Draft version January 23, 2023
Typeset using LaTeX default style in AASTeX631

nuance: Detection of planetary transits in the presence of correlated noise

Lionel J. Garcia, Daniel Foreman-Mackey, Fran Pozuelos

ABSTRACT

We present nuance, an algorithm to search for planetary transits in light curves featuring correlated noises, such as instrumental signals and stellar photometric variability. Simpler approaches consist in searching for transits in light curves cleaned from correlated noise, where only transit signals would remain. However, we show that commonly used detrending techniques strongly affect transits signal-to-noise, up to the point of no detection when the noise characteristics is close to that of the transits. Focusing on stellar variability, we explore the parameter space for which this degradation occur, and quantify its effect on variety of cases.

INTRODUCTION

Exoplanets, planets outside our solar system, are discovered at an ever-increasing rate. Beyond the study of their inner structure and atmosphere, they give a unique glimpse to extrasolar systems' formation, dynamics, as well as being probes to understand exoplanets' host stars. This is particularly true for systems whose orbital plane is aligned with our line of sight, leading to observable planetary transits. Although these transit signals can be seen in the apparent flux received from their stars, they are often mixed with other astrophysical and instrumental signals. If they can be disentangled from these nuisance signals, transits offer a powerful way to detect exoplanets. However, disentangling these signals comes with many challenges. Due to their nature we will refer to these signal as correlated noises

Widely used transit-search algorithms (Box-Least-Square algorithms **?**) are capable of detecting transits on light curves only containing transits and white noise. Hence, the simplest way to find periodic transit signals is to first clean a light curve from nuisance signals before performing the search. This strategy is widely adopted by the community, both using physically-motivated systematics models like Luger et al. (2016, 2018), or empirical filtering techniques, such as the ones described and implemented in Hippke et al. (2019). However, when correlated noises start resembling transits, this cleaning step (often referred to as *detrending*) strongly degrades their detectability. In such cases, the only alternative to search for transits is to perform a full-fledged modeling of the light curve, including both transits and correlated noises, and asses the likelihood of the model to the data on a wide parameter space, an approach largely avoided due to its untractable nature. Nonetheless, Kovács et al. 2016 ask the question: *Periodic transit and variability search with simultaneous systematics filtering: Is it worth it?*. The latter study discards the general benefit of using a full-fledged approach, but it fails at exploring the light-curves characteristics for which it becomes necessary.

We present nuance, an algorithm using linear models and Gaussian processes (GP) to simultaneously search for transits while modeling correlated noise in a tractable way, such as instrumental signals and stellar photometric variability. In section 1, we describe the issues inherent to the two-step approach described earlier, and study the parameter space for which commonly used detrending techniques degrade transit signals to the point of no detection. In section 2 we describe the tractable full-fledged approach of nuance, and its implementation in an open-source Python package. In section 3 we assess the performance of nuance by computing the recovery of planetary transits injected into simulated light curves. In section 4, we use nuance in a case study, to detect a known planet observed by the T E S S but undetected by the TESS pipeline due to its light curve characteristics. In section 5, we use nuance to search for transits in a list of .... Finally, we conclude this paper by providing avenues for the improvement of nuance, from its advantage to search for transits in ground-based observations, to its potential for the detection of multi-planetary systems affected by transit-times variations.

1. THE ISSUE WITH CORRELATED NOISE AND ITS DETRENDING

Two sources of correlated noise particularly justify the need for a detrending step before searching for transits: instrumental noises (such as telescope pointing errors) and stellar variability (induced by pulsations or starspots). In this section, we discuss the impact of such correlated noises and their detrending on transits detectability. For simplicity, we will model transit signals using an analytic empirical model described in Protopapas et al. 2005 (see Annexe).

1.1. *The effect of correlated noise on transits detectability*

To study transits detectability, we will focus on the signal-to-noise (SNR) of a unique event, reduced to the simplified expression (Pont et al. 2006, Equation 12):

$$SNR = \frac{df}{\sqrt{\frac{\sigma_w^2}{n} + \frac{\sigma_c^2}{N_{tr}}}} \tag{1}$$

where df is the relative transit depth, n is the number of points within transit, N_{tr} the number of transits (unity here since we consider a single transit), and σ_w and σ_c are the white and correlated noise standard deviations. To show the impact of correlated noise on transit detectability, we simulate a unique transit signal and compute its SNR using Equation 1, both in the absence and presence of correlated noise (Figure 1).

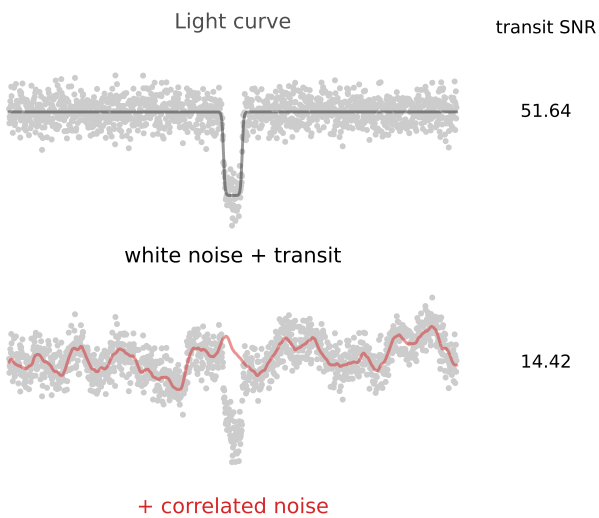


Figure 1. Illustration of the effect of correlated noise on a single transit signal-to-noise (SNR). A 1-hour transit of depth 1% is simulated on top of white noise as part of a 24-hours observation with an exposure time of 1 minute (top). Then, in the bottom plot, correlated noise is added to the transit signal and simulated using a Gaussian Process (GP) with a Matèrn-32 kernel of length-scale 1 hour and amplitude 0.2%. The SNR on the right of each light curve is computed using Equation 1.

As illustrated in Figure 1, the presence of correlated noise strongly decreases the transit signal SNR, limiting its detection. This issue rapidly motivated the development of systematics detrending algorithms such as the Trend Filtering Algorithm (TFA, Kovács et al. 2005, in its primary use case), SysRem (Tamuz et al. 2005), or Pixel Level Decorrelation (PLD, Deming et al. 2015; see also Everest from Luger et al. 2016, 2018). Most of these methods rely on the shared nature of instrumental signals among light-curves (or neighboring pixels) such that the correction applied should not degrade the transit signal. We note that, except for TFA, these algorithms are mostly applied to space-based continuous observations, that provide continuous stellar baselines and mostly reproducible systematic signals. This is not the case for the vast majority of ground-based observations, in addition subject to periodic daytime interruptions and varying atmospheric extinction.

1.2. *The effect of detrending on transits detectability*

Instrumental signals have the benefit to be shared among light curves of stars observed with the same instrument, strongly correlated with measurements from the experimental setup (like detector's temperature, pointing error, sky background or airmass), hence we will make the strong assumption that their detrending based on an incomplete model of the light curve, one that ignore transit signals (because unknowns), do not affect the latter. In opposition, stellar variability is generally unknown and harder to correlate with simultaneous measurements. This gave rise to two types of treatments in order to reconstruct and detrend stellar variability, and perform transit search on *flattened* light curves. One is physically-motivated and make use of Gaussian Processes (e.g. Aigrain et al. (2016)). The other is empirical and make use of filtering algorithms (a wide variety being described in Hippke et al. (2019)). In this section, we will study in details the effect of both approaches on transit detectability, by studying the degradation of a unique transit SNR for a wide variety of stellar variability characteristics.

Throughout this paper, we simulate stellar variability (and later model it) thanks to Gaussian Processes (GP), employing the physically-motivated stochastically-driven damped simple harmonic oscillator kernel (SHO) presented in Foreman-Mackey et al. 2017a. Extensive details about the following simulations can be found in Appendix A.

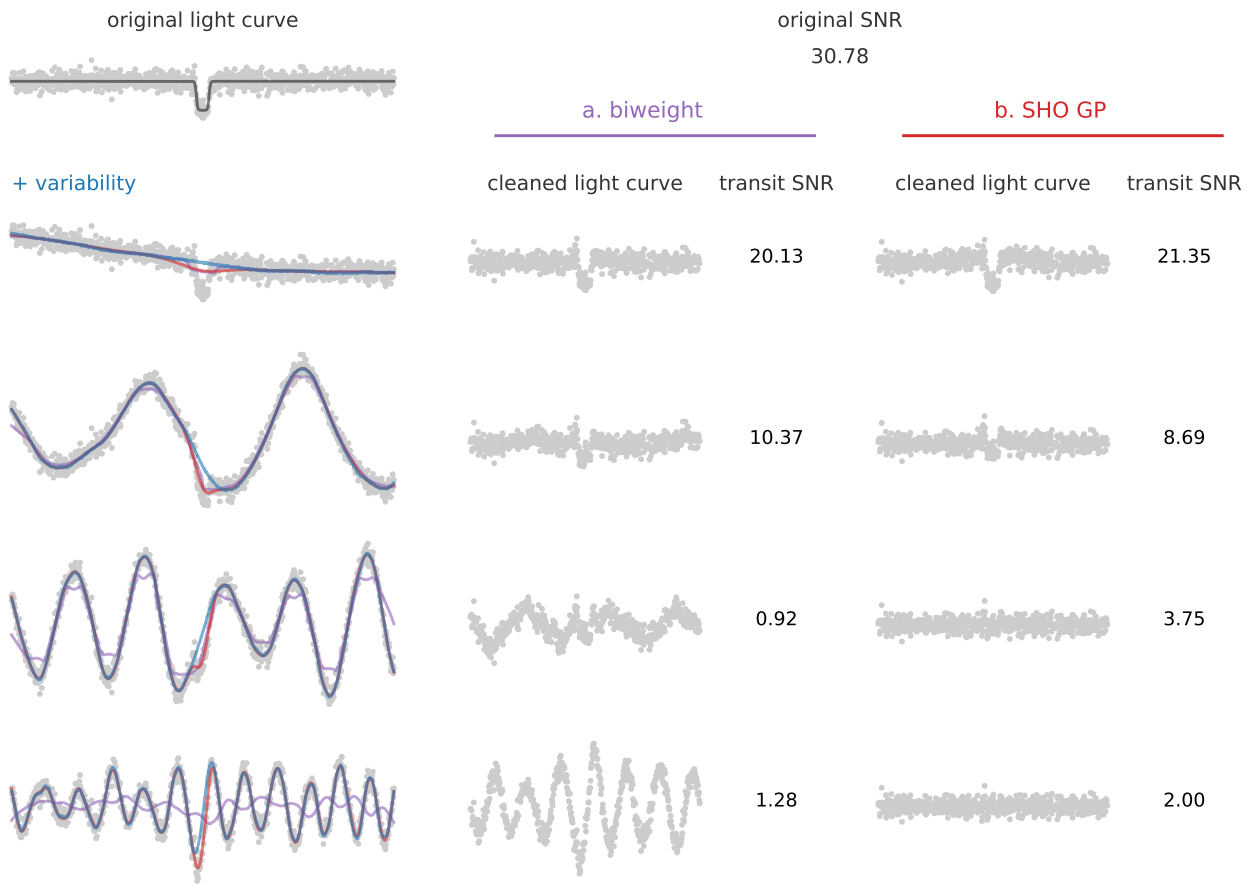


Figure 2.

In Figure 2, we simulate a light curve featuring a single transit (again using the model from Protopapas et al. 2005), in addition with stellar variability of different timescales and amplitudes. In case *a* (purple in Figure 2), we model and detrend these signals using the widely-adopted Tukey's biweight filtering method presented in Mosteller & Tukey 1977 and its implementatation from `wõtan`[1] (Hippke et al. 2019). In case *b* (red in Figure 2) we employ the same

[1] https://github.com/hippke/wotan

GP model used to simulate stellar variability to reconstruct and remove it from light curves ignoring the presence of potential transits (the case of unknown transits search). Once transits cleaned from stellar variability, we asses their remaining SNR using Equation 1.

Figure 2 clearly shows the effect of both detrending techniques on transits SNR, and suggests that this degradation is strongly dependant on the stellar variability characteristics encountered. In order to explore the parameter space for which detrending is the most problematic, we simulate light curves including a single transit and a much wider range of correlated noise characteristics defined with two parameters: τ_v the relative timescale of the variability with respect to the transit duration; and δ_v, the relative amplitude of the variability against transit depth, such that:

$$\tau_v \propto \frac{\text{variability timescale}}{\text{transit duration}} \quad \text{and} \quad \delta_v \propto \frac{\text{variability amplitude}}{\text{transit depth}} \tag{2}$$

To follow this parametrization, we simulate the photometric variability using a GP with an SHO kernel with hyper-parameters [2]:

$$\omega = \frac{\pi}{\tau_v \tau} \quad \sigma = \delta_v \frac{\delta}{2} \quad \text{and} \quad Q = 10 \tag{3}$$

with δ and τ the depth and duration of the simulated transit, similar in all light curves. We fix a relatively high value for the quality factor Q in order to restrict our simulations to strongly periodic variability signals. For $\tau_v = 1$ and $\delta_v = 1$, the expressions of ω and σ given in Equation B6 correspond to a periodic signal with a period half that of the transit duration, and a peak to peak amplitude two times that of the transit depth, i.e. strongly resembling the simulated transit signal. As in Figure 2, we reconstruct the variability using a biweight filter with a window length of 3τ (three times the transit duration, an optimal value according to Hippke et al. 2019). We then subtract the variability from each light-curve, estimate the resulting transit depth (using the in-transit minimum flux) and compute its SNR. Figure 3 shows the remaining SNR values of transits computed this way, after each variability signals with random (τ_v, δ_v) have been detrended.

Figure 3 shows that it exists an entire region of the (τ_v, δ_v) parameter space for which detrending degrades transit SNR to the point of no detection ($SNR < 6$). While this region might only represent a fraction of existing exoplanetary systems, we argue that these systems are of great value for the field of exoplanetary science. First, as variability is often linked to the presence of starspots, systems whose with the stellar photo in is likely to be observed on stars whose equatorial plane lines up with our line of sight, following recent studies on the dynamical evolution of planetary systems , hence being most likely to be transited. As a direct consequence, these same starspots are more likely to be occulted by planetary companions, events whose measurement benefits both the study of stellar atmospheres and their concerning impact on planetary atmospheric retrievals . Finally, the growing interest of the community for ultra-cool dwarf stars comes with observations done in the infrared, featuring enhanced red noise and variability, leading to lower transit SNR . Hence, commonly-used detrending technique (such as the biweight filter studied in this section) make transit-search blind to those pristine systems, motivating the need for a more informed transit search algorithm able to deal with correlated noises. In the next section we present such an algorithm: `nuance`.

[2]

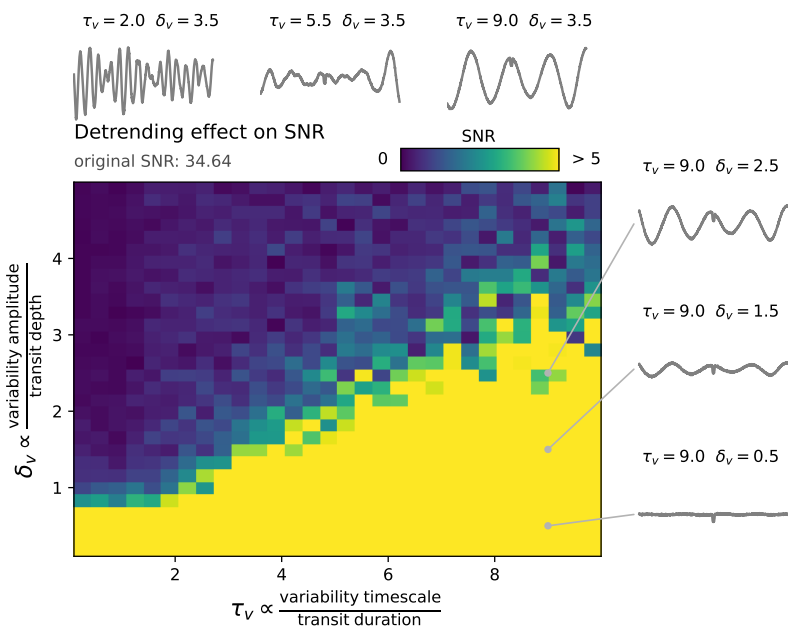


Figure 3.

2. NUANCE

In this section we describe `nuance`, an algorithm to search for planetary transits in light curves containing correlated noises, such as instrumental signals and stellar photometric variability.

Let a flux f of a star measured at time t be sampled and aranged in the $(1 \times N)$ column-vector $\boldsymbol{f}$. This flux contains instrumental signals, stellar variability and a periodic transit signal that we wish to uncover. An example of such signal is simulated and shown in Figure 4.

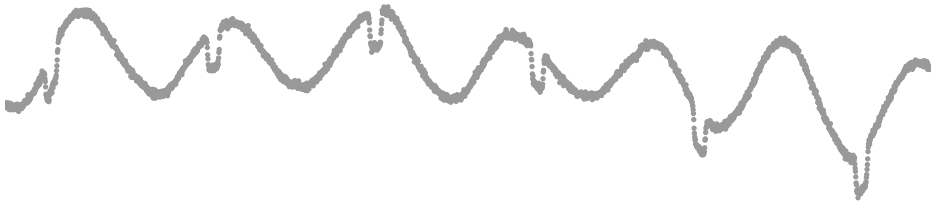


Figure 4.

Simultaenously to the flux measurement, we obtained a set of M measurements, aranged in the $(M \times N)$ matrix $\boldsymbol{X}$, that can be treated as explanatory variables for $\boldsymbol{f}$. Details about the simulation of these signals are provided in Appendix A.

Ideally, we would detect such a periodic transit signal by sampling the posterior likelihood of this data to a full-fledge model including stellar variability (more generally correlated noises), instrumental systematics (modeled with explanatory variables), and a periodic transit signal of period P, epoch T_0, duration D and depth Δ. We would then reduce the posterior likelihood to $p(\boldsymbol{f}|P)$, its marginalized version over all parameters except the period P, producing a transit search periodogram $\mathcal{Q}(P)$. However, this approach has two issues: It is highly intractable, and it may lead to multimodal distributions that are hard to interpret.

Given a period P, we instead want to compute the likelihood of a periodic transit signal at the maximum likelihood parameters $\hat{T}_0$, $\hat{D}$ and $\hat{\Delta}$, i.e the periodogram

$$\begin{aligned} \mathcal{Q}(P) &= p(\boldsymbol{f}|P, \hat{T}_0, \hat{D}, \hat{\Delta}) \\ \text{where} \quad (\hat{T}_0, \hat{D}, \hat{\Delta}) &= \underset{T_0, D, \Delta}{\arg\max}\, p(\boldsymbol{f}|P, T_0, D, \Delta) \end{aligned} \tag{4}$$

We will do that by adopting the strategy of Foreman-Mackey et al. (2015), and separate the transit search into two components: the *linear search* and the *periodic search*. During the *linear search*, the likelihood of a single non-periodic transit is computed for a grid of epochs and durations (but not depths, as we will see). Then, the *periodic search* consists in combining these likelihoods to compute the likelihood of the data given a periodic transit signal for a range of periods. These combined likelihoods yield a transit-search periodogram on which the periodic transit detection can be based. `nuance` differs by modeling the covariance of the light curve with a Gaussian Process, accounting for correlated noise (especially in the form of stellar variability) while keeping the model linear and tractable.

2.1. *The linear search*

During the *linear search*, the goal is to compute the likelihood $p(\boldsymbol{f}|T, D, \Delta)$ of our data given a single non-periodic transit signal of epoch T, duration D and depth Δ, for a grid of epochs, durations and depths.

To account for correlated noises, we model the light curve f as being drawn from a Gaussian Process such that

$$\boldsymbol{f} \sim \mathcal{N}(\boldsymbol{w}\boldsymbol{X}, \boldsymbol{\Sigma})$$

with mean $\boldsymbol{w}\boldsymbol{X}$ (i.e. a linear model of the M explanatory variables with coefficients $\boldsymbol{w}$) and covariance $\boldsymbol{\Sigma}$. To account for the presence of a single non-periodic transit of epoch T and duration D, we compute and append its signal as the last column of the design matrix $\boldsymbol{X}$, using the simple transit model from Protopapas et al. 2005 with a unitary depth. This way, the transit signal is part of the linear model and its depth Δ can be solved linearly. Under this assumption, the log-likelihood of the data given a single non-periodic transit is analytical

$$\ln p(\boldsymbol{f}|I) = -\frac{1}{2}(\boldsymbol{f} - \boldsymbol{w}\boldsymbol{X})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{f} - \boldsymbol{w}\boldsymbol{X}) - \frac{1}{2}|\boldsymbol{\Sigma}| - \frac{N}{2}\ln 2 \tag{5}$$

and can be computed on a grid of epochs and durations, the transit depth being linearly solved for any (T, D). This *linear search* lead to the set of log-likelihoods[3]

$$\ln \mathcal{L} = \{\ln p(\boldsymbol{f}|T_i, D_j, \Delta_{i,j})\}_{i,j}$$

were $\Delta_{i,j}$ is the depth linearly solved for a given (D_i, T_j). Figure 5 shows this likelihood grid computed for the simulated dataset shown in Figure 4, using the same Gaussian Process and design matrix $\boldsymbol{X}$ used to simulate the data (see annexe).

[3] This notation omits the vector $\boldsymbol{w}_{i,j}$ (except its last value $\Delta_{i,j}$) as it is linearly solved for any given value of (T_i, F_j) and irrelevant in what follows.

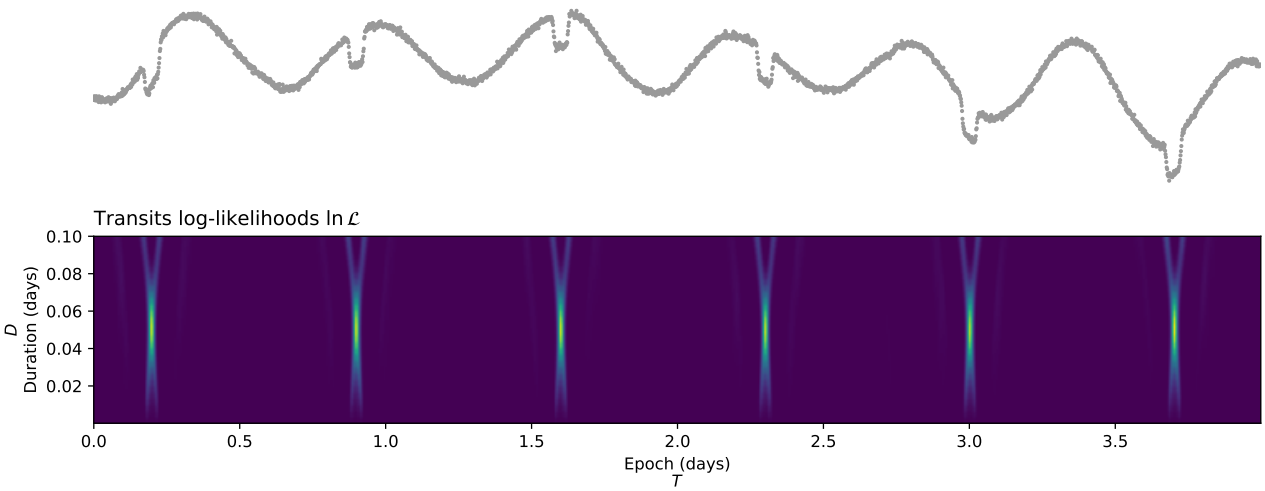


Figure 5. Principle and output of the *linear search*. First, a set of durations and depths $\{ T_i, D_j \}_{i,j}$ is generated. For each pair of index (i, j), the log likelihood $\ln p(\boldsymbol{f}|T_i, D_j, \Delta_{i,j})$ is computed, using the parameters from Equation 6 and the expression of Equation 5. This process yields the grid of log-likelihoods $\ln \mathcal{L}$ (bottom plot), as well as the $\{ \Delta_{i,j}, \sigma_{i,j} \}_{i,j}$ transit depths and errors.

To prepare the next step, the corresponding depths $\Delta_{i,j}$ linearly solved for any (T_i, D_j) are conserved, as well as their associated errors $\sigma_{i,j}$ obtained through the least-square solution

$$\begin{aligned} \boldsymbol{w} &= (\boldsymbol{X}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{f} \\ \boldsymbol{\sigma} &= (\boldsymbol{X}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{X})^{-1} \end{aligned} \tag{6}$$

where $\Delta_{i,j} = \boldsymbol{w}_M$ and $\sigma_{i,j} = \boldsymbol{\sigma}_{MM}$

2.2. *The periodic search*

We then need to combine the likelihoods computed from the *linear search* to obtain

$$p(\boldsymbol{f}|P, T_0, D, \Delta) \tag{7}$$

For a given transit duration D, any combination of (P, T_0) leads to K truly independent transits, for which it is tempting to write

$$p(\boldsymbol{f}|P, T_0, D, \Delta) = \prod_k^K p(\boldsymbol{f}|T_k, D, \Delta_k) \tag{8}$$

where $\{ T_k \}_k$ are the epochs matching (T_0, P) and $\{ \Delta_k \}_k$ the corresponding depths. So that

$$\ln p(\boldsymbol{f}|P, T_0, D, \Delta) = \sum_k^K \ln \mathcal{L}_k$$

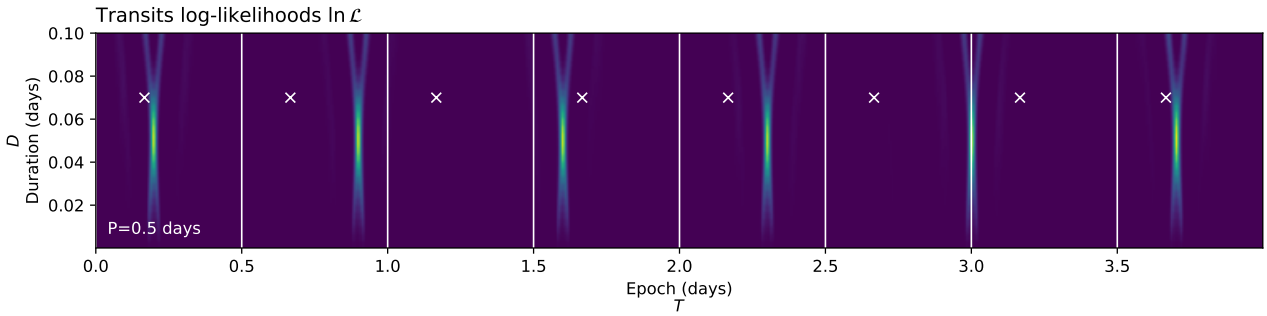
This is the joint log-likelihood of transits belonging to the same periodic signal but with varying depths $\{ \Delta_k \}_k$. However, individual transits from a periodic signal cannot be considered independent, and should instead share a common transit depth Δ. We show in Appendix B that there is an analytical expression for the joint log-likelihood of K individual transits with depths and errors $\{ \Delta_k, \sigma_k \}_k$ assuming a common depth Δ:

$$\begin{gathered} \ln p(\boldsymbol{f}|P, T_0, D, \Delta) = \sum_k^K \ln \mathcal{L}_k - \frac{K}{2} \ln(2\pi) - \frac{1}{2} \sum_k^K \left(\frac{\left(\Delta_k - \bar{\Delta}\right)^2}{\sigma_k^2 + \sigma^2} + \ln\left(\sigma_k^2 + \sigma^2\right) \right) \\ \text{with} \quad \frac{1}{\sigma^2} = \sum_k^K \frac{1}{\sigma_k^2} \quad \text{and} \quad \Delta = \sigma^2 \sum_k^K \frac{\Delta_k}{\sigma_k^2} \end{gathered} \tag{9}$$

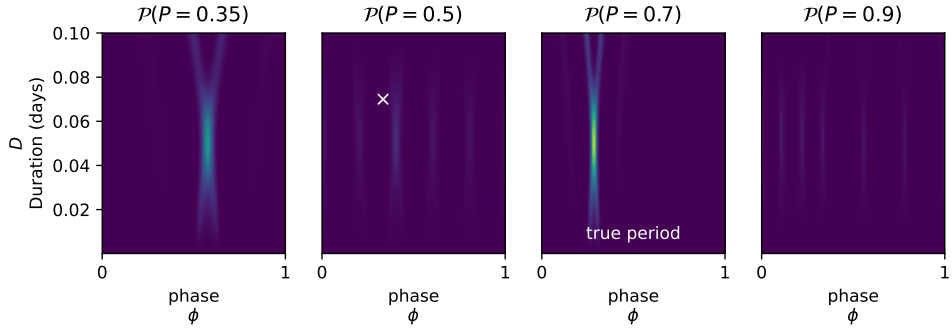
While this gives a solution for Equation 7[4], the individual epochs matching T_0 and P are not necessarily available in the grid of epochs $\{ T_k \}_k$. In Foreman-Mackey et al. (2015) this is solved by using the nearest neighbors in the epochs grid. Instead, to allow the efficient matrix computation of Equation 9, we interpolate the likelihood grid from $\{ T_i \}_i$ to a common grid of transit phases $\{ \phi_i \}_i$, leading to the *periodic search* log-likelihood

$$\ln \mathcal{P}(P) = \{ \ln p(\boldsymbol{f}|P, \phi, D) \}_{i,j}$$

shown for few periods in Figure 6.b, a notation where $\Delta_{i,j}$ is omitted since being interpolated from the *linear search* using ϕ_i and D_j, and $T_0 = 0$.



(a) $\ln \mathcal{L}$



(b) $\mathcal{P}(P)$

Figure 6. On the example dataset, this figure shows how the *periodic search* works at different periods P (including the true period $P = 0.7$ days). Given P and $T_0 = 0$, the log-likelihood $\ln \mathcal{L}$ shown in (a) is phase-folded and interpolated onto a common grid of phases shown in (b). As an example, the white lines in (a) mark the edges of each fold for a period of $P = 0.5$ days, and the white crosses the epochs $\{ T_k \}_{k \in \mathbb{K}}$ matching a particular phase in the grid (reported in (a) and (b)) on which the corresponding $\{ \ln \mathcal{L}_k \}_{k \in \mathbb{K}}$ are interpolated and combined using Equation 9. To allow the use of efficient matrix computations, this is done for every duration $\{ D_i \}_i$ so that $\mathcal{P}$ is computed on the full grid $\{ D_i, \phi_i \}_{i,j}$ at once. We understand from plots in (a) that a different choice of epoch T_0 may only shift the results in phase but do not affect the values of $\mathcal{P}$. For this reason, computing $\mathcal{P}$ for $T_0 = 0$ is sufficient. We also notice how the maximum value of $\mathcal{P}(P = 0.35)$ is lower than for $P = 0.7$ days, a result of combining the log-likelihoods using Equation 9 instead of Equation 8, in favor of individual transits matching a common depth Δ.

2.3. *The transit search periodogram*

Using Equation 9, we can now compute $\ln \mathcal{P}$ for a range of periods (see Figure 6) and build a transit search periodogram using Equation 4. But a final issue emerges, one that is fundamentally linked to our strategy. Each likelihood $p(\boldsymbol{f}|T, D, \Delta)$ estimated during the *linear search* is computed using N measurements. Hence, combining transits in the *periodic search*, through Δ_k, σ_k and the product of K likelihoods $\{ \mathcal{L}_k \}_k$ (Equation 9), artificially leads to a likelihood

[4] Different from the one found in Foreman-Mackey et al. (2015)

involving $N \times K$ measurements. This lead to a normalization issue when trying to compare the joint log-likelihoods $\mathcal{P}(P)$ from one period to another (like the ones in Figure 6 that have been normalized for visualization purpose), as the number of observed transits differs from one period to another. This motivates a final step to produce the transit search periodogram Q.

For any period P, instead of taking $\mathcal{Q}(P)$ as the maximum value of $\ln \mathcal{P}$, we retrieve the maximum likekihood parameters

$$(\phi_0, D) = \underset{\phi_i, D_j}{\arg\max} \{ \ln p(\boldsymbol{f} | P, \phi_i, D_j) \}_{i,j} \tag{10}$$

and define $\mathcal{Q}(P)$ as the SNR of the transit of period P, epoch $T_0 = \phi_0 P$ duration D and depth Δ, i.e.

$$\mathcal{Q}(P) = \frac{\Delta}{\sigma}$$

where Δ and σ are obtained using Equation 9 with the last column of X containing a periodic transit signal of period P, epoch T_0, duration D and depth 1. This process and the resulting periodogram $\mathcal{Q}$ are shown in Figure 7.

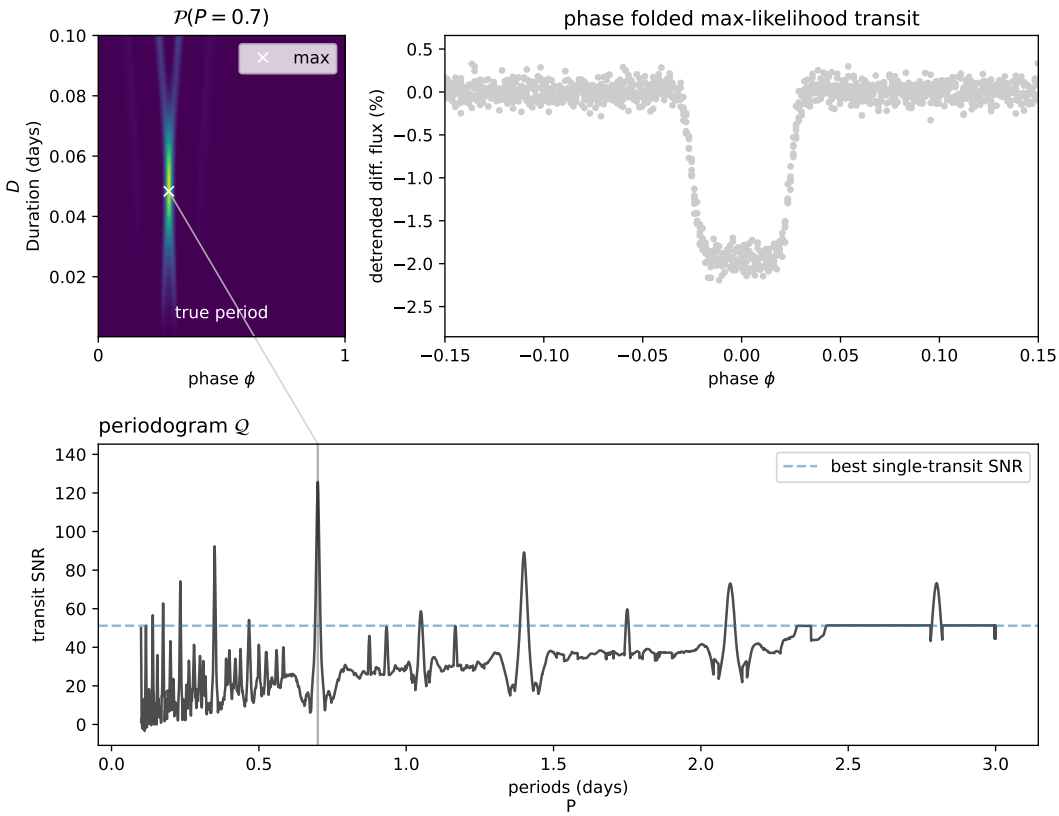


Figure 7.

The periodic transit of period P with the maximum SNR, i.e. maximizing $\mathcal{Q}$, is adopted as the best candidate, adapting our confidence through the minimum SNR for which we consider a signal to be real. The parameters of this transit are the period P, the epoch $T_0 = \phi_0 P$ and duration D (Equation 10), and the depth Δ with error σ (Equation 9).

2.3.1. *An open-source python package*

3. INJECTION-RECOVERY ON SIMULATED DATA

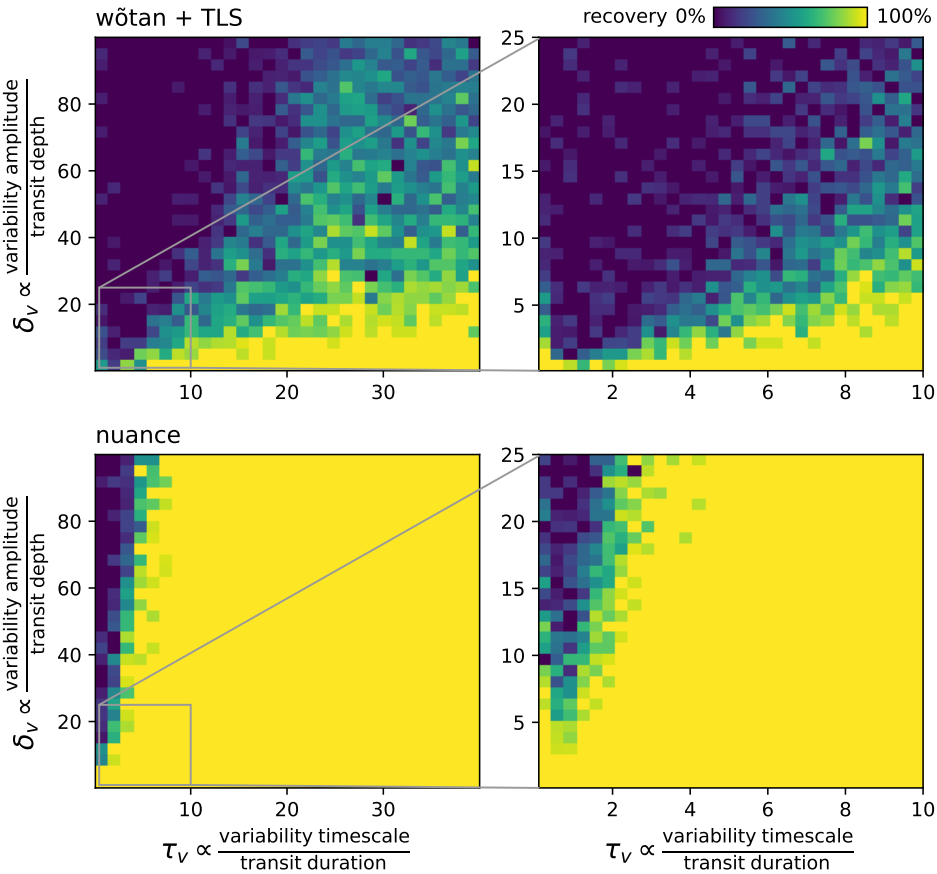


Figure 8.

4. COMPARISON WITH SHERLOCK ON TOI-540

5. PERFORMANCES AND LIMITATIONS

6. CONCLUSION

This simulated dataset constits in a light curve featuring a transit signal of depth 1%, duration 0.05 days and period 1.3 days using the simple Protopapas et al. 2005 analytical model. It also contains a simulated variability signal drawn from a Gaussian Process with a quasiseparable

APPENDIX

A. LIGHT CURVE SIMULATIONS

Let f be the simulated flux of a star sampled and aranged in the $(1 \times N)$ column-vector $\boldsymbol{f}$ associated to the column-vector of times $\boldsymbol{t}$. We construct $\boldsymbol{f}$ such that

$$\boldsymbol{f} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{C})$$

The mean $\boldsymbol{\mu}$ is built such that $\mu_i = \mathcal{T}(t_i) + \mathcal{S}(t_i)$ where $\mathcal{T}$ is a periodic transit signal and $\mathcal{S}$ an instrumental signal. The covariance matrix $\boldsymbol{C}$ is built such that $C_{i,j} = \nu(t_i, t_j)$ where ν is a covariance function accounting for correlated noise in the form of stellar variabilty with added white noise. An example of such signal is simulated and shown in Figure 9.

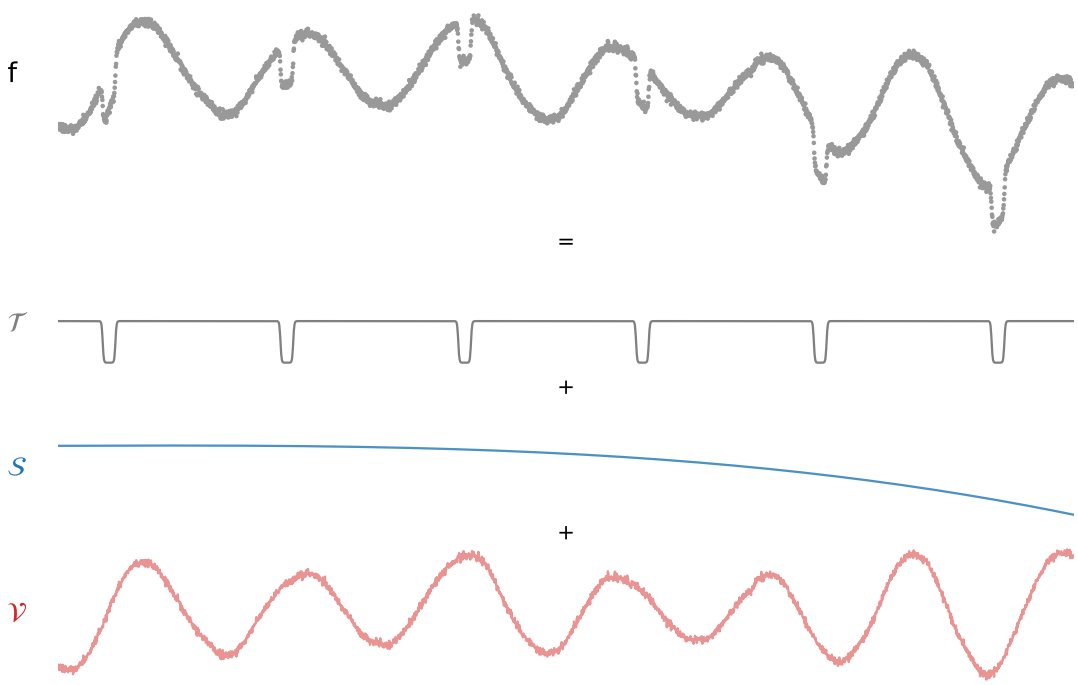


Figure 9. Example dataset sampled at $N = 2880$ times corresponding to an observation of 4 days with an exposure time of 2 minutes. The mean of this signal consists in a periodic transit signal plus instrumental signals. Stellar variability and white noise are simulated by modeling the covariance matrix of the signal with Gaussian Processes .

In what follows, we describe the models used to simulate the components of the signal f, i.e. the periodic transit $\mathcal{T}$, the instrumental systematics $\mathcal{S}$ and the covariance function ν describing stellar variability and white noise.

A.1. *Transit signal* $\mathcal{T}$

We simulate the periodic transit signal $\mathcal{T}$ using the simple model described in Protopapas et al. (2005) where a transit of period P, epoch T_0, duration D and unitary depth observed at time t is given by

$$\begin{aligned} \mathcal{T}_c(t, P, T_0, D) &= \frac{1}{2}\tanh\left(c\left(\theta - \frac{1}{2}\right)\right) - \frac{1}{2}\tanh\left(c\left(\theta + \frac{1}{2}\right)\right) \\ \text{with} \quad \theta &= \frac{P}{\pi D}\sin\left(\frac{\pi(t - T_0)}{P}\right) \end{aligned} \tag{A1}$$

where the dimensionless parameter c controls the roundness of the transit depth ($c \gg 1$ corresponding to a box-shaped transit as shown in Figure 10). This analytical model is fully empirical but easily differentiable. From this expression, we get the signal of the single non-periodic transit of unitary depth

$$\tau_c(t, T_0, D) = \lim_{P\to\infty} \mathcal{T}(t, P, T_0, D) = -\frac{1}{2}\tan\left(\frac{c}{2}(1 - T_0 + t)\right) - \frac{1}{2}\tan\left(\frac{c}{2}(1 + T_0 - t)\right) \tag{A2}$$

with epoch T_0 and duration D.

Figure 10. Simulations of the single non-periodic transit signal $\tau_c(t, T_0, D)$ (Equation A2) for different values of c

In our study, we simulate all transit signals with $c = 12$. The periodic transit signal $\mathcal{T}$ seen in Figure 9 corresponds to $\mathcal{T} = \mathcal{T}_{c=12}(\boldsymbol{t}, P = 0.7, T_0 = 0.2, D = 0.05)$, all parameters in unit of *days*.

A.2. *Instrumental signals* $\mathcal{S}$

We simulate instrumental signals as a linear model of M explanatory variables arranged in the $(M \times N)$ design matrix $\boldsymbol{X}$. Hence,

$$\mathcal{S} = \boldsymbol{w}\boldsymbol{X}$$

where the vector $\boldsymbol{w}$ are linear coefficients of the model. The simulated flux shown in Figure 9 contains a linear model where the $M = 4$ columns of the design matrix $\boldsymbol{X}$ are given by $\boldsymbol{X}_i = \boldsymbol{t}^i$ (i.e. $\boldsymbol{X}$ is the Vandermonde matrix order 3 of time t) and $\boldsymbol{w} = [1.0 \quad 0.0005 \quad -0.0002 \quad -0.0005]$.

A.3. *Stellar variability* ν

We simulate correlated noise thanks to Gaussian Processes. In our study, we are particularly interested in stellar variability and its effect on transit detection. For this reason, we use a physically-motivated kernel, describing the covariance of a stochastically-driven damped harmonic oscillator (SHO, Foreman-Mackey et al. (2017b); Foreman-Mackey (2018)) taking the form

$$k(\tau) = \sigma^2 \exp\left(-\frac{\omega\,\tau}{2\,Q}\right) \begin{cases} 1 + \omega\,\tau & \text{for } Q = 1/2 \\ \cosh(f\,\omega\,\tau/2\,Q) + \sinh(f\,\omega\,\tau/2\,Q)/f & \text{for } Q < 1/2 \\ \cos(g\,\omega\,\tau/2\,Q) + \sin(g\,\omega\,\tau/2\,Q)/g & \text{for } Q > 1/2 \end{cases} \tag{A3}$$
$$\text{where} \quad \tau = |t_i - t_j|, \quad f = \sqrt{1 - 4\,Q^2} \quad \text{and} \quad g = \sqrt{4\,Q^2 - 1}$$

We use this kernel through its implementation in the `tinygp`[5] Python package, providing an implententation of the quasi-separable kernel from Foreman-Mackey 2018 powered by `JAX`[6]. The stellar variability signal in Figure 9 has been sampled from a Gaussian Process with an SHO kernel of parameters $\omega = \pi/6D$ (i.e. a period equal to 12 times the duration D of the simulated transit), $Q = 45$ and $\sigma = \Delta$, the depth of the simulated transit. An extra term $\sigma_f^2 = 0.001^2$ is added to the covariance matrix, corresponding to the variance of the simulated measurement f and leading to the white noise observed in Figure 9.

[5] https://github.com/dfm/tinygp
[6] https://github.com/google/jax

B. COMBINING TRANSITS

From the *linear search*, we retain and index by k the parameters of the K individual transits whose epochs $\{T_k\}_k$ are compatible with a periodic signal of period P and epoch T_0. From the likelihoods of these transits (computed in subsection 2.1), we want an expression for

$$p(\boldsymbol{f}|P, T_0, D, \Delta) = \prod_{k\in\mathbb{T}} p(\boldsymbol{f}|T_k, D, \Delta) \tag{B4}$$

i.e., given a depth D, the likelihood of our data given a periodic transit signal of period P, epoch T_0 and a common depth Δ. Since only $\{p(\boldsymbol{f}|T_k, D, \Delta_k)\}_k$ is known (i.e. transits with different depths), we decompose

$$p(\boldsymbol{f}|T_k, D, \Delta) = \int p(\boldsymbol{f}|T_k, D, \tilde{\Delta}) p(\tilde{\Delta}|\Delta)\, d\tilde{\Delta} \tag{B5}$$

where $p(\boldsymbol{f}|T_k, D, \tilde{\Delta})$ is the probability of the k-th transit to have a depth $\tilde{\Delta}$ and $p(\tilde{\Delta}|\Delta)$ the probability to observe the depth $\tilde{\Delta}$ knowing the existence of a common depth Δ. In other words, Equation B5 involves the likelihood of the non-periodic transit k to be part of a periodic transit signal with a common depth Δ.

Since each depth Δ_k is found through generalized least square, each follow a normal distribution $\mathcal{N}(\Delta_k, \sigma_k^2)$, centered on Δ_k with variance σ_k^2, and with an amplitude $\mathcal{L}_k$, leading to the likelihood:

$$p(\boldsymbol{f}|T_k, D, \tilde{\Delta}) = \mathcal{L}_k G_k(\tilde{\Delta})$$

where G_k is the probability density function of $\mathcal{N}(\Delta_k, \sigma_k^2)$.

As for the common transit depth Δ, it can be estimated through the joint probability of all other transit depths than Δ_k, such that

$$\Delta \sim \prod_{i\neq k}^{K} \mathcal{N}(\Delta_i, \sigma_i^2) \quad \text{i.e.} \quad \Delta \sim \mathcal{N}(\Delta, \sigma^2)$$

a normal distribution defined by

$$\frac{1}{\sigma^2} = \sum_{i\neq k}^{K} \frac{1}{\sigma_i^2} \quad \text{and} \quad \Delta = \sigma^2 \sum_{i\neq k}^{K} \frac{\Delta_i}{\sigma_i^2} \tag{B6}$$

$$p(\tilde{\Delta}|\Delta) = \left(\prod_{i\neq k}^{K} \mathcal{L}_i\right) G(\tilde{\Delta})$$

where G is the probability density function of $\mathcal{N}(\Delta, \sigma^2)$. We can now rewrite Equation B5 as

$$p(\boldsymbol{f}|T_k, D, \Delta) = \mathcal{L}_k \left(\prod_{i\neq k}^{K} \mathcal{L}_i\right) \int G_k(\tilde{\Delta})\, G(\tilde{\Delta})\, d\tilde{\Delta} \tag{B7}$$

The product of two normal distributions is also a normal distribution, so that the integral in Equation B7 can be obtained analytically. We have

$$p(\boldsymbol{f}|T_k, D, \Delta) = \mathcal{L}_k \left(\prod_{i\neq k}^{K} \mathcal{L}_i\right) \frac{1}{\sqrt{2\pi\left(\sigma^2 + \sigma_k^2\right)}} e^{-\frac{1}{2}\frac{(\Delta-\Delta_k)^2}{\sigma^2+\sigma_k^2}}$$

And Equation 7 gives

$$\ln p(\boldsymbol{f}|P, T_0, D, \Delta) = \sum_{k}^{K} \ln \mathcal{L}_k - \frac{K}{2}\ln(2\pi) - \frac{1}{2}\sum_{k}^{K}\left(\frac{(\Delta_k - \Delta)^2}{\sigma_k^2 + \sigma^2} + \ln\left(\sigma_k^2 + \sigma^2\right)\right) \tag{B8}$$

the log-likelihood of our data given a periodic transit signal of period P, epoch T_0, duration D and common depth Δ. In order to reduce the number of times Equation B6 is computed, we adopt the biased estimates

$$\frac{1}{\sigma^2} = \sum_k^K \frac{1}{\sigma_i^2} \quad \text{and} \quad \Delta = \sigma^2 \sum_k^K \frac{\Delta_i}{\sigma_i^2}$$

so that Δ and σ are independent of k in the last sum of Equation B8.

REFERENCES

Aigrain, S., Parviainen, H., & Pope, B. J. S. 2016, MNRAS, 459, 2408, doi: 10.1093/mnras/stw706

Deming, D., Knutson, H., Kammer, J., et al. 2015, ApJ, 805, 132, doi: 10.1088/0004-637X/805/2/132

Foreman-Mackey, D. 2018, Research Notes of the American Astronomical Society, 2, 31, doi: 10.3847/2515-5172/aaaf6c

Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017a, AJ, 154, 220, doi: 10.3847/1538-3881/aa9332

—. 2017b, AJ, 154, 220, doi: 10.3847/1538-3881/aa9332

Foreman-Mackey, D., Montet, B. T., Hogg, D. W., et al. 2015, ApJ, 806, 215, doi: 10.1088/0004-637X/806/2/215

Hippke, M., David, T. J., Mulders, G. D., & Heller, R. 2019, AJ, 158, 143, doi: 10.3847/1538-3881/ab3984

Kovács, G., Bakos, G., & Noyes, R. W. 2005, MNRAS, 356, 557, doi: 10.1111/j.1365-2966.2004.08479.x

Kovács, G., Hartman, J. D., & Bakos, G. Á. 2016, A&A, 585, A57, doi: 10.1051/0004-6361/201527124

Luger, R., Agol, E., Kruse, E., et al. 2016, AJ, 152, 100, doi: 10.3847/0004-6256/152/4/100

Luger, R., Kruse, E., Foreman-Mackey, D., Agol, E., & Saunders, N. 2018, AJ, 156, 99, doi: 10.3847/1538-3881/aad230

Mosteller, F., & Tukey, J. W. 1977, Data analysis and regression. A second course in statistics

Pont, F., Zucker, S., & Queloz, D. 2006, MNRAS, 373, 231, doi: 10.1111/j.1365-2966.2006.11012.x

Protopapas, P., Jimenez, R., & Alcock, C. 2005, MNRAS, 362, 460, doi: 10.1111/j.1365-2966.2005.09305.x

Tamuz, O., Mazeh, T., & Zucker, S. 2005, MNRAS, 356, 1466, doi: 10.1111/j.1365-2966.2004.08585.x