The Astronomical Journal, 167:284 (20pp), 2024 June https://doi.org/10.3847/1538-3881/ad3cd6

nuance: Efficient Detection of Planets Transiting Active Stars

Lionel J. Garcia[1,2], Daniel Foreman-Mackey[1], Catriona A. Murray[3], Suzanne Aigrain[4], Dax L. Feliz[5], and Francisco J. Pozuelos[2,6]
[1] Center for Computational Astrophysics, Flatiron Institute, New York, NY, USA; lgarcia@flatironinstitute.org
[2] Astrobiology Research Unit, University of Liège, Liège, Belgium
[3] Department of Astrophysical and Planetary Sciences, University of Colorado, Boulder, CO, USA
[4] Subdepartment of Astrophysics, Department of Physics, University of Oxford, Oxford, UK
[5] American Museum of Natural History, New York, NY, USA
[6] Instituto de Astrofísica de Andalucía, Granada, Spain
Received 2024 February 9; revised 2024 March 29; accepted 2024 April 6; published 2024 May 24

Abstract

The detection of planetary transits in the light curves of active stars, featuring correlated noise in the form of stellar variability, remains a challenge. Depending on the noise characteristics, we show that the traditional technique that consists of detrending a light curve before searching for transits alters their signal-to-noise ratio and hinders our capability to discover exoplanets transiting rapidly rotating active stars. We present *nuance*, an algorithm to search for transits in light curves while simultaneously accounting for the presence of correlated noise, such as stellar variability and instrumental signals. We assess the performance of *nuance* on simulated light curves as well as on the Transiting Exoplanet Survey Satellite light curves of 438 rapidly rotating M dwarfs. For each data set, we compare our method to five commonly used detrending techniques followed by a search with the Box-Least-Squares algorithm. Overall, we demonstrate that *nuance* is the most performant method in 93% of cases, leading to both the highest number of true positives and the lowest number of false-positive detections. Although simultaneously searching for transits while modeling correlated noise is expected to be computationally expensive, we make our algorithm tractable and available as the JAX-powered Python package nuance, allowing its use on distributed environments and GPU devices. Finally, we explore the prospects offered by the *nuance* formalism and its use to advance our knowledge of planetary systems around active stars, both using space-based surveys and sparse ground-based observations.

Unified Astronomy Thesaurus concepts: Exoplanet detection methods (489); Stellar activity (1580); Time series analysis (1916); Gaussian Processes regression (1930); Computational methods (1965); GPU computing (1969)

1. Introduction

Transiting exoplanets are keystone objects for the field of exoplanetary science, but detecting transits in light curves featuring stellar variability and instrumental signals remains a challenge (e.g., Pont et al. 2006; Howell et al. 2016 or Clarice Yaptangco et al. 2024). For this reason, known transiting exoplanets tend to be found around quieter stars, or belong to the population of close-in giants whose transit signals dominate over stellar rotational variability (Simpson et al. 2023). However, transiting exoplanets around active stars are discoveries with significant scientific value. First, as younger stars are more active (Skumanich 1972), being able to detect planets transiting active stars will favor the discovery of young planetary systems (e.g., Newton et al. 2022). Second, as stellar variability may originate from surface active regions (such as starspots), transiting exoplanets can be used to map the photosphere of active stars (e.g., Morris et al. 2017), benefiting both the study of stellar atmospheres and the concerning impact of their nonuniformity on planetary atmosphere retrievals (Rackham et al. 2018). Overall, enabling the detection of transits in light curves with high levels of correlated noises will greatly benefit the study of terrestrial exoplanets around late M dwarfs, usually observed at lower signal-to-noise-ratio (S/N) and more likely to display photometric variability (e.g., Murray et al. 2020 or Petrucci et al. 2024).

Commonly used transit-search algorithms, such as the Box-Least-Squares (BLS) algorithm (Kovács et al. 2002), are capable of detecting transits in light curves containing only transit signals and white noise. Using this method, the simplest way to detect transits in a light curve featuring correlated noise (either astrophysical or instrumental) is to first clean it from nuisance signals before performing the search. This strategy is widely adopted by the community, both using physically motivated systematic models like Luger et al. (2016, 2018) or filtering techniques (Jenkins et al. 2010; Hippke et al. 2019). However, when correlated noise starts resembling transits, this cleaning step, often called *detrending*, is believed to degrade their detectability (see Section 4.3 of Hippke et al. 2019). In this case, the only alternative to search for transits is to perform full-fledged modeling of the light curve, including both transits and correlated noise, and to compute the likelihood of the data to the transit model on a wide parameter space, an approach largely avoided due to its intractable nature. Nonetheless, Kovács et al. (2016) ask: *Periodic transit and variability search with simultaneous systematics filtering: Is it worth it?* The latter study explores a handful of cases and generally discards the benefit of using a full-fledged approach.

In this paper, we identify regions of light curves morphological parameter space for which a full-fledged transit search is necessary, and we present *nuance*,[7] a method to search for transit signals while simultaneously modeling correlated noises

[7] Throughout the paper, *nuance* written in italics refers to the algorithm, while nuance in sans-serif refers to its implementation published as a Python package in its version 0.5.2 (see Section 3.5).

in a tractable way. In Section 2, we describe the effect of correlated noise on transit light curves and the effect of its detrending on transit signal detectability. In Section 3, we present *nuance*, and the two main steps on which this method is based: the linear search and the periodic search. In Section 4, we test the performance of *nuance* on a wide variety of cases and compare our method to commonly used transit-search algorithms. This includes transits injected in synthetic data sets, but also in the Transiting Exoplanet Survey Satellite (TESS) light curves of 438 rapidly rotating M dwarfs. Finally, in Section 5, we discuss the results and the limitations of *nuance*, before concluding in Section 6.

2. Motivation

A strong assumption when using the BLS algorithm to search for transits is that the searched data set only contains transit signals and white noise, justifying the need for detrending. In this section, we explore this effect by simulating light curves containing a transit signal and correlated noise in the form of stellar variability and study how detrending the latter affects the transit signal detectability depending on the light-curve morphological characteristics. Hence, we first present how transit light curves are simulated, using a stochastic model of stellar variability, and describe the detection metric we employ to quantify transit detectability in the presence of correlated noise.

2.1. Light-curve Simulations

To perform this study, we want to simulate realistic light curves that contain instrumental signals, transit signals, and correlated noise with characteristics that can be controlled using a set of interpretable parameters. To this end, we model light curves as realizations of a Gaussian process (GP; Rasmussen & Williams 2005; Aigrain & Foreman-Mackey 2023) with a mean containing the instrumental and transit signals, and a kernel allowing us to model different forms of correlated noise controlled by its hyperparameters.

Let f be the simulated differential flux of a star sampled and arranged in the vector $\boldsymbol{f}$ associated with the vector of times $\boldsymbol{t}$, such that

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

i.e., that $\boldsymbol{f}$ is drawn from a GP of mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. For practical reasons, we model $\boldsymbol{\mu}$ as a linear combination of M explanatory variables, such that

$$\boldsymbol{\mu} = \boldsymbol{X}\boldsymbol{w}, \tag{1}$$

where the first $M-1$ columns of $\boldsymbol{X}$ are contemporaneous instrumental time series measurements, such as the position of the star on the detector, the sky background, co-trending basis vectors, or any other explanatory variables. The last column of $\boldsymbol{X}$ is a box-shaped transit signal with a fixed epoch, duration, and period. This way, the transit signal is part of the mean linear model, meaning that once the design matrix $\boldsymbol{X}$ is constructed the transit is only parametrized by its depth Δ.

As we are interested in active stars whose fluxes feature correlated noise in the form of stellar variability, we choose a physically motivated GP kernel to model the covariance matrix $\boldsymbol{\Sigma}$, describing stellar variability through the covariance of a stochastically driven damped harmonic oscillator (SHO; Foreman-Mackey et al. 2017; Foreman-Mackey 2018) parametrized by its quality factor Q, its pulsation ω and the amplitude of the kernel function σ (the full expression of the kernel function is provided in Appendix A). This choice of mean and kernel function completely defines the GP, from which light curves with different levels of correlated noise can be drawn (see an example in Figure 1).

2.2. Transit Detectability

One way to quantify the detectability of a unique transit signal is to compute its S/N expressed as

$$\mathrm{S/N} = \frac{\Delta}{\sigma}\sqrt{n},$$

where Δ is the transit depth, σ is the measurement uncertainty (assuming homoscedasticity), and n is the number of points within transit. Although this metric is useful to assess the strength of the transit signal given a certain photometric precision, it does not account for the presence of correlated noise. However, instrumental and other astrophysical signals will necessarily affect the detectability of transits in realistic light curves (Pont et al. 2006). For this reason, the combined differential photometric precision (CDPP; Jenkins et al. 2010) metric was developed and used in the context of the Kepler mission to assess the level of correlated noise in light curves, affecting the detectability of transits with a given duration. The CDPP is computed by decomposing the data in the time-frequency domain using wavelets and measuring the significance of a distorted transit signal in the whitened data. For a given transit duration, the CDPP is then a measure of the noise remaining after filtering the light curve in each frequency band, taking the presence of nonstationary correlated noise of a given timescale into account (see Jenkins et al. 2010 for more details). Hence, we can estimate the significance of the transit signal accounting for the presence of correlated noise as

$$\mathrm{S/N} = \frac{\hat{\Delta}}{CDPP_D}, \tag{2}$$

where $CDPP_D$ is the CDPP computed for a given transit duration D and $\hat{\Delta}$ is the transit depth after detrending. For simplicity, the CDPP is computed using the method from Gilliland et al. (2011) implemented in the Lightkurve Python package (Lightkurve Collaboration et al. 2018) and $\hat{\Delta}$ is the depth obtained by solving Equation (1) with the last column of $\boldsymbol{X}$ containing a normalized transit model.

Figure 2 shows the S/N from Equation (2) computed for a unique transit observed in the absence (gray) and presence of correlated noise (red). As illustrated in this figure, the presence of correlated noise strongly decreases the transit signal S/N, which would ultimately limit its detectability.

2.3. Detrending Methods and Their Effects

The presence of instrumental correlated noise 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

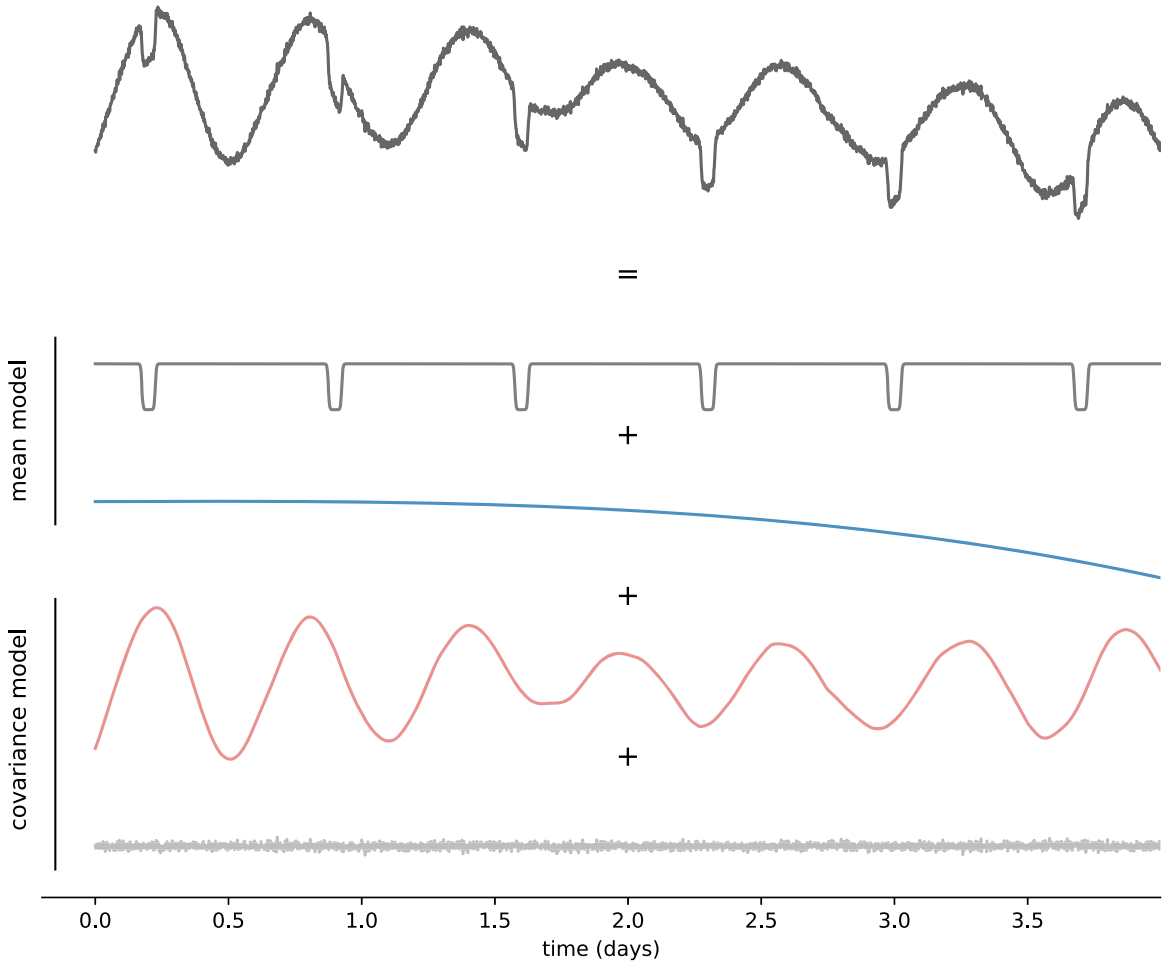

Figure 1. The top plot shows the simulated flux time series of an active star drawn from the GP model described in Section 2.1. This data set corresponds to an observation of 4 days with an exposure time of 2 minutes. The mean of this signal consists of a periodic transit signal of period $P = 0.7$ days, duration $D = 1.2$ hr, and depth $\Delta = 2\%$ (dark gray) plus an instrumental signal (blue). Correlated noise in the form of stellar variability is simulated with an SHO kernel of hyperparameters $\omega = \pi/6D$, $Q = 45$, and $\sigma = \Delta$. Finally, white noise with a standard deviation of 0.1% is added to the diagonal of the covariance matrix.

(or neighboring pixels) such that the correction applied should not degrade the transit signal and can be modeled using contemporaneous measurements (e.g., detector's temperature, pointing error, sky background, or airmass time series). But even after instrumental signals have been removed, stellar variability and other astrophysical signals remain, which gave rise to several approaches. Some of them are physically motivated and make use of GPs (e.g., Aigrain et al. 2016), others are empirical and make use of filtering and data-driven algorithms (Jenkins et al. 2010; Hippke et al. 2019). In this section, we show how these techniques impact transits detectability, depending on the morphological characteristics of light curves.

In Figure 3, we simulate a transit signal on top of which we add photometric stellar variability with different timescales, sampled from a GP with an SHO kernel described in Appendix A. For each light curve, we reconstruct and detrend stellar variability in two ways: one using the widely adopted Tukey's bi-weight filter, presented in Mosteller & Tukey (1977) and using the implementation from wõtan[8] (Hippke et al. 2019); the other using the same GP from which the data have been sampled. We then estimate the resulting transit depth and compute the remaining transit S/N using Equation (2). Figure 3 clearly shows the effect of both detrending techniques on transits S/N, and intuitively suggests that this degradation due to detrending is strongly dependent on the correlated noise characteristics encountered.

To explore the parameter space for which detrending is the most problematic, we employ the model described in Section 2.1 to simulate 10,000 differential light curves with different morphological characteristics and compute the remaining transit S/N after detrending. In order to place the stellar variability hyperparameters on a relative scale with the transit signal parameters, we reparametrize the SHO kernel with

$$\tau = \frac{\pi}{\omega D}, \qquad \delta = \frac{2\sigma}{\Delta} \quad \text{and} \quad Q = 10, \tag{3}$$

where τ is the relative timescale of the variability with respect to the transit duration and δ the relative amplitude of the variability against the transit depth, both being adimensional. Hence, for $(\tau, \delta) = (1, 1)$, the expressions of ω and σ given in Equation (3) correspond to a variability signal with a period half that of the transit duration, and a correlated noise

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

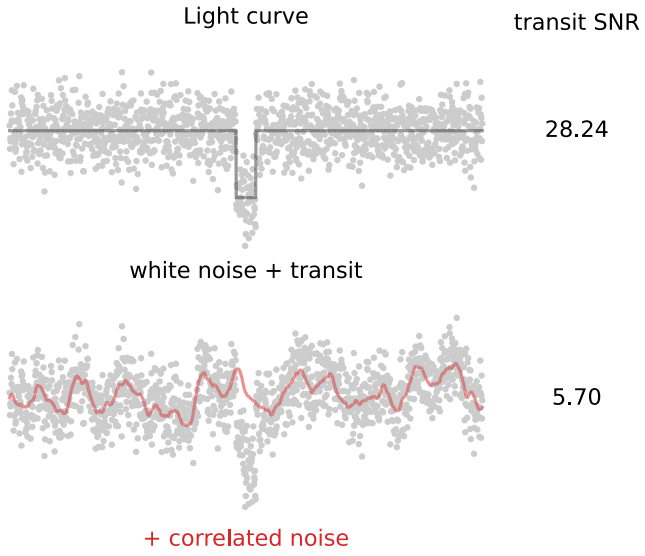

Figure 2. Illustration of the effect of correlated noise on a single transit S/N. An hour-long transit of depth 0.5% is generated on top of white noise (standard deviation of 0.15%) as part of a 24 hr observation with an exposure time of 1 minute (top). Then, in the bottom plot, correlated noise is added, generated using a GP with an SHO kernel of hyperparameters $\omega = 60$, $Q = 0.5$, and $\sigma = \Delta/4$. The S/N on the right of each light curve is computed using Equation (2). Models used to simulate these data are provided in Section 2.1.

amplitude comparable to the transit depth, i.e., strongly resembling the simulated transit signal.

For each of the 10,000 light curves generated, we separately reconstruct and detrend the variability signal using the two techniques employed in Figure 3, i.e., one using an optimal bi-weight filter with a window size 3 times that of the transit duration (Hippke et al. 2019) and the other using a GP with an optimal kernel. We then compute the detrended transit S/N using Equation (2).

Figure 4 shows that there exists an entire region in the (τ, δ) parameter space for which the bi-weight detrending degrades transit S/N to the point of no detection (S/N < 5). Although more robust, the same effect is observed when detrending with an optimal GP. Hence, detrending makes transit search blind to many systems, especially when the widely adopted bi-weight filter is used. We note that, outside this problematic parameter space, both techniques perform relatively well and such a degradation of the transit S/N should not be expected. Although this study should be extended to other detrending techniques, it highlights the need for a more informed transit-search algorithm able to deal with correlated noise, at least if present in the form of stellar variability.

3. nuance

When searching for exoplanets using an indirect method, such as using transits or radial velocities, the initial detection is often as important as the follow-up observations, ultimately leading to the confirmation of the planetary candidate. For that reason, the final product of most transit-search algorithms consists of a *periodogram*: a 1D detection metric given a set of trial periods that allows us to identify the presence of a potential candidate and the period and epoch at which it should be followed up.

If we assume that a transit is defined by its period P, epoch T_0, duration D, and depth Δ, we then wish to compute the likelihood of such a transit being present in the data for a set of periods P, leading to an interpretable periodogram. As we are interested in following up transits with specific parameters, such as a well-defined epoch, the likelihood we wish to compute must not be marginalized over all parameters other than P but rather specifically computed at the maximum-likelihood parameters $\hat{T}_0$, $\hat{D}$, and $\hat{\Delta}$, leading to the *profile likelihood*

$$\mathcal{Q}(P) = p(\boldsymbol{f}|P, \hat{T}_0, \hat{D}, \hat{\Delta}), \tag{4}$$

where $\boldsymbol{f}$ denotes the data. In this section, we present *nuance*, an algorithm to compute transit-search periodograms in a tractable way and detect planetary transits in light curves containing correlated noise such as instrumental signals and photometric stellar variability. In Section 3.1, we explain how our approach requires two separate steps in order to remain tractable and present these steps in Sections 3.2 and 3.3, leading to the transit-search periodogram described in Section 3.4. Finally, we present the nuance Python package in Section 3.5 and perform a control test of our implementation against BLS in Section 3.6.

3.1. The Approach

As in Section 2.1, let us assume that $\boldsymbol{f}$ is a vector of size N containing the flux of a star observed at times $\boldsymbol{t}$ such that

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

i.e., that $\boldsymbol{f}$ is drawn from a GP of mean $\boldsymbol{X}\boldsymbol{w}$ and covariance $\boldsymbol{\Sigma}$. Again, we set the first $M-1$ columns of the $(N \times M)$ design matrix $\boldsymbol{X}$ as contemporaneous measurements and the last column as a normalized box-shaped transit of period P, epoch T_0, and duration D. 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 the presence of a periodic transit signal of period P, epoch T_0, duration D, and a mean linear model with coefficients $\boldsymbol{w}$ is (Rasmussen & Williams 2005)

$$\ln p(\boldsymbol{f}|P, T_0, D, \boldsymbol{w}) = -\frac{1}{2}(\boldsymbol{f} - \boldsymbol{X}\boldsymbol{w})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{f} - \boldsymbol{X}\boldsymbol{w}) - \frac{1}{2}\ln|\boldsymbol{\Sigma}| - \frac{N}{2}\ln 2\pi. \tag{5}$$

Given the linearity of the mean model for P, T_0, and D fixed, this likelihood is maximized for the least-squares parameters

$$\boldsymbol{w} = (\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{f} \quad \text{with uncertainties } \boldsymbol{\sigma} = (\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}. \tag{6}$$

As we are only interested in the transit depth, we will omit $\boldsymbol{w}$ in the remainder of this paper and simply write likelihood expressions depending on its last value $\Delta = \boldsymbol{w}_M$, such as $p(\boldsymbol{f}|P, T_0, D, \Delta)$, where all other parameters of $\boldsymbol{w}$ are taken at their maximum-likelihood values.

Hence, given a period P, computing $\mathcal{Q}(P)$ boils down to a nonlinear optimization involving several evaluations of the likelihood given in Equation (5). While this is tractable for a few periods, it becomes highly untractable for the large set of trial periods required to properly sample a transit-search periodogram (on the order of tens of thousands).

To remain tractable, *nuance* employs an extension of the strategy used by Foreman-Mackey et al. (2015), which has

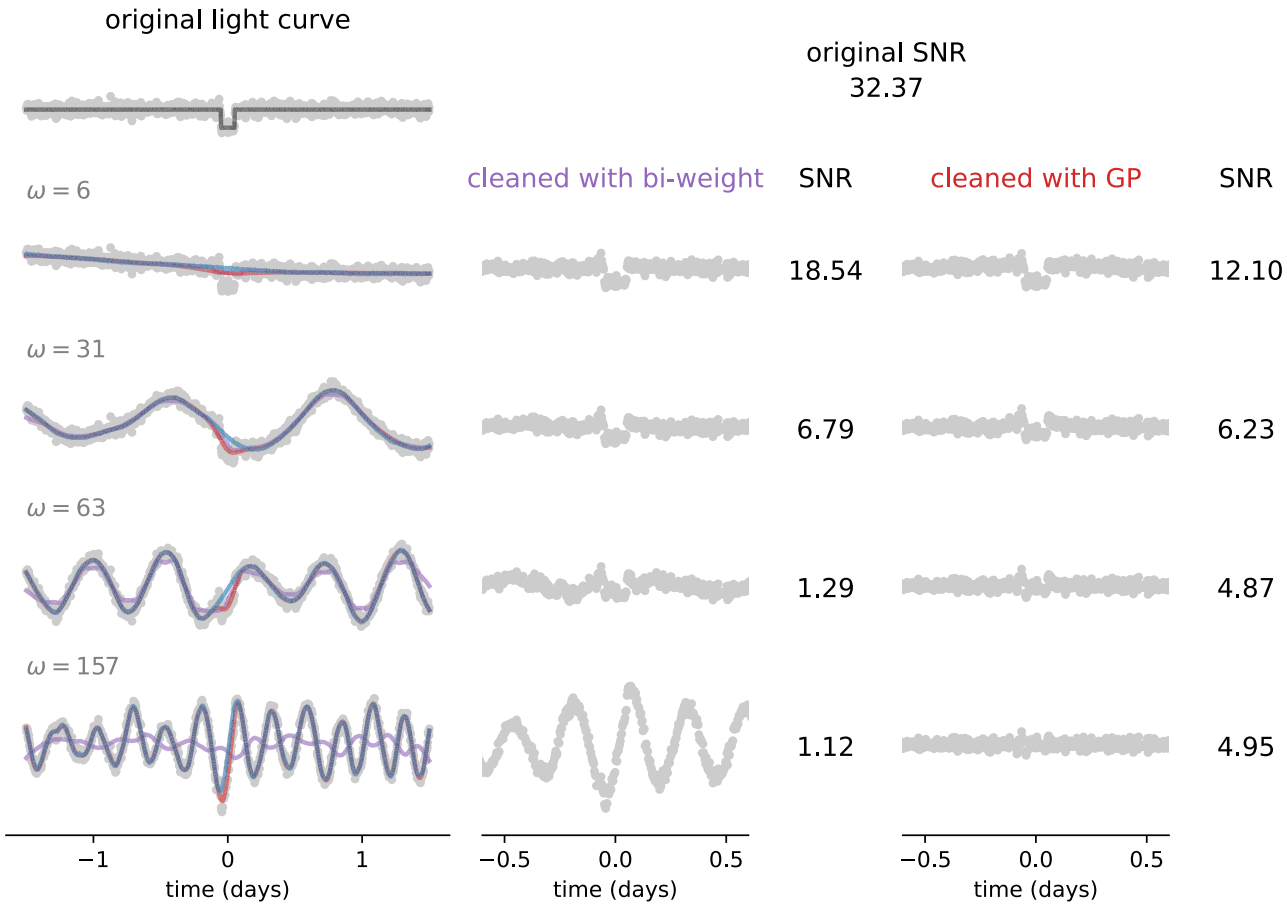

Figure 3. In each plot on the left, correlated noises simulated using a GP with an SHO kernel are added (blue line) to an original light curve containing a transit of depth $\Delta = 0.8\%$, duration $D = 0.1$ day, and white noise with a standard deviation of 0.15% (top left). The hyperparameters of the kernel are fixed to $Q = 10$ and $\sigma = 0.8\%$, with the pulsation ω increasing from top to bottom. The resulting stellar variability signals are then reconstructed using (in purple) Tukey's bi-weight filter with an optimal window size of $3 \times D$ (see Hippke et al. 2019), and (in red) a GP with the same kernel used to simulate the data. In each case the variability is reconstructed, subtracted, and the transit S/N computed using Equation (2).

significant intellectual overlap with the methods used by Aigrain & Irwin (2004) and Jenkins et al. (2010). This approach separates the transit search into two components: the linear search and the periodic search. During the linear search, the likelihood $p(\boldsymbol{f}|T_0, D, \Delta)$ of a single nonperiodic transit is computed for a grid of epochs and durations, each time solving for Δ linearly. Then, the periodic search consists of combining these likelihoods to compute the likelihood $p(\boldsymbol{f}|P, T_0, D, \Delta)$ of the data given a periodic transit signal for a range of periods P. These combined likelihoods yield $p(\boldsymbol{f}|P, \hat{T}_0, \hat{D}, \hat{\Delta})$, a transit-search periodogram on which the periodic transit detection is based. *nuance* differs from Foreman-Mackey et al. (2015) and other existing transit-search algorithms as it models the covariance of the light curve with a GP, accounting for correlated noise (especially in the form of stellar variability) while keeping the model linear and tractable. This way, *nuance* searches for transits while, at the same time, modeling correlated noise, avoiding the separate detrending step that degrades transit signals S/N (see Section 2).

3.2. The Linear Search

The goal of the linear search is to compute the likelihood $p(\boldsymbol{f}|T_0, D, \Delta)$ for a grid of epochs $\{ T_i \}_i$ and durations $\{ D_j \}_j$. For each pair (T_i, D_j) the transit depth $\Delta_{i,j}$ is linearly solved, which leads to the set of maximum likelihoods

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

An example of such a grid of likelihoods is shown in Figure 5. To prepare for the next step, uncertainties on the depths $\sigma_{i,j}$ are also computed using Equation (6) and stored.

3.3. The Periodic Search

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

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

i.e., the probability of a periodic transit of period P, epoch T_0, duration D, and depth Δ given the data $\boldsymbol{f}$. For a given transit duration D, any combination of (P, T_0) leads to K transits, for which it is tempting to write

$$p(\boldsymbol{f}|P, T_0, D, \Delta) = \prod_k^K p(\boldsymbol{f}|T_k, D, \Delta_k), \tag{7}$$

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

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

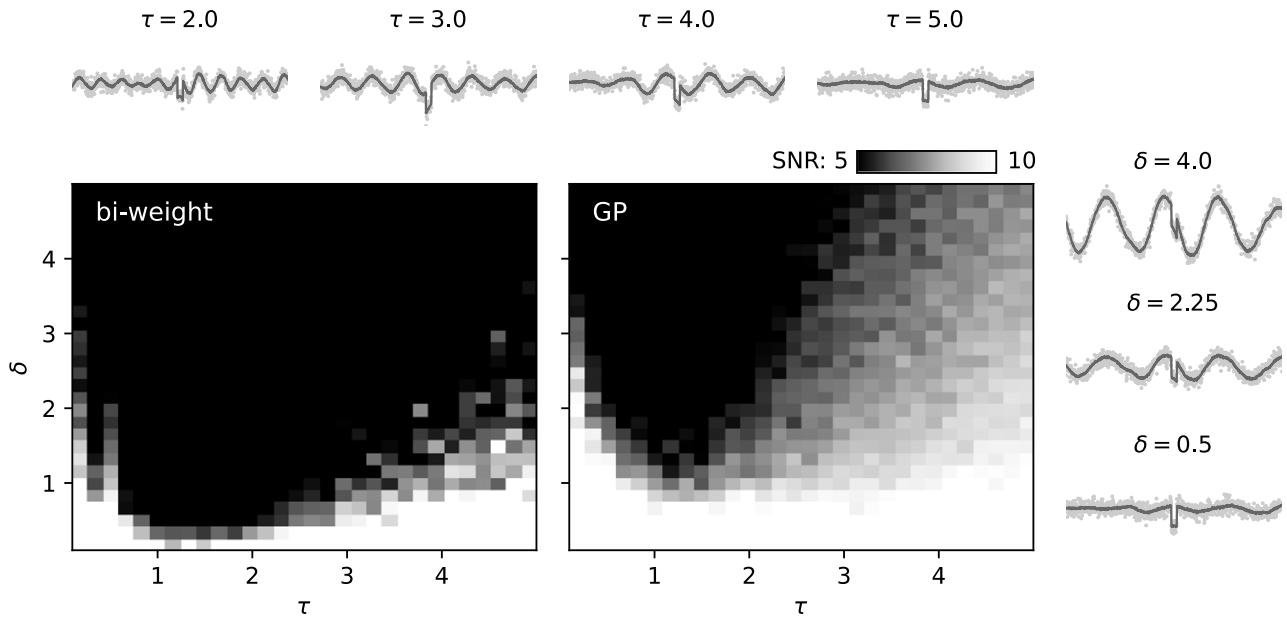

Figure 4. Transit S/N computed for 10,000 differential light curves with varying morphological characteristics. All light curves span 2.8 days with an exposure time of 2 minutes and contain the same transit signal of duration $D = 1$ hr, depth of 0.5%, and white noise with a standard deviation of 0.1%. Light curves at the top and right side of the central plot are shown with their corresponding τ and δ values, which again correspond to the relative timescale of stellar variability against the transit duration, and the relative amplitude of stellar variability against the transit depth. The two central plots show the transit S/N after detrending the light curves with one of two methods. On the left, light curves are detrended using Tukey's bi-weight filter with an optimal window size of $3 \times D$ (see Hippke et al. 2019), while on the right, light curves are detrended using a GP with the same kernel used to simulate the data.

This is the joint 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 be found periodically and share a common transit depth Δ. To this end, it can be shown (see Appendix B) that there is a closed form expression for the joint likelihood of K individual transits with depths and errors $\{\Delta_k, \sigma_k\}_k$ assuming a common depth Δ, corresponding to

$$\ln p(\boldsymbol{f}|P, T_0, D, \Delta) = \sum_k^K \ln \mathcal{L}_k - \frac{1}{2}\sum_k^K \times \left(\ln(\sigma_k^2) - \ln(\sigma^2 + \sigma_k^2) + \frac{(\Delta_k - \Delta)^2}{\sigma_k^2 + \sigma^2} \right)$$
$$\text{with } \frac{1}{\sigma^2} = \sum_k^K \frac{1}{\sigma_k^2} \text{ and } \Delta = \sigma^2 \sum_k^K \frac{\Delta_k}{\sigma_k^2}. \quad (8)$$

In order to compute this joint likelihood, we must assume that the likelihoods computed during the linear search are independent. However, the individual transits combined in the periodic search are not independent. Indeed, by using a GP, we assume that points belonging to separate transits may have a nonzero covariance. In practice, we notice that this covariance is small enough to consider each transit as independent, a reasonable assumption for most physically realistic data sets.

While Equation (8) takes a closed form, 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), a similar issue is solved by using the nearest neighbors in the epochs grid. Instead, to allow for the efficient matrix computation of Equation (8), 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

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

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

3.4. The Transit-search Periodogram

Using Equation (8), we can now compute $\ln \mathcal{P}$ for a range of periods and phases, and build a transit-search periodogram using Equation (4). This has two disadvantages: First, each likelihood $p(\boldsymbol{f}|T_0, D, \Delta)$ 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 $\{\mathcal{L}_k\}_k$ (see Equation (8)), artificially leads to a likelihood involving up to $N \times K$ measurements. For this reason, one has to normalize each likelihood $\ln \mathcal{P}$, by keeping track of the number of points used to compute each of them, which differs from one phase to another. Second, the maximum value of the likelihood $\ln \mathcal{P}$ is relative to a given data set, so a more intuitive and absolute metric must be used to relate to the transit signal detection (such as the Signal Detection Efficiency in Kovács et al. 2002). This motivates a final step to produce an interpretable transit-search periodogram $\mathcal{Q}$.

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

$$(\phi_0, D) = \underset{\phi_i, D_j}{\arg\max} \{\ln p(\boldsymbol{f}|P, \phi_i, D_j)\}_{i,j} \quad (9)$$

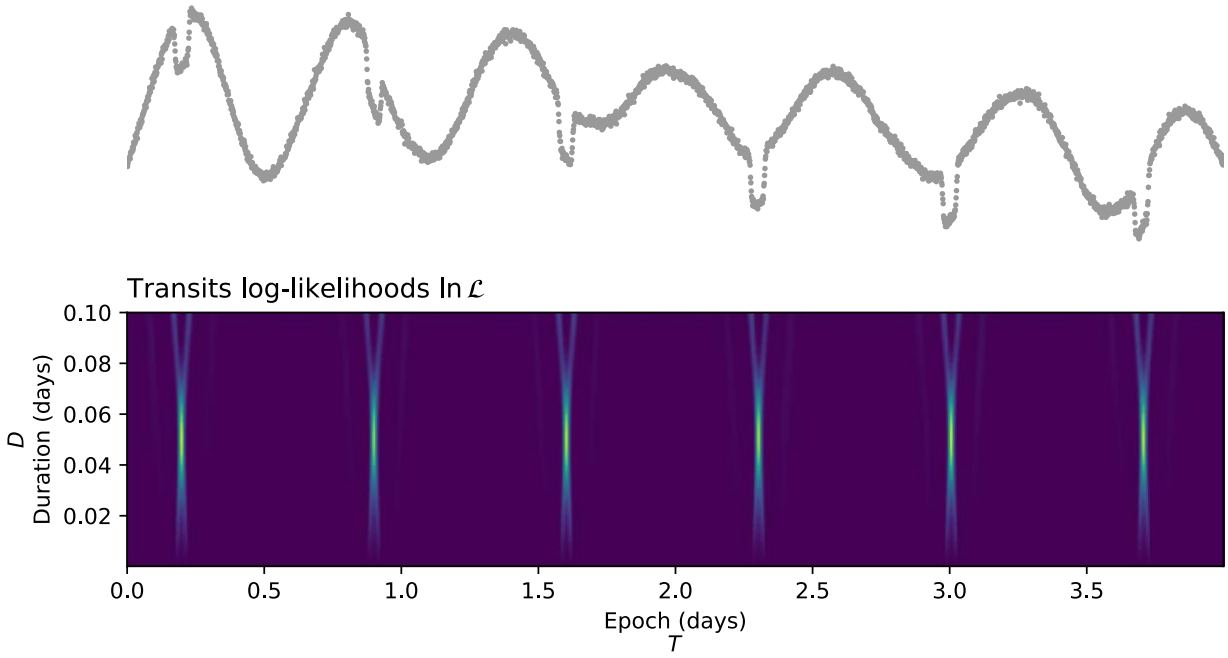

Figure 5. Principle and output of the linear search. The simulated data set (top) corresponds to the one shown and described in Figure 1. First, a set of durations and depths $\{ T_i, D_j \}_{i,j}$ is generated. For each pair of indices (i,j), the likelihood $\ln p(f|T_i, D_j, \Delta_{i,j})$ is computed 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 inferred linearly using Equation (6).

and define $\mathcal{Q}(P)$ as being equal to the S/N of the transit of period P, epoch $T_0 = \phi_0 \times P$, duration D, and depth Δ, i.e.,

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

where Δ and σ are obtained using Equation (8) 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.

Hence, periodic transit of period P with the maximum S/N, i.e., maximizing $\mathcal{Q}$, is adopted as the best candidate, basing the confidence in this signal through its S/N. The parameters of this transit are the period P, epoch $T_0 = \phi_0 P$, duration D (Equation (9)), and depth Δ with error σ (given by Equation (8)).

3.5. An Open-source Python Package

The methods presented in this paper are made available through the nuance open-source Python package hosted at https://github.com/lgrcia/nuance, released on the Python Package Index[9], and with documentation and tutorials hosted at https://nuance.readthedocs.io. All following mentions of nuance refer to version 0.6.0 (Garcia et al. 2024).

To instantiate a search, a user can start by creating a Nuance object with

```
from nuance import Nuance
nu = Nuance(time, flux, gp = gp, X = X)
```

where gp is a tinygp GP instance and X is the design matrix of the linear model. nuance exploits the use of tinygp,[10] a Python package powered by JAX,[11] allowing for custom kernels to be built and highly tractable computations. We can then define a set of epochs t0s and durations Ds and run the linear search with

```
import numpy as np


t0s=time.copy()
# a range of 10 durations
Ds=np.linspace(0.01, 0.2, 10)
nu.linear_search(t0s, Ds)
```

Finally, the periodic search is run with

```
# range of periods
periods=np.linspace(0.1, 5, 2000)
search=nu.periodic_search(periods)
```

From this search object, the best transiting candidate parameters can be computed (search.best), or the $\mathcal{Q}$ periodogram retrieved (search.Q_snr), together with valuable information about the transit search. The Nuance object also provides methods to perform transit search on light curves from multiplanetary hosts, the advantage of *nuance* being that the linear search only needs to be performed once and reused for the search of several transiting candidates (see Section 4.3). An extensive and maintained online documentation is provided at nuance.readthedocs.io.

3.6. Comparison with BLS

To start testing nuance against existing methods, a simple adimensional normalized light curve is simulated, consisting of pure white noise with a standard deviation of 5×10^{-4} spanning 6 days with an exposure time of 2 minutes. From this signal, we produce 4000 light curves, each containing box-shaped transits with periods randomly sampled from 0.3 to 2.5

[9] https://pypi.org/project/nuance/
[10] https://github.com/dfm/tinygp
[11] https://github.com/google/jax

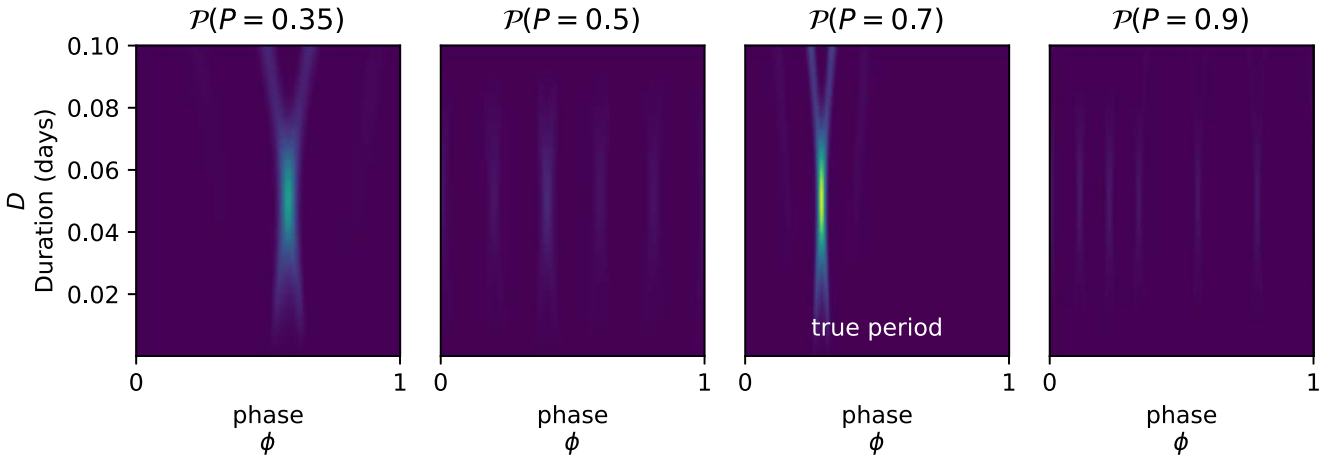

Figure 6. Periodic search likelihood $\mathcal{P}(P)$ computed for different trial periods P. Notice how the maximum value of $\mathcal{P}$ for the alias period $P = 0.35$ day (left plot) is lower than for $P = 2 \times 0.35 = 0.7$ day, a result of combining the log-likelihoods using Equation (8) instead of Equation (7), in favor of individual transits matching a common depth Δ.

days, durations of 50 minutes, and depths randomly sampled to lead to transit S/Ns ranging from 4 to 30. For each light curve, transits are searched using two different tools: nuance, using its implementation from the Python package described in the previous section; and the BLS algorithm from Kovács et al. (2002; using astropy's BoxLeastSquares[12] implementation). We note that the Transit-Least-Squares algorithm (Hippke & Heller 2019), which includes the effect of limb-darkening in the base transit template, could also be used here. However, as this effect has a negligible impact on transit detection compared to the effect of correlated noise, we choose to make our comparisons with BLS only.

For both methods, 3000 trial periods from 0.2 to 2.6 days are searched, with a single trial duration fixed to the unique known duration of 50 minutes. A transit signal is considered detected if the absolute difference between the injected and the recovered period is less than 0.01 day. To ease the detection criteria, orbital periods recovered at half or twice the injected ones (aka *aliases*) are considered as being detected. For this reason, detected transit epochs are not considered (although manually vetted). Results from this *injection–recovery* procedure are shown in Figure 8.

These results demonstrate the qualitative match between the detection capabilities of nuance and BLS on light curves with no correlated noise, where the BLS algorithm should be optimal. Explaining the subtle differences observed between the two methods when only white noise is present is beyond the scope of this paper, and we will assume that any differences observed in the following sections are due to the different treatments of correlated noise.

4. Performance

Figure 4 shows that nuance's full-fledged modeling capabilities may not always be necessary and may only be beneficial for certain noise characteristics, relative to the searched transit parameters. Here, we evaluate the performance of nuance in the relative parameter space (τ, δ) described in Equation (3) and probe when its specific treatment of correlated noise in the transit search becomes necessary.

We perform this study by comparing nuance to the approach that involves removing stellar variability from light curves before performing the search on a detrended data set. The following detrending strategies, each followed by a search with the BLS algorithm, are compared:

`bi-weight+BLS`: employs an optimal bi-weight filter implemented in the wõtan Python package with an optimal window size of $3 \times D$, i.e., 3 times the transit duration (e.g., Hippke et al. 2019 and Dransfield et al. 2024).

`GP+BLS`: employs a GP conditioned on the data (e.g., Lienhard et al. 2020). The kernel of the GP and its optimization is described on a case-by-case basis.

`Bspline+BLS`: employs a B-spline[13] for detrending, fitted using the scipy.interpolate.splrep function[14] (e.g., Hippke et al. 2019 and Canocchi et al. 2023).

`harmonics+BLS`: employs a linear harmonic detrending, where the light curve is modeled as a Fourier series including four harmonics of the stellar rotation period with coefficients found through ordinary least squares.

`iterative+BLS` iteratively detrends the light curve with a sinusoidal signal fitted to the data, each time using the dominant period of the residuals found using a Lomb–Scargle periodogram (five iterations).

Like in previous sections, we consider the planet detected if the recovered period is within 0.01 day of the true period or a direct alias (such as 2 times or half the true period). Again, we ignore the exact match between the injected and recovered transit epochs, although we visually vet that the found epochs are consistent with the ones injected. We consider a transit detectable if its original S/N is greater than 6. Hence we define *true positives* as detectable transits recovered with the correct period (or an alias) and a measured S/N greater or equal to 6, and *false positives* as nondetectable transits recovered with a measured S/N greater or equal to 6.

As we noticed that few methods were still affected by the remaining stellar variability after detrending, the grid of orbital periods being searched for does not contain the stellar rotation period P_* and its aliases. In practice, this is done by removing

[12] https://docs.astropy.org/en/stable/api/astropy.timeseries.BoxLeastSquares.html

[13] https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.BSpline.html

[14] https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.splrep.html

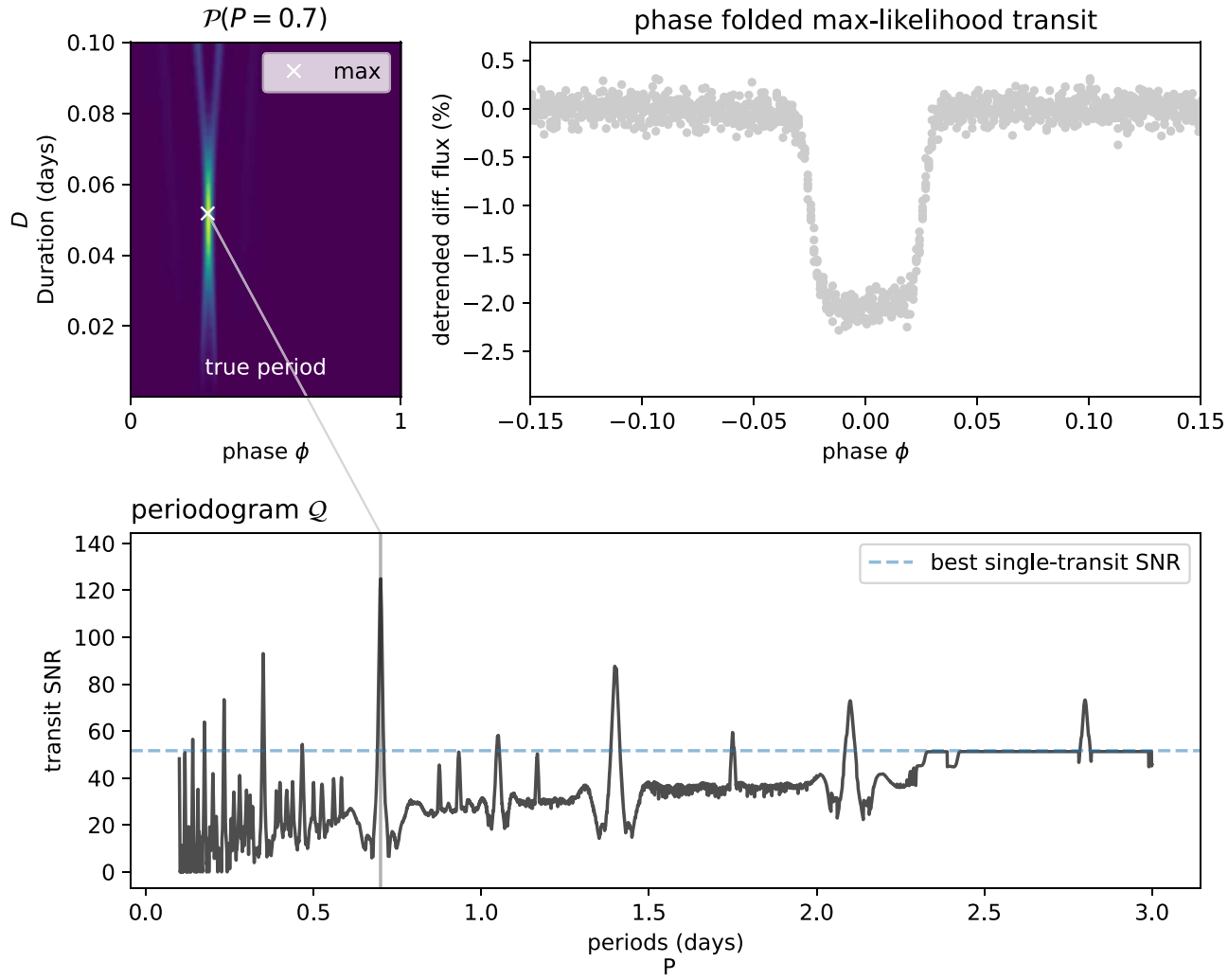

Figure 7. For each period P, the joint likelihood $\mathcal{P}(P)$ is computed using Equation (8), and the value of the maximum-likelihood transit S/N is retained as $\mathcal{Q}(P)$.

all orbital periods P from the search grid such that $dP = \frac{P}{P_*}$ is less than 2% from an integer value, i.e., $|dP - \lceil dP \rceil| < 0.02$.

4.1. Comparisons on Simulated Light Curves

Our first comparison data set consists of 4000 light curves simulated using the model described in Section 2.1. We simulate a common periodic transit added to all light curves, of period $P = 1.1$ days, epoch $T_0 = 0.2$ days, duration $D = 0.04$ days, and depth $\Delta = 1\%$. Each light curve consists of a 4 day observation with an exposure time of 2 minutes, leading to $N = 2880$ data points with a normal error of 0.1%.

For a given pair of (τ, δ), we simulate stellar variability using a GP with an SHO kernel of hyperparameters defined by Equation (3) computed with respect to the injected transit parameters D and Δ. The same kernel is used for the search with nuance and with the `GP+BLS` method, an optimal choice on equal footing with the optimal $3 \times D$ window size of the bi-weight filter employed in the `bi-weight+BLS` search. The 4000 pairs of (τ, δ) are drawn from

$$\tau \sim \mathcal{U}(0.1, 10), \quad \delta \sim \mathcal{U}(0.1, 25) \text{ and } Q \sim \mathcal{U}(10, 100),$$

where $\mathcal{U}(a, b)$ denotes a uniform distribution of lower bound a and upper bound b.

4.1.1. Results

The results of this injection–recovery procedure are shown in Figure 9 and highlight particularly well the benefit of nuance against most other methods. Except for the `GP+BLS` approach, nuance leads to a much higher rate of true positives for transits with relatively small depths compared to the stellar variability amplitude (i.e., $\delta > 2$), and duration comparable to the stellar variability period (i.e., $\tau < 2$). On the other hand, the `GP+BLS` strategy seems almost as performant as nuance, recovering most of the injected transits and only lacking detection in a relatively comparable portion of the (τ, δ) parameter space. We note that these empirical statements only concern simulated light curves with a given amount of white noise and may vary depending on the length of the observing window or the number of transits. For this reason, quantifying which values of (τ, δ) nuance outperform these methods would only apply to this specific but representative example. However, we verify that our conclusions remain qualitatively valid for various simulation setups.

This injection–recovery is done in a particularly optimal setup, on simulated light curves that are not all physically realistic and using an optimal GP kernel, hence demonstrating the performance of nuance only on a purely synthetic basis. Hence, we strongly emphasize that these tests do not reflect

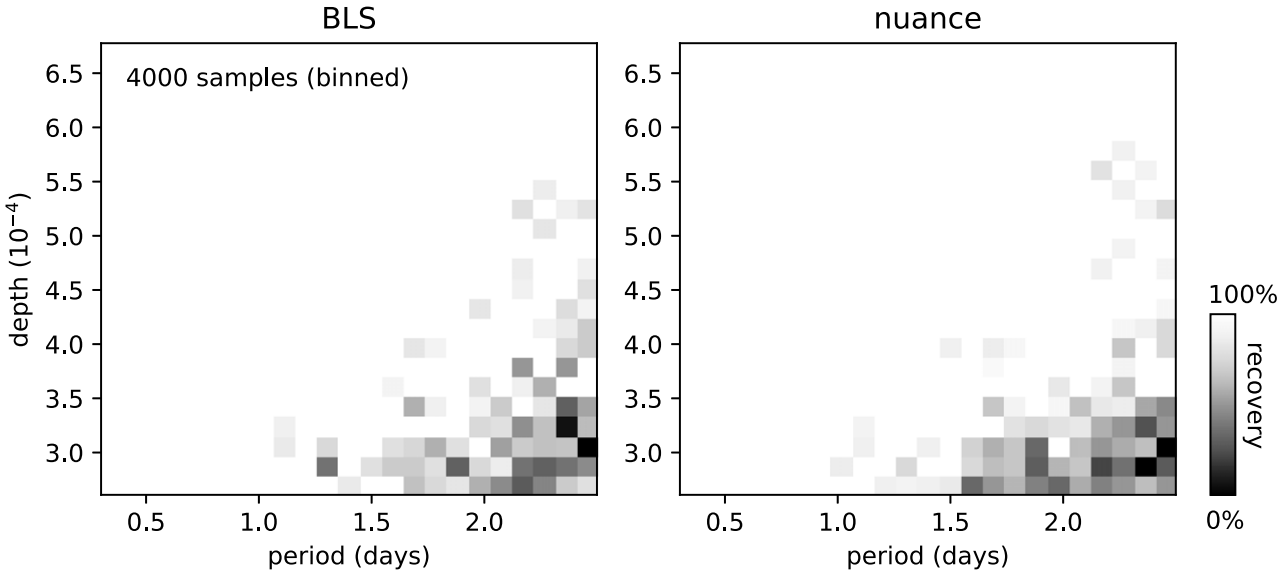

Figure 8. Binned statistics of the injection–recovery of 4000 transit signals in a flat light curve with only white noise using BLS and nuance. The color scale indicates the recovery of transits in the corresponding (period, depth) parameter space, white for a full recovery and black for no detection.

well the performance of each method on real data sets. In the next section, we perform transits injection–recovery on real space-based light curves.

4.2. Comparisons on Rapidly Rotating M Dwarfs TESS Light Curves

In order to assess the performance of *nuance* on real data sets, we inject and recover transits into light curves from TESS (Ricker et al. 2015). We focus this proof of concept on the light curves of 438 M dwarfs found to have detectable rotation signals with periods lower than one day (Ramsay et al. 2020), which lead to a parameter space justifying the use of *nuance*. For each of the 438 targets, transits are injected and recovered in the TESS 2 min cadence SPOC Simple Aperture Photometry and Pre-search Data Conditioning light curves (PDCSAP; Caldwell et al. 2020) of a single sector (the first being observed for each target) spanning on average 10 days. To our knowledge, none of these 438 targets have been searched for planetary transits before. However, we note that the presence of existing transit signals in these light curves before the injection of simulated ones is possible, but will not affect the relative comparison of one method to another. As described in the following sections, our experiment considers a total of 438×100 light curves, where transit signals are injected and recovered.

4.2.1. Light-curve Cleaning and Transits Injection

As some of the techniques compared to *nuance* can be affected by gaps in the data, we only use continuous measurements from half a TESS sector. We assume that all methods (including *nuance*) are based on an incomplete model of the data that does not account for stellar flares. For this reason, the light curve of each target is cleaned using an iterative sigma-clipping approach.

For each iteration, points 3 times above the standard deviation of the full light curve (previously subtracted by its median) are identified. Then, the 30 adjacent points on each side of the found outliers are masked. This way, large flare signals are masked, using a total of three iterations. At each iteration, the GP kernel hyperparameters are reoptimized. As PDCSAP light curves often start with a ramp-like signal, the first 300 points (as well as the last 300 points) of each continuous observation are masked. Finally, each light curve is normalized by its median value. Such a cleaned light curve is shown in Figure 10. We note that the gaps left after sigma clipping may be problematic for some of the detrending techniques (such as `bspline+BLS`). However, adopting this flare-cleaning step and analyzing light curves with a few small gaps is a practice commonly found in the literature.

For each of the 438 light curves considered, transits of planets with 10 different orbital periods combined with 10 planetary radii are individually considered, for a total of 100 periodic transits injection–recovery per target. Orbital periods P are sampled on a regular grid between 0.4 and 5 days, and planetary radii R_p are sampled on a regular grid designed to yield a minimum transit S/N of 2 and a maximum of 30. Using Equation (2) with $\sigma_r = 0$, the planetary radius leading to a transit with a desired S/N s is given by

$$R_p = R_\star \, n^{-\frac{1}{4}} \sqrt{\sigma s},$$

where σ is equal to the mean uncertainty estimated by the SPOC pipeline, $R_\star$ is the radius of the star reported by Ramsay et al. (2020), and n is the number of points in transit computed using a transit duration assuming a circular orbit. In total, 438×100 light curves are produced and searched for transits, corresponding to planets radii ranging from $0.46 R_\oplus$ to $12.29 R_\oplus$ and a median radius of $2.40 R_\oplus$.

4.2.2. Stellar Variability Kernel

In the `GP+BLS` and nuance methods, we model light curves using a GP with a mixture of two SHO kernels of period $P_\star$ and $P_\star/2$, where $P_\star$ is the rotation period of the star. This model is representative of a wide range of stochastic variability in stellar time series[15] (e.g., David et al. 2019; Gillen et al. 2020). In order to account for additional correlated noises, we

[15] https://celerite2.readthedocs.io/en/latest/api/python

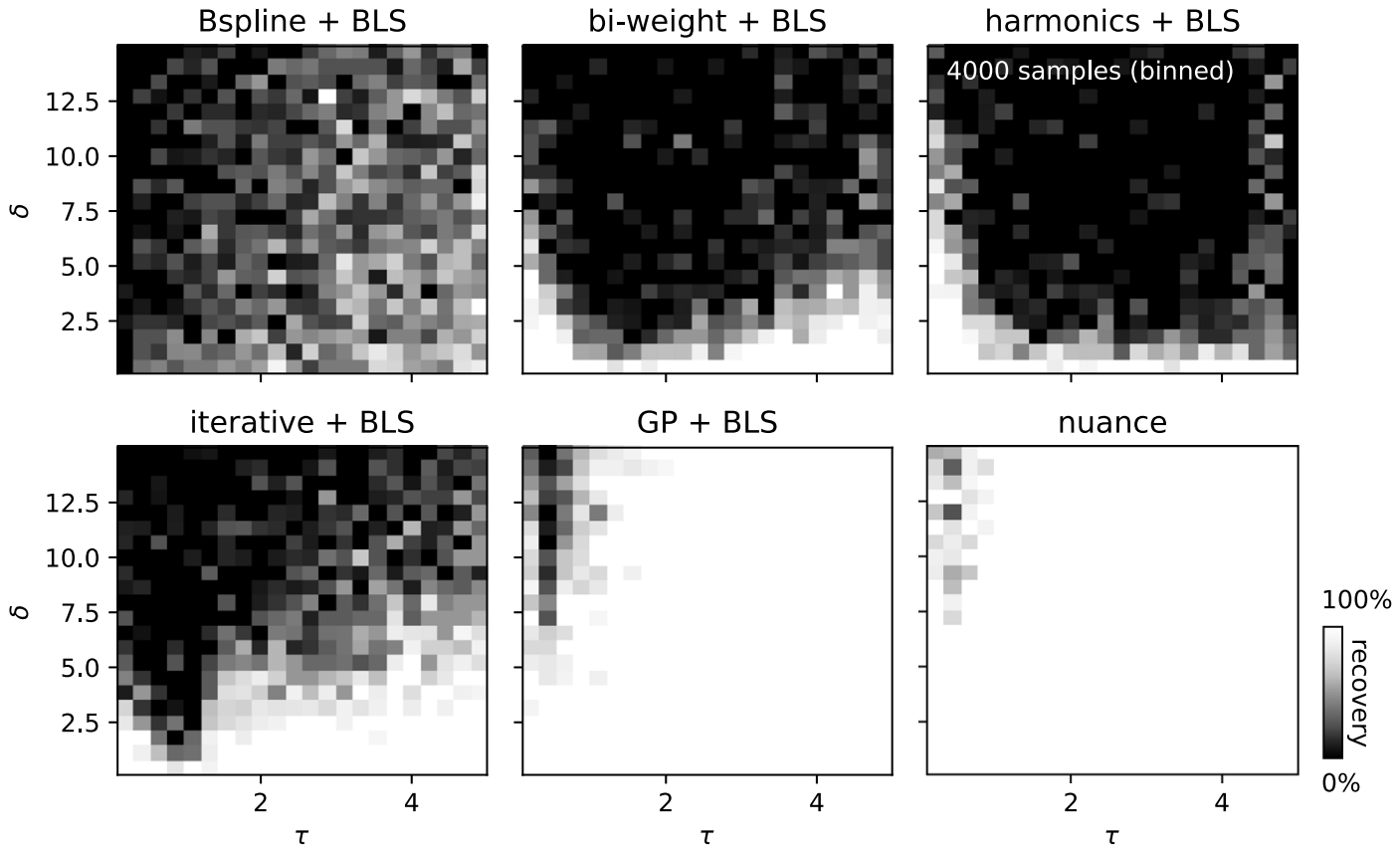

Figure 9. Rate of true positives for the 4000 simulated light curves. The color scale represents the fraction of true positives, white if all injected transits are recovered in a given portion of the (τ, δ) parameter space, and black if none are recovered.

complement this kernel with a short and a long-timescale exponential term, so that the full kernel can be expressed as

$$k = k_1 + k_2 + k_3 + k_4$$

with

1. k_1 a SHO kernel with hyperparameters

$$Q_1 = 1/2 + Q_0 + \delta Q,$$
$$\omega_1 = \frac{4\,\pi\,Q_1}{P\,\sqrt{4\,Q_1^2 - 1}} \text{ and } S_1 = \frac{\sigma^2}{(1+f)\omega_1\,Q_1}.$$

2. k_2 a SHO kernel with hyperparameters

$$Q_2 = 1/2 + Q_0,\ \ \omega_2 = 2\omega_1 \text{ and } S_2 = \frac{f\,\sigma^2}{(1+f)\omega_2\,Q_2},$$

where Q_0 is the quality factor for the secondary oscillation, δQ is the difference between the quality factors of the first and the second modes, f is the fractional amplitude of the secondary mode compared to the primary, and σ is the standard deviation of the process. Kernels k_3 and k_4 are expressed as

$$k(t, t') = \sigma^2 \exp\left(-\frac{|t - t'|}{\ell}\right),$$

with ℓ and σ being the scale and standard deviation of the process, respectively. These are meant to model short and long-timescale nonperiodic correlated noise. In total, the rotation kernel k has eight hyperparameters.

The hyperparameters of this kernel are optimized on trimmed and cleaned light curves containing the injected transits, using the scipy.optimize.minimize wrapper provided by the jaxopt Python package,[16] and taking advantage of the JAX implementation of tinygp and its quasi-separable kernels[17] (Foreman-Mackey et al. 2017). As correlated noise is expected to affect the light-curve uncertainty estimates performed by SPOC, the diagonal of the full covariance matrix of the data (i.e., their uncertainty, assuming homoscedasticity) is held free, increasing the number of optimized parameters to nine. The optimization is performed using the BFGS algorithm (Fletcher 1987), minimizing the negative log-likelihood of the data as expressed in Equation (5) (without transit), i.e., accounting for a linear systematic model of the data in addition to stellar variability. For simplicity, and to adopt a uniform treatment for all target light curves, a design matrix $\boldsymbol{X}$ with a single constant column is adopted, such that the systematic model only consists of a single parameter corresponding to the mean value of the differential flux (expected to be close to 1) solved linearly. Our motivations for choosing this very simplistic baseline, despite the capability of *nuance* to account for more complex linear models, is discussed in Section 5.2.

4.2.3. Search Parameters and Transit Detection Criteria

For all techniques, we only search for transits with a duration fixed to the known duration of the injected transits. This is mainly done for computational efficiency and allows for a

[16] https://jaxopt.github.io/stable/_autosummary/jaxopt.ScipyMinimize.html
[17] https://tinygp.readthedocs.io/en/latest/api/kernels.quasisep.html

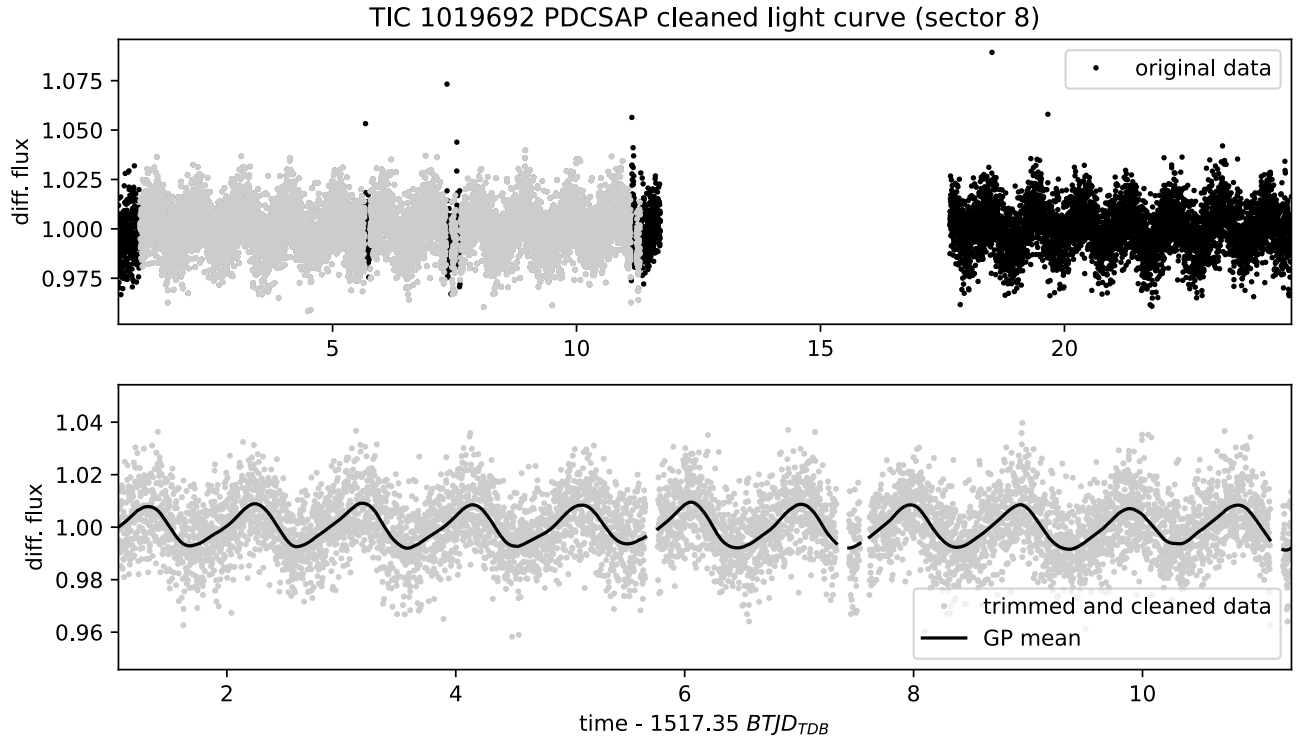

Figure 10. Trimmed and cleaned single-sector light curve of the target TIC 1019692. The top plot shows how much of the data is truncated and sigma clipped, resulting in a quasi-continuous light curve shown in the bottom plot. On this bottom plot, the black line corresponds to the mean of the GP model (with hyperparameters optimized here on the cleaned light curve).

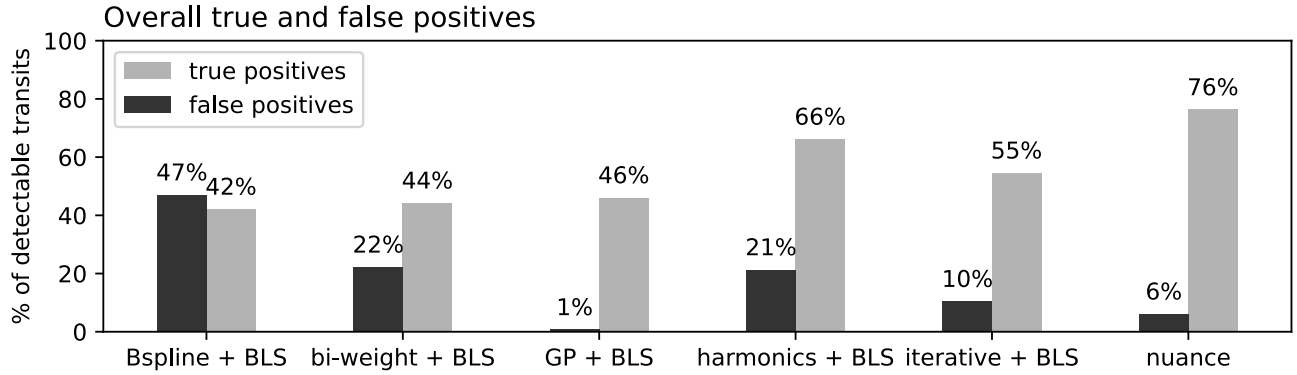

Figure 11. Rate of true and false positives for all methods as a fraction of detectable transit signals.

narrower comparison between *nuance* and all BLS-based techniques. Finally, the search is done over 4000 trial periods linearly sampled from 0.3 to 6 days. A realistic transit search on a wider parameter space (e.g., multiple trial durations) is discussed in the next section.

4.2.4. Results

An example of the transit injection–recovery and its result is shown in (Figure 19) for the target TIC 1019692, a representative example of the kind of stellar variability encountered in our data set. Figure 11 shows the global results of the injection–recovery for all targets, while Figure 12 shows the true and false positives plotted against the relative parameters τ and δ (defined in Equation (3)). These figures are a synthesis of Figure 13, which shows the rate of true and false positives binned in the full parameter space (τ, δ).

Compared to other techniques, we find that nuance leads to the highest number of true positives, with successful detection of 76% of the 7008 detectable transits injected (Figure 11). While the performances of other methods strongly depend on the characteristics of the variability, nuance is the best technique in 93% of cases, leading to both the highest number of true positives and the lowest number of false positives (Figure 11).

From Figure 11, we note that the number of true positives of the `GP+BLS` method is much worse than what could be anticipated from the results presented in Figure 9, where the GP kernel was fully optimal (given it was also used to simulate the data). In that case, our kernel might not be optimal, whether because of its form or because of the values of its hyperparameters. In any case, the fact that the same kernel performs significantly better when used with nuance shows that our method is less sensitive to the choice of kernel and its optimization compared to detrending with a GP.

Finally, in Figure 14, we observe a dependence between the rate of true positives and the injected planet radii. Overall `nuance`

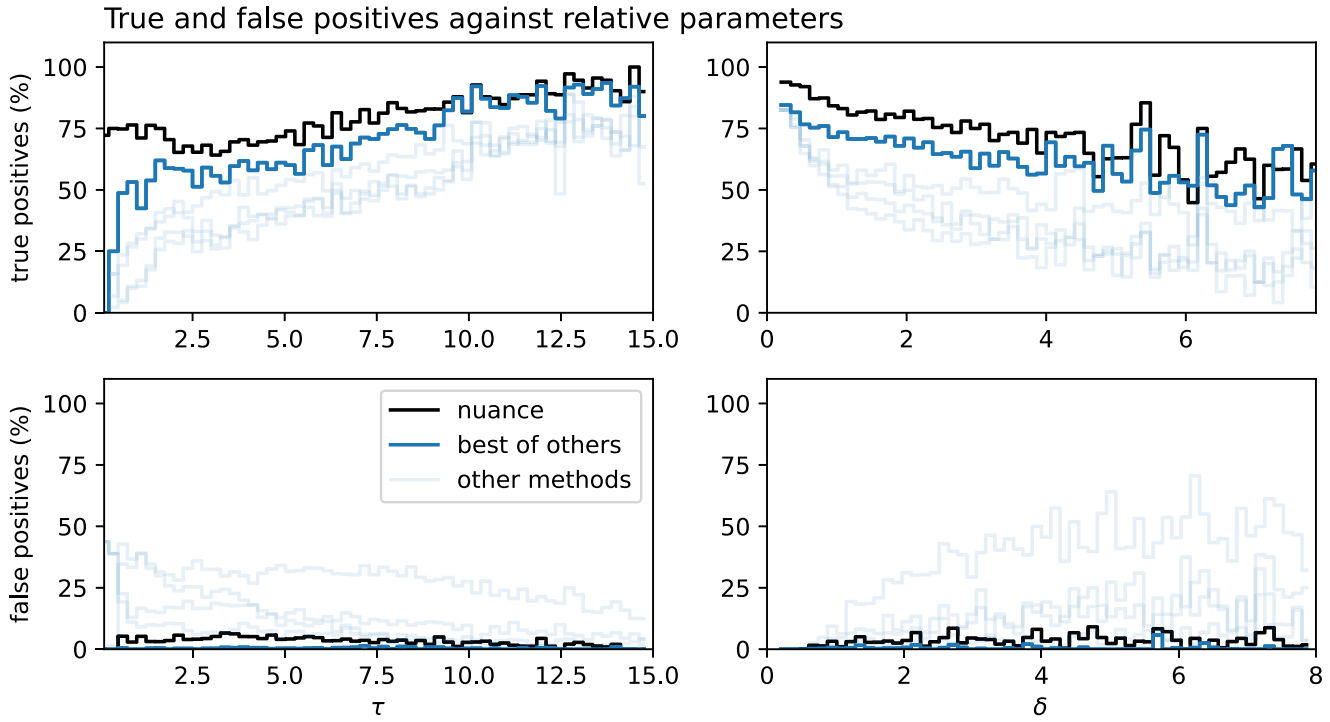

Figure 12. Rate of true and false positives of nuance compared to other methods, as a function of the relative parameters τ and δ. The solid blue line corresponds to the maximum of true positives among all methods on the top panel (dominated by the `harmonics+BLS` method), and the minimum of false positives on the bottom panel (waned by the `GP+BLS` method, all other methods being above nuance).

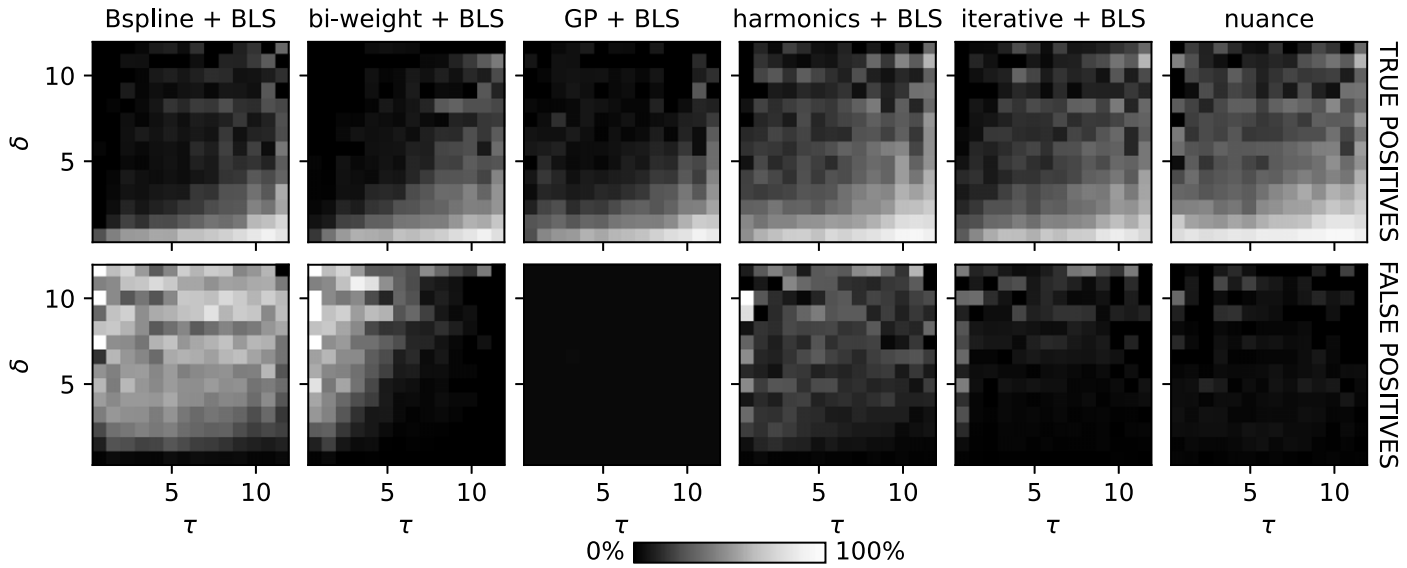

Figure 13. Rate of true and false positives binned over the full parameter space (τ, δ) for all methods and all targets.

leads to significantly higher true positives in all radii bins compared to other methods, except for `BLS+harmonics` (which dominates other methods), which shows a similar yield for planets with radii less than $2R_\oplus$. Hence, we note that the advantage of `nuance` is particularly significant for planets larger than $2R_\oplus$. This can be explained by the fact that, for a transit signal to be comparable to a high-amplitude stellar variability signal, i.e., when `nuance` shows an advantage, a larger transit depth, hence a larger planetary radius, is required. Although this particular result highly depends on the variability present in the studied data set, we conclude that the performance of *nuance* is not sensitive to the searched planets' radii.

When analyzing the light curves of the 438 targets from Ramsay et al. (2020), our search solely focused on the injected transit signals. Although nuance seems particularly well suited to search for real transit signals in this data set, we did not conduct such a study. To be done properly, this task would simply require a search on a wider and finer trial epochs and durations grid, using all available TESS sectors; a project that we highly encourage.

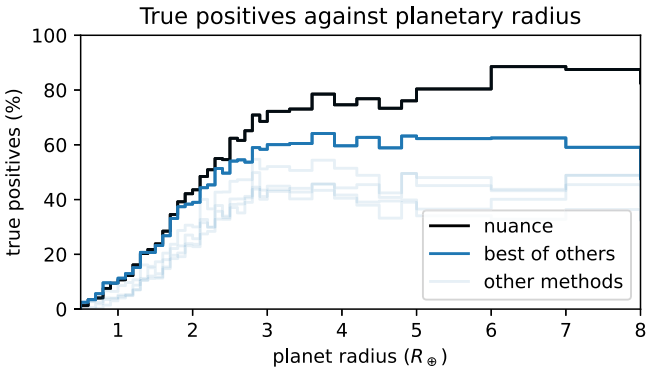

Figure 14. Rate of true and false positives of nuance compared to other methods, as a function of the injected planet's radius. The solid blue line corresponds to the maximum of true positives among all other methods than *nuance* (dominated by the `harmonics+BLS` method).

4.3. Comparison for a Multisector TESS Candidate: TOI-540

In order to further validate nuance on a realistic data set, we focus this section on the multisector TESS light curves of TOI-540, and the search for its Earth-like companion TOI-540 b (Ment et al. 2021). As we will see, *nuance* does not have significantly superior performance for this candidate. However, its light curves feature a very typical stellar variability signal (see Figure 15), that would commonly be detrended before searching for planetary transits.

We downloaded the 2 minute cadence SPOC PDCSAP light curves of TOI-540 observed in five sectors (4, 5, 6, 31, and 32). Like in the previous section, we use a Lomb–Scargle periodogram and identify the 0.72 day rotation period of the star, which we use as an initial value to optimize the kernel described in Section 4.2.2 on each sector independently. Here again, we employ an upper sigma clipping to mask flares out of the data. The resulting light curve for sector 4 and its mean model are shown in Figure 15.

For each sector, we perform the linear search of nuance on the cleaned light curve, using the original times as the trial epochs and 10 trial transit durations linearly sampled from 15 minutes to 1.5 hr. We then perform the periodic search on all sectors combined, using a concatenation of all linear searches. By adopting this by-sector GP modeling of the light curve, the linear search of *nuance* scales linearly with the number of sectors being processed.

This approach is adopted for efficiency but also to encapsulate the changing properties of stellar variability from one sector to another, often separated by year-long gaps. The periodic search is done on 20,000 trial periods ranging from 0.5 to 10 days.[18] This search, using nuance, is compared to the more traditional approach that consists of detrending each sector with a bi-weight filter and then searching for transits with the BLS algorithm (denoted `bi-weight+BLS` and described in Section 4.2) on all sectors combined. Since we do not know the transit duration a priori, we perform the detrending and BLS search using 15 filtering window sizes sampled from 30 minutes to 5 hr and retain the search that leads to the highest transit S/N peak in the periodogram (as done e.g., in the SHERLOCK transit-search pipeline described in Pozuelos et al. 2020). The results of this comparison are shown in Figure 16.

After a first periodic search, trial epochs in windows of widths $2 \times D$ centered on the detected periodic transits are masked. In practice, this is done by masking the linear search products $\{\ln \mathcal{L}_{i,j}\}_{i,j}$, $\{\Delta_{i,j}\}_{i,j}$, and $\{\sigma_{i,j}\}_{i,j}$ (defined in Section 3.2).

As seen in Figure 16, the S/N periodogram using `bi-weight+BLS` and `nuance` are very similar, with the known transiting exoplanet TOI-540 b detected with an orbital period P = 1.24 days. This is well expected as the relative parameter τ equals 13 for TOI-540 b,[19] which lies outside the range where *nuance* is expected to be beneficial (see Figure 4). Nonetheless, the *nuance* periodogram of TOI-540 features less spurious S/N peaks, largely due to the penalty naturally occurring when single transits with different depths are periodically combined. In the second search (right panel) of Figure 16, we also notice a higher number of peaks that would lead to false positive detections of transits in the `bi-weight+BLS` case. The proper treatment of correlated noise in *nuance*, as observed in Section 4.2, makes these peaks nonsignificant, avoiding a large number of false detections.

We note that finding a known TESS candidate that displays characteristics for which *nuance* is expected to be significantly beneficial proved to be very challenging during the writing of this paper, as transits with such characteristics are expected to be missed by current state-of-the-art techniques (see, e.g., Figure 9). Nonetheless, we verify that *nuance* is capable of finding a large number of already known transiting exoplanet candidates, in light curves featuring various forms of correlated noises, at least as efficiently as with commonly used techniques. We reserve the search for new transit signals to a follow-up publication.

5. Discussion

In the previous section, we demonstrated the capability of *nuance* to search for synthetic or known transit signals, in simulated or real data sets. Here, we discuss the caveats of this algorithm, the advantages and limitations of the nuance implementation, and future prospects for its extension.

5.1. Processing Time

In Figure 17, the processing time of nuance linear and periodic search are recorded against the number of points in a simulated light curve, assuming a simple nonoptimized GP with a squared exponential kernel. These are compared to the processing time of `bi-weight+BLS` (see Section 4.1) separated into the bi-weight filtering step and the BLS search.

As seen in Figure 17, most of the computational costs of nuance and the `bi-weight+BLS` method come from the linear search and the bi-weight filtering step. This is not always true and depends on the size of the trial durations and period grids. One advantage of *nuance* is that the linear search can be performed separately on different continuous observations, and then combined in the periodic search. Hence, if searching for transits in separate observations with approximately similar durations, such as different TESS sectors or different ground-based nightly observations, the computational cost of nuance

[18] We acknowledge that the grid of trial periods and durations used in this study is nonoptimal for a real transit search. However, this simple choice is sufficient for our periodogram comparison against BLS. Optimal grids of parameters for realistic transit searches are discussed in Hippke & Heller (2019).

[19] Using Equation (3) with the stellar rotation $P = 0.72$ days and the known transit duration $D = 29.5$ minutes of TOI-540 b (Ment et al. 2021).

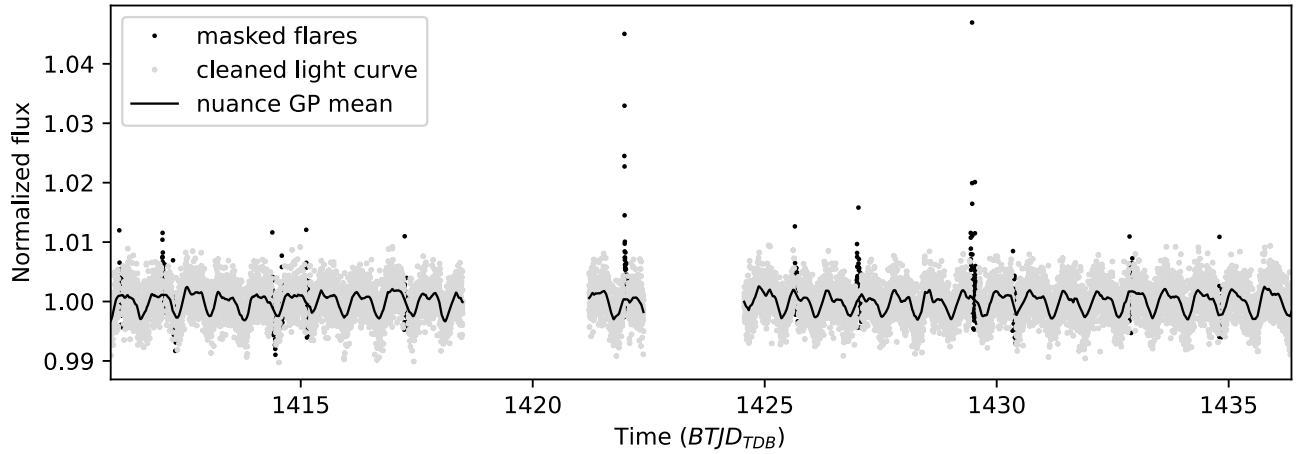

Figure 15. Sector 4 light curve of TOI-540. The cleaned signal (gray points) has been masked for flares (black points), and the black line corresponds to the mean of the GP model.

grows linearly with the number of observations (see gray line in Figure 17). Nonetheless, considering a variety of other detrending algorithms, nuance is expected to perform between one and 2 orders of magnitude slower than more conventional techniques associated with BLS.

Nonetheless, searching for TOI-540 b transits with the parallelized implementation of nuance (see Section 4.3) on the 12 cores of an Apple M2 Max chip, took only 5 minutes and 35 s, 1 minute 50 s for the five sectors linear searches (around 22 s per TESS sector), and 3 minutes 45 s for the combined periodic search. In comparison, the brute force search with `bi-weight+BLS`, consisting of trying 15 bi-weight windows, took a total of 1 minute 49 s (only 7 s each).

Because of its computational cost, we do not recommend using nuance in the general case, but rather when light curves contain correlated noise with specific characteristics. If employing a bi-weight filter for detrending, these characteristics correspond to the ones discussed in Section 2. But as these strongly depend on the type of detrending technique employed (see Figure 9), we do not provide general guidelines as to when *nuance* should be preferred over a specific detrending technique. To aid users in making an informed choice of algorithm, extensive benchmarks, and guidelines are reserved for future developments and will be progressively shared on nuance's online documentation.[20]

5.2. Systematics Modeling

Throughout the paper, a single-column design matrix $\boldsymbol{X}$, corresponding to the mean differential flux (ideally unitary), was employed, hence assuming that the instrumental systematics signals were nonexistent. In practice, *nuance* has been developed to linearly model systematic signals through more complex design matrices (as in Foreman-Mackey et al. 2015), in addition to its capability to model correlated noise while searching for transits. This feature is intentionally unexploited in the comparisons presented in Section 4, as detrending light curves assuming a linear systematics model, such as PLD co-trending vectors (Deming et al. 2015), is highly incomplete if applied to data while ignoring the presence of other astrophysical signals. Comparisons involving more complex design matrices would also be sensitive to the choice of linear components and would have unwanted repercussions on their results.

As an illustration, the NEMESIS pipeline (Feliz et al. 2021) starts processing the differential light curves by employing a linear systematics detrending using a least-squares fit of the data with a reduced PLD basis, before smoothing the signal from stellar variability using an approach similar to the one employed in the `bi-weight+BLS` approach (see Section 4.2), hence detrending the systematics with an incomplete model that does not account for stellar variability. To account for stellar variability while fitting the linear systematics model to the data, a step further would be to use a GP, such as done in the EVEREST (Luger et al. 2018) pipeline. However, this would also involve some potential degradation of the transit signals (see, e.g., Figure 3), with a hardly distinguishable origin. For these reasons, and to keep our comparisons as targeted as possible, we do not compare commonly used systematics detrending approaches and decided to focus our comparisons solely on stellar variability detrending techniques (although these two aspects often overlap in the literature, e.g., in Luger et al. 2016).

Although not being demonstrated here, modeling systematics signals while searching for transits on data acquired sparsely is extremely promising for the search of transiting exoplanets, including for ground-based observations that usually suffer from daily interruptions. In this respect, we note the similarity of our linear search (see Section 3.2) to the one presented in Berta et al. (2012), which focused on the detection of single eclipses in the MEarth light curves (Irwin et al. 2009). Similarly, *nuance* would highly benefit the search for transiting exoplanets around M dwarf type stars, such as the ones observed by the SPECULOOS survey (Sebastian et al. 2021) whose monitoring suffers from both increased red noise (due to atmospheric and instrumental thermal effects discussed in, e.g., in Berta et al. 2012 and Pedersen et al. 2023) and enhanced stellar variability (Murray et al. 2020). We reserve this promising application for a future study.

5.3. The GP Kernel

While not being discussed in our study, the efficiency of *nuance* to detect transits in correlated noise might be dependent on the design of its GP kernel. In the ensemble comparison of Section 4.2, the goal was to choose a kernel and an

[20] https://nuance.readthedocs.io/en/latest/

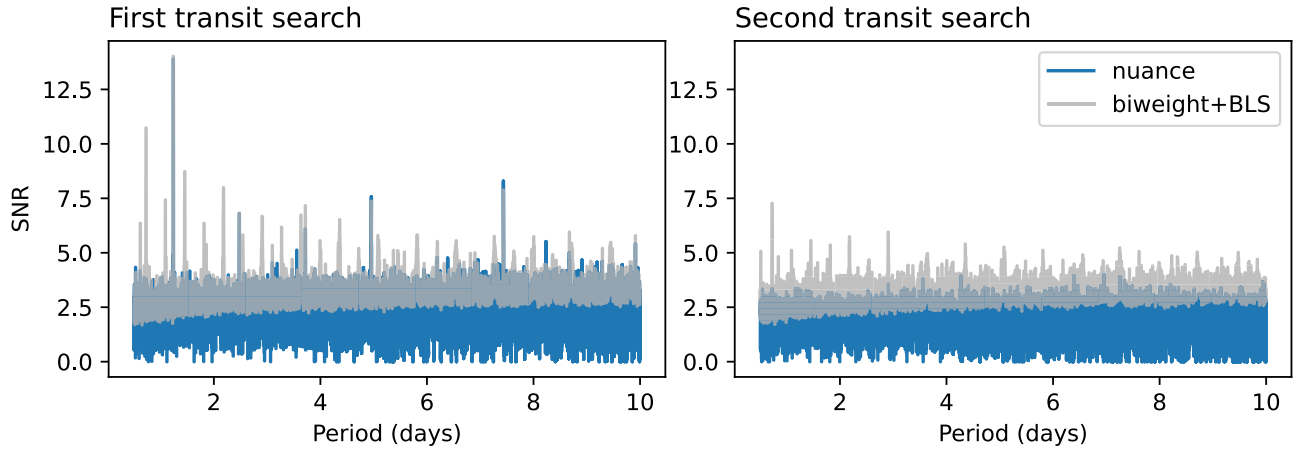

Figure 16. Transit-search S/N periodograms of TOI-540 using `bi-weight+BLS` and *nuance*. After a first periodic search (left panel), the epochs corresponding to the maximum-S/N transit are masked before the second search is performed (right panel).

optimization strategy suited to most of the studied light curves, leading to a few outliers in the results, which were indicative of a badly designed and/or optimized kernel. An alternative, recommended for more realistic blind searches, is to perform a model comparison on well-selected kernels and to adapt the optimization strategy to each data set.

When using nuance on TESS light curves for example, it must be noted that the observed light-curve variability might originate from contamination due to several nearby blended stars, so a physically interpretable GP kernel representing a single star activity is not necessarily appropriate. On the other hand, a single squared exponential GP kernel might also be sufficient for some applications; an aspect we intend to explore in future applications.

Something to emphasize is that *nuance* cannot be used to produce reliable detrended light curves, as an optimal GP is often flexible enough to partially model transits (see Figure 3). In contrast, the idea behind *nuance* is rather to compute the likelihood of data against a model containing both transits and correlated noise, without ever trying to disentangle both signals. In practice, it means that it can be very hard to actually verify the presence of transits found by *nuance* visually, so that transits may be detected but remain hidden in correlated noise. This is particularly true for stars displaying very-high-frequency photometric pulsations (see the example in Figure 18).

5.4. Prospects

The present implementation of *nuance* has the potential to be extended to be used beyond the search of periodic box-shaped transits. Here are ideas of possible use and extensions, from the most straightforward to the most ambitious:

1. In order to compare nuance to BLS-based methods, we injected and retrieved only box-shaped transits. However, similarly to the Transit-Least-Squares algorithm from Hippke & Heller (2019), limb-darkened transits can serve as a base model in the linear search and are expected to improve the transit search in the same way TLS provided an improvement over BLS.
2. The linear search of *nuance* is a single-event detection algorithm that can be used to search for single transit events, but also detect transiting exocomets and flares, by simply changing the base astrophysical model in the last column of the design matrix $\boldsymbol{X}$. While not being tested and benchmarked, nuance already integrates this feature. An example of the detection of known transiting exocomets in β Pictoris TESS light curves is shown in Figure 18.
3. During the periodic search, no prior about the transit duration related to the orbital period of the planet was used. This was done in order to allow for the detection of grazing transiting exoplanets that would produce shorter-timescale transits compared to what is expected from a circular orbit with a null impact parameter. However, adding such priors might produce fewer spurious periodogram peaks and be very beneficial for the automatic search of transits in large data sets. Another idea, similar to the one employed in Foreman-Mackey et al. (2015), is to leverage model comparison in order to reject transits that are better described by the GP model alone. Both ideas come at no cost given our modeling approach.
4. Finally, the formalism of *nuance* could be adapted and used to search for exoplanets featuring transit time variations (TTVs). Indeed, this application only requires a modification of the periodic search, as the maximum likelihood peaks close to linearly predicted transit epochs may be considered. This could be done either with a special nearest-neighbor algorithm or with a convolution of the computed likelihood grid with a Gaussian kernel. To maximize efficiency and interpretability, we would recommend these approaches to be explored analytically, rather than using a data-driven treatment of the linear search products.

6. Conclusion

This paper presents *nuance*, an algorithm designed to detect planetary transits in light curves featuring correlated noise in the form of instrumental signals and stellar variability. In this context, a conventional approach involves detrending a light curve before searching for transits using the BLS algorithm. However, we show that this approach degrades transit S/Ns down to the point of not being detectable. Adopting commonly used detrending strategies, we explore the extent of this

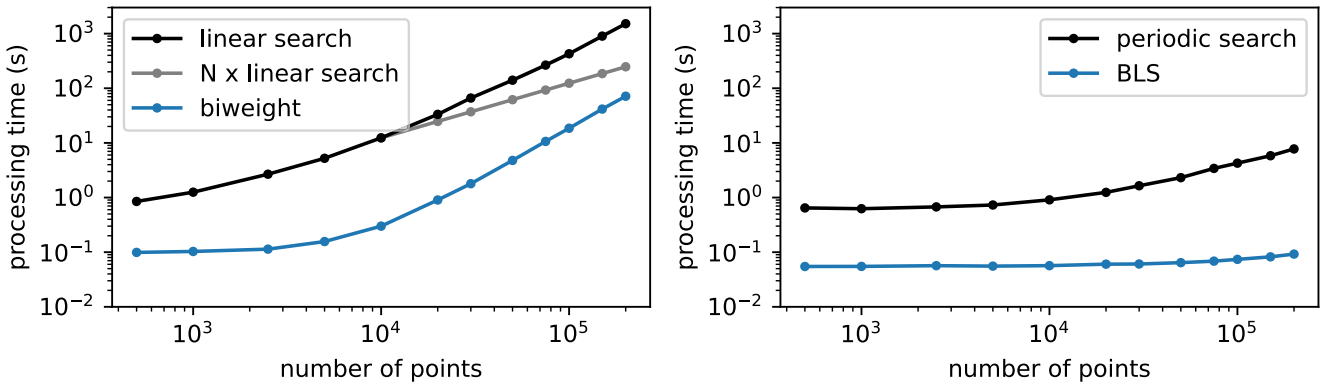

Figure 17. Study of nuance (black line) processing time against `bi-weight+BLS` (see Section 4.1, blue line). The gray curve shows the performance of nuance linear search when applied to chunks of 10,000 points continuous observations, instead of considering these observations all together. This study is performed on a single CPU core of an Apple M2 Max chip. While not being shown, we verify that both the BLS algorithm and *nuance* processing times scale linearly with the number of trial transit durations and orbital periods.

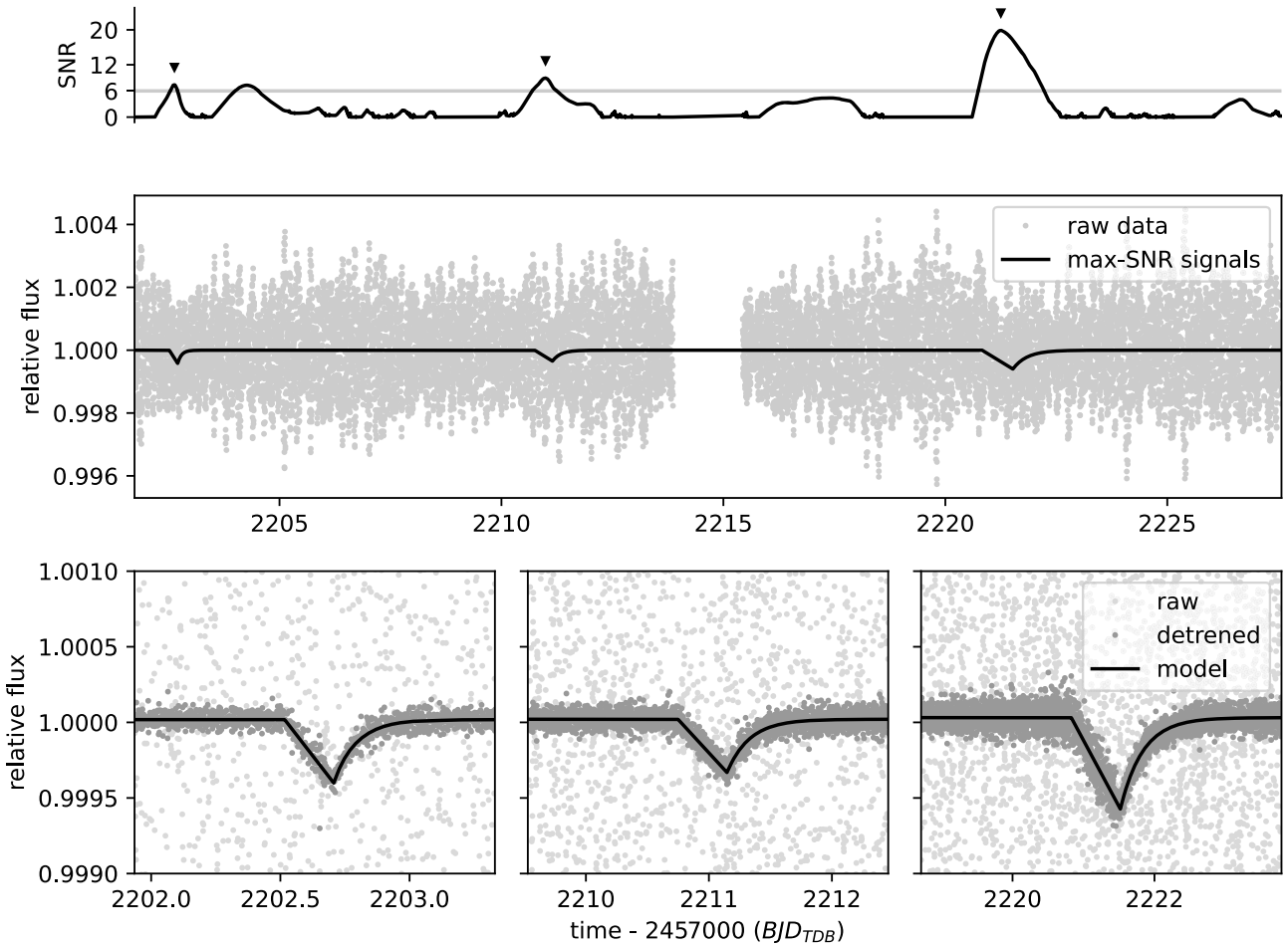

Figure 18. Demonstration of the search of transiting exocomets in the TESS light curves of the star β Pictoris (2 minutes PDCSAP data from sector 33). This star is known to display rapid δ Scuti type photometric variations with a period of about 30 minutes (Lecavelier des Etangs et al. 2022). The GP kernel and its hyperparameters are chosen and optimized as in Section 4.2. Here, we simply used the linear search of nuance with a different base model, one that mimics the shape of transiting exocomets (with 20 trial durations), to compute the S/N time series of the signal over one sector. The maximum-S/N events are displayed at the bottom of the figure and match with the ones found by Lecavelier des Etangs et al. (2022).

degradation on simulated light curves, and its dependence on the photometric stellar variability characteristics, showing the need for a full-fledged transit-search method like *nuance*.

The effectiveness of *nuance* is tested using a synthetic data set and further validated on real TESS light curves of 438 rapidly rotating M dwarfs. These injection–recovery tests reveal that *nuance* consistently outperforms commonly used transit-search techniques, especially when the timescale of stellar variability is less than twice that of the transit duration. In all cases, *nuance* not only leads to a higher number of true positive detections but also

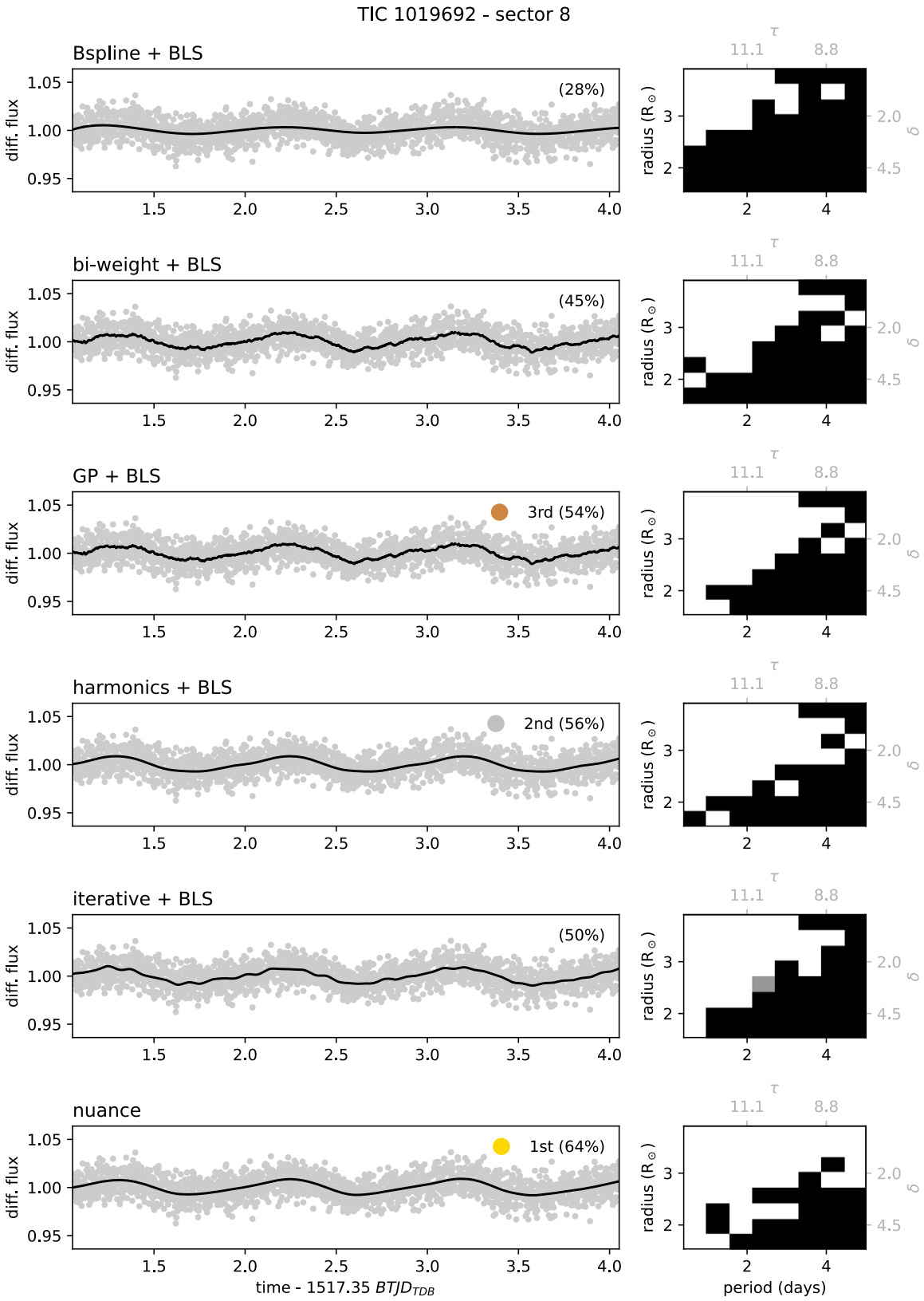

Figure 19. Results of the transits injection–recovery on TIC 1019692 half-sector light curve. Left: cleaned light curve with computed trend overplotted in black (except for nuance where it corresponds to the mean of the GP model). Right: results of the transit search where a black square denotes a transit signal not detected, gray a signal detected at an alias period ($P/2$ or $2P$), and white a signal detected with the correct period. On the right plots, secondary axes show the (τ, δ) relative parameter space. For each method, the upper right legend on the left plot indicates its ranking based on the percent of recovered transit signals (where a transit with an aliased period counts as being detected).

minimizes false positives, demonstrating its robustness and reliability.

We make *nuance* publicly available through the nuance open-source Python package, developed with JAX to allow its use on distributed computing environments and GPU devices. Overall, we acknowledge the limitations of *nuance* and its increased computational cost compared to more conventional techniques. Hence, *nuance* should be used as an alternative to more traditional techniques only in the presence of substantial correlated noise. As guidelines for choosing our method over other techniques are highly dependent on the type of detrending algorithm employed, we reserve this study for future work.

Finally, we suggest future improvements and extensions of the algorithm, including its application for detecting single transits, exocomets, flares, and exoplanets featuring TTVs, underscoring its versatility and potential for broader impact in astronomical research.

The software presented in this work is open source under the MIT License and is available at https://github.com/lgrcia/nuance, with documentation and tutorials hosted at https://nuance.readthedocs.io.

Acknowledgments

We would like to thank Julien De Wit and Prajwal Niraula for meaningful discussions at the beginning of this project, Germain Garcia for his useful insights about numerical optimization, Michaël Gillon for his overall support, the member of the Exotic group at the University of Liège, and the member of the Astronomical Data Group at the Center for Computational Astrophysics for many enriching discussions and feedback. We would also like to thank the journal referee for their thorough and encouraging report, bringing some very meaningful additions to the paper. This publication benefits from the support of the French Community of Belgium in the context of the FRIA Doctoral Grant awarded to L.J.G. F.J.P acknowledges financial support from grant CEX2021-001131-S funded by MCIN/AEI/ 10.13039/ 501100011033 and through projects PID2019-109522GB-C52 and PID2022-137241NB-C4. S.A. acknowledges funding from the European Research Council (ERC) under grant agreement 865624 (GPRV) and from the UK Science and Technology Facilities Council (STFC) under grants ST/S000488/1 and ST/R004846/1. This paper includes data collected by the TESS mission. Funding for the TESS mission is provided by NASA's Science Mission Directorate and NASA Explorer Program. We acknowledge the use of public TESS data from pipelines at the TESS Science Office and at the TESS Science Processing Operations Center. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center for the production of the SPOC data products. TESS data were obtained from the MAST data archive at the Space Telescope Science Institute (STScI). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. The Center for Computational Astrophysics at the Flatiron Institute is supported by the Simons Foundation.

Software: numpy (Harris et al. 2020), scipy (Virtanen et al. 2020), matplotlib (Hunter 2007), jax (Bradbury et al. 2018), jaxopt (Blondel et al. 2021), tinygp (Foreman-Mackey et al. 2024), lightkurve (Lightkurve Collaboration et al. 2018), astropy (Astropy Collaboration et al. 2013, 2018, 2022), pandas (McKinney 2010; pandas development team 2020), wotan (Hippke et al. 2019), and nuance version 0.6.0 (Garcia et al. 2024).

Appendix A
SHO Kernel

In order to model stellar variability and its effect on transit detection, we employ a simple physically motivated GP kernel, describing stellar variability through the covariance of an SHO (Foreman-Mackey et al. 2017; Foreman-Mackey 2018) taking the form

$$k(\tau) = \sigma^2 \exp\left(-\frac{\omega\,\tau}{2\,Q}\right) \times \begin{cases} 1 + \omega\,\tau & \text{for } Q = 1/2 \\ \cosh(f\,\omega\,\tau/2\,Q) + \sinh(f\,\omega\,\tau/2\,Q)/f & \text{for } Q < 1/2 \\ \cos(g\,\omega\,\tau/2\,Q) + \sin(g\,\omega\,\tau/2\,Q)/g & \text{for } Q > 1/2 \end{cases}$$
$$\text{where} \quad \tau = |t_i - t_j|, \quad f = \sqrt{1 - 4\,Q^2}$$
$$\text{and} \quad g = \sqrt{4\,Q^2 - 1}\,, \tag{A1}$$

where Q is the quality factor of the oscillator, ω is its pulsation, and σ is the amplitude of the kernel function. GP computations in this paper use the implementation from `tinygp`,[21] a Python package exposing the quasi-separable kernels from Foreman-Mackey (2018) and powered by JAX.[22]

Appendix B
Proof for the Periodic Search Expression

From the linear search presented in Section 3.2, 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 Section 3.2), we want an expression for

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

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

$$p(\boldsymbol{f}|T_k,\, D,\, \Delta) = \int p(\boldsymbol{f}|T_k,\, D,\, \tilde{\Delta}) p(\tilde{\Delta}|\Delta) d\tilde{\Delta}, \tag{B1}$$

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 (B1) involves the likelihood of the nonperiodic 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 squares, each follows a normal distribution $\mathcal{N}(\Delta_k, \sigma_k^2)$, centered on Δ_k with variance σ_k^2 and an amplitude $\mathcal{L}_k$, leading to the likelihood function

$$p(\boldsymbol{f}|T_k,\, D,\, \tilde{\Delta}) = \mathcal{L}_k \exp\left(-\frac{(\tilde{\Delta} - \Delta_k)^2}{2\sigma_k^2}\right).$$

[21] https://github.com/dfm/tinygp
[22] https://github.com/google/jax

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

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

with

$$\frac{1}{\sigma^2} = \sum_{i \neq k}^{K} \frac{1}{\sigma_i^2} \quad \text{and} \quad \Delta = \sigma^2 \sum_{i \neq k}^{K} \frac{\Delta_i}{\sigma_i^2}. \tag{B2}$$

Hence

$$p(\tilde{\Delta}|\Delta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\tilde{\Delta} - \Delta)^2}{2\sigma^2}\right).$$

We can now rewrite Equation (B1) as

$$p(\boldsymbol{f}|T_k, D, \Delta) = \frac{\mathcal{L}_k}{\sqrt{2\pi\sigma^2}} \int \exp\left(-\frac{(\tilde{\Delta} - \Delta_k)^2}{2\sigma_k^2}\right) \exp\left(-\frac{(\tilde{\Delta} - \Delta)^2}{2\sigma^2}\right) d\tilde{\Delta}.$$

The integral in this equation is a product of Gaussian integrals that can be obtained analytically, leading to

$$p(\boldsymbol{f}|T_k, D, \Delta) = \mathcal{L}_k \sqrt{\frac{\sigma_k^2}{\sigma^2 + \sigma_k^2}} \exp\left(-\frac{1}{2}\frac{(\Delta_k - \Delta)^2}{\sigma_k^2 + \sigma^2}\right).$$

Finally,

$$\ln p(\boldsymbol{f}|P, T_0, D, \Delta) = \sum_{k}^{K} \ln \mathcal{L}_k - \frac{1}{2} \sum_{k}^{K} \times \left(\ln(\sigma_k^2) - \ln(\sigma^2 + \sigma_k^2) + \frac{(\Delta_k - \Delta)^2}{\sigma_k^2 + \sigma^2}\right), \tag{B3}$$

the log-likelihood of the 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 (B2) is computed, we adopt the biased estimates

$$\frac{1}{\sigma^2} = \sum_{k}^{K} \frac{1}{\sigma_i^2} \quad \text{and} \quad \Delta = \sigma^2 \sum_{k}^{K} \frac{\Delta_i}{\sigma_i^2}, \tag{B4}$$

so Δ and σ are independent of k in the last sum of Equation (8).

ORCID iDs

Lionel J. Garcia https://orcid.org/0000-0002-4296-2246
Suzanne Aigrain https://orcid.org/0000-0003-1453-0574
Dax L. Feliz https://orcid.org/0000-0002-2457-7889
Francisco J. Pozuelos https://orcid.org/0000-0003-1572-7707

References

Aigrain, S., & Foreman-Mackey, D. 2023, ARA&A, 61, 329
Aigrain, S., & Irwin, M. 2004, MNRAS, 350, 331
Aigrain, S., Parviainen, H., & Pope, B. J. S. 2016, MNRAS, 459, 2408
Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Berta, Z. K., Irwin, J., Charbonneau, D., Burke, C. J., & Falco, E. E. 2012, AJ, 144, 145
Blondel, M., Berthet, Q., Cuturi, M., et al. 2021, arXiv:2105.15183
Bradbury, J., Frostig, R., Hawkins, P., et al. 2018, JAX: composable transformations of Python+NumPy programs, v0.3.13, http://github.com/google/jax
Caldwell, D. A., Tenenbaum, P., Twicken, J. D., et al. 2020, RNAAS, 4, 201
Canocchi, G., Malavolta, L., Pagano, I., et al. 2023, A&A, 672, A144
Clarice Yaptangco, D., Ballard, S., & Dittmann, J. 2024, arXiv:2402.00115
David, T. J., Petigura, E. A., Luger, R., et al. 2019, ApJL, 885, L12
Deming, D., Knutson, H., Kammer, J., et al. 2015, ApJ, 805, 132
Dransfield, G., Timmermans, M., Triaud, A. H. M. J., et al. 2024, MNRAS, 527, 35
Feliz, D. L., Plavchan, P., Bianco, S. N., et al. 2021, AJ, 161, 247
Fletcher, R. 1987, Practical Methods of Optimization: V. 1-2 (2nd ed.; Hoboken, NJ: Wiley)
Foreman-Mackey, D. 2018, RNAAS, 2, 31
Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, AJ, 154, 220
Foreman-Mackey, D., Montet, B. T., Hogg, D. W., et al. 2015, ApJ, 806, 215
Foreman-Mackey, D., Yu, W., Yadav, S., et al. 2024, dfm/tinygp: The tiniest of Gaussian Process libraries, v0.3.0, Zenodo, doi:10.5281/zenodo.10463641
Garcia, L., Foreman-Mackey, D., Murray, C., et al. 2024, nuance: Efficient detection of planets transiting active stars, v0.6.0, Zenodo, doi:10.5281/zenodo.10832081
Gillen, E., Briegal, J. T., Hodgkin, S. T., et al. 2020, MNRAS, 492, 1008
Gilliland, R. L., Chaplin, W. J., Dunham, E. W., et al. 2011, ApJS, 197, 6
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Hippke, M., David, T. J., Mulders, G. D., & Heller, R. 2019, AJ, 158, 143
Hippke, M., & Heller, R. 2019, A&A, 623, A39
Howell, S. B., Ciardi, D. R., Giampapa, M. S., et al. 2016, AJ, 151, 43
Hunter, J. D. 2007, CSE, 9, 90
Irwin, J., Charbonneau, D., Nutzman, P., & Falco, E. 2009, in IAU Symp. 253, Transiting Planets, ed. F. Pont, D. Sasselov, & M. J. Holman (Cambridge: Cambridge Univ. Press), 37
Jenkins, J. M., Chandrasekaran, H., McCauliff, S. D., et al. 2010, Proc. SPIE, 7740, 77400D
Kovács, G., Bakos, G., & Noyes, R. W. 2005, MNRAS, 356, 557
Kovács, G., Hartman, J. D., & Bakos, G. Á. 2016, A&A, 585, A57
Kovács, G., Zucker, S., & Mazeh, T. 2002, A&A, 391, 369
Lecavelier des Etangs, A., Cros, L., Hébrard, G., et al. 2022, NatSR, 12, 5855
Lienhard, F., Queloz, D., Gillon, M., et al. 2020, MNRAS, 497, 3790
Lightkurve Collaboration, Cardoso, J. V. d., Hedges, C., et al. 2018, Lightkurve: Kepler and TESS time series analysis in Python, Astrophysics Source Code Library, ascl:1812.013
Luger, R., Agol, E., Kruse, E., et al. 2016, AJ, 152, 100
Luger, R., Kruse, E., Foreman-Mackey, D., Agol, E., & Saunders, N. 2018, AJ, 156, 99
McKinney, W. 2010, in Proc. 9th Python in Science Conf., ed. S. van der Walt & J. Millman (Austin, TX: SciPy), 56
Ment, K., Irwin, J., Charbonneau, D., et al. 2021, AJ, 161, 23
Morris, B. M., Hebb, L., Davenport, J. R. A., Rohn, G., & Hawley, S. L. 2017, ApJ, 846, 99
Mosteller, F., & Tukey, J. W. 1977, Data Analysis and Regression. A Second Course in Statistics (Boston, MA: Addison-Wesley)
Murray, C. A., Delrez, L., Pedersen, P. P., et al. 2020, MNRAS, 495, 2446
Newton, E. R., Rampalli, R., Kraus, A. L., et al. 2022, AJ, 164, 115
pandas development team, T 2020, pandas-dev/pandas: Pandas, 2.2.2, Zenodo, doi:10.5281/zenodo.3509134
Pedersen, P. P., Murray, C. A., Queloz, D., et al. 2023, MNRAS, 518, 2661
Petrucci, R. P., Gómez Maqueo Chew, Y., Jofré, E., Segura, A., & Ferrero, L. V. 2024, MNRAS, 527, 8290
Pont, F., Zucker, S., & Queloz, D. 2006, MNRAS, 373, 231
Pozuelos, F. J., Suárez, J. C., de Elía, G. C., et al. 2020, A&A, 641, A23
Rackham, B. V., Apai, D., & Giampapa, M. S. 2018, ApJ, 853, 122
Ramsay, G., Doyle, J. G., & Doyle, L. 2020, MNRAS, 497, 2320
Rasmussen, C. E., & Williams, C. K. I. 2005, Gaussian Processes for Machine Learning, Adaptive Computation and Machine Learning Series (Cambridge, MA: MIT Press)
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, JATIS, 1, 014003
Sebastian, D., Gillon, M., Ducrot, E., et al. 2021, A&A, 645, A100
Simpson, E. R., Fetherolf, T., Kane, S. R., et al. 2023, AJ, 166, 72
Skumanich, A. 1972, ApJ, 171, 565
Tamuz, O., Mazeh, T., & Zucker, S. 2005, MNRAS, 356, 1466
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261