, which is shown in figure **??**. The impact parameter is ranged from 0.1 to 0.8, and the p values for the four lines are 0.01, 0.02, 0.04, and 0.1 respectively.

Finally, we calculate the exponential best fit line for the four lines and express the equation in terms of p and b, which is presented below.

$$SNR \approx 406.8 \cdot (bp)^{\frac{3}{2}} \tag{7}$$

The R Squared values are shown in table **??**.

.

2.3 Kepler modelling

We analyze the statistical properties of Kepler planets and compute the distribution of $\frac{dF}{db}$ with respect to the number of planets in the system. For the artificial planets, we randomly choose the radius from all the Kepler planets. We then generated the mass of the planets based on the radius by assuming the density is approximately equal to the earth's density and applied the equation below (Zhu et al. 2018).

$$m = 3.0 \cdot 10^{-6} \cdot r^{2.06} \tag{8}$$

We then randomly choose the period of the first planet in the system and generate the rest based on the period of the first through log uniform distribution with a lower limit of 1.3 and an upper limit of 4.0. The semi-major axis of the planets is calculated based on the Kepler 3rd law below, in which M is the mass of the sun, m is the mass of the planet, and T stands for the period of the planet.

$$a = \sqrt[3]{T^2\,(M_* + m)} \tag{9}$$

We generate the inclination angle by randomly choosing from the two-dimensional normal distribution, where both real parts and imaginary parts center at zero, with the standard deviation based on the number of planets in the system given by Zhu's equation below Zhu et al. (2018).

$$\sigma_I = 0.7(\frac{N}{5})^{-4} \tag{10}$$

We then calculate the orbital frequency, ω_{jk}, between planets j and planets k based on the formulas below. m_j and m_k stands for the mass of the two planets, while $a_>$ is the semi-major axis of the planets that far away from the sun, and $a_<$ is the one that close to the sun. The equations are given by Murray & Dermott (2000).

$$\omega_{jk} = \frac{G m_j m_k a_<}{a_>^2 L_j} b_{3/2}^{(1)}(\alpha) \tag{11}$$

where

$$L_j = m_j \sqrt{G M_* a_j} \tag{12}$$

$$b_{3/2}^{(n)}(\alpha) = \frac{1}{2\pi}\int_0^{\pi} \frac{\cos(nt)}{(\alpha^2 + 1 - 2\alpha\cos t)^{3/2}} \mathrm{d}t \tag{13}$$

We then calculated the changing of inclination with respect to the time of each two planets and sum up all pairs of planet-planet interactions. For j-th planet,

$$\frac{\mathrm{d}I_j}{\mathrm{d}t} = \sum_{k\neq j} \omega_{jk}(I_j - I_k) \tag{14}$$

Finally, we calculated the $\frac{\mathrm{d}b}{\mathrm{d}t}$ based on the equation below.

$$\frac{\mathrm{d}b}{\mathrm{d}t} = \frac{a}{R_*}\left|\frac{\mathrm{d}I}{\mathrm{d}t}\right| \tag{15}$$

We generate plenty of systems with different planet numbers, and plot the histogram for each system, while all the graphs are presented in the appendix. We also calculate the number of planets that obtained significant changing of impact parameters ($\frac{db}{dt} > 0.1$) for each system. All of the data are presented in table 5.

In the table, the first column displays the number of planets in each system, the second column stands for the total number of the artificial system we have generated and simulated. The third column means the total number of $\frac{\mathrm{d}b}{\mathrm{d}t}$ generated in the simulation. The fourth column means the percentile of $\frac{\mathrm{d}b}{\mathrm{d}t}$ that has a value bigger than 0.1 per year.

We found the number of planets in the system (N) has a strong exponential correlation, R-Squared equals to 0.9903, with the percentile of $\frac{db}{dt}$ that has the value bigger than 0.1 per year which is given by the equation below.

$$\frac{\mathrm{d}b}{\mathrm{d}t}\% = 13.5889 \cdot 0.7762^N \tag{16}$$

Moreover, we also generate the graph of $\frac{db}{dt}$ with respect to the mass and period of Kepler Planets, which is shown in figure 2. We construct an artificial two planets system in which the star mass and radius and a default planet with five day period and three-earth mass. The tested planets obtain mass ranging from two earth masses to three hundred

Table 3. $\frac{db}{dt}$ distribution with the possibility of significance for various planets' number system.

Model	Multiplicity Distribution	Relative Frequency
Zhu et al.(2018)	$N = 2$	8.598%
Zhu et al.(2018)	$N = 3$	6.108%
Zhu et al.(2018)	$N = 4$	4.761%
Zhu et al.(2018)	$N = 5$	3.806%
Zhu et al.(2018)	$N = 6$	3.069%
Lissauer et al.(2011)	$N \sim \mathrm{Pois}(3)$	2.016 %

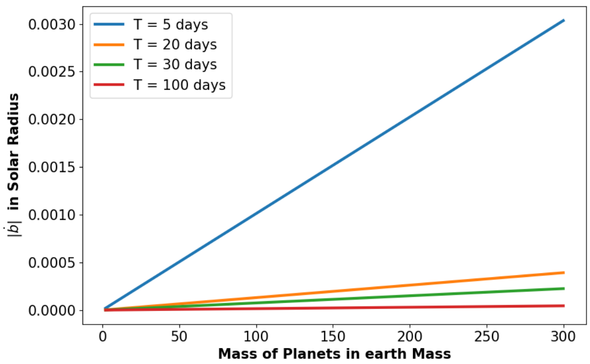

Figure 2. | $\frac{db}{dt}$ | respect to the mass of the planets. This graph presented the expected absolute value of $\frac{db}{dt}$ in the unit of the solar radius with respect to the mass of the planets in the unit of earth mass with four different periods. The colorful lines represent the results of different periods respectively.

Table 4. R Squared Value of Four Mass to $\dot{b}$ lines

Period	R-Squared
5	1.0000
20	1.0000
30	0.99999
100	1.0000

masses with periods of five days, twenty days, thirty days, and one hundred days respectively. The inclination angle is set at 2 degrees.

Finally, we calculate the best fit line for the four lines and express the equation in terms of period and mass, which are presented below.

$$\dot{b} = 0.75\frac{m}{T} \tag{17}$$

The R Squared values are shown in table 4.

2.4 Sample Selection

Theoretical SNR calculation reveals the signal of our model is directly proportionally to the planet star radius ratio, p, impact parameters, b, and $\dot{b}$, which is later proven in the Kepler modeling that obtained an inverse relationship with the period of the planet, T, proportional relationship with the mass of the planet, and more likely exist in the system with less number of the planet, N. Due to the reasons above, we selected the candidates that obtained planet-star radius ratio bigger than 0.03, but less than 0.1 due to the constraints of the limb darkening model that the light density from the star will not change inside the planet. The period of the planets is smaller than but has to be larger than in order to get enough data points for each transit as the short cadence of the Kepler telescope is thirty minutes. The mass of the planets is selected as larger than four earth masses. Finally, all the candidates are in the system that obtained four or fewer confirmed planets to increase the likelihood of the existence of significance of $\dot{b}$. The relationship of SNR with impact parameter, period, the mass of the planets, period, and the number of planets in the system is presented below.

$$SNR \propto (b, p, m, \frac{1}{N}, \frac{1}{T}) \tag{18}$$

3 MODEL

This chapter presents the detail of our model. 3.1 reveals the light curve Detrending process for the raw flux from the telescope and outlier removal. 3.2 demonstrates the parameters of the model and the generation of the model light curve. 3.3 presents the production of new parameters with the analysis of the accuracy with the detailed process of the model and refit the light curve.

3.1 Light Curve Detrending

We download the original data from Kepler and conducted the polynomial fitting for the light curve from the Kepler mission in order to reduce the noise and remove the outlier. We first scan all the light curves to locate the transit. We then calculated the relative time of the transit. In the equation below, $\tilde{t}$ is the time in hours from nearest transit, t means any data point from the time function, t_0 is the transit Epoch, and T means the period of the planet.

$$\tilde{t} \equiv (t - t_0 + \frac{T}{2}) \mod (T - \frac{T}{2}) \tag{19}$$

We extract the transit by blocking the transit and apply a third-degree polynomial to fit the transit, and then divide the blocked-out transiting part by the trend. After the third-degree polynomial, the linear and quadratic equations are also applied.

In addition, we smooth the flux function by Savitzky–Golay filter. The window of the Savitzky–Golay filter, integer W, is defined based on the equation below, in which cad is the time difference between each data point, the cadence of the transits.

$$W = \left[(\frac{1}{cad})^2 + 1 \right] \tag{20}$$

The filter flux $\bar{F}$, is generated by Savitzky–Golay filter based on the window above with a three-degree polynomial. We then remove iteratively all the data points in the original flux function, F_o, that is 5 σ away from filter flux, $\bar{F}$, until no more outliers are detected. Then the detrending flux function, F, is given by dividing the observed signal with a smoothed version of the trend.

$$F = \frac{\bar{F}}{F_0} \tag{21}$$

Assuming all the noise follows the gaussian distribution, we then applied the savgol factor to suppress the effect of noise on the data point. We weigh the value of the data point based on its distance from the neighbor data points and set the weights following the third-degree polynomial. We first set the window = 21. The savgol factor, SF, is given by the equation below, in which W is the window.

$$\mathrm{SF} = 1 - \sqrt{\frac{3(3W^2 - 7)}{W(W^2 - 4)}} \tag{22}$$

We then scaled the standard deviation by savgol factor and get the times function and the flux function.

3.2 Light Curve Generation

The light curve is generated based on seven variables: Transit Epoch which means the center of the first detected transit in Barycentric Julian Day (BJD) minus a constant offset of 2,454,833.0 days(t_0), the normalized impact parameters, b, which set into 0, the period of the planet which came from the kplr module(T), the planet star radius ratio, p, which approximated from Transit Depth, D, $p \approx \frac{\sqrt{D}}{10^3}$, the two limb darkening parameter u_1 and $u_2$set into 0.5, and the density of star, ρ which set into 1.

We then construct a vector composed of that seven variables and named them as parameters which are indicated in the equation below.

$$\vec{\Theta} =< t_0, b, T, p, u_1, u_2, \rho > \tag{23}$$

The star radius semi-major axis ratio, $\frac{a}{R_*}$, is approximated based on the density of the star, ρ, and the period of the planet, T, where G means the gravitational constant.

$$\frac{a}{R_*} = \sqrt[3]{\frac{GT^2\rho}{3\pi}} \tag{24}$$

We then generate the normalized separation of the centers, z, as a function of time.

$$z(t) = \sqrt{b^2 + (\frac{t - t_0}{T} 2\pi \frac{a}{R_*})^2} \tag{25}$$

After that, we generated the model light curve, $\bar{F}$, by using the equations below from the Agol Mandel and the batman package from Laura Friedberg. I(u) is the brightness of the star as a function of angle, and $u = \cos\theta$ is the polar angle of location on the star, and the function of u depends on the limb darkening.

$$u = \sqrt{1 - r^2} \tag{26}$$

r is the normalized radial coordinate of the star and $0 \leq r \leq 1$.

We assumed the planet is small enough that the no-limb darkening takes effect inside the area of the planets, and the limb darkening of the star follows the quadratic equation indicated below. Mandel & Agol (2002)

$$I(u) = 1 - u_1(1 - u) - u_2(1 - u)^2 \tag{27}$$

The flux of the star $\bar{F}$, with respect to normalized center distance, z, planet star ratio, p, and time t is presented below. Mandel & Agol (2002)

$$\hat{F}(z, p, t) = 1 - \frac{p^2 I^*(z)}{4\Omega} \tag{28}$$

which

$$I^*(z) = (4zp)^{-1} \int_{z+p}^{z-p} I(r) 2r \mathrm{d}r \tag{29}$$

and

$$\Omega = \frac{u_1 + u_2}{6} - \frac{u_2}{8} \tag{30}$$

3.3 Light Curve Modeling

This subsection introduced how we generated the new parameters with the analysis of the accuracy after the model and refit the light curve. The first part stated the statistical principle behind our optimization. The second part demonstrates the new parameters generated by the model with and without the changing of impact parameters after the removal of Transit time variation. The last part applied statistical power to analyze the accuracy of the model with $\frac{db}{dt}$.

3.3.1 *Statistic Power*

We intend to find the model that best explains the data from the telescope. We apply the Bayes inference to find the probability of the model we generated from $\overrightarrow{\Theta}$ and given the presence of the data from the Kepler telescope, $P(\text{model}|\text{data})$, which can be calculated based on the equation below, in which F means the detrending flux from the Kepler telescope and $\hat{F}$ means the model flux generated by $\overrightarrow{\Theta}$

$$P(\mathrm{F}|\hat{\mathrm{F}}(\overrightarrow{\Theta}, \mathrm{t})) = \frac{P(\hat{\mathrm{F}}(\overrightarrow{\Theta}, \mathrm{t})|\mathrm{F})P(\mathrm{F})}{P(\hat{\mathrm{F}}(\overrightarrow{\Theta}, \mathrm{t}))} \tag{31}$$

We arrange the equation and obtained the following:

$$P(\hat{\mathrm{F}}(\overrightarrow{\Theta}, \mathrm{t})|\mathrm{F}) = P(\mathrm{F}|\hat{\mathrm{F}}(\overrightarrow{\Theta}, \mathrm{t})) \cdot P(\hat{\mathrm{F}}(\overrightarrow{\Theta}, \mathrm{t})) \cdot \frac{1}{P(\mathrm{F})} \tag{32}$$

The data from the telescopes are all constant and can not manipulate, thus P(F) is a constant. Because $P(\mathrm{F}|\hat{\mathrm{F}}(\overrightarrow{\Theta}, \mathrm{t}))$ is the prior, it is reasonable to assume all underlying models are equally likely to be generated, thus can be set equal to 1, which is also constant. Because of the reason above, $P(\hat{\mathrm{F}}(\overrightarrow{\Theta}, \mathrm{t})|\mathrm{F})$ is directly proportional to $P(\mathrm{F}|\hat{\mathrm{F}}(\overrightarrow{\Theta}, \mathrm{t}))$.

the formula of the log, the estimate came from MLE, MLE is based on the first derivate of the log likelihood, hessian calculate variance

We let L denoted the probability of data F given parameters $\overrightarrow{\Theta}$ and model $\hat{F}(\overrightarrow{\Theta}, t)$

$$L = P(\mathrm{F}|\hat{\mathrm{F}}(\overrightarrow{\Theta}, \mathrm{t})) \tag{33}$$

We assume the signal is only composed of the true signal and the random noise. Because the noise is completely random, thus the distribution is gaussian, thus the $\log L$ can be obtained from the equation below and named $G(\overrightarrow{\Theta})$. In this case, the input time function, t, means $\tilde{t}$ is the time in hours from nearest transit, which is a constant. The equation is constraints under four conditions: $|b_0| < 1$, both of u_1 and $u_0 > 0$, $u_1 + u_0 < 0$, and $R_x > 0$.

$$G(\overrightarrow{\Theta}, t) = \log L(\overrightarrow{\Theta}, t) = -\sum \frac{(\hat{F}(\overrightarrow{\Theta}, t) - F)^2}{2\sigma^2} \tag{34}$$

in which the standard deviation of the flux, σ, is given by

$$\sigma = \sqrt{\frac{\sum \mid F - \hat{F}(\Theta, t) \mid}{n}} \tag{35}$$

For $G(\overrightarrow{\Theta})$, minimized the square error $(\hat{F}(\overrightarrow{\Theta}) - F)^2$ is the same as maximizing log L. Therefore, the least square is the maximum likelihood.

3.3.2 *Optimization though TTV and TDV*

We remove the effects of transit time variation by adding or subtracting a certain number that the center of the transit is equal to 0 in "folded" time. To find the maximum likelihood of transit time variation, we defined another log probability function $G_t(\overrightarrow{\Theta})$. The only difference between $G_t(\overrightarrow{\Theta})$ and $G(\overrightarrow{\Theta})$ is that the input time function t, equals $\tilde{t}$+ TTV, in which TTV is a variable that needs to be fit.

We also define the equation for chi-square value for the model that involves $\frac{db}{dt}$, the impact parameters b(t) function respect to time is given by

$$b(t) = b_0 + \frac{\mathrm{d}b}{\mathrm{d}t}\Delta t \tag{36}$$

Thus the new vector, which we name $\overrightarrow{\Theta}_b$, obtained a new variable $\dot{b}$, compare with $\overrightarrow{\Theta}$ as b is changing respect to time.

$$\vec{\Theta}_b =< t_0, b_0, P, R_x, u_1, u_2, \rho, \dot{b} > \tag{37}$$

The corresponding Log probability function, G_b, is presented below

$$G_b(\vec{\Theta}_b, t) = -\sum \frac{(\hat{F}(\vec{\Theta}_b, t) - F)^2}{2\sigma^2} \tag{38}$$

We apply the basinhopping from scipy to find the global minimum of $G_t(\overrightarrow{\Theta}, t)$ and its corresponding $\overrightarrow{\Theta}$value and t value, named it $\overrightarrow{\Theta}^{(1)}$ and $t^{(1)}$. We then consider the involvement of $\dot{b}$ and plug in $\overrightarrow{\Theta}^{(1)}$into G_b. After the same process, we got $\overrightarrow{\Theta}_b^{(1)}$ and $t^{(2)}$. We then compare the model flux, $\hat{F}(\overrightarrow{\Theta_b^{(2)}}, t^{(2)})$, with the actual flux from Kepler telescope, F, and remove all the data points that obtain more than 15 residuals from the predicted model. We applied basinhopping again to refit the model after the outlier is removed with G and G_b separately and obtain $\overrightarrow{\Theta}^{(2)}$ and $\overrightarrow{\Theta}_b^{(2)}$.

3.3.3 Accuracy

The fisher information of $\vec{\Theta}_b$ is calculated by

$$I(\vec{\Theta}_b) = E[-\frac{\mathrm{d}^2}{\mathrm{d}\vec{\Theta}_b^{\,2}} G_b\left(\vec{\Theta}_b, t\right)] \tag{39}$$

Therefore, the difference of the estimating parameters between the two models will form a normal distribution that centers at zero with variance inverse proportional to the fisher information, $\vec{\Theta}_b$, via

$$\sqrt{n}(\vec{\Theta}_b - \vec{\Theta}) \rightarrow N(0, \frac{1}{I(\vec{\Theta}_b)}) \tag{40}$$

The corresponding z-score of the difference, $Z_{\Delta G}$ is then computed from the normal distribution presented within equation 1. Similarly, the z-score of $\dot{b}$, $Z_{\dot{b}}$ is derived from equations above. Lastly, we calculate the signal-to-noise ratio, , S_r, from equation (15) from Pont et al. (2006)

$$S_r = \alpha^{1/2} p^2 N^{1/2} \delta_t^{1/2} T^{-1/2} \sigma_d^{-1} \tag{41}$$

with p stands for the planet-star radius ratio, N is the number of data points, δ_t is the short cadence time of Kepler telescope, T is the period of the planets, and σ_d is the uncertainty that calculated within equation 40. α is limb darkening parameters which we set into 1. where $\Theta_b^{(2)}$ and $\Theta^{(2)}$ from iteration 2.

4 DISCUSSION AND CONCLUSION

The stacked plot and folded plot of each planet were drawn based on the data calculated above, in which each transit is represented by the different color of data points, and the outliers of the transits are represented by the cross.

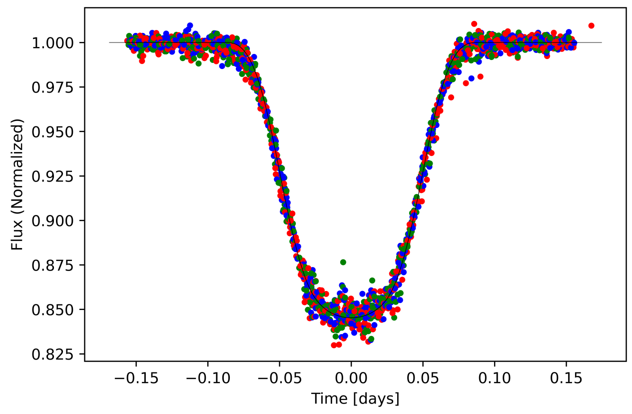

Figure 3. Deterending Lightcurve for Koi 6187.01. The x-axis represents the last of each transit in the unit of days, the y-axis represents the normalized flux from the star. Every dot represents a data point from the Kepler telescope and different colors marked the data points from different transit. The dark lines are the best fit model for each transit.The outliers that are 5 σ away are marked as cross (x).

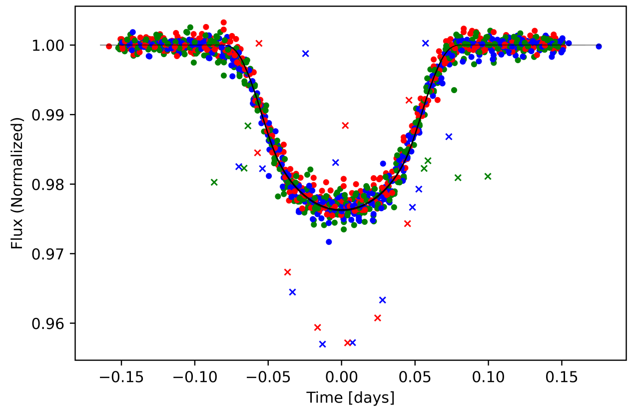

Figure 4. Deterending Lightcurve for Koi 6236.01. The x-axis represents the last of each transit in the unit of days, the y-axis represents the normalized flux from the star. Every dot represents a data point from the Kepler telescope and different colors marked the data points from different transit. The dark lines are the best fit model for each transit.The outliers that are 5 σ away are marked as cross (x).

4.1 Table of Candidates

We searched 3344 planets and we found planets koi 6187.01, koi 6236.01, koi 3627,01, 3659.01, and 5616.01 have a clear transit shape. Among those, the system has demonstrated a clear Transit Depth Variation. All of the statistics and data are provided in the table below.

4.2 Discussion

Our research is focused on the application of impact parameters. The distribution of $\frac{dF}{db}$ could help future scientists to choose a possible candidate in this method. We find there are five planets have significant long-term linear changes in impact parameters, which are koi 6187.01, koi 6236.01, koi 3627.01, koi 3659.01, and koi 5615.01.

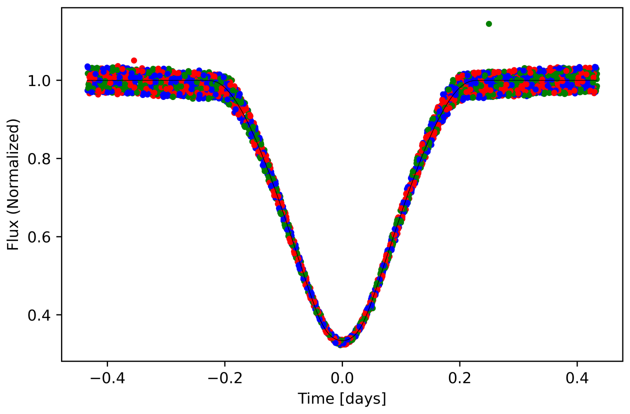

Figure 5. Deterending Lightcurve for Koi 3627.01. The x-axis represents the last of each transit in the unit of days, the y-axis represents the normalized flux from the star. Every dot represents a data point from the Kepler telescope and different colors marked the data points from different transit. The dark lines are the best fit model for each transit.The outliers that are 5 σ away are marked as cross (x).

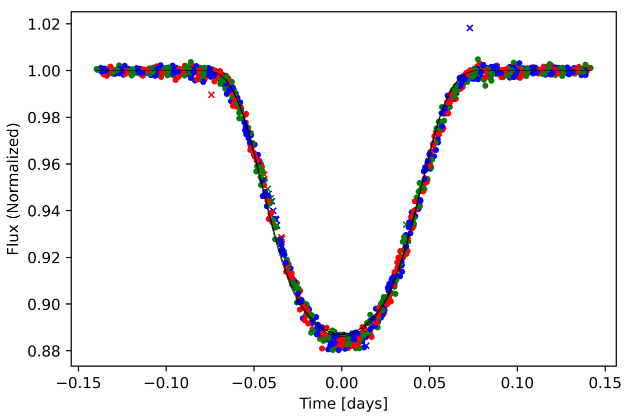

Figure 6. Deterending Lightcurve for Koi 3659.01. The x-axis represents the last of each transit in the unit of days, the y-axis represents the normalized flux from the star. Every dot represents a data point from the Kepler telescope and different colors marked the data points from different transit. The dark lines are the best fit model for each transit.The outliers that are 5 σ away are marked as cross (x).

Table 5. The Statistical Significance of Satisfied Planets

Planet	Num Data Points	Num Transits	$Z_{\Delta G}$	$Z_{\dot{b}}$	SNR
koi 6187.01	438	16	-17.1	11.77	13.96
koi 6236.01	300	20	11.3	13.11	80.53
koi 3627.01	259	32	17.2	23.34	244.89
koi 3659.01	124	13	12.9	10.39	58.27
koi 5615.01	212	6	11.7	19.48	126.59

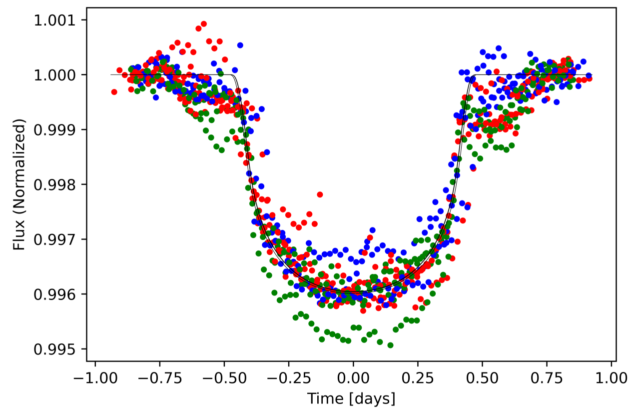

Figure 7. Deterending Lightcurve for Koi 5619.01. The x-axis represents the last of each transit in the unit of days, the y-axis represents the normalized flux from the star. Every dot represents a data point from the Kepler telescope and different colors marked the data points from different transit. The dark lines are the best fit model for each transit.The outliers that are 5 σ away are marked as cross (x).

This process will allow us to increase the precision of their inclination, eccentricity, and mass, which is crucial for searching habitable planets. We could also used the long term changes of planetary orbit to suspect the existence of non-transit planets and partially resolve the Kepler dichotomy.

The bulk of planets in the research has only been searched with a long cadence due to computing power limitations. Therefore, we suggest using a short cadence for better statistical significance and more accurate parameters. For instance, we shouldn't plot outliers that are too far from the model and should only look for confirmed planets, not all potential planets. To find the optimal parameters and get rid of noise, we employ Bayesian inference and the log maximum likelihood function in the study.For future studies, we recommend applying more advanced model for light curve detrending, eg deep learning, along consideration of polynomial fitting for impact parameters instead of an elementary linear approach.

REFERENCES

Batalha N. M., et al., 2013, ApJS, 204, 24
Borucki W. J., et al., 2010, Science, 327, 977
Borucki W. J., et al., 2011, ApJ, 736, 19
Fabrycky D. C., et al., 2014, ApJ, 790, 146
Fang J., Margot J.-L., 2012, ApJ, 761, 92
Hadden S., Lithwick Y., 2014, ApJ, 787, 80
Hamann A., Montet B. T., Fabrycky D. C., Agol E., Kruse E., 2019, AJ, 158, 133
Judkovsky Y., Ofir A., Aharonson O., 2020, AJ, 160, 195
Mandel K., Agol E., 2002, The Astrophysical Journal, 580, L171–L175
Millholland S. C., He M. Y., Ford E. B., Ragozzine D., Fabrycky D., Winn J. N., 2021, AJ, 162, 166
Murray C. D., Dermott S. F., 2000, Solar System Dynamics. Cambridge University Press, doi:10.1017/CBO9781139174817
Pont F., Zucker S., Queloz D., 2006, Monthly Notices of the Royal Astronomical Society, 373, 231
Tremaine S., Dong S., 2012, AJ, 143, 94
Zhu W., Dong S., 2021, ARA&A, 59, 291
Zhu W., Petrovich C., Wu Y., Dong S., Xie J., 2018, The Astrophysical Journal, 860, 101

APPENDIX A: DERIVATION OF SNR EQUATION

$F_b(x)$ is the flux function that regard $\dot{b} = 0$, while the $F(x)$ is the flux function that treat $\dot{b}$ as a finite non-zero number. The independent variable of the function is defined as $z = \frac{d}{R_*}$, which d means the distance between the center of the planet and the center of the star, and R_*stands for the radius of the star. x is ranged from -1 to 1. At the first transit, two functions obtain the same result, and we applied linear approximation to estimate $F_b(x)$ after N orbits, which is presented below.

$$F_b(z) \approx F(z) + \frac{\mathrm{d}F(z)}{\mathrm{d}b}\Delta b \tag{A1}$$

Δb is given by the equation below and N_T means the completion of the nth orbit which is also the number of transits and T is the orbital period.

$$\Delta b = \dot{b} N_T T \tag{A2}$$

The mean difference of $F_b(x)$ and $F(x)$, $\left\langle \delta F^2 \right\rangle$, is given by the equation below

$$\left\langle \delta F^2 \right\rangle \approx \frac{1}{(x_{max} - x_{min})} \int_{x_{min}}^{x_{max}} (\frac{\mathrm{d}F(x)}{\mathrm{d}b}\Delta b)^2 \mathrm{d}x \tag{A3}$$

Plugging in the corresponding values, we get the equation below, and p stands for the planet-star ratio.

$$\approx \frac{1}{2(1+p)} \int_{-1-p}^{1+p} (\frac{\mathrm{d}F(x)}{\mathrm{d}b}\frac{\mathrm{d}b}{\mathrm{d}t} T N_T)^2 \mathrm{d}x \tag{A4}$$

The signal-to-noise ratio will be the sum of the difference of all data points in the two models, which is given by the equation below.

$$\mathrm{SNR}^2 = \sum \frac{(F_b(z) - F(z))^2}{\sigma_b^2} \tag{A5}$$

The σ is relative to the cadence of we applied as an accurate measurement means less variation of the data points, so we applied $\frac{\tau}{cad}$to fix the value of σ and obtained the equation below.

$$SNR^2 = \left\langle \delta F^2 \right\rangle \cdot (\frac{\tau}{cad}) \cdot \frac{1}{2\sigma_b^2} \tag{A6}$$

APPENDIX A: HISTOGRAM OF IMPACT PARAMETER FOR MULTIPLE NUMBERS OF PLANET SYSTEM

If you want to present additional material which would interrupt the flow of the main paper, it can be placed in an Appendix which appears after the list of references.

This paper has been typeset from a TEX/LATEX file prepared by the author.

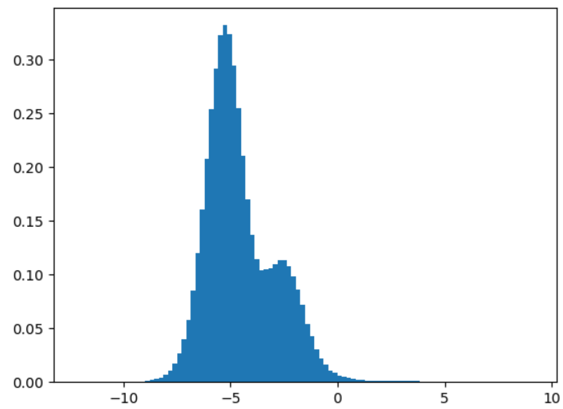

Figure A1. The histogram of $\frac{db}{dt}$ for two-planet system. The x-axis represented the value of $\dot{b}$ in the $\log 10$ scale, and the y-axis represented their relative frequency in the population.

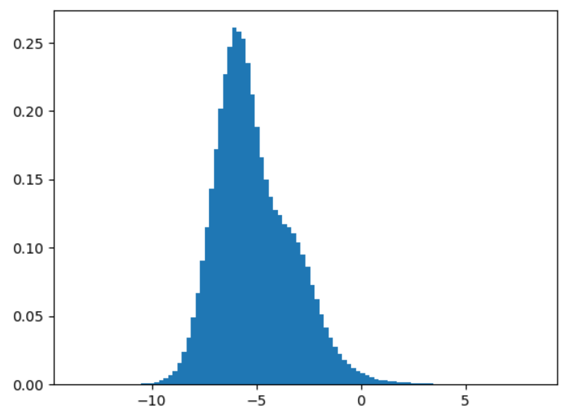

Figure A2. The histogram of $\frac{db}{dt}$ for three planet system.The x-axis represented the value of $\dot{b}$ in the $\log 10$ scale, and the y-axis represented their relative frequency in the population.

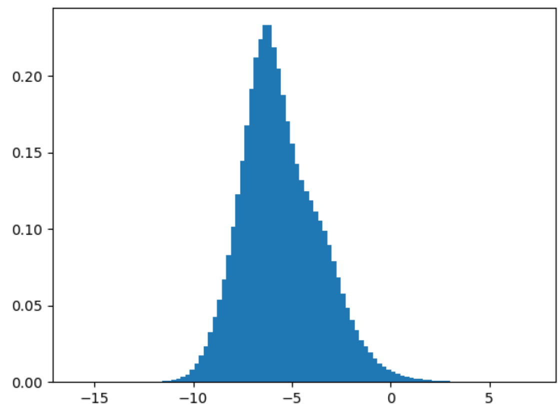

Figure A3. The histogram of $\frac{db}{dt}$ for four-planet system.The x-axis represented the value of $\dot{b}$ in the log 10 scale, the y-axis represented their relative frequency in the population

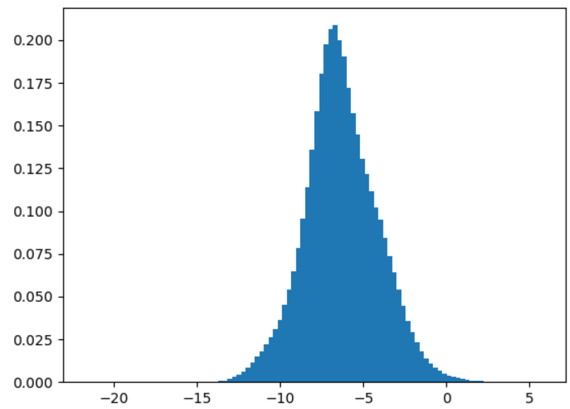

Figure A5. The histogram of $\frac{db}{dt}$ for six planet system.The x-axis represented the value of $\dot{b}$ in the log 10 scale, the y-axis represented their relative frequency in the population

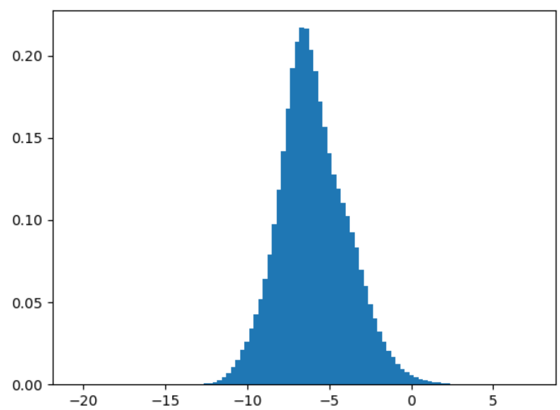

Figure A4. The histogram of $\frac{db}{dt}$ for five planet system. The x-axis represented the value of $\dot{b}$ in the log 10 scale, y-axis represented their relative frequency in the population