

# PIPS, an advanced platform for period detection in time series – I. Fourier-likelihood periodogram and application to RR Lyrae stars

Yukei S. Murakami<sup>1,2,3★</sup>, Connor Jennings,<sup>1</sup> Andrew M. Hoffman,<sup>1,2,4</sup> Arjun B. Savel,<sup>5</sup> James Sunseri,<sup>1,2</sup> Raphael Baer-Way,<sup>1</sup> Benjamin E. Stahl<sup>1</sup>, Ivan Altunin,<sup>1</sup> Nachiket Girish<sup>2</sup> and Alexei V. Filippenko<sup>1</sup>

<sup>1</sup>*Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA*

<sup>2</sup>*Department of Physics, University of California, Berkeley, CA 94720-7300, USA*

<sup>3</sup>*Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA*

<sup>4</sup>*Institute for Astronomy, University of Hawai'i, 2680 Woodlawn Drive, Honolulu, HI 96822, USA*

<sup>5</sup>*Department of Astronomy, University of Maryland, College Park, MD 20742, USA*

Accepted 2022 May 25. Received 2022 May 24; in original form 2021 July 29

## ABSTRACT

We describe the Period detection and Identification Pipeline Suite (PIPS) – a new, fast, and statistically robust platform for period detection and analysis of astrophysical time-series data. PIPS is an open-source Python package that provides various pre-implemented methods and a customizable framework for automated, robust period measurements with principled uncertainties and statistical significance calculations. In addition to detailing the general algorithm that underlies PIPS, this paper discusses one of PIPS' central and novel features, the Fourier-likelihood periodogram, and compares its performance to existing methods. The resulting improved performance implies that one can construct deeper, larger, and more reliable sets of derived properties from various observations, including all-sky surveys. We present a comprehensive validation of PIPS against artificially generated data, which demonstrates the reliable performance of our algorithm for a class of periodic variable stars (RR Lyrae stars).

**Key words:** methods: data analysis – methods: statistical – stars: variables.

## 1 INTRODUCTION

Photometric time-series data are highly effective in allowing astronomers to examine the behaviour of distant objects. By studying the pattern, duration, and degree of change in the brightness of such objects as they vary over time, deep insights can be made regarding the astrophysical engines that drive these variations. For instance, gravitational microlensing can be used to study black holes or other significant gravitational sources (Paczynski 1986), and the explosions of stars (i.e. supernovae) afford insights into the intense conditions of stellar interiors (e.g. Branch & Wheeler 2017, and references therein). A particularly interesting class of such astrophysical objects is periodic variable stars, which exhibit large, periodic changes in their luminosity. Most commonly discovered are pulsating variable stars (see e.g. Percy 2007, for a review). For such objects, a host of properties, including metallicity, mass, luminosity, and size are encoded in their brightness variations (e.g. Leavitt & Pickering 1912; Christy 1968; Sandage & Tammann 1968; Jurcsik & Kovacs 1996). It is, therefore, valuable to determine from a light curve both the pulsation period and ‘shape’ of brightness variations as a function of time.

The most straightforward method for performing such analyses is to conduct a high-cadence observing campaign that covers the entire

pulsation period. The period can then be calculated by measuring the consecutive times (i.e. epochs) of light-curve extrema. This method can be found in the literature dating back several centuries, including Goodricke (1786). However, such high-cadence observations require a great deal of telescope time and are rarely feasible for variable stars.

In many instances, ranging from dedicated observations to modern all-sky surveys (e.g. Filippenko & Simon 1981; Holl et al. 2018, respectively), the data are often sparse and unevenly sampled, obscuring the features of the light curve in the raw time-series data (Lomb 1976). These data also often lack consistent observations of maxima, making it difficult to determine the corresponding period using the simple (high-cadence) method. This data sparsity necessitates a more involved approach for deriving the period that utilizes robust mathematical tools to recover the continuous shape and features of the light curve by phase-folding the data (Budding & Demircan 2007).

Many methods have been proposed to perform this task. For observations at a fixed cadence, the discrete Fourier transform (DFT) provides the spectral power in frequency space from which one can estimate the period. Most observations, however, are not obtained with uniform temporal sampling, and thus more advanced tools are required; for example, the method proposed by Kurtz (1985). A practical implementation of least-squares spectral analysis was proposed by Lomb (1976) and Scargle (1982), and it has since been widely accepted as a standard method (the Lomb–Scargle periodogram; hereafter LS periodogram) for finding periodic

★ E-mail: [sterling.astro@berkeley.edu](mailto:sterling.astro@berkeley.edu)

signals. For more specific cases where the target light curve is well known, the method of template fitting [e.g. template Fourier fitting (TFF) described by Kovács & Kupi 2007], which finds the period that minimizes the deviation of resulting phase-folded light curves from the known shape, can also be effective. Alternatively, template-independent approaches have also been proposed, such as phase dispersion minimization (PDM; Stellingwerf 1978). While this method has an obvious advantage over the methods introduced above in that it can avoid the possible bias due to the choice of template, it has been shown that PDM is not as sensitive to the signal as the least-squares method owing to the embedded parametrization in phase binning (Schwarzenberg-Czerny 1999). A variety of signal-finding tools implement the methods above, with unique improvements to enhance the reliability and reduce computational costs (e.g. SigSpec, Reegen 2007; PERIOD04, Lenz & Breger 2005).

Unfortunately, least-squares spectral analysis and template fitting are only capable of inferring the period of a light curve, and optimization or uncertainty estimation of the period and best-fitting analytic light curve require separate, additional analysis techniques, such as Markov Chain Monte Carlo (MCMC) or linear/non-linear regression (e.g. Eyer et al. 2017). While these methods can provide reliable and accurate results, they are also extremely sensitive to initial conditions and the range of parameters explored.

Moreover, the method employed to determine the statistical significance of detected signals is not well-established. Thus, improving this aspect of the analysis – i.e. automatically and statistically informing the selection of detected signals – is especially important in the current era of large all-sky surveys.

In this work, we present an all-in-one, fast, and statistically robust period-determination approach. Our method utilizes statistical information that has been embedded in the periodogram but not previously used to simultaneously (i) determine the period, (ii) estimate the uncertainty of the period detection, and (iii) compute the statistical significance of the detected signal without additional analyses. This method, along with other features, is implemented in the publicly available Python package PIPS (Period Detection and Identification Pipeline Suite).<sup>1</sup> We specifically focus herein on its most basic and flexible method to perform period detection with a Fourier-likelihood (FL) periodogram, and we also showcase its application to the study of RR Lyrae stars. With pre-optimized parameters that are further customizable, PIPS determines the period and analyses the shape of a light curve simultaneously, minimizing template bias and improving overall accuracy and reliability.

We begin our discussion in Section 2 by reviewing the concepts of the  $\chi^2$  periodogram and several associated issues that necessitate additional analyses. We then propose a solution to said issues in Section 3, and follow-up by describing the design and algorithm of our Python implementation, PIPS (Section 4). Section 5 serves to evaluate the performance of PIPS using simulated data. In addition, we discuss, in Section 6, a variety of additional features in PIPS, such as classification and long-term modulation analysis. Finally, we visit a variety of individual topics related to our work, as well as possible areas of further applications and developments, in Section 7.

<sup>1</sup>PIPS is available at <https://github.com/SterlingYM/PIPS>. The basic usage is described in Appendix A.

## 2 CHI-SQUARE PERIODOGRAM

### 2.1 Concept

As we introduce our improvements to period detection with our periodogram process, a solid understanding of the traditional and most commonly used  $\chi^2$  interpretation of the periodogram is essential. In this  $\chi^2$  framework, the signal strength in period space is represented by  $\chi^2$  values normalized to the best-fitting curve fitted at each period. A visualization of this concept is provided in Fig. 1. As introduced in Section 1, the most standard form of  $\chi^2$  periodogram is the LS (Lomb 1976; Scargle 1982), for which a sinusoidal signal is fitted to the data. This method is a direct application of the DFT for unevenly sampled data, and it has numerous statistical benefits [e.g. calculating false-alarm probability (FAP); for a review, see VanderPlas (2018)].

More generally, the  $\chi^2$  periodogram is created by fitting a selected model function  $\mathcal{F}$  to the phase-folded data,

$$y_{\text{fit}}(x; P_{\text{test}}) = \mathcal{F}(\phi; P_{\text{test}}, \theta_1, \theta_2, \dots, \theta_n). \quad (1)$$

Here,  $\phi \equiv (x \bmod P_{\text{test}})$  is the time-series data phase-folded at the test period,  $P_{\text{test}}$ . The variables  $\theta_1, \theta_2, \dots, \theta_n$ , are free parameters determined via fitting.

This best-fitting function serves as a ‘template’ for the measurement of the spectral energy density (SED). The goodness-of-fit can be evaluated with the reduced  $\chi^2$  value:

$$\sum \chi_v^2(P_{\text{test}}) = \frac{1}{\nu} \sum_j \left( \frac{y_j - y_{\text{fit}}(x_j, P_{\text{test}})}{\sigma_{y_j}} \right)^2, \quad (2)$$

where  $\nu = N - n_\theta$  is the number of degrees of freedom and the subscript  $j$  denotes the data point. In studies such as those of Bretthorst (1988) or Palmer (2009), the list of such  $\chi^2$  values across the search range is used to construct the periodogram. VanderPlas (2018) defines the general case of a normalized periodogram as

$$\mathcal{P} = 1 - \frac{\chi^2}{\chi_0^2}, \quad (3)$$

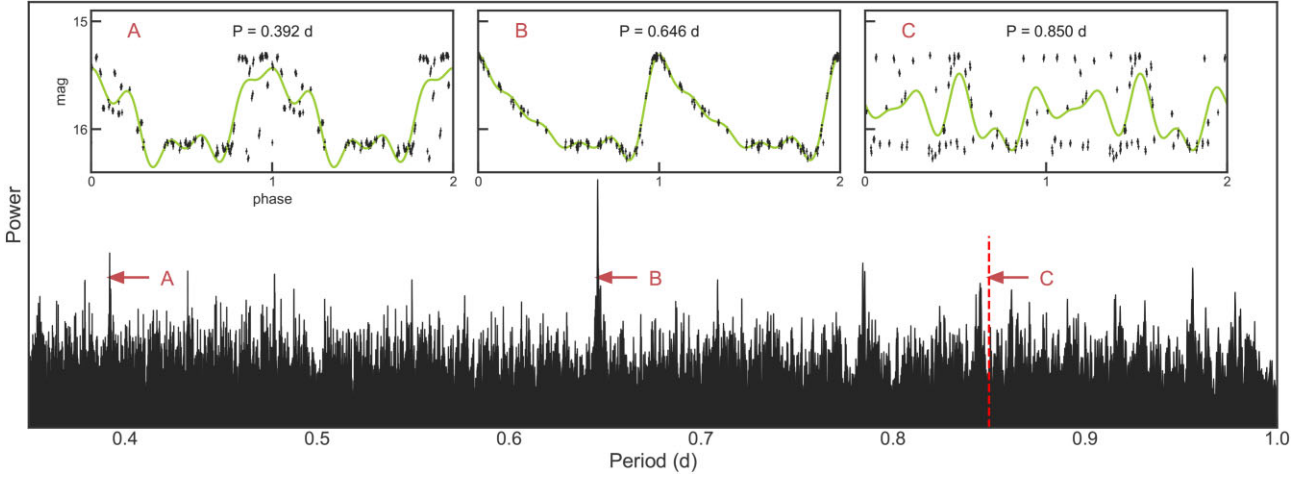
where the  $\chi^2$  values are normalized with the non-varying  $\chi^2$  value  $\chi_0^2 = \sum_i (y_i - \bar{y})^2 / \sigma_{y,i}^2$ , which represents the maximum possible  $\chi^2$  value for the data. In this formulation of the periodogram, the normalization  $1/\nu$  is cancelled.

### 2.2 Challenges and current solutions

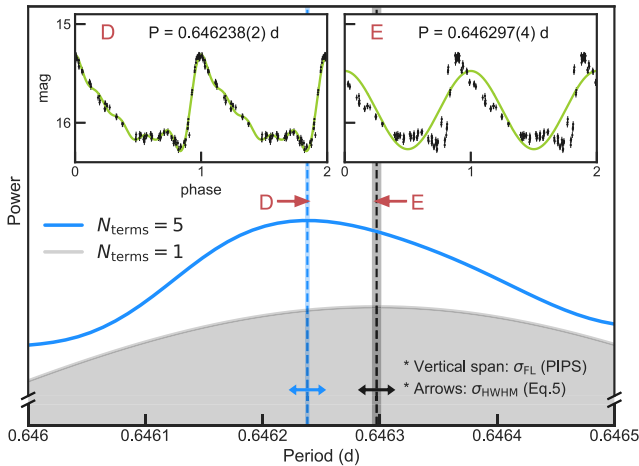
In applications to modern astronomy, the  $\chi^2$  periodogram poses three challenges to users – namely template bias, uncertainty estimation, and significance estimation. Each of these has been discussed in the literature, and several solutions have been proposed.

#### 2.2.1 Template bias

Despite its popularity, results from an LS periodogram are known to deviate from true periods (e.g. Schwarzenberg-Czerny 1989), mainly owing to the fact that LS is a correlation function to a sinusoid. Its detected period can therefore deviate from the true value when the signal is non-sinusoidal. Similarly, other template-fitting methods (e.g. TFF) only guarantee the best match of the data to a template, resulting in performance being dependent upon the completeness of the template library. In general, this bias occurs when the light-curve model used to construct the periodogram ( $\mathcal{F}$ ) differs from the model one wishes to use for the final analysis. For pulsating variable stars, a multiterm Fourier series is most commonly used for



**Figure 1.** A visual demonstration of the  $\chi^2$  periodogram in which the template function is fitted to phase-folded data (i.e. V-band observations of the RRab star dubbed V008 by Hoffman et al. 2021) at each period and the resulting goodness-of-fit measured by  $\chi^2$  is translated into the strength spectrum (black). Inset A shows the secondary peak (false signal) in the periodogram and the phase-folded light curve at that period. The data show some structure, but the dispersion is still large around the best-fitting curve. Inset B demonstrates the primary peak (true signal), which is in agreement with the value in the literature derived by various independent methods (see Hoffman et al. 2021, for a detailed comparison). Inset C shows a randomly selected location (noise) in the periodogram. The phase-folded data at this period lack any structure and are broadly dispersed.

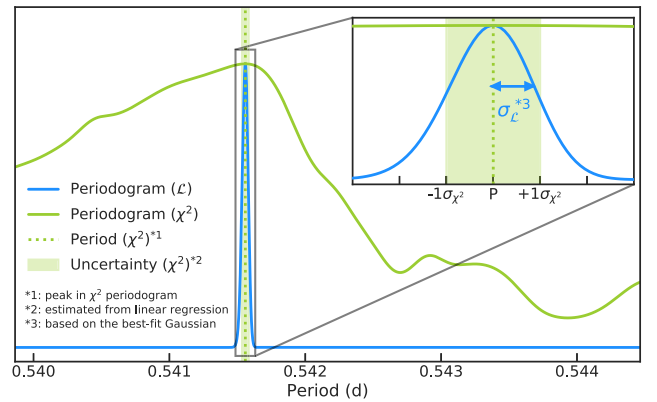


**Figure 2.** A visual demonstration of potential issues with the  $\chi^2$  periodogram. Two periodograms (blue and grey) generated with different templates (single- and five-term Fourier series) are shown. The difference of peak locations and heights (D and E) illustrate the template bias. The difference in sizes between the half width at half-maximum (HWHM)-based uncertainty (equation 5) and PIPS uncertainty (which is approximately equivalent to MCMC-based and linear regression-based uncertainty values) illustrates the difficulty of directly estimating uncertainty from the periodogram itself.

the final (astrophysical) analysis, and using the same model for the periodogram can solve this issue. This effect is visually demonstrated in Fig. 2.

### 2.2.2 Uncertainty

In addition to the issue with template bias, we note that in many publications, period-uncertainty quantification is not treated robustly – nor, in some cases, reported at all. This surprising omission is dismayingly common in both older studies which use the traditional method to calculate the period (e.g. Bailey 1902; Wemple 1932)



**Figure 3.** Likelihood periodogram (blue curve) and  $\chi^2$  periodogram (light-green curve), normalized to equal height at the peak. Unlike the traditional  $\chi^2$  periodogram, the likelihood periodogram can be directly interpreted as a probability density after normalization, and the width of the peak accurately represents the uncertainty of the detected period.

and more recent studies with periodogram-based determinations (e.g. Siegel et al. 2015). A clear understanding and treatment of uncertainties is of great importance for modern studies, as long-term data can be used to study the evolution of stars (e.g. Lee 1991). In the current state of affairs (with numerous measurements that lack uncertainty estimates), it is difficult to assess the quality of conclusions that rely on such historical data points. Clear understanding of the uncertainty in period is also essential for studying other parameters that covary with the period in the fitting process or that are affected by a propagated uncertainty from the period. A demonstration of this issue and improvements that result from using PIPS are presented in Section 6.4.

Statistically speaking, the uncertainty in frequency space is often defined by the width of the periodogram peak, or more specifically, the HWHM height. This value is a function of the span of the data, and proper weight must be added considering the quality and size of

the data according to Gregory (2001):

$$\sigma_f \approx f_{1/2} \sqrt{\frac{2}{N(\text{SNR})^2}}. \quad (4)$$

Hence,

$$\sigma_P = \left| \frac{-(1/f_{\text{peak}})\sigma_f}{f_{\text{peak}}} \right| \approx \frac{f_{1/2}}{f_{\text{peak}}^2} \sqrt{\frac{2}{N \cdot \text{rms}[(y_j - \bar{y})/\sigma_j]^2}}. \quad (5)$$

Although this formula is reasonable considering the effect of period folding,<sup>2</sup> observations (e.g. Fig. 2) clearly show that this method does not provide an accurate view of the uncertainty. VanderPlas (2018) also suggests that this method should be avoided and introduces various methods to calculate FAP, with a conclusion that there is no universal method to directly estimate uncertainties from the periodogram.

Instead, as the second step after the peak detection in a periodogram, one can perform a full-parameter optimization. This process treats the period as a free parameter, rather than as a fixed value at the test location. The goal is not only to determine the period, but to also measure its uncertainty. This full-parameter fitting can be achieved by either linear regression or more computationally expensive Bayesian methods, such as a MCMC exercise. For the linear regression method, one can estimate the size of the uncertainty from the covariance matrix  $\Sigma_{ij}$  provided by linear regression methods as  $\sigma_P \approx \sqrt{\Sigma_{P,P}}$ . Similarly, the uncertainties of other parameters can be approximated by taking the square root of each element in  $\Sigma_{ii}$ , where  $i$  denotes the index of each parameter. This full-parameter optimization is commonly employed in the literature when the uncertainty values are reported (e.g. Eyer et al. 2017), but the computational cost becomes increasingly high as the data grow.

### 2.2.3 Significance

The methods described in Section 2.1 and Section 2.2.2 to detect the period and estimate its uncertainty do not distinguish between physical signals and false positives in the noise. A method that provides a statistical measure of the proposed signal's reliability is required, and PIPS thus provides signal-detection efficiency (SDE) as a statistic.

What qualifies as a 'significant' SDE is not generally derived empirically. For instance, a common significance threshold is the SDE below which 99 percent of false-positive signals fall (e.g. Hippke & Heller 2019a). Hence, an SDE threshold is in principle algorithm-specific.

Following, for example, Kovács, Zucker & Mazeh (2002) and Hippke & Heller (2019a), SDE is evaluated to measure the relative strength of a signal with respect to the background noise level:

$$\text{SDE}(P) = \frac{\text{SR}(P) - \langle \text{SR}(P) \rangle}{\sigma_{\text{SR}}}, \quad (6)$$

where SR (signal residual) is the periodogram power in equation (3) normalized to 1 [i.e.  $\text{SR}(P) = \mathcal{P}(P)/\mathcal{P}_{\text{max}}$ ]. The values of  $\langle \text{SR}(P) \rangle$  and  $\sigma_{\text{SR}}$  are the mean and standard deviation of the SR evaluated over the period search range  $P_{\text{min}} < P < P_{\text{max}}$ , respectively. As pointed out by Kovács et al. (2002), finite evaluation range may bias the SDE, and

<sup>2</sup>The deviation due to the incorrect period is multiplied by  $M$  if the data set spans over  $M$  pulsation cycles. One may thus assume the uncertainty should become linearly smaller as the baseline of the data becomes longer and  $M$  becomes larger.

it is therefore suggested that one use the SDE under the same set of parameters (e.g. the same period search range) and only use the SDE as a relative statistic for each data point in a data set. A larger SDE value indicates a larger statistical significance, and a smaller value is often associated with noise, false positives for non-varying objects, or astrophysical false positives fitted with an incorrect template.

## 3 A NEW APPROACH: LIKELIHOOD PERIODOGRAM

In Sections 2.2.2 and 2.2.3, we discussed important issues with the traditional  $\chi^2$  representation of the periodogram, namely that its spectral density does not directly translate into scientifically important quantities such as signal amplitude, uncertainty, or statistical significance. We discussed the existing solutions to these issues, such as more generalized or specialized templates, independent Monte Carlo sampling in the light-curve model's parameter space near the peak of the periodogram, and evaluation of SDE. These solutions are the current standard for scientific analysis, but have their limitations due to the computational cost and lack of statistical understandings. To address the existing issues and expand the possibility set of transient analyses in the era of large, data-rich all-sky surveys, we propose a new approach, the *Likelihood periodogram*.

Likelihood periodogram is a new representation of the  $\chi^2$  periodogram, and is achieved by converting the  $\chi^2$  values to Gaussian likelihoods for each data point, and then taking the product of the likelihoods across all data points. The likelihood is thus defined as

$$\mathcal{L}(P_{\text{test}} | x, y, \sigma_y) = \prod_i^{N_{\text{data}}} \frac{1}{\sigma_i \sqrt{2\pi}} \exp \left[ -\frac{(y_i - y_{i,\text{fit}})^2}{2\sigma_i^2} \right] \quad (7)$$

$$\propto \exp \left( -\frac{1}{2} \sum_i \chi_i^2 \right). \quad (8)$$

This form, inspired from Bayesian statistics, comes with great benefits thanks to its own statistical properties. We discuss these benefits below, demonstrating how this new approach can address the issues with the current standard.

### 3.1 Concept

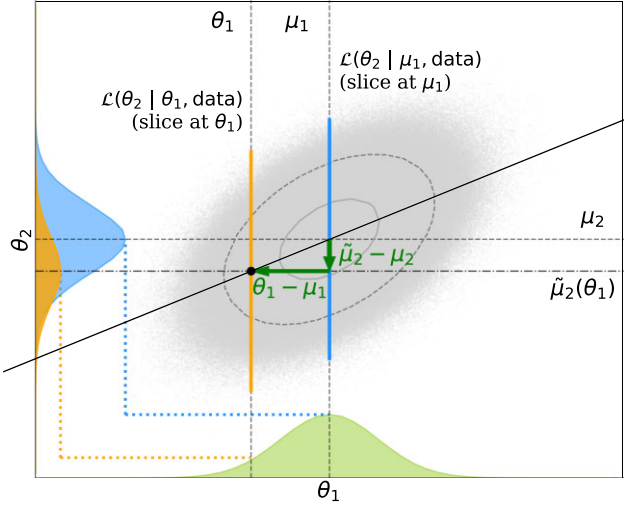
In model fitting – which we often employ in studies of RR Lyrae stars, Cepheids, eclipsing binaries, or exoplanet transits – data quality, model parametrization, and robust sampling algorithms create near-Gaussian parameter posterior distributions. This is often assumed, and thus the uncertainty is commonly reported using one standard deviation ( $1\sigma$ ) of the Gaussian posterior probability density function (PDF).<sup>3</sup> This is an important but often-overlooked fact in the period analysis of variable objects in astronomy; it will enable us to expand our understanding of the periodogram significantly.

For the multivariate posterior PDF, the likelihood that a parameter  $\theta_i$  (e.g. period) is consistent with the observed data ( $x$ ,  $y$ , and  $\sigma_y$ ) is calculated (often from MCMC samples) by marginalizing the posterior PDF over all other parameters. For the bivariate (two-parameter) case, for instance, we obtain the likelihood of  $\theta_1$  by integrating over  $\theta_2$ :

$$\mathcal{L}(\theta_1 | x, y, \sigma_y) = \int \mathcal{L}(\theta_1, \theta_2 | x, y, \sigma_y) d\theta_2. \quad (9)$$

<sup>3</sup>One can tell that this is a common practice by the fact that only one value of uncertainty is generally reported in the literature.





**Figure 4.** A visualisation of equation (10). The marginalized distribution of the parameter of interest  $\theta_1$  (green) for the given two-dimensional (2D) posterior samples (grey dots and contour) can be calculated without intense sampling if one can find the optimal value of  $\theta_2$  ( $\equiv \tilde{\mu}_2$ , the centre of distribution for each slice) at each  $\theta_1$ . The height of normal distributions (blue and orange curves, centred at  $\mu_2$  and  $\tilde{\mu}_2$ , respectively) is proportional to the marginalized likelihood for  $\theta_1$ .

When the posterior PDF  $\mathcal{L}$  is exactly or well approximated by a multivariate Gaussian, the likelihood of  $\theta_2$  at fixed  $\theta_1$  can be written as

$$\mathcal{L}(\theta_2 | \theta_1, x, y, \sigma_y) \propto \alpha(\theta_1) \exp \left[ -\frac{(\theta_2 - \tilde{\mu}_2)^2}{2\beta^2} \right], \quad (10)$$

where the local mean  $\tilde{\mu}_2$  is the best-fitting value of  $\theta_2$  at fixed  $\theta_1$  (see Fig. 4). The height of the Gaussian  $\alpha$  is a function of  $\theta_1$ , and  $\beta$ , the standard deviation of  $\theta_2$  PDF sliced at fixed  $\theta_1$  is independent from  $\theta_1$  or  $\theta_2$ . A more detailed derivation of equation (10) is provided in Appendix B.

In other words, the shape and the relative size of the marginalized likelihood  $\mathcal{L}(\theta_1 | x, y, \sigma_y)$  can be obtained by the value of  $\alpha$  (arbitrarily normalized likelihood) at  $(\theta_1, \theta_2) = (\theta_1, \theta_{2,\text{fit}})$ . This feature is beneficial because the value of  $\theta_{2,\text{fit}}$  can be obtained by the maximum-likelihood fitting in the  $\theta_2$  space at each  $\theta_1$ . Generalizing this concept to the  $N$ -parameter case, we obtain

$$\mathcal{L}(\theta_1 | x, y, \sigma_y) \propto \mathcal{L}(x, y, \sigma_y | \theta_1, \theta_{2,\text{fit}}, \theta_{3,\text{fit}}, \dots), \quad (11)$$

and therefore

$$\log \mathcal{L}(\theta_1 | x, y, \sigma_y) = -\frac{1}{2} \sum_i \chi_i^2 + \gamma \quad (12)$$

with a constant offset  $\gamma$ .

Our formulation suggests that the periodogram provides all the statistical tools we need to assess the posterior PDF. This added statistical value in the periodogram now allows us to address the existing issues directly in more informed and far more cost-effective ways.

### 3.2 Uncertainty

In Fig. 3, we present an example of the likelihood periodogram, normalized to the same height as the traditional  $\chi^2$  periodogram. From the comparison between the estimated uncertainty (light-green fill) and the size of the standard deviation in the likelihood

periodogram's Gaussian shape (blue arrows), it is evident that the likelihood periodogram accurately portrays the (relative) statistical significance.

This example and the equations in Section 3.1 suggest, more importantly, that the statistical significance of the peak can be directly calculated from the likelihood periodogram. This is a strong advantage of the likelihood periodogram over the  $\chi^2$  periodogram. Computing the likelihood periodogram is, therefore, highly impactful in a time-series analysis, especially when one aims to detect a new candidate for previously unknown variable stars.

The significantly narrower (and hence more accurate) peak causes a problem with computation. As we discussed in Section 4.1, the sampling width has to be smaller than the width of the peak, and a significantly narrower peak requires a large number of calculations. We address this issue by carefully spacing our samples in the periodogram as described in Section 4.1. When the likelihood periodogram is enabled in PIPS, it (by default) uses the period and uncertainty detected using the original  $\chi^2$  periodogram to pinpoint the location to sample the likelihood periodogram. PIPS determines the sampling resolution and range using the fact that the size of the estimated uncertainty is comparable to the half width of the peak.

### 3.3 Statistical significance

Once the likelihood periodogram is calculated, the most probable period and its uncertainty can be directly calculated by fitting a Gaussian function to the periodogram. A statistical test can then be performed to quantify the significance of the detected peak. For instance, a likelihood-ratio (LR) test (e.g. Buse 1982) can be performed, with the LR defined as

$$\text{LR}_{\text{peak}} = \ln \left[ \frac{\mathcal{L}(P_{\text{peak}})}{\langle \mathcal{L}_{\text{noise}}(P) \rangle} \right] = \ln [\mathcal{L}(P_{\text{peak}})] - \ln [\langle \mathcal{L}_{\text{noise}}(P) \rangle], \quad (13)$$

where  $\mathcal{L}(P_{\text{peak}})$  is the likelihood value at the peak, and  $\langle \mathcal{L}_{\text{noise}}(P) \rangle$  is the mean of likelihood values at the background (i.e. noise). Calculating the noise likelihood can be accomplished by bootstrapping: Shuffling the magnitudes while keeping the uncertainty values attached to each data point and randomly sampling the likelihood values across the period search-range serves as a simulation of the background noise. Comparison of this  $\text{LR}_{\text{peak}}$  value to the noise level  $\sigma_{\text{LR, noise}}$  quantifies the significance of the detected signal. The statistical significance, represented by the Z-score, is thus

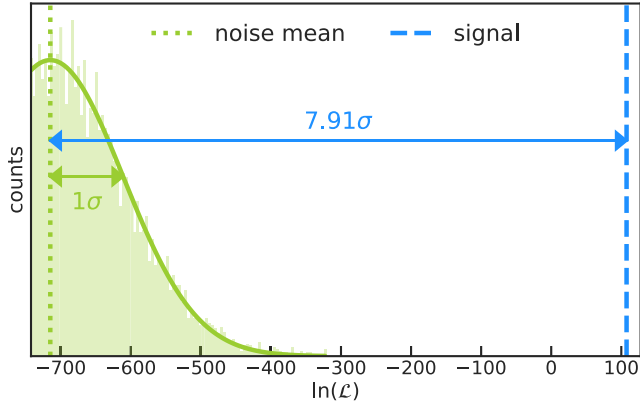
$$Z_{\text{peak}} = \frac{\text{LR}_{\text{peak}}}{\sigma_{\text{LR, noise}}}. \quad (14)$$

Fig. 5 shows the significance analysis for the peak value in Fig. 3. The Z-score is 7.91 (i.e. a  $7.91\sigma$  detection), or  $p < 0.00001$ ,<sup>4</sup> and the detected peak is said to be statistically significant.

## 4 PIPS: ALGORITHM AND IMPLEMENTATION

We implement the above methods for period detection in PIPS – an advanced, open-source Python platform for the period analysis of astronomical data. PIPS provides an all-in-one workflow for automated period detection, which we rigorously test in the next section. In addition, PIPS gives highly optimizable and customizable tools for advanced users, thereby enabling tightly-controlled analyses for specific needs. In this section, we discuss the methods for automated

<sup>4</sup>Here, we are referring to the statistical  $p$ -value, not to be confused with the exponent in the SGM in equation (20).



**Figure 5.** An example significance test in which the log-likelihood value of the detected signal (blue dashed line) is compared to the background noise (light green). The noise values are simulated by shuffling the data and measuring the log-likelihood at a random location (i.e. bootstrapping). The separation of the signal from mean noise is significantly (7.9 times) larger than the dispersion of the noise, indicating that the detected signal is statistically significant.

period detection implemented in PIPS, a number of light-curve templates that we implement by default, and additional features related to period detection. A selection of additional features for further analysis, such as stellar-parameter estimations or long-term period-change analysis, are discussed in later sections.

#### 4.1 Peak selection and refinement

Detecting the period of time-series data requires a robust algorithm to select a peak in the periodogram and its location as precisely as possible. As commonly known in many optimization problems (for review, see Kochenderfer & Wheeler 2019), sampling size in  $\chi^2$  space can have a significant impact on the results – sampling that is too coarse will result in a missed peak spike, while increasing the sampling size directly impacts the computational cost. Our peak-selection algorithm provides a good compromise between these two, successfully detecting the period for most cases without excessively increasing the computing cost.

PIPS initially generates the periodogram<sup>5</sup> within a specified period-search range, which is sampled densely enough so that the spacing of the sampling grid is smaller than the expected peak width. The expected peak width is calculated following the discussion by VanderPlas (2018). Peak size in frequency space can be estimated using the baseline length  $T$ :  $\Delta f = 1/T$ . Defining both edges of the width at peak  $f$  by  $f_1 = f - \Delta f/2$  and  $f_2 = f + \Delta f/2$ , the conversion between period and frequency  $p = 1/f$  gives the expected width size in period space:

$$\Delta p = p_1 - p_2 = \frac{1}{f_1} - \frac{1}{f_2} = \frac{f_2 - f_1}{f_1 f_2} = \frac{\Delta f}{f^2 - (\Delta f)^2/4}. \quad (15)$$

Combining this result with  $\Delta f = 1/T$ , we obtain

$$\Delta p = \frac{1/T}{(1/p)^2 - (1/4T^2)} = \frac{p^2 T}{T^2 - p^2/4}. \quad (16)$$

<sup>5</sup>Fitting of light-curve models is performed by a manually implemented linear problem solver or by linear regression with the `curve_fit` function in SciPy (Virtanen et al. 2020).

Oversampling with a grid size smaller than this peak width – by a factor of  $N_0$  – will ensure that the peak is detected, so PIPS uses this as the grid size,<sup>6</sup>

$$(\Delta p)_{\text{grid}} = \frac{\Delta p}{N_0} = \frac{p^2 T}{N_0(T^2 - p^2/4)}, \quad (17)$$

and the resolution of periodogram  $R_{\text{all}}$  within the specified search range between  $p_{\text{min}}$  and  $p_{\text{max}}$  becomes

$$R_{\text{all}} = \frac{p_{\text{max}} - p_{\text{min}}}{(\Delta p)_{\text{grid}}} = \frac{N_0(p_{\text{max}} - p_{\text{min}})(T^2 - p^2/4)}{p_{\text{min}}^2 T}. \quad (18)$$

Once the periodogram with resolution  $R_{\text{all}}$  is generated, the top  $N_{\text{peak}}$  independent peaks are selected from the initial periodogram. The independence of these peaks is checked by ensuring a distance of at least  $(\Delta p)_{\text{min}}$  between peaks. The periodogram only around those peaks is then densely sampled (with a sample resolution  $R_{\text{peak}} \gg N_0$  within  $\Delta p_{\text{peak}}$  at each peak). Usually this ensures a fine sampling around the true peak where  $(\Delta p)_{\text{grid}} \ll \sigma_p$  (for the estimation of  $\sigma_p$ , see Section 2.2.2), and when this condition is not satisfied, PIPS raises a warning to suggest an increased peak resolution,  $R_{\text{peak}}$ .

#### 4.2 Light-curve models

We designed PIPS as a general platform for period detection with customizable models. That is, while we implement a few fundamental generative functions, any template for light curves can be used. For instance, for objects whose light curves have well-known analytic forms, using the true physical model will minimize template bias – assuming that any long-term instrumental trends and outliers have already been removed.<sup>7</sup> Details for using custom functions can be found in the online documentation.<sup>8</sup> In this section, we describe the three functions we choose as pre-implemented models for general-purpose use.

##### 4.2.1 Fourier series

By default and following the convention in the literature (e.g. Holl et al. 2018), PIPS uses a multiterm Fourier series of the form

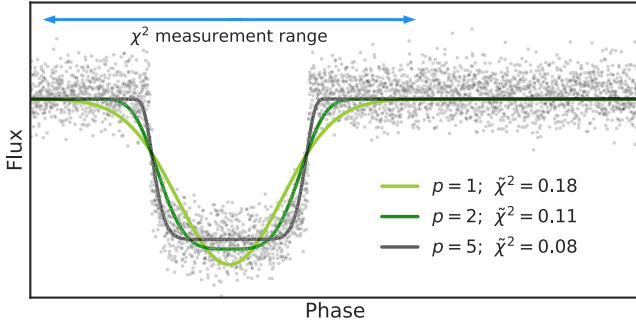
$$y_{\text{fit}}(x; P_{\text{test}}) = A_0 + \sum_{k=1}^{K_{\text{max}}} \left[ a_k \cos\left(\frac{2\pi k}{P_{\text{test}}} x\right) + b_k \sin\left(\frac{2\pi k}{P_{\text{test}}} x\right) \right], \quad (19)$$

to approximate light curves. Here,  $y_{\text{fit}}$  is the fitted magnitude at given time  $x$ . The constant offset  $A_0$  makes the fit independent of the distance modulus and/or the luminosity of the object under study, and the coefficients  $a_k$  and  $b_k$  control the amplitude of each  $k^{\text{th}}$  harmonic. These free parameters (i.e.  $A_0$ ,  $a_k$ , and  $b_k$ ) are optimized so that the phase-folded light curve at each test period  $P_{\text{test}}$  best matches the observed magnitudes. The series length,  $K_{\text{max}}$ , controls the number of harmonics in the light curve: A larger value of  $K_{\text{max}}$  results in a better approximation to sharper peaks and smaller features in a light

<sup>6</sup>We formulate this grid size in period space because it is less sensitive to the numerical error with finite grid size. In most cases, this can be approximated by adopting an equally spaced grid in frequency space and taking the reciprocal, but for long-period objects (i.e.  $f \ll 1$ ), this can cause a significant numerical error, and caution is needed.

<sup>7</sup>Outliers can be removed by user-defined criteria using a function in PIPS. Future versions of PIPS will include long-term detrending with `lightcurve`: <https://docs.lightcurve.org/>.

<sup>8</sup><https://pips.readthedocs.io/en/latest/>



**Figure 6.** Best-fitting super-Gaussian (SG) curves to artificially generated exoplanet transit data (grey dots). SG with  $p = 5$  produces a better fit to the data than  $p = 1$  (i.e. Gaussian). The  $\chi^2$  values are measured only near the transit phase (indicated by an arrow) to illustrate the difference between different choices of  $p$ .

curve, while making the fit more sensitive to the noise and possibly lowering the signal detection rate. In our first use of this algorithm with RR Lyrae stars by Hoffman et al. (2021), we set  $K_{\max} = 5$ , based on a cross-validation test using *Gaia* data.

#### 4.2.2 Super-Gaussian

For some types of periodically varying objects, such as transiting exoplanets or binary stars, the light curve can have a narrow, sharp feature, which cannot be well approximated by a Fourier series without massively increasing the number of free parameters (i.e. setting  $K_{\max}$  to a large value). To account for these types of features, we experimentally implement the Super-Gaussian model (SGM), which we formulate as

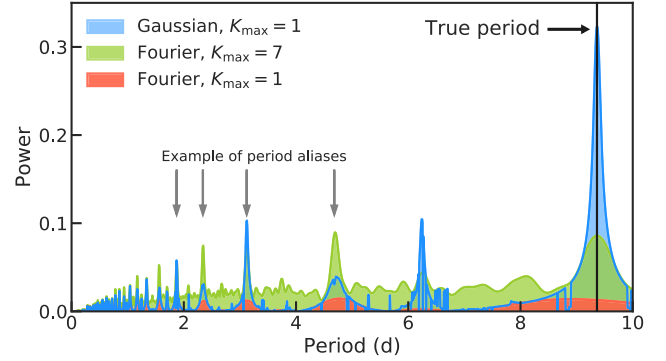
$$y_{\text{fit}}(x; P_{\text{test}}) = B_0 - B_1 \exp \left[ - \left( \frac{(\phi - \mu)^2}{2\sigma^2} \right)^p \right]. \quad (20)$$

In equation (20),  $\phi = (x \bmod P_{\text{test}})$  is the phase-folded time,  $B_0$  is the mean of the non-varying (i.e. flat) part of the light curve, and three parameters ( $B_1$ ,  $\mu$ , and  $\sigma$ ) respectively determine the height, location, and width of the feature. The power parameter,  $p$ , forces the original Gaussian function to a more square-like shape. The effect of changing the  $p$ -value is demonstrated in Fig. 6, in which the normalized  $\chi^2$  ( $\bar{\chi}^2 = \chi^2/\chi_0^2$ ) to the best-fitting curve around the transit well is calculated.

This model can be converted to a log-linear form, which is computationally less expensive,

$$\begin{aligned} \log(y_{\text{fit}} - B_0) &= \log(-B_1) - \frac{1}{(2\sigma^2)^p} (\phi - \mu)^{2p} \\ &= \theta_0 + \theta_1\phi + \theta_2\phi^2 + \theta_3\phi^3 + \dots + \theta_{2p}\phi^{2p}. \end{aligned} \quad (21)$$

This form has a computational advantage because the ‘exact’ solution can be calculated as a linear-algebra problem. It should be noted, however, that using linear algebra to find the solution *does not* guarantee that the  $\chi^2$  value is minimized (for a review, see e.g. Al-Nahhal et al. 2019). Moreover, this generalized form in equation (21) is only equivalent to the original form in equation (20) when all free parameters ( $\theta_0, \theta_1, \dots$ ) satisfy the corresponding function of the original parameters  $B_1, \mu$ , and  $\sigma$ . This means that the log-linear form in equation (21) produces a higher level of noise with less-sharp peaks, even though it is computationally much more efficient to produce the periodogram. Because of this issue, we use non-linear, iterative, least-squares fitting by default.



**Figure 7.** Benchmarking the effectiveness of different PIPS models in determining the period of a transiting exoplanet. The single-term Gaussian model outperforms the Fourier model both in better constraining the true period and in deweighting the statistical importance of its harmonics.

#### 4.2.3 Gaussian mixture

We also implement another well-known generative function: the Gaussian mixture model (GMM), which can take multiple narrow independent peaks (or dips). As a result, we expect this model to work better than Fourier models when the input light curve has a combination of wide and narrow features. The GMM can be expressed as

$$y_{\text{fit}}(x; P_{\text{test}}) = C_0 + \sum_{k=1}^{K_{\max}} c_k \exp \left( - \frac{(\phi - \mu_k)^2}{2\sigma_k^2} \right), \quad (22)$$

where the constant offset ( $C_0$ ), amplitude ( $c_k$ ), location ( $\mu_k$ ), and width ( $\sigma_k$ ) are fitted as free parameters for each  $k^{\text{th}}$  term up to the specified number of terms,  $K_{\max}$ . For convenience, we combine the SGM and GMM, and implement the Super-GMM in PIPS, as

$$y_{\text{fit}}(x; P_{\text{test}}) = C_0 + \sum_{k=1}^{K_{\max}} c_k \exp \left[ - \left( \frac{(\phi - \mu_k)^2}{2\sigma_k^2} \right)^p \right]. \quad (23)$$

Setting  $p = 1$  reverts equation (23) back to the GMM (equation 22), while taking  $K_{\max} = 1$  reverts it back to the SGM (equation 20).

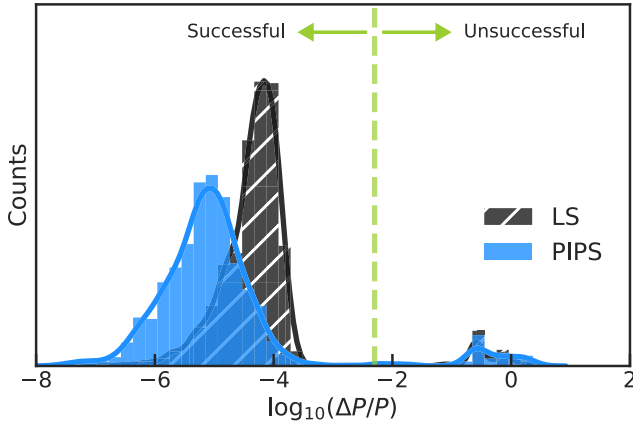
#### 4.2.4 The impact of model choice

As shown in Figs 2 and 6, light curves with discrepant morphologies are best fit by models that are appropriate for their specific morphologies. While a basic, two-term Fourier model would be able to provide a period estimate for any periodic data, the signal accuracy and strength will be severely diminished for a pathological choice of light-curve morphology (e.g. a very shallow exoplanet transit). Increasing the number of Fourier terms will likely provide a better fit to the data, but at the expense of computation time and an increased risk of overfitting. Hence, it is crucial to identify the optimal model choice for light-curve fitting; working with a subpar model will only provide incremental fitting improvement.

This intuition is made more evident in Fig. 7. Here, the advantage in choosing better models for a given data set can clearly be seen – a narrow periodogram peak near the true period, in addition to weaker aliasing at harmonics of the true period.

### 4.3 Multiperiodic analysis

Some objects, such as double-mode pulsators or transiting binaries with one or both being a variable star on its own, exhibit a light curve



**Figure 8.** The same plot as Fig. 4 or Hoffman et al. (2021) but in the relative deviation. The abscissa is in log base 10. Both LS and PIPS result in a gap between ‘successful’ and ‘unsuccessful’ estimates. A more detailed distribution of the two populations in various parameter spaces is visualized in Fig. C1.

that is a linear combination of two or more variations at different frequencies. Analysis of such objects requires decomposing each variation without losing the light-curve shape, and PIPS achieves this by iteratively performing period analysis and removing the dominant component at each detected period (pre-whitening; e.g. Blomme et al. 2011).

The result of this analysis can be represented as an amplitude spectrum. As described in Hoffman et al. (2021), this amplitude spectrum is a representation of consecutive pre-whitening and an improved and generalized version of the non-linear DFT method:

$$y = A_0 + \sum_i^{i_{\max}} A_i \mathcal{F}_i(x, P_i; \{\theta_j\}_i), \quad (24)$$

where  $\mathcal{F}_i$  is the normalized model function (see Section 4.2) and  $\{\theta_j\}_i$  is a set of best-fitting model parameters at each period. The period  $P_i$  at which the function is evaluated is determined by our main period-search method described in Section 4.1. This method is effective at detecting the pulsation modes in the resulting amplitude spectrum, since the harmonic components of the target frequencies at each pulsation mode are fitted (i.e. removed) at each iteration.

Note that  $F_i$  has a subscript because the model can be changed for each period. For instance, one could search for the transiting exoplanet around a single-mode variable star by specifying  $F_0 = \text{Fourier}$  and  $F_1 = \text{SGM}$ . Since each model has a unique behaviour that enables more optimized search for specific targets, this should greatly enhance the detecting capability of scientifically impactful objects beyond what general Fourier DFT allows us to do.

## 5 PERIOD-DETECTION VALIDATION

As described in Section 4, PIPS provides an automated period-detection algorithm. This algorithm is developed with the analyses conducted by large surveys in mind, and therefore understanding the performance of PIPS as a function of various data quality and input parameters is essential. In the following sections, we describe our approach and results for validating the performance of PIPS. These tests are focused on validating the period and uncertainty determination capabilities of PIPS, and we discuss best practices and optimal data structures as indicated by our results.

### 5.1 Data construction

In order to test the accuracy of PIPS for data of varying quality, cadence, and quantity, we generate artificial data based on samples from OGLE (Soszyński et al. 2009). These samples provide highly complete and dense phase-folded light curves, which allow us to generate artificial data with realistic light-curve shapes. For the OGLE samples, we take the top 20 results (16 RRAb and 4 RRc stars) with the lowest reported period uncertainty.

We generate a high-cadence set of observations for a single period by phase-folding the original OGLE data using the period provided by OGLE. We then extrapolate to generate a high-cadence set of observations over the desired period of time.<sup>9</sup> To create an artificial data set for a specified number of observations ( $N$ ), days between observations (DBO), and signal-to-noise ratio (SNR), the extrapolated data set is sampled at the desired observation times. Gaussian noise of the desired size is then added to each observation’s magnitude, and we define the uncertainty of the constructed data by adding the original uncertainty and the noise in quadrature. The size of the resulting uncertainty is used to measure the SNR. We define a parameter, days between observations-to-window ratio (DWR), as the ratio of the average number of DBO, and the length of the observation window ( $W$ ) for a single observation within which the observation’s target time was randomly chosen. For instance, if a star is visible for 6 h ( $=1/4$  d) during the night and an observation was taken at a random time during that  $W$  each day, then  $\text{DBO} = 1$ ,  $W = 1/4$ , and  $\text{DWR} = 4$ . This is used to determine when to sample the artificial data points. Target observation times for the new data set are then generated, which follow the desired values of  $N$ , DBO, DWR, and SNR.

The new data are then analysed by PIPS, and the period determined by PIPS is compared to the one reported by OGLE. It is worth noting that, although the period reported by OGLE is only assumed to be the true period for the purpose of phase folding, it is the ‘true’ period for the data given to PIPS, since it is the period used to extrapolate the data during generation. Deviation in the assumed period results in decreased SNR, but the period of the new set, and hence the difference between that period and the one calculated by PIPS, is exact.

### 5.2 Evaluation metrics

We evaluate the performance of PIPS using two metrics.

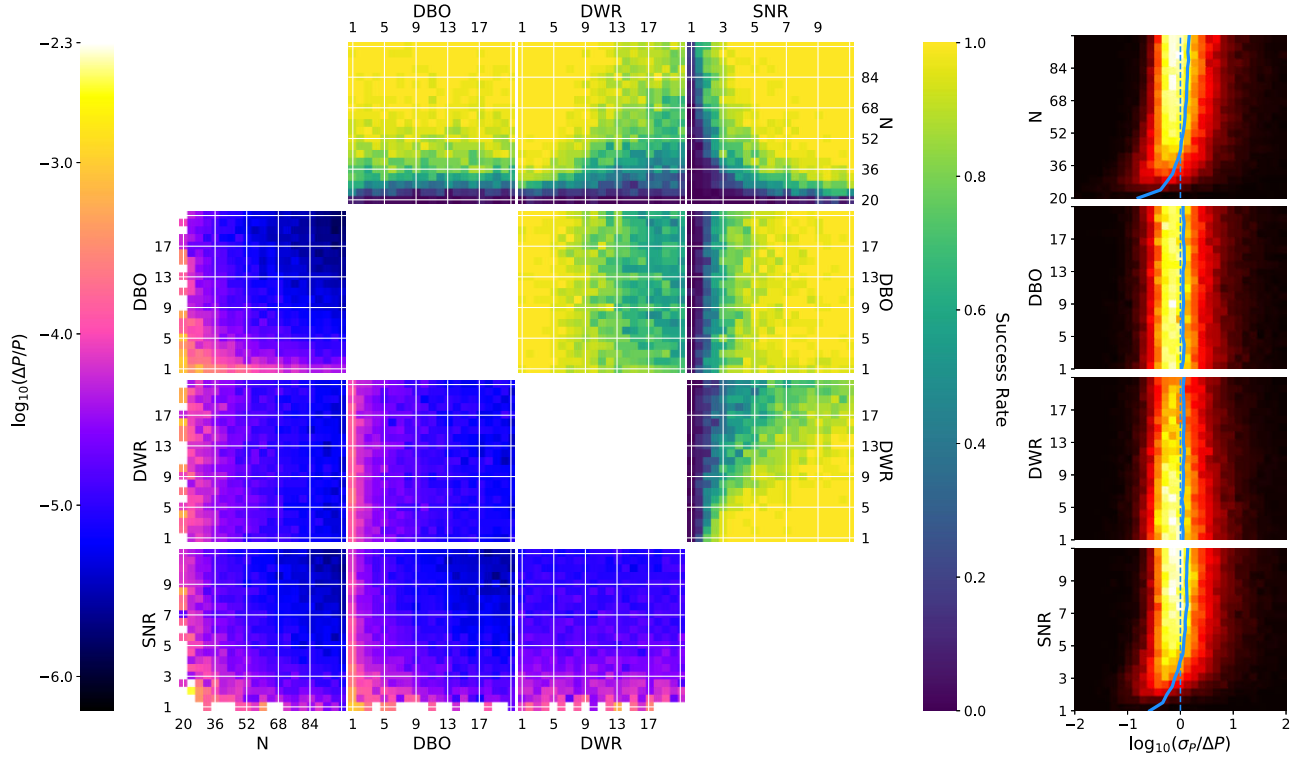
- (i) Success rate – the proportion of trials where PIPS achieves a measurement within 0.5 per cent<sup>10</sup> of the true period.
- (ii) Relative deviation – the deviation relative to the accepted period for successful tests.

The threshold value of 0.5 per cent deviation from the true period is chosen to separate our results owing to the bimodal nature of the estimate’s deviation from accepted values, as seen in Fig. 8 (and also in Fig. C1).

<sup>9</sup>This extrapolation allows us to generate the artificial data without analytic functions. The method provides sufficient data for this work since our tests mainly focus on poorly sampled, noisy data. One can also generate artificial data by using high-order Fourier series fitted to data sets from much more precisely calibrated surveys, such as *Kepler* (Kolenberg et al. 2010) when more physically ‘true’ mock data are desired.

<sup>10</sup>We choose this threshold based on a clear separation as seen in Fig. C1.





**Figure 9.** Left-hand panel: The success rate and relative deviation from the known period using PIPS on generated data. When not specified,  $N = 50$ ,  $\text{DBO} = 7$ ,  $\text{SNR} = 5$ , and  $\text{DWR} = 8$ . Relative deviation values are calculated only for the ‘successful’ population. The empty pixels indicate that the success rate is 0, so the relative deviation cannot be determined. Right-hand panel: 2D histograms of the ratio between estimated uncertainty and the deviation, with a  $\log_{10}$  scale. The parameter ranges on the vertical axes are the same as in the left-hand panel, and the data only include the ‘successful’ population.

### 5.3 PIPS behaviour by observation parameters

In our investigation, we identified four key attributes ( $N$ ,  $\text{DBO}$ ,  $\text{DWR}$ , and  $\text{SNR}$ ) that can influence the results of a period detection, be it derived via LS, PIPS, or otherwise. Here, we discuss how these attributes influence the results of PIPS. Our data-construction methods for a given set of attributes to be tested are outlined in Section 5.1.

#### 5.3.1 Number of observations

Fig. 9 shows that increasing the number of observations increases both the success rate and the deviation size. This is expected, since the larger number of degrees of freedom in fitting should result in more accurate results.

#### 5.3.2 Days between observations

Fig. 9 shows that this has no effect on the success rate but does decrease the deviation size. This may be due to increasing the total time over which observations are taken, thereby magnifying small errors in the period which in turn allows PIPS to further constrain its results. Note that, while  $N$  also increases the total time, it has an additional effect on the deviation size. For instance, in Fig. 9  $N = 20$  and  $\text{DBO} = 4$  will cover the same total time as  $N = 80$  and  $\text{DBO} = 1$ , but the deviation size at the latter point is less than the deviation size at the former.

#### 5.3.3 DWR

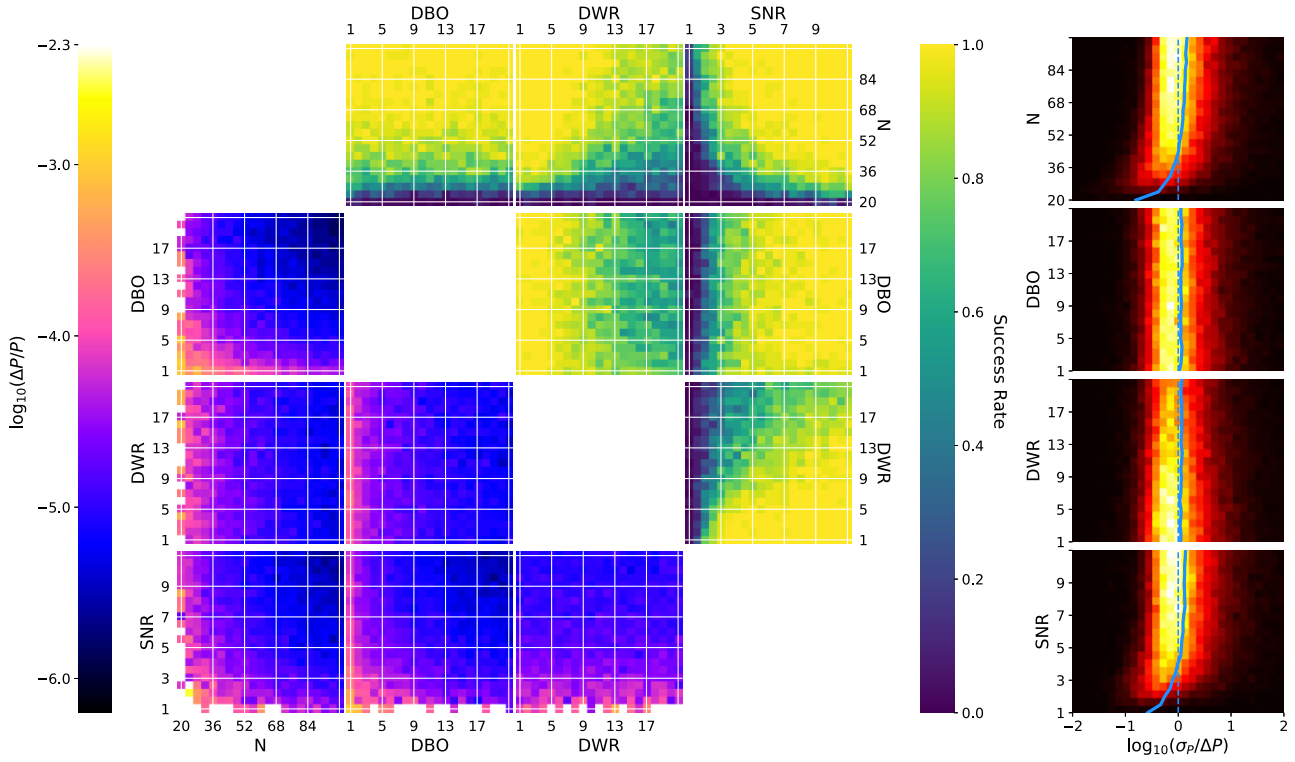
For the tests run in this section, target times are determined using the number of observations and the  $\text{DBO}$ . A random value between  $-W/2$  and  $W/2$  is added to each target time, where  $W$  can be interpreted as the size of the  $W$  for a single observation. For a given  $\text{DBO}$ ,  $\text{DWR}$  is determined by

$$\text{DWR} = \text{DBO}/W. \quad (25)$$

As  $\text{DWR}$  increases, the cadence becomes more regular, with observations becoming equally spaced as  $\text{DWR}$  approaches infinity. Fig. 9 shows that  $\text{DWR}$  has a negative correlation with success rate, and no relation with deviation size. The effects of low  $\text{DWR}$  can also be achieved by varying the time between observations using a non-random method, such as by advancing the time of observation by a set amount on consecutive nights, looping when necessary. When regularity of observation causes issues with aliasing, employing  $W$  functions may reduce the effect. We discuss the usage of  $W$  functions in the documentation.

#### 5.3.4 SNR

Higher  $\text{SNR}$  increases the success rate and the deviation size. While it is difficult to control for the  $\text{SNR}$  when collecting data, it is possible to overcome an  $\text{SNR}$  of as low as 2 by focusing on other parameters, such as  $\text{DWR}$ , as seen in Fig. 9.



**Figure 10.** Left-hand panel: The success rate and relative deviation from the known period using PIPS on generated data. When not specified,  $N_{\text{terms}} = 5$ ,  $N_{\text{peak}} = 7$ ,  $R_{\text{peak}} = 500$ , and  $N_0 = 5$ . The generated data have values of  $N = 35$ ,  $\text{DBO} = 7$ ,  $\text{DWR} = 8$ , and  $\text{SNR} = 3$ , the effects of which are discussed in Section 5.3. Right-hand panel: 2D histograms of the ratio between estimated uncertainty and the deviation, in  $\log_{10}$  scale. The parameter ranges on the vertical axes are the same as in the left-hand panel, and the data only include the ‘successful’ population.

#### 5.4 PIPS behaviour by internal parameters

PIPS is shipped with optimized internal variables (for detail, see, Section 4). These variables – which we detail below – are  $N_{\text{terms}}$ ,  $N_{\text{peak}}$ ,  $R_{\text{peak}}$ , and  $N_0$ , and each can be set when calling the fully automatic period-detection function `get_period` in PIPS. The effect each of these internal variables has on the performance of PIPS is tested similarly to Section 5.3, and we choose the default values we employ in PIPS based on our observation. Using the method described in Section 5.1, data are generated from 20 different initial stars to have values of  $N = 35$ ,  $\text{DBO} = 7$ ,  $\text{DWR} = 8$ ,  $\text{SNR} = 3$ , the effects of which are discussed in Section 5.3.

##### 5.4.1 $N_{\text{terms}}$

$N_{\text{terms}}$  determines the number of terms in the Fourier series PIPS uses to approximate a star’s light curve. In addition to affecting relative deviation and success rate,  $N_{\text{terms}}$  determines the lower limit on the number of observations needed for PIPS to be able to run. When  $N_{\text{terms}} = 1$ , PIPS is nearly identical to LS. Based on the results shown in Fig. 10, the ideal value of  $N_{\text{terms}}$  is either 4 or 5, with 5 providing a slightly larger success rate and 4 providing a slightly smaller deviation size, for the characteristics of data we fed to PIPS in this test.

##### 5.4.2 $R_{\text{peak}}$

$R_{\text{peak}}$  determines the number of samples taken when PIPS closely investigates a peak. The only drawback to increasing this value is an

increase in the computation time. The default value for  $R_{\text{peak}}$  is 500, but if computation time is a concern, it can be lowered to 250 with minimal quality degradation.

##### 5.4.3 $N_{\text{peak}}$

$N_{\text{peak}}$  determines the number of peaks to closely investigate after PIPS creates an initial periodogram. As with  $R_{\text{peak}}$ , the only drawback to increasing this value is more computation time. Increasing this value slightly improves the success rate, but has no effect on relative deviation.

##### 5.4.4 $N_0$

$N_0$  refers to the number of points sampled in the periodogram for each peak. Increasing  $N_0$  slightly increases the success rate and decreases the relative deviation. The only drawback to increasing this value is more computation time.

#### 5.5 Accuracy of reported uncertainties

PIPS is capable of reporting uncertainties that closely match the actual deviation from the true value of the period. When compared, the reported uncertainty and deviation from the true period are within one order of magnitude of each other 92 per cent  $\pm$  1 per cent of the time, as seen in Fig. 9. This does not appear to hold true when  $N < 25$ , or when  $\text{SNR} < 1.5$ , but is otherwise true regardless of parameters.

### 5.6 Performance by star subtypes

Of the RR Lyrae star subtypes tested, PIPS performs better on RRC stars than on RRab, as the light curves of RRC stars are closer to sinusoids. Owing to the multimodal nature of RRd stars, data could not be generated in the way described above, so they were not used for validation. Cepheid variable stars in only the fundamental mode can be studied as well – although without any adjustments, we see an increased deviation size. Results for each star type can be found in Fig. C2 for reference.

## 6 ADDITIONAL MODULES FOR APPLICATION

Beyond the period detection, we provide a wide variety of modules to enable a smooth workflow in the scientific analysis of periodically varying light curves. In the current version (PIPS-0.3.0), we have focused on pulsating variable stars. Below is a list of implemented modules.

### 6.1 Classification

PIPS provides a set of tools related to the type-classifier for pulsating variable stars. This is the first step in our plan to implement a wide variety of classifiers with different methods, and we select a simple machine-learning algorithm and its pre-trained classifier for the current version. We used OGLE (Udalski et al. 2008; Udalski, Szymanski & Szymanski 2015; Soszyński et al. 2009, 2011) and *Gaia* (Prusti et al. 2016; Eyser et al. 2017; Brown et al. 2018; Holl et al. 2018) for training and validation, achieving 99 per cent training accuracy and 94 per cent testing accuracy.

The classifier included in PIPS is based on the random forest (RF) classification algorithm (Ho 1995), which itself relies on the decision tree (D-Tree) algorithm. D-Trees and RFs have been used for numerous classification tasks on various data sets with considerable success (Ball & Brunner 2010; Hinners, Tat & Thorp 2018). The D-Tree algorithm works by hierarchically splitting the data, based on the given features, into subsets for each class such that each subset is maximally mutually dissimilar, while achieving most similarity within each group. This criterion of similarity is quantified by metrics such as the Gini index and entropy, which are the cost functions used to optimize the classifier (Rosa 2016). The RF algorithm is an improvement upon the D-Tree classifier that randomly resamples the training data and creates multiple D-Trees for these samples in a process known as bootstrapping. Each D-Tree uses a different, randomly chosen hierarchy of feature nodes (i.e. order of splitting) and makes its own decision on the classification. In the end, all the trees vote on the outcome, and the class with a plurality of votes is assigned to the data being classified (Sarica, Cerasa & Quattrone 2017).

We use the RF classifier implemented in the *scikit-learn* Python package (Pedregosa et al. 2011). 121 907 stars in total (10 281 classical Cepheids, 1856 Type II Cepheids, 88 411 RRab stars, and 21 359 RRC stars) are extracted from the OGLE variable-star data base (Udalski et al. 2008, 2015) to train the algorithm. In total, 900 stars from each class are then chosen as our training data, and the rest are earmarked as testing data to determine the accuracy of the model on the pre-classified data. This is done to ensure that no bias is introduced owing to any mismatch in the sizes of the training data of the different classes. The OGLE data base contains data on the stars' periods and the Fourier parameters. The model is trained on these data using the period and the  $r_{21}$ ,  $r_{31}$ ,  $\phi_{21}$ , and  $\phi_{31}$  parameters. A 100-tree RF is used with a depth of 10. The model returns 99 per cent

training accuracy. Upon deploying this model on the held-out testing data, the model returns a 98 per cent accuracy score.

To further test the validity of this model, we use the *Gaia* variable-star data base (Prusti et al. 2016; Brown et al. 2018; Holl et al. 2018) as an additional testing set. Since these data were obtained from a completely different, independently compiled source, a good test score on them would make a strong case that the model is not biased toward the OGLE data sample used to train it. After removing objects that have missing information, we obtain 271 classical Cepheids, 43 Type II Cepheids, 910 RRab stars, and 445 RRC stars in our *Gaia* testing set. Upon running the algorithm on these data, we obtain a 94 per cent testing accuracy, suggesting that the model is not biased toward OGLE data and is indeed applicable across data sets.

Since the light-curve shape and amplitude vary between different passbands, a more accurate classification may be achieved by modelling such variations and including the passband information in the training. We plan to include this update in our next work that utilizes multiband observation of RR Lyraes (Jennings et al., in preparation).

### 6.2 Phase-offset analysis

We implement a supplementary module to PIPS to calculate phase offsets (Observed – Calculated;  $O - C$ ) over years of observations. Following Jurcsik et al. (2001, 2012) and Szeidl et al. (2011), we define the  $O - C$  value to be

$$(O - C)_i = \mu_i, \quad (26)$$

where  $\mu_i$  is optimized so that

$$\mu_i = \argmin_j \sum \frac{[\mathcal{F}(x_{ij} - \mu_i; P_{\text{mean}}, \theta_{\text{template}}) + b_i - y_{ij}]}{\sigma_{y,ij}}. \quad (27)$$

Note that the data set-unique offset  $b_i$  is added to relax the possible differences in the calibration of telescopes. The template light-curve parameters are defined as the best-fitting light curve at the reference period:

$$\theta_{\text{template}} = \argmin_j \sum \frac{[\mathcal{F}(x_{\text{template},j}; P_{\text{template}}, \theta_{\text{template}}) - y_{\text{template},j}]^2}{\sigma_{y,\text{template},j}^2}. \quad (28)$$

Here,  $\mathcal{F}$  is the Fourier series and  $\theta$  are Fourier parameters for the template light curve,  $j$  denotes each exposure data point, and  $i$  denotes the set of data for a historical data set (see Jennings et al., in preparation).

### 6.3 Temperature estimation

When studying pulsating variable stars, the temperature is often regarded as a model-dependent value to be derived from other observables (see Section 6.4). However, statistical methods are capable of inferring the temperature from multiband photometry at a given snapshot. This adds another highly beneficial quantity to the observed values, possibly enabling more precise model constructions. In PIPS, we provide a module that performs temperature-dependent SED optimization for any given multiband photometry to create a time series of temperature values.

This analysis is performed under the assumption that the stellar flux can be approximated as blackbody radiation. In wavelength

space, Planck's law

$$F(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{\exp\left(\frac{hc}{\lambda k_B T}\right) - 1}, \quad (29)$$

gives us the estimated flux for a given temperature. For Vega-calibrated objects, the magnitude in any given band with the response function  $\phi_{\text{band}}(\lambda)$  then becomes

$$m_{\text{sim}}(T, b, \text{band}) = -2.5 \log_{10} \left( \frac{\int \phi_{\text{band}}(\lambda) F(\lambda, T) d\lambda}{\int \phi_{\text{band}}(\lambda) F(\lambda, T_{\text{Vega}}) d\lambda} \right) + b. \quad (30)$$

Here, the temperature for Vega (Kinman & Castelli 2002) is a fixed value, and the object's temperature  $T$  and the constant offset<sup>11</sup> vary. Using this  $m_{\text{sim}}$  as simulated magnitudes, PIPS performs MCMC to find the best-fitting temperature, according to the maximum-likelihood function

$$p(T, b, \sigma_{\text{int}} | \text{band}, m, \sigma_m) \propto p(T, b, \sigma_{\text{int}}) p(m | T, b, \text{band}, \sigma_m, \sigma_{\text{int}}). \quad (31)$$

In the equation above,  $m$  is the observed apparent magnitude of the object in a given band,  $\sigma_m$  is the uncertainty of the magnitude, and  $\sigma_{\text{int}}$  is the intrinsic scatter for the magnitudes across all given bands. We choose wide, uniform priors for the fitted parameters. For instance, the following are used for RRab stars at  $\sim 16$  mag:

$$T \sim \text{Uniform}(2000, 10000)$$

$$b \sim \text{Uniform}(0, 30)$$

$$\sigma_{\text{int}} \sim \text{HalfNormal}(1).$$

A visualized example<sup>12</sup> of this process is shown in Fig. 11.

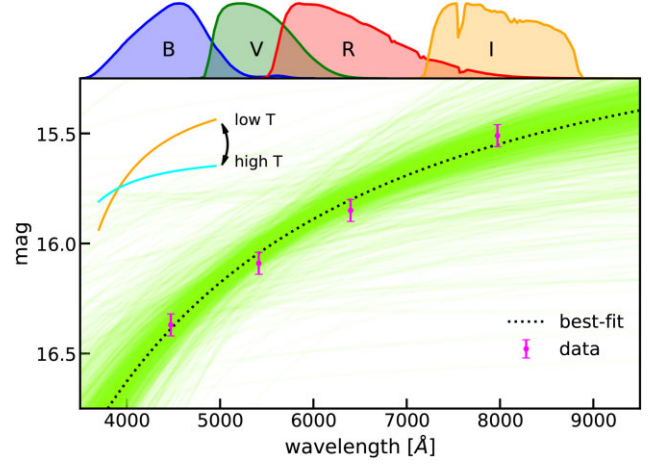
#### 6.4 Stellar properties

For many objects (including RR Lyrae stars), the light-curve parameters recovered from the period-determination process can be combined with theoretical models to extract stellar properties such as mass, luminosity, colour, and metallicity (Kovács & Walker 2001). PIPS implements a stellar parameter module that is integrated with its standard period-determination capabilities. As with period determination, this module is designed to accept a user-specified model as a flexible workflow, enabled by a robust and automatic error-propagation chain.

While the choice of model is intended to be open to the user, we have also implemented several stellar-parameter models as described by Kovács & Buchler (1994), Jurcsik (1998), and Cacciari, Corwin & Carney (2005). A point of concern that we intend to address with PIPS is the lack of clear uncertainty treatment in any of these previous works. Perhaps because of this, we find that our results suggest that the current models may be somewhat inaccurate and require refinement. For further discussion, see Jennings et al. (in preparation).

<sup>11</sup>This constant offset  $b$  absorbs the distance modulus and thus does not require an absolute magnitude in the fitting process.

<sup>12</sup>For this example, we use four-band ( $B$ ,  $V$ ,  $R$ , and  $I$ ) photometry data for RR Lyrae stars in globular cluster M3 (Jennings et al., in preparation), since the M15 data of Hoffman et al. (2021) are limited to two bands ( $B$  and  $V$ ). We utilized the *Katzman Automatic Imaging Telescope* at Lick Observatory (Filippenko et al. 2001) for the M3 observations, and the response functions are obtained from Stahl et al. (2019).



**Figure 11.** A visualization of the temperature-estimation method in PIPS. The Vega-calibrated flux (light-green curves; dotted black for the most probable case) is being fitted to the observed multiband photometric data (pink points) using the response function specific to the telescope (coloured curves at the top). The posterior analysis yields the fitted parameters of  $T = 6078 \pm 279$  K and  $b = 14.2 \pm 0.2$  mag. The intrinsic scatter  $\sigma_{\text{int}}$  is consistent with zero, with the 68<sup>th</sup> percentile at 0.13 mag.

## 7 DISCUSSION

### 7.1 False-alarm probability

In PIPS, we do not implement method to estimate the FAP. In the literature, the FAP has been considered to be an effective method to distinguish between true signal and false positives (VanderPlas 2018). It is known, however, that the analytic form of the FAP is only available for the classical LS periodogram (i.e. a Fourier model with a single term); models other than that require bootstrapping to compute the FAP. Since we provide a more statistically robust interpretation of the periodogram (likelihood periodogram) and the statistical significance can be directly derived from it (Section 3), we consider our solution to be more fundamental.

### 7.2 Planned development

#### 7.2.1 Visualization

Currently, our team is beginning development of a visualization module that will allow users to explore the functionality of PIPS. These visualization tools will serve as a more intuitive way for users to utilize and interact with PIPS with a minimal amount of extra code. Furthermore, we hope it will aid users in understanding how PIPS works, enabling them to determine if the code is working properly and confirming that the outputs make sense. Such visualization will also aid in user-input debugging, enabling users to understand why they might not be arriving at the results they are expecting. This module will initially make use of the `Ipython.widgets`<sup>13</sup> library for Python, but it may eventually evolve into a custom GUI for PIPS. Initial visualization tools will allow users to explore the periodogram, light-curve folding, and  $N_{\text{term}}$  fitting process. Further implementations will potentially include a GUI allowing for the importing of data files and the setting of initial parameters via text box and buttons rather than lines of code.

<sup>13</sup><https://ipywidgets.readthedocs.io/en/latest/>



### 7.2.2 Exoplanets

A long-term goal for this codebase is for it to maintain a flexible collection of light-curve models, suitable for a variety of periodic fitting needs. One possible extension could be with respect to exoplanets – e.g. by implementing the analytic light-curve models of Mandel & Agol (2002) to take the place of our Fourier and Gaussian methods. As shown in Section 4.2.4, more appropriate models lead to better fits to a given light curve’s period by a number of metrics. Therefore, creating an exoplanet-specific model may lead to further improvement over applying the standard PIPS Fourier approach to exoplanet data. Rigorous testing against standards in the field such as box least squares (Kovács, Zucker & Mazeh 2002, 2016), transit least squares (Hippke & Heller 2019b), or wavelet-based methods (Régulo et al. 2007) would be required to understand where in parameter space PIPS could serve as an effective and useful tool. Comparative studies of transit-finding algorithms indicate that different approaches tend to complement one another (Moutou et al. 2005), so we expect even in the best-case scenario that PIPS would serve as an addition to an existing set of tools available to exoplanet researchers.

### 7.3 Application to Rubin Observatory Legacy Survey of Space and Time

While the finalized plan for the Rubin Observatory Legacy Survey of Space and Time (LSST) is expected to be released later this year, we use information provided by LSST Science Collaboration et al. (2017) about the baseline observing strategy (minion\_1016) to assess it as an application for PIPS. LSST plans to take a large number of observations (199 in the  $r$  band and 201 in the  $i$  band) of every point in the southern sky over 10 yr. Using the definitions from Section 5.3,  $N = 199$  and  $\text{DBO} = 18.35$  for the  $r$  band. The periodogram purity function, a measure of the phase completeness, varies considerably depending on location in the sky. DWR is related to phase completeness, indicating that DWR will be well suited to PIPS in some regions, but not in others. For variable stars close enough to have high SNR and that occupy regions with good phase coverage, PIPS can be used to detect the periods of stars from LSST data. Since the value of  $N$  is so large for the  $r$  and  $i$  bands, it will likely be possible to divide observations into 4 or 5 subsets based on when they were taken for  $N$  values of  $\sim 50$  or  $\sim 40$  (respectively), which may allow PIPS to detect period changes over the 10 yr of LSST observations.

## 8 CONCLUSION

We have presented our new approach, the FL periodogram, for fast and statistically robust period detection in astronomical time-series data. We have shown that, in many cases, the FL periodogram unifies the currently separate processes of period detection, uncertainty estimation, light-curve fitting, and statistical significance estimation, without the help of additional, computationally expensive calculations. In particular, the improved statistical significance estimation provides opportunities for a fast and efficient search for periodic signals in large data sets.

To enhance the advantage of the FL periodogram, we have developed our new Python package, PIPS, which is designed to be an advanced platform for analysis tools of astronomical time-series data. The algorithm and mathematical/statistical models implemented in PIPS were discussed, with primary focus on the period detection with FL periodogram. The main features include:

- (i) unlimited light-curve models,
- (ii) automatic period detection,
- (iii) uncertainty estimation,
- (iv) various representations of periodogram (including FL),
- (v) statistical significance calculation, and
- (vi) automatic multiperiod detection.

The performance of PIPS for period detection was tested with artificially generated data based on the true light curves, and we discussed the optimal range of data attributes and internal parameters that make the performance of PIPS valid for scientific analysis.

PIPS also includes other modules, such as:

- (i) stellar parameter model helper,
- (ii) type classification,
- (iii)  $O - C$  (phase offset) analysis tools, and
- (iv) multiband temperature estimation.

The detailed use of these modules and scientific results for analysis of RR Lyrae stars will be presented in our forthcoming paper (Jennings et al., in preparation).

We also realize that the generalized light-curve fitter in PIPS is possibly capable of complementing the existing methods for other types of variable objects, such as exoplanets. We plan to discuss this application in the near future (Savel et al. 2022, in preparation).

In the era of all-sky surveys and space telescopes, we are certain that with its advanced features and user-friendly interface, PIPS will serve as a critical platform in the analyses of variable astronomical objects. We plan to continue implementing and updating a variety of tools to enhance the user experience and widen the range of roles that PIPS can play.

## ACKNOWLEDGEMENTS

We are grateful to the anonymous reviewer for providing critical and constructive suggestions that led to a substantial improvement of this paper. We thank Daniel Weisz, Joshua Bloom, and Kareem El-Badry for stimulating our interest in period analysis that initiated this project and providing feedback during the early development of PIPS. Furthermore, we thank Keto Zhang for continuously providing his knowledge of statistics. YSM thanks D’Arcy Kenworthy for his comments on statistical methods. WeiKang Zheng is acknowledged for his support in scheduling and conducting the observations. Finally, YSM thanks Alex Ho for providing great support when remote work was needed.

This research made use of ASTROPY,<sup>14</sup> a community-developed core Python package for astronomy (Astropy Collaboration et al. 2013; Price-Whelan et al. 2018). Also, it used data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

We acknowledge generous support from Marc J. Staley (whose fellowship partly funded BES as a graduate student whilst contributing to this work), the Christopher R. Redlich Fund, the U.C. Berkeley Miller Institute for Basic Research in Science (in which AVF was a Miller Senior Fellow), and many individual donors to AVF’s group. We also thank the U.C. Berkeley student observers involved with the research group: Samantha Stegman, Julia Hestenes,

<sup>14</sup><http://www.astropy.org>

Keto Zhang, Teagan Chapman, Matthew Chu, Asia deGraw, Romain Hardy, Evelyn Lu, Emily Ma, Emma McGinness, Shaunak Modak, Derek Perera, Druv Punjabi, Jackson Sipple, Kevin Tang, Sergiy Vasylyev, Jeremy Wayland, and Abel Yagubyan.

*Katzman Automatic Imaging Telescope* and its ongoing operation were made possible by donations from Sun Microsystems, Inc., the Hewlett-Packard Company, AutoScope Corporation, Lick Observatory, the National Science Foundation, the University of California, the Sylvia & Jim Katzman Foundation, and the TABASGO Foundation. Research at Lick Observatory is partially supported by a generous gift from Google. We appreciate the expert assistance of the staff at Lick Observatory.

## DATA AVAILABILITY

PIPS is available at the public software repository PyPI (<https://pypi.org/project/astroPIPS/>), GitHub (<https://github.com/SterlingYM/astroPIPS>), or by request to the authors. For detailed instructions and tutorials, visit the online documentation (<https://pips.readthedocs.io/en/latest/>). The raw data used in our validation will be shared upon request to authors. The photometry of RR Lyrae stars used in Fig. 1 is available at <https://github.com/SterlingYM/M15data2020>.

## REFERENCES

- Al-Nahhal I., Dobre O. A., Basar E., Moloney C., Ikki S., 2019, *IEEE Signal Process. Mag.*, 36, 157
- Astropy Collaboration et al., 2013, *A&A*, 558, A33
- Bailey S. I., 1902, *Ann. Harv. Coll. Obs.*, 38, 1
- Ball N. M., Brunner R. J., 2010, *Int. J. Mod. Phys. D*, 19, 1049
- Blomme R. et al., 2011, *A&A*, 533, A4
- Branch D., Wheeler J. C., 2017, *Supernova Explosions*. Springer, Berlin
- Brethorst G. L., 1988, *Bayesian Spectrum Analysis and Parameter Estimation*. Springer, Berlin
- Brown A. G. A. et al., 2018, *A&A*, 616, A1
- Budding E., Demircan O., 2007, *Introduction to Astronomical Photometry*. Cambridge Observing Handbooks for Research Astronomers, Vol. 6. Cambridge Univ. Press, Cambridge
- Buse A., 1982, *Am. Statistician*, 36, 153
- Cacciari C., Corwin T. M., Carney B. W., 2005, *AJ*, 129, 267
- Christy R. F., 1968, *QJRAS*, 9, 13
- Eyer L. et al., 2017, preprint ([arXiv:1702.03295](https://arxiv.org/abs/1702.03295))
- Filippenko A. V., Simon R. S., 1981, *AJ*, 86, 671
- Filippenko A. V., Li W. D., Treffers R. R., Modjaz M., 2001, in Paczynski B., Chen W.-P., Lemme C., eds, *ASP Conf. Ser. Vol. 246*, IAU Colloq. 183: Small Telescope Astronomy on Global Scales. Astron. Soc. Pac., San Francisco, p. 121
- Goodricke J., 1786, *Phil. Trans. R. Soc.*, 76, 48
- Gregory P. C., 2001, in Mohammad-Djafari A., ed., *AIP Conf. Ser. Vol. 568*, Bayesian Inference and Maximum Entropy Methods in Science and Engineering. Am. Inst. Phys., New York, p. 557
- Hinners T. A., Tat K., Thorp R., 2018, *AJ*, 156, 7
- Hippke M., Heller R., 2019a, *A&A*, 623, A39
- Hippke M., Heller R., 2019b, *A&A*, 623, A39
- Ho T. K., 1995, *Proc. IEEE*, 1, 278
- Hoffman A. M., Murakami Y. S., Zheng W., Stahl B. E., Filippenko A. V., 2021, *MNRAS*, 502, 818
- Holl B. et al., 2018, *A&A*, 618, A30
- Jurcsik J., 1998, *A&A*, 333, 571
- Jurcsik J., Kovacs G., 1996, *A&A*, 312, 111
- Jurcsik J., Clement C., Geyer E. H., Domsa I., 2001, *AJ*, 121, 951
- Jurcsik J. et al., 2012, *MNRAS*, 419, 2173
- Kinman T., Castell F., 2002, *A&A*, 391, 1039
- Kochenderfer M., Wheeler T., 2019, *Algorithms for Optimization*. The MIT Press, Cambridge
- Kolenberg K. et al., 2010, *ApJ*, 713, L198
- Kovács G., Buchler J., 1994, *A&A*, 281, 749
- Kovács G., Kupi G., 2007, *A&A*, 462, 1007
- Kovács G., Walker A., 2001, *A&A*, 371, 579
- Kovács G., Zucker S., Mazeh T., 2002, *A&A*, 391, 369
- Kovács G., Zucker S., Mazeh T., 2016, *Astrophysics Source Code Library*, record ascl:1607.008
- Kurtz D. W., 1985, *MNRAS*, 213, 773
- Leavitt H. S., Pickering E. C., 1912, *Harv. Coll. Obs. Circ.*, 173, 1
- Lee Y.-W., 1991, *ApJ*, 367, 524
- Lenz P., Breger M., 2005, *Commun. Asteroseismology*, 146, 53
- Lomb N. R., 1976, *Ap&SS*, 39, 447
- LSST Science Collaboration et al., 2017, preprint ([arXiv:1708.04058](https://arxiv.org/abs/1708.04058))
- Mandel K., Agol E., 2002, *ApJ*, 580, L171
- Moutou C. et al., 2005, *A&A*, 437, 355
- Paczynski B., 1986, *ApJ*, 301, 503
- Palmer D. M., 2009, *ApJ*, 695, 496
- Pedregosa F. et al., 2011, *J. Mach. Learn. Res.*, 12, 2825
- Percy J. R., 2007, *Understanding Variable Stars*. Cambridge Univ. Press, Cambridge
- Price-Whelan A. M. et al., 2018, *AJ*, 156, 123
- Prusti T. et al., 2016, *A&A*, 595, A1
- Reegen P., 2007, *A&A*, 467, 1353
- Régulo C., Almenara J., Alonso R., Deeg H., Cortés T. R., 2007, *A&A*, 467, 1345
- Rosa R. D., 2016, preprint ([arXiv:1604.03278](https://arxiv.org/abs/1604.03278))
- Sandage A., Tammann G. A., 1968, *ApJ*, 151, 531
- Sarica A., Cerasa A., Quattrone A., 2017, *Front. Aging Neurosci.*, 9, 329
- Scargle J. D., 1982, *ApJ*, 263, 835
- Schwarzenberg-Czerny A., 1989, *MNRAS*, 241, 153
- Schwarzenberg-Czerny A., 1999, *ApJ*, 516, 315
- Siegel M. H., Porterfield B. L., Balzer B. G., Hagen L. M. Z., 2015, *AJ*, 150, 129
- Soszyński I. et al., 2009, *Acta Astron.*, 59, 1
- Soszyński I. et al., 2011, *Acta Astron.*, 61, 1
- Stahl B. E. et al., 2019, *MNRAS*, 490, 3882
- Stellingwerf R. F., 1978, *ApJ*, 224, 953
- Szeidl B., Hurta Z., Jurcsik J., Clement C., Lovas M., 2011, *MNRAS*, 411, 1744
- Udalski A., Szymanski M. K., Soszynski I., Poleski R., 2008, *Acta Astron.*, 58, 69
- Udalski A., Szymanski M. K., Szymanski G., 2015, *Acta Astron.*, 65, 1
- VanderPlas J. T., 2018, *ApJS*, 236, 16
- Virtanen P. et al., 2020, *Nature Methods*, 17, 261
- Wemple L., 1932, *Harv. Coll. Obs. Bull.*, 889, 9

## APPENDIX A: BASIC USAGE

In tandem with this paper, we release the stable version of PIPS (v0.3.0), which contains all of the features and modules described here. Most of the period-analysis features are implemented in the `photdata` class, which is to be initialized with photometric time-series data. In Hoffman et al. (2021), for instance, we initialized the `photdata` object for each RR Lyrae star, using `time` (time in MJD), `mag` (observed apparent magnitudes), and `mag_err` (uncertainty in each magnitude). The minimal working example of period detection with PIPS in this case takes the following form:

```
import PIPS
star = PIPS.photdata([time, mag, mag_err])
P, P_err, Z = star.get_period(return_Z = True)
```

The returned values `P`, `P_err`, and `Z` are the most probable period, period uncertainty, and statistical significance (Z-score), respectively. The function `get_period()` is a direct implementation of the likelihood-periodogram method described in Sec-

tion 3. When the statistical significance is not needed, an additional argument `repr_mode = 'chi2'` reverts the method back to the  $\chi^2$  periodogram-based method (Section 2), which is slightly faster. A variety of settings can be applied as optional arguments. Available options include, but are not limited to, period search range, light-curve model (Section 4.2), number of terms ( $K_{\max}$  in equation 19 and equation 22), model-specific parameters (e.g.  $p$  in equation 20), and sampling parameters (e.g.  $N_0$ ,  $R_{\text{peak}}$  in Section 3). An example of code to specify these items is given below.

```
P, P_err = star.get_period(
    repr_mode = 'chi2',
    p_min = 0.1,
    p_max = 1.0,
    method = 'custom',
    model = 'Gaussian',
    Nterms = 1,
    p = 5,
    N0 = 15,
    R_peak = 500,
)
```

Here, `method = 'custom'` is passed to specify that a non-linear model (currently, anything except Fourier model) has been selected.

Furthermore, it is made easy in PIPS to visualize the periodogram. For instance, the line of code

```
star.periodogram(model = 'Fourier', Nterms = 5).plot()
```

automatically plots the periodogram, while simply calling the `periodogram()` function returns arrays of periods and periodogram powers. By default, periodogram is in  $\chi^2$  interpretation for easier representation, and passing an argument `repr_mode = 'likelihood'` forces the likelihood interpretation to be used.

For more detailed usage, including methods to implement custom light-curve models, visit the online documentation at <https://pips.readthedocs.io/en/latest/>.

## APPENDIX B: MARGINALIZATION OF BIVARIATE GAUSSIAN

Here we provide a more detailed derivation of equation (10). To begin, we define  $\tilde{\mu}_2$ , a function of  $\theta_1$  along the best-fitting line to the distribution:

$$\tilde{\mu}_2 = \mu_2 + \rho \frac{\sigma_2}{\sigma_1} (\theta_1 - \mu_1), \quad (\text{B1})$$

where  $\mu$  and  $\sigma$  are the mean and standard deviation values of parameters, respectively, and  $\rho$  denotes the correlation between two parameters,  $\theta_1$  and  $\theta_2$ . Then,

$$\begin{aligned} \mathcal{L}(\theta_2 | \theta_1, x, y, \sigma_y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2(1-\rho^2)} \left[ \left(\frac{\theta_1 - \mu_1}{\sigma_1}\right)^2 - 2\rho \left(\frac{\theta_1 - \mu_1}{\sigma_1}\right) \left(\frac{\theta_2 - \mu_2}{\sigma_2}\right) + \left(\frac{\theta_2 - \mu_2}{\sigma_2}\right)^2 \right]\right) \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2(1-\rho^2)} \left[ \left(\frac{\theta_1 - \mu_1}{\sigma_1}\right)^2 - 2 \left(\frac{\tilde{\mu}_2 - \mu_2}{\sigma_2}\right) \left(\frac{\theta_2 - \mu_2}{\sigma_2}\right) + \left(\frac{\theta_2 - \mu_2}{\sigma_2}\right)^2 \right]\right) \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2(1-\rho^2)} \left[ \left(\frac{\theta_1 - \mu_1}{\sigma_1}\right)^2 - \rho^2 \left(\frac{\theta_1 - \mu_1}{\sigma_1}\right)^2 + \left(\frac{\theta_2 - \tilde{\mu}_2}{\sigma_2}\right)^2 \right]\right) \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{(1-\rho^2)}{2(1-\rho^2)} \left(\frac{\theta_1 - \mu_1}{\sigma_1}\right)^2\right] \\ &\quad \times \exp\left[-\frac{(\theta_2 - \tilde{\mu}_2)^2}{2(1-\rho^2)\sigma_2^2}\right] \\ &\equiv \alpha(\theta_1) \exp\left[-\frac{(\theta_2 - \tilde{\mu}_2)^2}{2\beta^2}\right], \end{aligned} \quad (\text{B2})$$

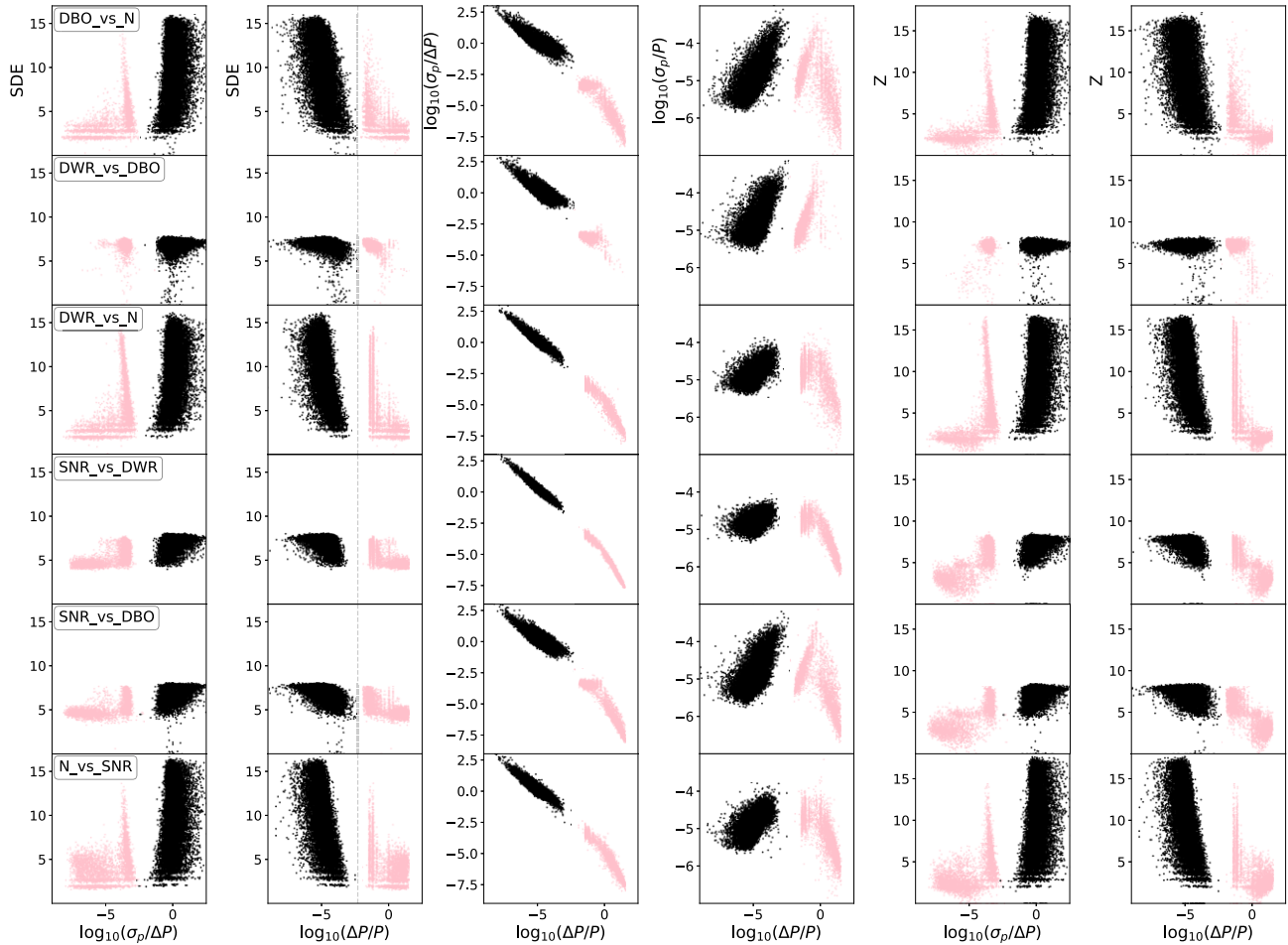
This is a slice of the multivariate Gaussian at fixed  $\theta_1$  along the  $\theta_2$  direction. The height of the Gaussian ( $\alpha$ ) is a function of  $\theta_1$ , and the standard deviation  $\beta$  is determined by the variance-covariance matrix only (i.e. independent of  $\theta_1$ ). This suggests that

$$\begin{aligned} \int \mathcal{L}(\theta_1, \theta_2 | x, y, \sigma_y) d\theta_2 &= \alpha(\theta_1) \int \exp\left(\frac{-[\theta_2 - \mu_2(\theta_1)]^2}{2\beta^2}\right) d\theta_2 \\ &\equiv \alpha(\theta_1) \cdot I(\theta_2 - \mu_2), \end{aligned} \quad (\text{B3})$$

$$\equiv \alpha(\theta_1) \cdot I(\theta_2 - \mu_2), \quad (\text{B4})$$

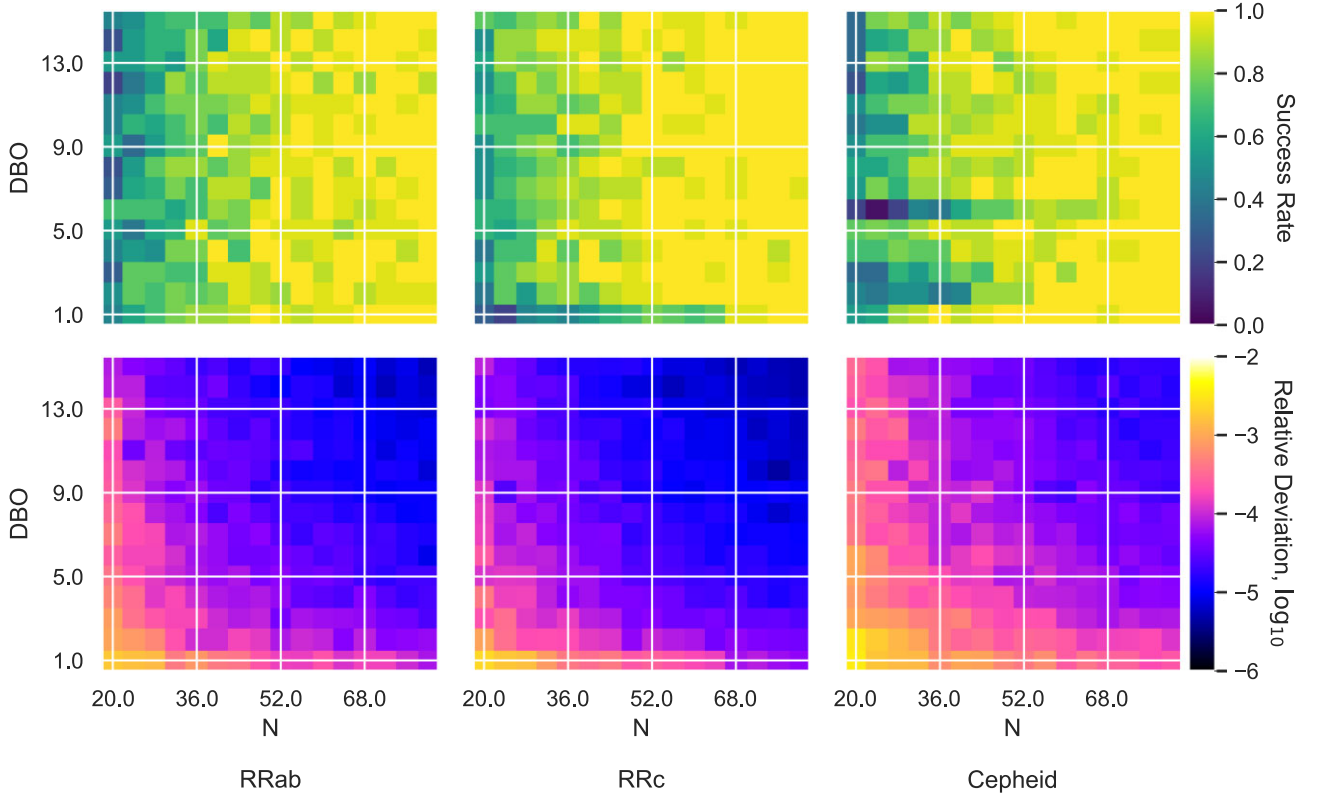
and the integral term  $I$  becomes independent of the exact value of  $\theta_1$  and  $\mu_1$ .

## APPENDIX C: SUPPLEMENTAL FIGURES



**Figure C1.** Distribution of the ‘successful’ (black) and ‘unsuccessful’ (light pink) populations in various parameters. Each row corresponds to the results in each panel in Fig. 9. It is evident that there is a clear separation at  $\log_{10}(\Delta P/P) \approx -2.3$  (grey dashed line), which corresponds to 0.5 per cent error. This results in clear separations in all other visualized parameters, such as  $\log_{10}(\sigma_p/P)$ . The statistical metrics, SDE and Z-value, show nearly identical distributions for the successful populations, while the typical Z-values for unsuccessful populations tend to be smaller than SDE. This indicates that the likelihood-based Z-value is a superior metric to evaluate the significance of the detected period.





**Figure C2.** The success rate (top panels) and relative deviation from the known period (bottom panels) with PIPS data generated using fundamental mode Cepheids at given number of observations ( $N$ ) and DBO, in comparison to RRab and RRc stars. SNR = 5 and DWR = 8 are kept constant. It is apparent that PIPS yields the better results in both the success rate and the relative deviation for RRc stars. The relative deviation for Cepheids suggests that PIPS requires larger  $N$  and DBO to achieve results similar to those of RRab/c stars, in contrast.

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.