The SHERLOCK pipeline: new exoplanet candidates in the WASP-16, HAT-P-27, HAT-P-26, and TOI-2411 systems

Martín Dévora-Pajares [1,2]★ Francisco J. Pozuelos,[3]★ Antoine Thuillier,[4,5] Mathilde Timmermans,[6] Valérie Van Grootel,[4] Victoria Bonidie,[7,8] Luis Cerdeño Mota[9] and Juan C. Suárez [1]

[1]*Dpto. Física Teórica y del Cosmos, Universidad de Granada, E-18071 Granada, Spain*
[2]*Avature Machine Learning, Madrid, 28004, Spain*
[3]*Instituto de Astrofísica de Andalucía (IAA-CSIC), Glorieta de la Astronomía, s/n, E-18008 Granada, Spain*
[4]*Space Sciences, Technologies and Astrophysics Research (STAR) Institute, Université de Liège, Allée du 6 Août 19C, B-4000 Liège, Belgium*
[5]*Institut d'Astronomie et d'Astrophysique, Université Libre de Bruxelles (ULB), CP 226, B-1050 Bruxelles, Belgium*
[6]*Astrobiology Research Unit, Université de Liège, Allée du 6 Août 19C, B-4000 Liège, Belgium*
[7]*Department of Physics and Astronomy, University of Pittsburgh, 3941 O'Hara Street, Pittsburgh, PA 15260, USA*
[8]*Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC), University of Pittsburgh, Pittsburgh, PA 15260, USA*
[9]*Promociones y Construcciones Pryconsa SL, Madrid, 28020, Spain*

ABSTRACT

The launches of NASA *Kepler* and *Transiting Exoplanet Survey Satellite* (*TESS*) missions have significantly enhanced the interest in the exoplanet field during the last 15 yr, providing a vast amount of public data that are being exploited by the community thanks to the continuous development of new analysis tools. However, using these tools is not straightforward, and users must dive into different codes, input–output formats, and methodologies, hindering an efficient and robust exploration of the available data. We present the SHERLOCK pipeline, an end-to-end public software that allows the users to easily explore observations from space-based missions such as *TESS* or *Kepler* to recover known planets and candidates issued by the official pipelines and search for new planetary candidates that remained unnoticed. The pipeline incorporates all the steps to search for transit-like features, vet potential candidates, provide statistical validation, conduct a Bayesian fitting, and compute observational windows from ground-based observatories. Its performance is tested against a catalogue of known and confirmed planets from the *TESS* mission, trying to recover the official *TESS* Objects of Interest (TOIs), explore the existence of companions that have been missed, and release them as new planetary candidates. SHERLOCK demonstrated an excellent performance, recovering 98 per cent of the TOIs and confirmed planets in our test sample and finding new candidates. Specifically, we release four new planetary candidates around the systems WASP-16 (with $P \sim 10.46$ d and $R \sim 2.20$ R$_\oplus$), HAT-P-27 (with $P \sim 1.20$ d and $R \sim 4.33$ R$_\oplus$), HAT-P-26 (with $P \sim 6.59$ d and $R \sim 1.97$ R$_\oplus$), and TOI-2411 (with $P \sim 18.75$ d and $R \sim 2.88$ R$_\oplus$).

Key words: software: data analysis – planets and satellites: detection – planetary systems – techniques: photometric.

1 INTRODUCTION

The search for exoplanets has experienced one the most significant developments in modern astronomy, where space missions such as *Kepler* (Borucki et al. 2010) and *Transiting Exoplanet Survey Satellite* (*TESS*; Ricker et al. 2016) have emerged as beacons of discovery confirming thousands of new worlds and providing with thousands of planetary candidates waiting to be confirmed. In addition, these missions' public nature has produced hundreds of thousands of light curves, allowing the community to conduct alternative data exploration. To explore and exploit this vast amount of data, many software packages emerged; some designed for the analysis of light curves, including data loading and visualization tools such as LIGHTKURVE (Lightkurve Collaboration 2018) and ELEANOR (Feinstein et al. 2019), some others to conduct data detrending

using a variety of functions and algorithms to remove instrumental systematics that produce photometric variability such as EVEREST (Luger et al. 2016, 2018), K2SSF (Vanderburg & Johnson 2014), and WŌTAN (Hippke et al. 2019), others to conduct a transit search such as the Box Least Squares (BLS; Kovács, Zucker & Mazeh 2002), the Quasiperiodic Automated Transit Search (QATS; Carter & Agol 2013), and the Transit Least Squares (TLS; Hippke & Heller 2019), others to perform results vetting such as VETTER (Zink et al. 2020) and ROBOVETTER (Coughlin et al. 2016), others dedicated to validate planetary candidates by determining whether a transit-like signal is statistically likely to be caused by a planet or a background eclipsing binary such as VESPA (Morton 2012), PASTIS (Díaz et al. 2014), and TRICERATOPS (Giacalone et al. 2021), others to conduct a Bayesian fitting of the data to refine the planet parameters such as ALLESFITTER (Günther & Daylan 2021), JULIET (Espinoza, Kossakowski & Brahm 2019), and EXOPLANET (Foreman-Mackey et al. 2021), and others to compute observational windows such as TAPIR (Jensen 2013), to name a few.

★ E-mail: mdevorapajares@correo.ugr.es (MD-P); pozuelos@iaa.es (FJP)

While most of these packages are open source, their different methodologies and input–output formats hinder their correct usage and implementation, relegating them to independent uses in most cases instead of an efficient workflow where all the procedures are connected. The lack of connection between these packages drives the community to rely mostly on the alerts issued by the official pipelines such as Science Processing Operations Center (SPOC; Jenkins et al. 2016) and the Quick Look Pipeline (QLP; Huang et al. 2020), i.e. the so-called *Kepler* Objects of Interest (KOIs) and *TESS* Objects of Interest (TOIs). However, many planets remain unnoticed due to the detection thresholds of these pipelines, the lack of data exploration, or the poor photometric quality. A proof of that is the number of alternative catalogues providing hundreds of new planetary candidates using *Kepler* data (see e.g. Ofir & Dreizler 2013; Coughlin et al. 2016), *K2* data (see e.g. Barros, Demangeon & Deleuil 2016; Mayo et al. 2018; Kruse et al. 2019), and *TESS* data (Cloutier 2019; Montalto et al. 2020; Feliz et al. 2021; Montalto 2023), highlighting the importance of the parallel analyses of the data provided by the space-based missions by multiple teams. A brief review of some of these alternative catalogues can be found in Melton et al. (2024), where the authors presented their own catalogue using *TESS* data from Cycle 1. None the less, the pipelines producing these extra candidates are mostly non-public; only a few are open source, such as DAVE (Kostov et al. 2019a), which is designed to search and vet planetary candidates using *K2* data. Then, the community is restricted to solely focus on those alerts, impeding the freedom to concentrate on specific scientific cases designed by independent groups. This paper presents the SHERLOCK (Searching for Hints of Exoplanets fRom Light curves Of spaCe-based seeKers) pipeline. This public PYTHON package allows the users to conduct their own planetary searches, examining robustly the data from space-based missions. The tool was first introduced at its initial development stages with its core search functionality in Pozuelos et al. (2020) and Demory et al. (2020), and has been used in a number of studies. The iterative development has led SHERLOCK to become an end-to-end pipeline where a special effort has been made to reduce user intervention and sort all the steps in an efficient workflow. With only a few command lines, the users can explore the existing data, conduct the transit search, vet the most promising signals, compute a statistical validation, conduct a Bayesian fitting to refine planetary parameters and ephemerides, and trigger a follow-up camping using ground-based facilities. With the implementation of its six modules, SHERLOCK can be used in any imaginable exoplanetary survey that relies on space-based data. For example, SHERLOCK is being used in the SPECULOOS (Search for habitable Planets EClipsing ULtra-cOOl Stars) project (Gillon 2018; Sebastian et al. 2021), which aims at finding transiting planets orbiting nearby ultracool dwarfs using a network of six 1-m ground-based telescopes located in different observatories such as ESO Paranal Observatory (Atacama Desert, Chile; Burdanov et al. 2018; Delrez et al. 2018), Teide Observatory (Canary Islands, Spain; Burdanov et al. 2022), and National Astronomical Observatory of Mexico (San Pedro Mártir, Mexico; Demory et al. 2020). Due to the small sizes of these stars, transiting planets orbiting them are prime targets for atmospheric characterization using, for example, the *JWST* (Gillon et al. 2017, 2020). In this context, SHERLOCK is exploring those stars in the SPECULOOS target list that have been observed by *TESS*, searching for transit-like features that might be attributable to planets. While the SPECULOOS telescopes have better photometric precision than *TESS* for these stars (see e.g. Delrez et al. 2022), a robust inspection of the *TESS* data can help to speed up the process of planetary searching, saving telescope time. In addition, SHERLOCK is the main pipeline of the FATE (forecasting optical turbulence to push

the Very Large Telescope to its full potential) project (Van Grootel et al. 2021), which aims at finding the first observational evidence of transiting planets orbiting hot subdwarf stars, which are direct post-asymptotic giant branch objects that have lost most of their envelopes. In particular, the project objective is to shed light on understanding what happens to planetary systems once their host stars leave the main sequence and engulf close-in planets, presumably perturbing the dynamical architecture of the system. In this context, SHERLOCK systematically searches for transiting-like features in the *Kepler*, *K2*, and *TESS* data (Thuillier et al. 2022).

Hence, the idea behind SHERLOCK is not producing a vast amount of new planetary candidates as other existing pipelines do, but empowering independent researchers to carry on their own designed surveys. The paper is organized in the following way. Section 2 presents the key concepts and the workflow behind SHERLOCK. In Section 3, we use SHERLOCK to recover the planets of some known *TESS* systems, evaluate the pipeline performance while also presenting the detection of four new candidates within those systems. In Section 4, we present a new project in which SHERLOCK will be used to search for extra planets that remain unnoticed. Finally, Section 5 presents our discussions, conclusions, and prospects.

2 THE SHERLOCK PIPELINE

SHERLOCK is an end-to-end pipeline that allows the users to explore the data from space-based missions to search for planetary candidates. It can be used to recover alerted candidates by the automatic pipelines such as SPOC (Jenkins et al. 2016) and QLP (Huang et al. 2020), the so-called *Kepler* KOIs and *TESS* TOIs, and to search for candidates that remain unnoticed due to detection thresholds, lack of data exploration, or poor photometric quality. To this end, SHERLOCK has seven different modules to (1) acquire and prepare the light curves from their repositories, (2) search for planetary candidates, (3) vet the interesting signals, (4) perform a statistical validation, (5) carry out planetary system stability tests, (6) model the signals to refine their ephemeris, and (7) compute the observational windows from ground-based observatories to trigger a follow-up campaign. To execute all these modules, the user only needs to fill in an initial yaml file with some basic information such as the star ID, the cadence to be used, etc., and use sequentially one command line to pass from one step to the next. Alternatively, the user may provide SHERLOCK with the light curve in a csv file, where the time, the normalized flux, and the flux error need to be given. In the following subsections, we describe in detail each of these steps. The complete list of parameters that can be used in the yaml file, along with a description of each one, can be found in GitHub's SHERLOCK repository.[1] Keeping in mind that future implementations and releases might require new parameters; hence, we suggest regularly checking the list.

2.1 Light-curve acquisition and preparation

SHERLOCK makes use of the LIGHTKURVE package (Lightkurve Collaboration 2018) to obtain the data of the space-based missions *Kepler*, *K2*, and *TESS*, from NASA *Mikulski Archive for Space Telescope* (*MAST*). In most cases, the user only needs to provide the star's *Kepler* Input Catalog (KIC), *K2* Ecliptic Plane Catalog (EPIC), or *TESS* Input Catalog (TIC) identification number and the cadence to be used, i.e. 20, 120, 200, 600, or 1800 s for *TESS* and

---

[1] https://github.com/franpoz/SHERLOCK/blob/master/sherlockpipe/properties.yaml

120 or 1800 s for *Kepler/K2*. In the case of the 600 and 1800 s for *TESS*, i.e. the full-frame images, the data can also be retrieved via the ELEANOR package (Feinstein et al. 2019). By default, the Pre-search Data Conditioning Simple APerture (PDC-SAP) is used, which has some level of processing with long-term trends removed using co-trending basis vectors. However, the PDC-SAP light curves are often affected by some other trends or systematics that need to be corrected before starting any planetary search. For example, some light curves present noisy regions with a high level of data dispersion, some others are affected by strong variability due to a fast rotation or stellar pulsations, and some others present short-term trends induced by the instrumentation. To efficiently deal with all these scenarios, SHERLOCK accounts with different semi-automatic tools.

2.1.1 Masking noisy regions

The light curves often show regions with high data dispersion where the photometric quality is significantly reduced, strongly affecting the performance of the search algorithm. These noisy regions are sometimes related to the spacecraft's motion, light scattering produced by the Earth, background moving objects, or other unidentified sources. Hence, to clean the light curve of such nuisance data, SHERLOCK can automatically compute the root-mean-square (rms) of the data set by 4-h binning. Afterward, all the light curve time gaps longer than 3 d are identified to split the data into segments. The mean rms of each data segment is then computed, and each data point that falls into an rms bin above $2\times$ the mean rms of its respective data segment is removed (see Fig. 1). The threshold of 2 is applied by default, but the user can adjust it if needed in the yaml file. As shown in Fig. 1, SHERLOCK also plots three rms threshold suggestions (1.25, 1.5, and $1.75\times$ the rms) alongside the user selection to provide insight into the typical values used for this mask. In this figure, some potential transit signals from TOI-144 appear to be masked when low rms threshold values are used, whereas a threshold of $1.75\times$ rms seems to be optimal for masking the entire noisy region without affecting the transits. In cases where the automatic high-rms-masking algorithm is not precise enough due to the complexity of the noisy features, the user can freely enter a set of time ranges to be masked.

2.1.2 Stationary stellar pulsations extraction

For quasi-stationary pulsations, we identified many cases (thanks to injection and recovery tests) where short-period transits might be confused with them, and then our targets would be removed using the automatic rms-based noise removal algorithm. For such targets where these types of pulsations are identified, we developed a custom variation of the Barceló Forteza et al. (2015) steps to reduce their effect on the light curve:

(i) Clear transits masking: An initial batch of searches with TLS is done to mask transits with low residuals, which are typically appreciable in the light-curve periodogram and might significantly alter the results of the next steps. That would happen because the subharmonics of the main transit period get narrower, and their shape in the periodogram might get close to the ones offered by the stationary pulsations that we would like to reduce.

(ii) Computation of highest peak signal-to-noise ratio (S/N): Once the highest peak is identified in the periodogram, its S/N is computed by dividing its power by the median of powers of its surrounding periods within a window of 4 h. We used the median instead of the mean as it is typically a more robust statistical estimator of baseline values, which is what we try to extract here.
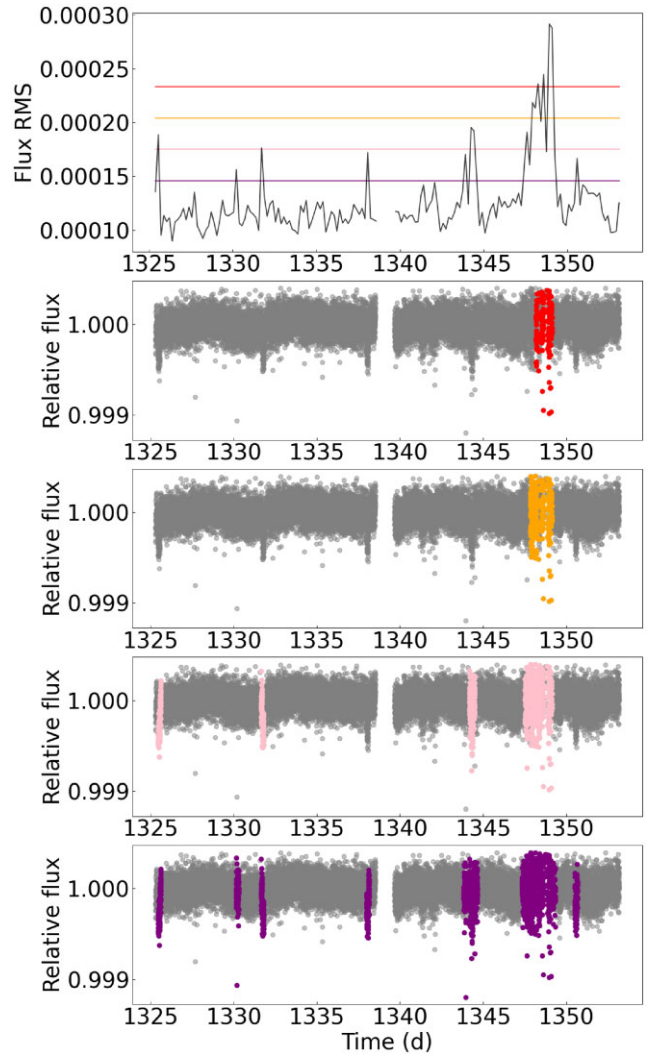

Figure 1. Output from SHERLOCK high-rms masked regions for the *TESS* light curve corresponding to TIC 261136679, observed during the Sector 1. The first panel shows the binned rms values and the computed thresholds. From above to the bottom, the rest of the panels show the curve flux values masked by 2 x rms, 1.75 x rms, 1.5 x rms and 1.25 x rms respectively.

(iii) Model the stationary signal: In case the S/N is higher than a given threshold (defaults to 4), the signal is modelled with a non-linear least-squares algorithm with an initial estimation of the amplitude and frequency through a fit of a cosine wave.

(iv) Check model validity: If the model shows an amplitude uncertainty lower than 20 per cent and a phase uncertainty lower than 0.2 rad, it will be extracted from the current light-curve flux. If the resulting signal shows lower rms, it will be accepted, and the resulting flux will be stored. We return to step (ii) and apply the algorithm iteratively whenever a model is accepted.

To test the capabilities of the Stationary Stellar Pulsations Extraction (SSPE) algorithm, we selected the TIC 169285097, catalogued as a hot subdwarf pulsating in *p* and *g* modes by Sahoo et al. (2020). These targets are complex hosts to search for small planets producing transits. Indeed, the strong rms induced by the several pulsation modes would make such planets undetectable, increasing their residuals and reducing their S/N. Also, the typical transit detection algorithms often fail when searching for transits around
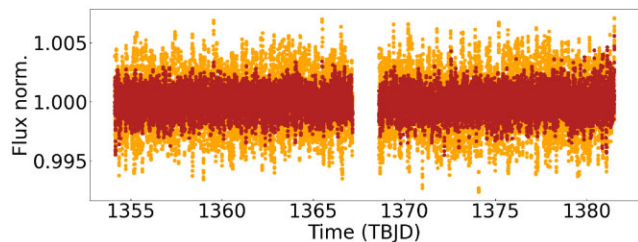
Figure 2. Results of applying the SSPE algorithm to the TIC 169285097 PDC-SAP flux data obtained from *TESS* Sector 2. The corrected light curve obtained after the removal of pulsations by our algorithm is represented over the more scattered original light curve.

these targets because they usually match sine-shaped pulsations to transit models, and their noise can be strongly correlated. As we are trying to diminish these effects, we tried our algorithm on the selected target to inspect the results that could be achieved on one of the available sectors from the *TESS* mission. We display the resulting flux in Fig. 2, which was obtained after subtracting 44 pulsation frequencies. The most powerful frequencies are shown in Fig. A1, where their corresponding models are overplotted to the phase-folded light curves. To properly investigate the capabilities of the algorithm in terms of transiting exoplanet signals recovery and to ensure that the procedure does not remove transit features, we tried our implementation on a public tool that we also developed in-house and published in a public repository under a free usage license: The MATRIX toolkit[2] built on top of the same code base as SHERLOCK, it contains the same algorithmic capabilities as the former and is conceived as an injection and recovery engine that includes a stage to generate a grid of periods, radii, and epochs to be injected in the original target's light curve. Those synthetic curves are later used in a recovery stage to search for the known injected transits. MATRIX will store a csv file to keep track of which models were found and which were not, generating at the end a heat map showing the matches for each period, radius, and epoch from the grid.

To assess properly the improvement brought by our SSPE algorithm, we first tried the injection and recovery scenario for a simple search without using it. Later, we tried the same scenario with the inclusion of the algorithm. The results can be appreciated in Fig. 3, where we can see that the first algorithm could add for this target an increment in the detectability of 1 $R_\oplus$ for the best cases. Altogether this study case leads us to an outstanding improvement in our detection rates for our current transit surveys like the one started in Van Grootel et al. (2021) and upcoming ones in the near future.

An improved version of this iterative algorithm would include the entire set of already selected signals on each iteration to improve the model-fitted parameters as it is done with FELIX in Charpinet et al. (2010) and Zong et al. (2016). However, this is not included in the SHERLOCKSSPE procedure yet.

2.1.3 Removing short trends

SHERLOCK uses a multidetrend approach to deal with short trends caused, for example, by instrumental drifts and intrinsic stellar variability. This strategy is implemented via the WŌTAN package (Hippke et al. 2019). By default, SHERLOCK detrends the flux using the biweight method. This choice is based on its good speed–performance ratio compared with other methods, such as the Gaus-
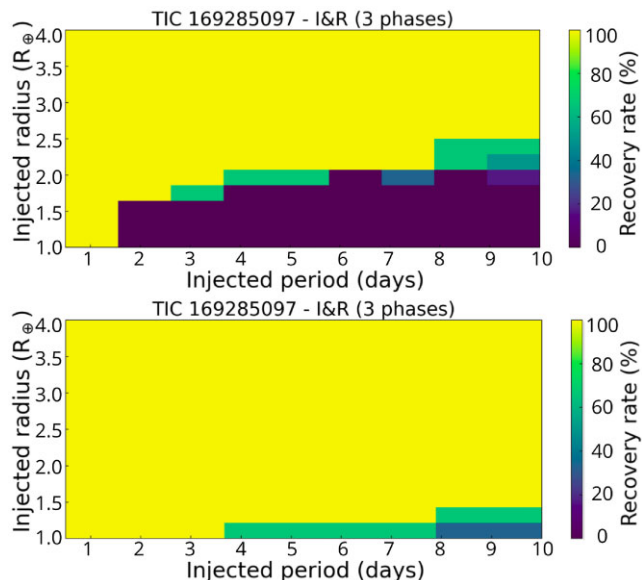
Figure 3. Injection and recovery results for the hybrid pulsator TIC 169285097 without and with the usage of SSPE. Top: The injection and recovery execution with a standard biweight detrend and search. Bottom: The same scenario execution including a previous SSPE run to reduce the quasi-stationary periodic pulsations effect on the light curve.

sian process (GP), which is also available. The biweight method is a time-windowed slider, where short windows can efficiently remove short trends, but there is an associated risk of removing any actual transit signal. Hence, to prevent this issue, the pipeline explores, by default, 10 cases with different window sizes. The minimum value of the window size is obtained by computing the transit duration (T_{14}) of a hypothetical Earth-size planet with an orbital period of 10 d. A minimum window size of $3 \times T_{14}$ is set to protect such a transit. On the other hand, the maximum value explored is $20 \times T_{14}$, which is enough to remove the variability in most of the cases. However, if preferred, the user may set the number of detrends to be applied and the minimum and maximum values of the window size in the yaml file. Hence, from the original PDC-SAP flux, SHERLOCK generates a new set of detrended light curves, each differing only in the window size used for detrending.

The transit search, described later in Section 2.2, is then conducted over all the detrended light curves jointly with the original PDC-SAP light curve. If a signal is detected in only one (or very few) of the light curves, it is considered 'detrend dependent', suggesting it may be an artefact produced by the detrending algorithm. Conversely, if a signal is consistently found in multiple (or all) light curves, it is considered 'detrend independent', thus increasing its credibility. By comparing the S/N and signal detection efficiency (SDE) of each detection across these light curves, we can determine the most appropriate window size for detrending. This method enhances the reliability of our planet detections by minimizing the risk of false positives due to the detrending process, ensuring that any promising signals found are not artefacts of a specific detrending configuration.

2.1.4 Correcting strong variability

Many stars suffer significant changes in brightness, which might be caused, for example, by stellar spots or intrinsic stellar variability. In such cases, the PDC-SAP light curves show intense sinusoidal trends of high amplitudes. To deal with these scenarios, SHERLOCK com-

[2]https://github.com/PlanetHunters/tkmatrix

putes the Lomb–Scargle periodogram (LSP) and extracts the value of the strongest peak using the LIGHTKURVE package (Lightkurve Collaboration 2018). By default, a factor of this value is used to remove the sinusoidal trend by fitting a sum of sines and cosines employing the cosine function provided by the WŌTAN package (Hippke et al. 2019). Our calibration experiments showed that the most appropriate value for such a factor is 0.25. However, the user may adjust it to any other value if wanted in the yaml file. There are some limitations in this procedure; for example, the presence of a hot Jupiter may induce a strong peak in the LSP that can be confused by a sinusoidal trend. Hence, a visual inspection is always desirable before applying this correction.

2.2 Search for planetary candidates

By default, SHERLOCK performs the transit search by using a modified version of the TLS package (Hippke & Heller 2019). Since TLS uses an analytical transit model taking into account the stellar parameters, it offers a better S/N and SDE than classical algorithms based on BLS (Kovács et al. 2002). This stellar-model optimization allows the detection of shallow periodic transits, which are more challenging to find. However, if desired, the user may set the BLS algorithm in the yaml file.

The procedure followed by SHERLOCK is to search for transits in all the detrended light curves jointly with the original one (PDC-SAP flux). Then, the results obtained for each individual light curve are stored; i.e. the orbital period (P) and its error in days, mid-transit time for the first identified transit (T_0) in TBJD (*TESS* Barycentric Julian Date) or BJD (Barycentric Julian Date), the transit depth in ppt (parts per thousand), the transit duration (D) in minutes, the number of transits, the S/N, the SDE, and the false-alarm probability (FAP). Once all the light curves are analysed, the results are combined to select the best transiting signal. To this end, SHERLOCK provides first a BORDER-CORRECTION algorithm. This algorithm computes a synthetic SDE value for each signal, penalizing the transits that are found in the border of the data set, which often corresponds to regions where there are more uncorrected systematics, as follows:

$$\text{SDE}_{\text{synthetic}} = \text{SDE} \left(1 - \frac{N_{\text{tb}}}{N_{\text{t}}} \right),$$

where N_{tb} is the number of transits close to borders, and N_{t} is the total number of transits. Secondly, the signal selection is performed by a QUORUM algorithm, where signals found in more light curves are favoured in a voting process over those found on less light curves. Once a signal is selected as the best candidate, SHERLOCK masks it in the PDC-SAP light curve using its P, T_0, and $3 \times D$, and starts a new searching run. The user can set how many times this process is repeated in the yaml file; however, our experience points out that results found beyond five runs are less reliable. This issue is coming due to the accumulated gaps in a given light curve after many mask-and-run iterations. Hence, by default, this process is repeated five times. In addition, SHERLOCK establishes minimum thresholds for the S/N and the SDE of a selected signal to pass from one run to the next. By default, both S/N and SDE are set to five, but the user may change these minimum thresholds in the yaml file.

Besides the QUORUM algorithm, SHERLOCK has two other more simple alternative selection methods. These two methods are based on the selection of the signal with larger S/N and the larger S/N$_{\text{synthetic}}$, respectively. Moreover, SHERLOCK provides a flexible module to use any other user-defined selection algorithm.

Once the searching stage is finished, SHERLOCK produces as many folders as the number of runs executed, containing plots showing for

Table 1. S/N for three known transiting planets with long time gaps given by SHERLOCK searches with the standard TLS version and with our modified version.

Planet	Sectors	S/N$_{\text{TLS}}$	S/N$_{\text{TLS-modified}}$
LHS-1140 b	3, 30	18.065	41.962
TOI-178 d	2, 29	11.723	35.317
TOI-736 c	9, 36	13.940	43.312

each light curve (original PDC-SAP flux and detrended light curves) the best signal found. Moreover, SHERLOCK provides two log files: (1) report.log, where the user can find all the information about the searching process, and (2) candidates.log, where is listed a summary of the main signals found. Moreover, the user will have access to extra folders named /fov, /detrends, /rms, and /sa. In /fov (from field of view) is stored the target pixel file (TPF) with the aperture used to extract the photometry overplotted along with the nearby stars from *Gaia* Data Release 2 catalogue, obtained via the TPFPLOTTER package (Aller et al. 2020), and the TPF with the light curves corresponding to each individual pixel using the LIGHTKURVE package (Lightkurve Collaboration 2018). In /detrends, the user can visually inspect all the detrends applied to the original PDC-SAP light curve as described in Section 2.1.3. In /rms, a collection of plots shows the regions of the light curve that are masked due to their high rms values, if any, as explained in Section 2.1.1. In case the SSPE module is activated (Section 2.1.2), the extracted sinusoidal trends are plotted in the /sa directory. Finally, the LSP and the value of the strongest peak are displayed as described in Section 2.1.4. Combining these outputs, SHERLOCK gives the user the information needed to evaluate the results' reliability and the signals listed as promising.

2.2.1 *TLS metrics fine-tuning*

The original TLS algorithm presented by Hippke & Heller (2019) reported some inconsistencies with the SDE and S/N in some scenarios. Hippke & Heller (2019) compute the transit duration by using a custom factor called 'curve stretch', supposedly taking into account the time gaps within the curve. Despite, this does not seem to be working properly and the error created in the transit duration propagates to the S/N calculation, which is based on this value. The S/N is especially hindered in curves with long time gaps, which are a typical case in the *TESS* mission. That is why we decided to rewrite the transit duration calculation by only taking the best model fit provided by TLS, measuring the period fraction where it falls below 1. This fix to the algorithm seemed to repair the S/N computations and the package began to return very stable S/N and some compared outputs can be explored in Table 1.

In addition, the SDE is a measurement that is expected to remain at scale among use cases and, hence, be comparable. We remind the reader that SDE is a metric indicating the statistical significance of a given period compared to the other periods tested in the transit search. In particular, an SDE value of x for any given period P means that the statistical significance of this period is $x \times \sigma$ compared to the mean significance of all other periods. [We refer the reader to Hippke & Heller (2019) and references therein for further details.] Then, this metric allows us to evaluate and compare the prominence of detected signals across different periods. However, the computation of SDE by TLS implies the calculation of the residuals, which in some edge cases are set to a default baseline that can lead to an overestimation of the SDE. Specifically, when many residuals are set to the baseline value,
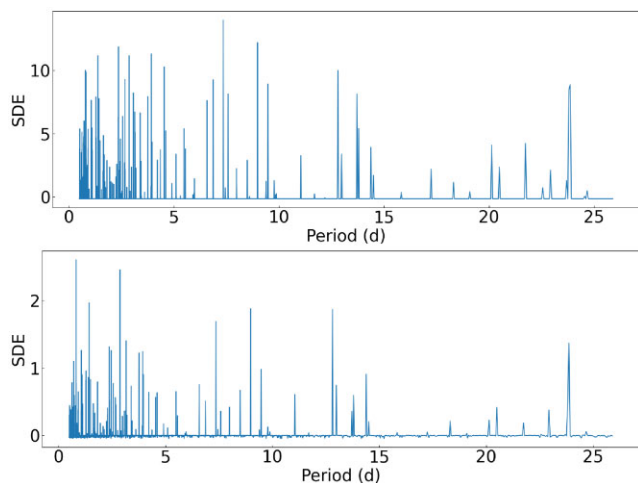
Figure 4. Signal detection efficiency (SDE) periodogram for TIC 10837041 (TOI-2411) obtained during the third run, i.e. the transits corresponding to TOI-2411 b and another candidate found during the searching process (Section 3) have been removed from the light curve. Due to the short orbital period of TOI-2411 b, which is only 0.78 d, the resulting light curve used in the third run contains many gaps after removing the transits. These numerous gaps impact the computation of the SDE conducted by TLS, leading to overestimated values, as seen in the top panel. The bottom panel displays the proposed correction by SHERLOCK, where the SDE periodogram now displays more reliable values, indicating the very low detection significance of the found signals (see text for details).

the few significant spikes can cause abnormally high SDE values, for example, when a given light curve contains many gaps. To mitigate this, we rescale each final SDE value based on the proportion of local residuals that are above the baseline. This adjustment ensures that the SDE values more accurately reflect the true statistical significance of the detected signals. As an example, Fig. 4 shows the difference between an original SDE spectrum with many baseline values and its rescaled counterpart. The last problem that we observed affecting the SDE is the difficulty of establishing fixed thresholds based on FAPs. The original TLS implementation calibrated a bijective relationship between SDE and FAP values. However, this calibration was carried out only for scenarios with time spans up to 28 d (the *TESS* typical length for one sector data) and does not provide good FAP values for longer data contexts, where SDEs tend to be higher on average even for false positives.

2.2.2 The Savitzky–Golay digital filter

SHERLOCK allows applying the Savitzky–Golay digital filter (Savitzky & Golay 1964) to the original PDC-SAP flux before any detrend and starting the transit search. This filter is implemented by using the SCIPY package (Jones et al. 2001), which performs a data convolution by fitting subsets of adjacent data points with a low-degree polynomial using the linear least-square method. Hence, the filter correlates nearby points in the data reducing their local white noise without distorting the signal trend, leading to half the initial rms value in the light curves. The Savitzky–Golay digital filter is handy when dealing with threshold-crossing events (TCEs) with low S/N. While the usage of this filter has the potential to find signals buried in the noise, it also has the associated risk of obtaining more false-positive detections that the user needs to vet carefully. In Table 2, we show the S/N and SDE values for TOI-270 (Günther et al. 2019) and TOI-736 (Crossfield et al. 2019; Peterson et al. 2023) using a pure-

Table 2. Transit parameters for TOI-270 and TOI-736 known planets given by SHERLOCK searches without and with Savitzky–Golay filter.

Planet	P (d)	D (min)	Depth (ppt)	S/N	S/N (SG)	SDE	SDE (SG)
TOI-270 b	3.36	75.3	0.8	14.3	31.6	61.1	40.4
TOI-270 c	5.66	94.7	2.9	49.1	112.4	39.8	46.6
TOI-270 d	11.38	115.9	2.0	28.1	70.2	44.7	44.0
TOI-736 b	0.95	35.8	3.8	8.9	17.4	18.5	18.7
TOI-736 c	4.99	59.1	14.4	17.7	41.2	24.4	25.0
TOI-736 d	2.75	49.2	2.15	5.192	12.2	6.6	6.6

TLS search and applying first the Savitzky–Golay digital filter. As can be noticed, the S/N reported by TLS is improved by a factor of 2, and the SDE is sometimes improved and sometimes worsened without a noticeable impact on the search. The S/N and SDE reported by SHERLOCK are extracted directly from the TLS algorithm, calculated with a naive estimation that cannot be compared with values from the SPOC pipeline. Therefore, these values are only used internally to compare the goodness of selected candidates and choose the clearest one. A recent application of this filter was used to recover the low S/N candidate TOI-4306.01 by Delrez et al. (2022), for which the pure-TLS search found the candidate with S/N below the typical detection threshold of $S/N \geq 5$.

2.3 Semi-automatic vetting

As described in Section 2.2, once the searching step is done, SHERLOCK provides several outputs that allow the user to perform a preliminary vetting. However, many other effects might be causing the detected signals that need careful inspection.

The finished *Kepler* mission (Borucki et al. 2003) provided more than 30 000 TCEs (Thompson et al. 2018), which are candidate signals that were found by the SPOC main search pipeline. However, different sources could generate these signals or even be instrumental artefacts induced into a target's light curve. To rule out these scenarios, the SPOC implemented the data validation reports (DVRs; Twicken et al. 2018), which are documents showing different metrics to qualify or discard an analysed candidate. These scenarios are (1) transit shape model fit, (2) odd–even transits checks, (3) centroids shifts, (4) optical ghost effects, (5) transit source offsets, (6) rolling band contamination, (7) secondary depth significance, and some additional numerical tests. Based on these tests SHERLOCK implements a vetting procedure that is explained in the next subsections.

While these tests may offer a good vetting for promising signals, there are some limitations, mainly when dealing with shallow transits with very low S/N, where the results of the vetting might be inconclusive. SHERLOCK uses the vetting stage by importing the WATSON package, which we also developed and published as an independent open-source tool,[3] offered in the official PYTHON repositories[4] with its own documentation.[5]

2.3.1 Single transit inspection

For each vetted candidate, SHERLOCK computes the light curve with the official pipeline aperture and a smaller one and displays them together for each single transit. The official pipeline aperture's

---

[3] https://github.com/PlanetHunters/watson
[4] https://pypi.org/project/dearwatson/
[5] https://dearwatson.readthedocs.io/en/latest/

centroids, motion, and background flux data are also displayed. In addition, the TPF is folded and plotted focused on each single transit. No metrics are computed from this data, but the visual inspection of these plots is very useful for finding differences or problematic features in single transits.

2.3.2 Cadence-based curves signal reliability

By using the period, epoch, and duration of a promising signal found in the transit search module, SHERLOCK computes the S/N for a box-shaped transit for each available cadence. If the transit is not significant enough (S/N $\geq$ 5) in any of the cadences or the S/N is substantially different from the nominal cadence, SHERLOCK will warn about it in the report.

2.3.3 Odd–even and secondary events

The transit-like signals might be originated from astrophysical sources that might not be planets. This is the case of eclipsing binaries or nearby eclipsing binaries, which typically show secondary eclipses with different transit signatures to the primary one. SHERLOCK computes the S/N for odd and even transits depths and also for secondary eclipses. A red flag is reported when secondary eclipses are detected with S/N $\geq$ 3 or in case either odd or even transits show S/N $\leq$ 5.

2.3.4 Transit source offset: images

The difference imaging method proposed in Bryson et al. ([2017](#)) was the selected approach by the SPOC to accept or reject signals based on the offset found on the computed transit source from pixel images. The method consists of computing the centre of mass of the subtraction of the flux measured within the TPF for in-transit and out-of-transit per pixel data. This way the most probable source of the transit events can be estimated. We have implemented a simplified version of this test in SHERLOCK. In addition, we have also included a second method to compute this metric by running a per pixel BLS search with a fixed period on the folded TPFs. We then compute the so-called centre of mass from each of the per pixel BLS residuals and, finally, average the output of both methods to get the final transit source position.

2.3.5 Transit source offset: centroids shift

An alternative method to estimate scenarios where the transit signal might have originated outside the target is the centroids diagnosis, also presented by Bryson et al. ([2017](#)). In our case, we subtract the mission-computed motion values (polynomials relating the sky and the focal plane for each cadence) from the mission-computed centroid data (target position on the TPF given by the pixel response function obtained from the out-of-transit control image; Bryson et al. [2013](#)). This method measures the deviation of the target position variations on the TPF against the focal plane alignment variations to the sky. If these values report an S/N $>$ 3, SHERLOCK will be placing a red flag for the analysed signals.

2.3.6 Optical ghost diagnostic

Sometimes the transit-like signals are caused by artificial sources. That is the case of optical ghosts, which can be created by light reflections from high-intensity sources between the optical system and the CCD as shown by Caldwell et al. ([2010](#)). To find false positives caused by optical ghosts, the *Kepler* DVR performs a test that measures the correlation between core (mean flux from all pixels within the target aperture) and halo (mean flux from all the pixels around the target aperture) fluxes and the transit signal. When the signal is affected by a ghost, the halo flux will be more correlated than the core one.

SHERLOCK implements a similar test. In our case, we compute and subtract the core and halo fluxes. In case the result still shows a transit signal with an S/N $>$ 3, SHERLOCK will reject the candidate.

2.4 System stability

In some cases, several promising signals are selected during the transit search, but their orbital periods are very close to each other. In other cases, the selected signal is very close to a known confirmed planet. The system's architecture might be unstable in these scenarios, revealing that the candidates found are likely false positives.

To shed light on these scenarios, SHERLOCK incorporates a module to assess the system dynamics stability. This analysis can be conducted at two points: during the initial vetting of a signal to assess whether there are possible system configurations for our candidates and as part of the validation scenario to get stability indicators for realistic or known configurations. That is, SHERLOCK can run a stability analysis in two modes. The first consists of giving user inputs, which create a linear grid of possible orbital parameters for each candidate. The second one loads the results from the Bayesian fit (see Section [2.6](#)) for the selected signal, allowing the addition of some constraints to the most likely system configuration.

To compute the system stability, SHERLOCK evaluates the fast chaos index Mean Exponential Growth factor for Nearby Orbits (MEGNO; Cincotta & Simó [1999](#), [2000](#); Cincotta, Giordano & Simó [2003](#)), which is widely used in this context (see e.g. Jenkins et al. [2019](#); Delrez et al. [2022](#); Pozuelos et al. [2023](#)). SHERLOCK uses the *N*-body numerical package REBOUND (Rein & Liu [2012](#)), which provides a MEGNO implementation by employing the Wisdom–Holman WHFAST code (Rein & Tamayo [2015](#)). The MEGNO mean value over time, denoted as $\langle Y(t) \rangle$, enhances stochastic fluctuations. This characteristic facilitates the differentiation of trajectories into quasi-periodic, that is to say, stable orbits, where $\langle Y(t \rightarrow \infty) \rangle$ approaches 2, and chaotic orbits, where $\langle Y(t \rightarrow \infty) \rangle$ tends towards infinity for an infinitive integration time or diverging from 2 for finite integration times.

This highly flexible module allows users to investigate as many parameters as desired to test the system's stability. Indeed, the user can explore the system's stability by varying the stellar mass, the planetary period, eccentricity, mass, argument of periastron, and inclination. Since we are dealing with transiting planets, not always the planetary masses are known; then, for these cases, SHERLOCK obtains the planetary mass relying on mass–radius relationships from Chen & Kipping ([2017](#)), which provides differentiation between small and large planets in a similar way than Weiss et al. ([2013](#)) and Bashi et al. ([2017](#)). In all the cases, for each parameter SHERLOCK uses uniform distributions, where the user needs to provide the nominal value, the associate uncertainty, and the number of bins that want to explore within that range. The results are then stored in csv files that the user can employ to construct stability maps. It is important to notice that while SHERLOCK offers the possibility of exploring a number of the parameters, not all of them equally affect the system stability. Indeed, as stated in Cranmer et al. ([2021](#)), the planetary semimajor axes (or the orbital periods), the masses, and

the eccentricities are the parameters that most affect the stability of a given system. Out of these three parameters, usually are the masses and the eccentricities the most unconstrained and, hence, the ones used to explore system architecture (see e.g. Demory et al. 2020; Murgas et al. 2023).

To correctly sample the planetary positions over their orbits, the scenarios are integrated using a time-step of 5 per cent of the shortest orbital period in the system. Finally, running the integration for a time of 10^5–10^6 times the largest planetary orbital period in the system is recommended to start getting convergent and reliable results. In any case, this analysis does not intend to replace a more robust dynamical study of a given system but a preliminary view of the credibility of planetary candidates and the potential system architecture.

2.5 Statistical validation

Kepler/*K2* and *TESS* found many candidates that, due to the faint nature of their host stars and/or the shallowness of their transits, cannot be confirmed from ground-based observatories as true planets. To mitigate this situation were developed the first statistical validation tools during the *Kepler* era (see e.g. Torres et al. 2011; Morton 2012; Morton et al. 2016), which allow us to calculate the false-positive probability (FPP) of a given detection to validate it statistically.

This procedure is nowadays widespread to support the planetary interpretation of a given signal even when the candidate can be reached from ground-based observatories (see e.g. Kostov et al. 2019b; Quinn et al. 2019). SHERLOCK allows the user to perform a statistical validation automatically connecting the results with the TRICERATOPS[6] package (Giacalone et al. 2021, 2022). Originally, TRICERATOPS was designed to vet and validate *TESS* candidates only; however, we have contributed to developing this open-source package to manage also *Kepler*/*K2* candidates. In short, TRICERATOPS computes the probabilities of a wide range of transiting-producing astrophysical scenarios such as a transiting planet and eclipsing binary (in many different configurations) using the primary transit of the candidate, the prior knowledge of the host and nearby stars, and the current understanding of planet occurrence and stellar multiplicity. Hence, TRICERATOPS computes the location of a given candidate in the FPP–nearby FPP (NFPP) plane, where the candidate will be classified as a validated planet if FPP < 0.015 and NFPP $< 10^3$, as likely planet if FPP < 0.5 and NFPP $< 10^{-3}$, and as likely false positive if NFPP $> 10^{-1}$. Due to the probabilistic nature of the procedure to calculate the values of FPP and NFPP by TRICERATOPS, SHERLOCK performs five computations and then averages the results to provide the final values. In addition, SHERLOCK uses the results from Lissauer et al. (2012), which estimated the fraction of planet candidates that are false positives when in systems with multiple planet candidates based on *Kepler* data. Then, this prior probability is multiplied by the FPP and NFPP previously computed (see e.g. Demory et al. 2020).

2.6 Bayesian model to refine the ephemeris

Once a candidate has been vetted and validated, it can be desirable to trigger a follow-up campaign to firmly confirm the planet in the target star using ground-based observatories. However, the results from the searching stage via TLS are not optimal for scheduling an observation due to the large uncertainty associated with the found orbital period (ΔP), whose value directly depends on the period grid density. Indeed, the transit time (T) and its uncertainty (ΔT) of future observational windows are given by (see e.g. Mallonn et al. 2019)

$$T = T_0 + n P,$$
$$\Delta T = \sqrt{\Delta T_0^2 + (n \, \Delta P)^2},$$

with T_0 and P previously defined in Section 2.2, and n the number of cycles since T_0. Hence, error propagation is strongly dominated by the term $n \, \Delta P$, which may produce ΔT of the order of several hours if the detection is performed in *Kepler*/*K2* data or early *TESS* sectors. This situation might hinder any ground-based confirmation.

In this context, SHERLOCK allows the user to refine the ephemeris by performing a Bayesian fit using the ALLESFITTER[7] package (Günther & Daylan 2019). ALLESFITTER is an open-source code that performs a Bayesian fit to infer the star and exoplanet properties given a set of photometric and/or radial velocity data. Then, SHERLOCK automatically injects the results from the searching stage and the stellar parameters in ALLESFITTER as priors and implements two steps fitting. (1) The candidate signals are masked, and the out-of-transit correlated noise estimation is performed using a GP with a 3/2 Matérn kernel with a Markov chain Monte Carlo fit on the PDC-SAP flux light curve, which can describe smooth long-term trends and short-term stochastic variations by its two hyperparameters: the amplitude scale (σ) and length scale (ρ). (2) The candidate signals priors are fit taking into account the previously derived GP parameters using normal priors, modelling the radius ratio of the planet to host, $R_p/R_\star$, the sum of the stellar and companion radii divided by the semimajor axis, $(R_\star + R_p)/a_b$, the cosine of the orbital inclination, $\cos i_p$, the epoch/transit mid-time in days, T_0, the orbital period of the planet in days, P_p, the transformed limb-darkening coefficients following a quadratic law, q_1–q_2, and the natural logarithm of the white noise, $\log \sigma_w$. By default, the Bayesian model used by SHERLOCK is the nested sampling, but the user might switch to MCMC if desired in the `yaml` file. All the posterior distributions and the derived parameters are stored in a new folder named `/fit`. At the end of this stage, the user obtains a robust inference of the planetary parameters. In particular, the ΔP is reduced considerably, and the new values can be used to schedule ground-based observations with much more accurate predictions.

2.7 Ground-based observational windows

Once the planetary properties are optimal from Section 2.6, the user may safely plan a follow-up campaign from ground-based observatories to confirm the planetary nature of a given candidate. This step is particularly critical for results obtained using *TESS* data. Indeed, the pixel scale of *TESS*, of 21 arcsec, is remarkably larger than *Kepler*'s, of 4 arcsec, moreover the *TESS*' point spread function (PSF) might be as large as 1 arcsec. Hence, due to dilution effects, the pixel scale, and the large PSF, the probability of contamination by nearby eclipsing binaries, whose deep eclipses might mimic shallow transits detected in the target star, is not null and, ideally, has to be checked from ground-based facilities. This ground-based confirmation is the primary goal of the seeing-limited photometry working group of *TESS* (see e.g. Kostov et al. 2019b). To schedule these observations, the community typically uses the *TESS* Transit Finder tool, which is a customized version of the TAPIR package (Jensen 2013).

---

[6]https://github.com/stevengiacalone/triceratops

[7]https://github.com/MNGuenther/allesfitter

In this context, SHERLOCK automatically loads the Bayesian fit results and a list of observatories provided by the user from where the transit windows will be computed. The observational windows from each observatory will consider some user-defined constraints, such as the minimum altitude of the object, distance from the moon in relationship to its illumination phase, and the minimum transit coverage. Then, SHERLOCK displays the observational windows in csv and pdf files, including the ingress, mid-transit, and egress times in UT for each event and plots with the airmass and altitudes of the host star during the entire event. This module was built by using the ASTROPY package (Astropy Collaboration 2013, 2018) and was tested against the *TESS* Transit Finder tool to verify its good performance.

3 SHERLOCK PERFORMANCE

Previously, we have outlined several projects leveraging the SHER-LOCK pipeline, such as SPECULOOS and FATE. Moreover, many other studies employed this pipeline to retrieve alerts from the *TESS* mission and known planets, as well as to identify new planetary candidates (see e.g. Wells et al. 2021; Schanche et al. 2022; Murgas et al. 2023; Dransfield et al. 2024; Dreizler et al. 2024; Gillon et al. 2024). In this section, our focus shifts to demonstrating the performance of the SHERLOCK pipeline through a specifically designed experiment to provide a level of reliability.

In particular, we built a catalogue of 100 stars that fulfil the following conditions in the Exoplanet Follow-up Observing Program (ExoFOP) list:[8] (1) the targets should, at least, contain one previously known planet or a TOI confirmed planet; (2) one of these planets should have orbital periods shorter than 10 d; (3) the maximum number of sectors should be two; and (4) the produced list with previous conditions is sorted by descending brightness and limited to the first 100 stars. Filters (1), (2), and (3) are applied to reduce the computational cost of this experiment. Appendix B lists the catalogue obtained following these criteria providing the TIC-IDs. The main goal of this test is to run SHERLOCK over this catalogue and record the successful recovery and non-recovery of planets and TOIs. Hence, to easily illustrate the performance, we introduce the recall parameter defined as $TP/(TP + FN)$, where TP are the true positives, i.e. the number of correct planets or TOIs found, and FN are the false negatives, i.e. the number of known planets or TOIs that we do not recover. Then, the performance of the code is optimal when this parameter approaches 1, indicating that SHERLOCK successfully identifies the expected planets without omissions.

From the 100 stars composing our catalogue, there were 104 TOIs that might correspond to previously confirmed planets, i.e. known planets for which *TESS* also provided an official TOI, or new planetary candidate alerts released by *TESS* that the scientific community has confirmed as planets. The number of multi-alert stars was much reduced because of the filters used to build the target list. Indeed, our filters favoured targets with large planets in short orbital periods, such as hot Jupiters (periods up to 10 d and radius from 6 R⊕). Population studies suggested that nearby planetary companions rarely accompany these planets. However, thanks to the extreme photometric precision of space-based missions and dedicated radial velocity surveys, nearby planetary companions to hot Jupiters have emerged in recent years (Wu, Rice & Wang 2023). These recent discoveries suggest that the formation of hot Jupiters via quiescent mechanisms is more efficient than previously thought,

when the main formation channel invoked was the dynamical high-eccentricity migration (Dawson & Johnson 2018). Hence, in this context, our target list not only serves to compute the recall factor and establish the goodness of the SHERLOCK pipeline but also offers an interesting scientific case searching for low S/N signals that might hint at the existence of planetary companions to these hot Jupiters in our list, supporting this new emergence of planets.

Once our target list was well established, we ran SHERLOCK over it. In general, after the execution of the first module, i.e. the exploration of the available data for a given candidate (see Section 2.1), we triggered a standard search with no initial fine-tuning. However, we had to refine the search parameters for some targets due to several factors, such as poor photometric quality and lack of data from the official pipelines. We provide a brief explanation for these particular cases.

(i) TIC 146520535 and TIC 464646604: In our multidetrend procedure, the window size ranges from 0.2 to 1.2 d by default, which we found optimal for most cases. However, due to the strong short-term variability of these light curves, we had to adapt the minimum and maximum window sizes to lower values from 0.25 to 0.75 d.

(ii) TIC 189380158, TIC 281408474, and TIC 33595516: The data for these targets were not available in *MAST*, impeding us from gathering them using LIGHTKURVE, which is the default route to get the data. Hence, we switch to ELEANOR long-cadence data by specifying the author in the yaml file.

SHERLOCK recovered 102 out of 104 transiting planets in the catalogue, making a count of two false negatives. These two were Kepler-89 e (TOI-4581.02), which was present in the *TESS* data as a single transit and hence, SHERLOCK mismatched it with a different period, and K2-285 c (TOI-4538.01), which was entirely missed in favour of a different non-promising signal. With this data, the final recall made up a value of 0.98 for our catalogue. However, SHERLOCK was also able to recover two other known planets that were not alerted with a TOI: the TOI-942 c and Kepler-89 c. For the TOI-942 system, *TESS* alerted one candidate with an orbital period of 4.3 d (TOI-942 b), and the follow-up of the system conducted by Carleo et al. (2021) found an outermost transiting planet with an orbital period of 10.16 d (TOI-942 c), which corresponds to our detection. In the case of the Kepler-89 system (Marcy et al. 2014), *TESS* alerted one candidate that corresponded to Kepler-89 d with an orbital period of 22.34 d, and a single transit that belonged to Kepler-89 e. In our experiment, SHERLOCK also found a planetary candidate with an orbital period of 10.42 d, which corresponded to Kepler-89 c.

On top of this, SHERLOCK spotted four interesting signals that could be attributable to planets and for which there is not any previous information. These new candidates were found in the systems WASP-16, HAT-P-27, HAT-P-26, and TOI-2411. We followed all the steps to verify their planetary nature by using the modules available in SHERLOCK as described in Section 2. In the following subsections, we describe these new candidates and the confirmation steps that we conducted to claim them as planetary candidates with high probabilities of being genuine planets. Table 3 shows the summary of these new candidates, and the planetary systems architectures are presented in Fig. 5. We release these candidates as Community *TESS* Objects of Interest (CTOIs), and will be listed in the ExoFOP CTOIs list.[9]

Hence, despite the small size of our sample, these results allow us to confirm the high performance of SHERLOCK as a planetary search

---

[8]https://exofop.ipac.caltech.edu/tess/view_toi.php; retrieved in 2023 October.

[9]https://exofop.ipac.caltech.edu/tess/view_ctoi.php

Table 3. New planetary candidates found during the performance survey.

System	CTOI-ID	Period (d)	$T_0 - 245\,7000$ (BJD$_{\mathrm{TDB}}$)	Duration (h)	Depth (ppt)	Radius (R$_\oplus$)	T_{eq} (K)
WASP-16	46096489.02	$10.457^{+0.018}_{-0.028}$	$1618.725^{+0.012}_{-0.014}$	$3.67^{+0.32}_{-0.45}$	0.54 ± 0.1	2.20 ± 0.23	810 ± 36
HAT-P$-$27	461239485.02	1.1994 ± 0.0002	$2705.510^{+0.002}_{-0.001}$	$1.27^{+0.11}_{-0.10}$	2.13 ± 0.20	4.33 ± 0.44	1426 ± 87
HAT-P$-$26	420779000.02	$6.594^{+0.009}_{-0.007}$	$2683.160^{+0.014}_{-0.010}$	$2.71^{+0.19}_{-0.41}$	0.59 ± 0.11	1.97 ± 0.20	785 ± 15
TOI-2411	10837041.02	18.750 ± 0.004	$2141.315^{+0.004}_{-0.003}$	$3.56^{+0.10}_{-0.08}$	1.77 ± 0.12	2.88 ± 0.15	430 ± 15
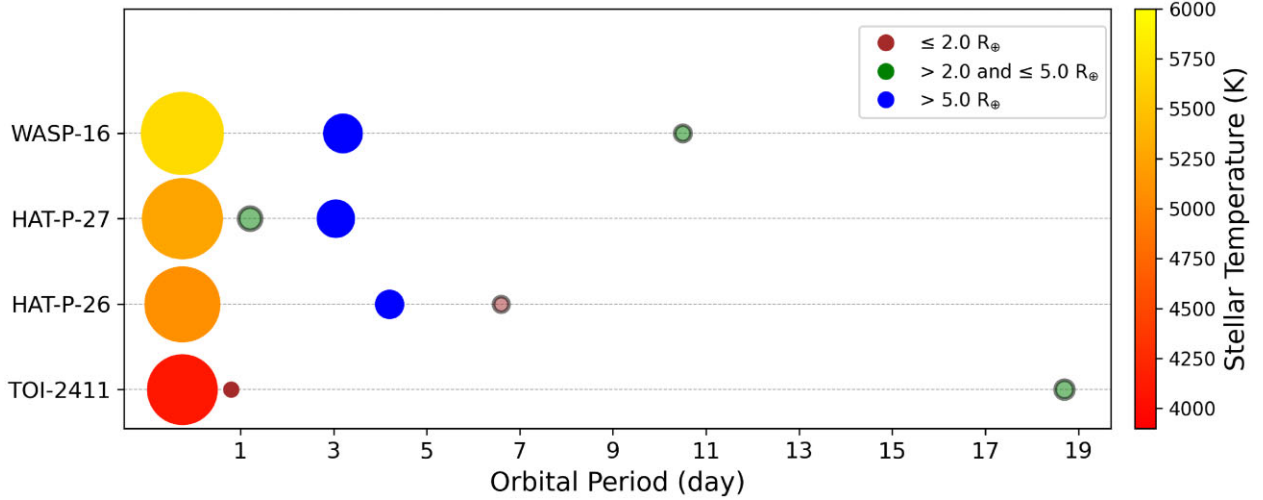
Figure 5. The planetary architectures of the systems WASP-16, HAT-P-27, HAT-P-26, and TOI-2411 are shown from top to bottom. The sizes of the elements in this figure scale with the sizes of the represented stars and planets. The sizes of the stars are scaled relative to each other, and the planets are scaled relative to each other, but the relative scales between stars and planets are not accurate, with planets oversized for clearer visualization. Known planets in the systems are displayed together to the candidates found in this study (whose properties are listed in Table 3), which are highlighted with a halo.

engine that can be used to explore space-based data efficiently, not only to recover known planets but also to search for new ones that remained unnoticed by the community.

3.1 WASP-16

WASP-16 is a Sun-like star with $T_{\mathrm{eff}} = 5700 \pm 150$ K, $R_\star = 0.946 \pm 0.054$ R$_\odot$, and $M_\star = 1.022 \pm 0.101$ M$_\odot$ observed by the Wide Angle Search for Planets (WASP) project (Pollacco et al. 2006), in particular using WASP-South, between 2006 and 2008. These observations hinted at the presence of a giant planet that was later confirmed using spectroscopic observations with the CORALIE spectrograph in 2008 and 2009. The results confirmed the planetary nature of this signal, yielding $R = 1.008\,R_{\mathrm{Jup}}$, $M = 0.855\,M_{\mathrm{Jup}}$, and with an orbital period of 3.12 d. This planet is the hot Jupiter WASP-16 b (Lister et al. 2009).

WASP-16 was observed by the *TESS* mission in Sector 11 between 2019 April and May, using the 120 and 1800 s cadences (TIC-46096489). The search performed with SHERLOCK was conducted using the standard parameters, and we recovered in the first run the signal corresponding to WASP-16 b, with S/N = 97.9 and SDE = 18.6 in the 11 light curves (see Section 2.1.3). During the second run, SHERLOCK did not spot any promising signal, but events that happened close to the borders of the light curves, which clearly hinted at a systematic origin. During the third run, we found a weak signal with S/N = 4.7 and SDE = 13 with a period of 10.44 d. While this signal is just below our default detection threshold of S/N $\geq$ 5, it was found in all the 11 light curves processed, which reveals a non-detrend dependence. Hence, we vetted the signal and excluded the

possibility of it being produced by any systematics (see Section 2.3). Then, we conducted the statistical validation (see Section 2.5), and we got FPP = 0.26 and NFPP = 0.0036 values, slightly above the likely planet area defined by Giacalone et al. (2021). Subsequently, we conducted the Bayesian fitting (see Section 2.6) and compared the Bayes factor ($\Delta \ln Z$) of two models: (1) only with the WASP-16 b planet, i.e. the null hypothesis, and (2) including the second potential planet found by SHERLOCK, i.e. a more complex scenario with two planets. Then, we got $\Delta \ln Z = \Delta \ln Z_{\text{2-planet}} - \Delta \ln Z_{\text{1-planet}} \sim 6.7$, what favours the two-planet scenario (Kass & Raftery 1995). Considering the global results of these analyses, we concluded that the signal found with a 10.44 d is a strong planetary candidate named CTOI-46096489.02, which would correspond to a mini-Neptune of $\sim$2.2 R$_\oplus$. Its main parameters from the fitting can be consulted in Table 3. However, it is important to notice that the Bayes factor is suggestive but not decisive; hence, extra observations using high-precision photometry would be needed to firmly confirm the planetary nature of the CTOI-46096489.02 candidate. Due to the shallowness of its transit of $\sim$0.54 ppt, ground-based instrumentation is inadequate for detecting such a subtle eclipse. Indeed, the limited sensitivity and atmospheric interference associated with terrestrial telescopes prevent accurate measurement of these weak signals. Therefore, to obtain precise and reliable data that allow us to confirm the planet, it is essential to utilize space-based instrumentation, such as *CHaracterising ExOPlanet Satellite* (*CHEOPS*; Benz et al. 2021), which can bypass atmospheric distortions and provide the necessary sensitivity to observe these shallow transits effectively (see e.g. Delrez et al. 2021). Unfortunately, since the last observational data available come from *TESS* Sector 11, the current uncertainty in the

transit time, ΔT (see Section 2.6), is ~100 h, which makes any time-critical observation impracticable. Another *TESS*-like campaign would be needed to mitigate this lack of accuracy in the current transit times. According to the *TESS*-point web tool,[10] WASP-16 will be reobserved in 2025 April, what will offer the opportunity to recover this candidate and trigger an accurate follow-up campaign that will allow us to confirm or refute its planetary nature.

The two-planet model is displayed in Fig. 8, the posterior distribution is provided in Appendix C, Table C1, and the convergence of the model is visualized in Fig. C1.

3.2 HAT-P-27

HAT-P-27 is a G8 dwarf star with a $T_{\mathrm{eff}} = 5250 \pm 100$ K, $R_\star = 0.90 \pm 0.05$ R$_\odot$, and $M_\star = 0.94 \pm 0.04$ M$_\odot$ monitored on a nightly basis by The Hungarian-made Automated Telescope Network (HAT-Net; Bakos et al. 2011) from 2009 January to August. These photometric observations were combined with nine spectroscopic measurements using the High Resolution Echelle Spectrometer (HIRES) instrument on the Keck I telescope from 2009 December to 2010 June. The results were presented in Béky et al. (2011), reporting the discovery of HAT-P-27 b, a hot Jupiter with a size of $1.038\,R_{\mathrm{Jup}}$, a mass of $0.66\,M_{\mathrm{Jup}}$, and an orbital period of 3.04 d.

HAT-P-27 was observed by the *TESS* mission in Sector 51 in 2022 May, using the 120 and 600 s cadences (TIC-461239485). We used the 120 s data in our search with SHERLOCK, allowing us to recover in all the light curves a transiting signal corresponding to the HAT-P-27 b planet with S/N = 37.3 and SDE = 12.6. During the second run, SHERLOCK spotted a transit-like feature with a period of 1.20 d, with S/N = 10.3 and SDE = 29.8, also detected in all the light curves. We executed the vetting module for this signal, which allowed us to rule out its systematic origin. We noticed that SPOC identified a TCE for a candidate with an orbital period of 2.40 d, i.e. the first harmonic of the signal that SHERLOCK spotted. Hence, during the vetting stage, we carefully examined the depths of the odd/even eclipses and transit shapes in short and long cadences and found no remarkable difference among them. Moreover, the candidate did not show any problematic metric either on the transit source or the optical ghost checks, reinforcing our findings with a period of 1.20 d as the more plausible solution. The TCE found by SPOC did not become a TOI due to its low S/N (<10) and slightly centroids offset. Next, we launched the statistical validation module, yielding values of FPP = 0.18 and NFPP = 0.00054, which fall on the likely planet area.

Following the same procedure as in the case of the WASP-16 system, we conducted the Bayesian fitting and compared the Bayes factor between two models: (1) only with the HAT-P-27 b planet (null hypothesis) and (2) including the second potential planet found by SHERLOCK. We got $\Delta \ln Z = \Delta \ln Z_{\text{2-planet}} - \Delta \ln Z_{\text{1-planet}} \sim 64.6$, which strongly favours the two-planet scenario over the simplest with a single planet. This candidate is then the CTOI-461239485.02, which corresponds to a hot Neptune-size planet of ~4.3 R$_\oplus$. Its parameters from the fitting can be consulted in Table 3. The transit depth found for this candidate is ~2.1 ppt, which is deep enough for small- and medium-sized ground-based telescopes. Moreover, its orbital period is 1.20 d, which eases scheduling and follow-up efforts due to the large number of possible observable events. However, the uncertainty in the transit time is $\Delta T \sim 3.3$ h, which is challenging to
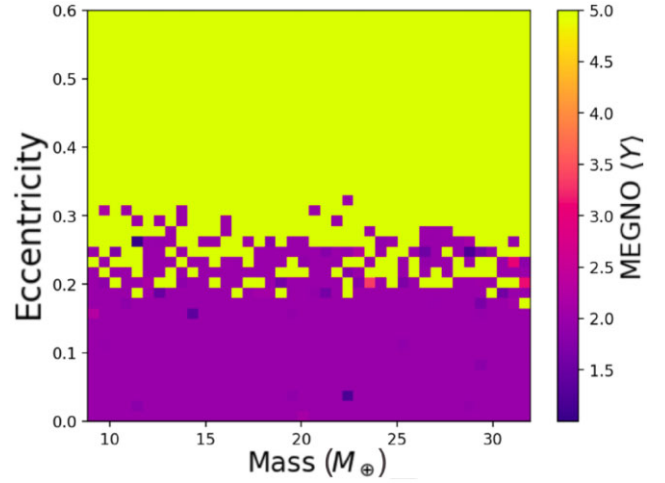
Figure 6. Stability map of the system HAT-P-27 composed of the two planets: HAT-P-27 b and the planetary candidate CTOI-461239485.02. The map contains 40×40 pixels, exploring the e_{p}–M_{p} parameter space for the candidate. The stability of each scenario is evaluated using the MEGNO parameter, where $\langle Y(t) \rangle \to 2$ represents quasi-periodic orbits, and chaotic systems when $\langle Y(t) \rangle \to 5$.

conduct time-critical observations from the ground but still possible. An alternative approach would be to conduct long observations from space-based telescopes such as *CHEOPS*, allowing for scanning up to 3σ the predicted time without the day–night cycles, making a putative detection easier. All these considerations make CTOI-461239485.02 an excellent candidate to be confirmed or refuted with relatively low effort.

The two-planet model is displayed in Fig. 8, the posterior distribution is provided in Appendix C, Table C1, and the convergence of the model is visualized in Fig. C2.

Despite the vetting and validation tests passed by CTOI-461239485.02, the resulting planetary configuration with a hot Jupiter at 3.04 d and an innermost candidate at 1.20 d might be questionable due to stability concerns. For this reason, we executed the stability test (see Section 2.4) assuming the nominal values for the star and HAT-P-27 b planet as reported by Béky et al. (2011), and for the new candidate using a grid of 40 eccentricities from 0 to 0.6 and 40 planetary masses from 9.5 to 32 M$_\oplus$ derived from the radius using a mass–radius relationship ($17^{+13.8}_{-8.1}$ M$_\oplus$; Chen & Kipping 2017). This configuration allowed us to build a stability map exploring 1600 scenarios in the e_{p}–M_{p} parameter space. Then, the MEGNO value is computed six times in each scenario by modifying randomly the orbital parameters. Hence, in total, the stability map explored 9600 scenarios. Each scenario was integrated over ~10^7 orbits of the outermost planet. The stability map in Fig. 6 shows that the HAT-P-27 system can be stable for eccentricities up to 0.19 for the full range of masses explored for the planetary candidate. For larger eccentricities, the system becomes unstable for any of the masses within our explored grid. This analysis allowed us to conclude that despite the unusual architecture of the two-planet solution proposed here, consisting of a hot Jupiter and an inner hot Neptune, the system has ample room for stable configurations, resembling the WASP-84 system, where recently, an inner mini-Neptune has been found accompanying the previously known hot Jupiter WASP-84 b (Maciejewski et al. 2023).
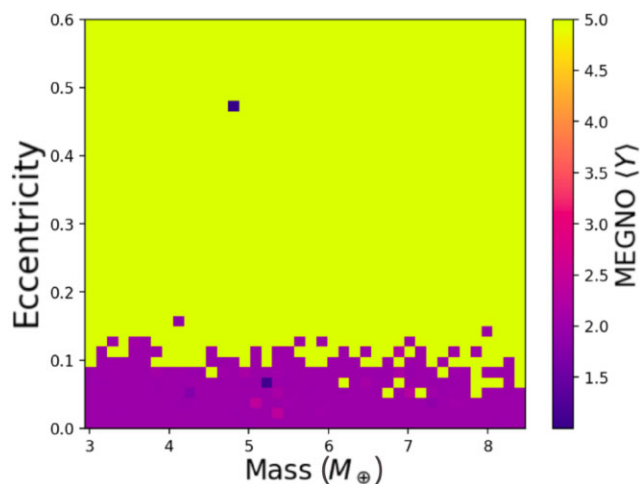
Figure 7. A stability map similar to the one displayed in Fig. 6 but for the system HAT-P-26 composed of the two planets: HAT-P-26 b and the planetary candidate CTOI-420779000.02.

3.3 HAT-P-26

HAT-P-26 is a K1 star with a $T_{eff} = 5079 \pm 88$ K, $R_{\star} = 0.79^{+0.10}_{-0.04}$ R$_{\odot}$, and $M_{\star} = 0.82 \pm 0.03$ M$_{\odot}$ observed nightly by the HAT-5 and HAT-6 telescopes of the HATNet project between 2009 January and August, which hinted at the presence of a transiting planet with an orbital period of 4.23 d. The planetary origin of this signal was later confirmed using the HIRES spectrograph on the Keck I telescope from 2009 December to 2010 June. All these observations together allowed the confirmation of the low-density Neptune-mass planet HAT-P-26 b (Hartman et al. 2011) with a size of 0.565 R_{Jup}, a mass of 0.059 M_{Jup}, and an orbital period of 4.23 d.

This star was observed by the *TESS* mission in Sector 50 in 2022 May, using the 120 and 600 s cadences (TIC-420779000). During our search with SHERLOCK, we fully recovered the signal corresponding to HAT-P-26 b with S/N = 42.2 and SDE = 15.6 in the first run. During the second run, SHERLOCK spotted an interesting signal with a period of 6.59 d. While this signal was slightly below our detection threshold, with S/N = 4.9 and SDE = 7.4, it was confidently found in all the light curves. This motivated us to run the vetting module, where we did not find any solid potential systematic origin that explained this signal; however, it is important to note that due to its very low S/N, some of the tests were not conclusive. Still, we decided to conduct the statistical validation, which yielded FPP = 0.24 and NFPP = 0.0075, falling in the likely planet area.

We then executed the fitting module to refine the planet parameters and compare the Bayes factor of the two models as for the other candidates. We found $\Delta \ln Z = \Delta \ln Z_{2\text{-planet}} - \Delta \ln Z_{1\text{-planet}} \sim 4.7$, which slightly favours the two-planet scenario over the single-planet model. The candidate, named CTOI-420779000.02, would correspond to a mini-Neptune of ~ 2.0 R$_{\oplus}$ (see planet parameters in Table 3), producing shallow transits of only ~ 0.6 ppt, which hinders any confirmation using ground-based telescopes. However, its confirmation would be feasible, as stated for the CTOI-46096489.02 in the WASP-16 system, by employing space-based telescopes such as *CHEOPS*. The uncertainty in the transit time is $\Delta T \sim 22.7$ h, which is too large for time-critical observations, and unfortunately, according to the *TESS*-point web tool, this star will not be revisited in future cycles, preventing us from reducing this uncertainty using upcoming *TESS* data. However, the orbital period of this candidate of only 6.59 d would favour a filler follow-up camping to recover the

planet by accumulating observations. This strategy is also applied in the FATE project using *CHEOPS* observations (Van Grootel et al. 2021) to find unknown transiting planets with orbital periods shorter than 10 d. Following this strategy, it would be possible to confirm or refute the existence of this planet.

The two-planet model is displayed in Fig. 8, the posterior distribution is provided in Appendix C, Table C1, and the convergence of the model is visualized in Fig. C3.

As for HAT-P-27, the candidate's orbital period was close to the known planet HAT-P-26 b. That architecture might be physically impossible and discarded by exploring the system stability. We, hence, followed the same strategy as for the HAT-P-27 system, building a stability map to explore 1600 scenarios in the e_p-M_p parameter space, displayed in Fig. 7. In this analysis, the candidate's eccentricity must be lower than 0.05 to ensure the system's stability for all the planetary masses explored, while the scenarios with the candidate's mass below 6 M$_{\oplus}$ show stable solutions for eccentricities up to 0.09 in most scenarios.

A particularly interesting point of the HAT-P-26 system is that von Essen et al. (2019) reported the detection of a drift in the spectroscopic data alongside a weak curvature in the timing residuals in the photometric data. The authors conducted a follow-up of the planet and obtained a ~ 4 min amplitude of transit timing variations (TTVs). While the authors prudently did not attribute these TTVs to a planetary companion, they gathered enough information to rule out the spot-induced origin of such TTVs. In a more recent study, A-thano et al. (2023) revisited the transit timings of HAT-P-26 b over 7 yr, reporting a TTV amplitude of ~ 2 min, which might be attributable to the presence of an additional planet in the system of ~ 6.36 M$_{\oplus}$ with an orbital period of 8.47 d, i.e. in a first-order mean-motion resonance 2:1. However, there is not any hint of a transiting planet under this scenario. On the contrary, our solution points to the transiting nature of a potential planetary candidate of ~ 1.96 R$_{\oplus}$ (predicted mass from mass–radius relationships of $4.8^{+3.5}_{-2.0}$ M$_{\oplus}$) with an orbital period of 6.59 d, i.e. within ~ 6 per cent of the first-order resonance 3:2. The findings presented here can significantly encourage the scientific community to continue monitoring the HAT-P-26 system, to ultimately unravel its planetary architecture.

3.4 TOI-2411

TOI-2411 (TIC-10837041) is a K7 dwarf star with $T_{eff} = 4099 \pm 123$ K, $R_{\star} = 0.68 \pm 0.02$ R$_{\odot}$, and $M_{\star} = 0.65 \pm 0.02$ M$_{\odot}$. This star was monitored by the *TESS* mission in Sector 3 (2018 October) using 1800 s cadence and in Sector 30 (2020 October) using 120 and 600 s cadences. Orbiting TOI-2411 was alerted in 2020 November a planetary candidate, which was validated by Giacalone et al. (2022) by combining high-resolution images collected using the SOAR/HRCam, Palomar/PHARO, and Keck/NIRC2 instruments, ground-based multiband photometry (i' and r') taken employing the 1-m Sinistro telescope at Las Cumbres Observatory and the 0.7-m CDK telescope at the Mt. Kent Observatory, and statistical arguments. The planet, TOI-2411 b, is a super-Earth of 1.68 ± 0.11 R$_{\oplus}$ in an ultrashort period orbit of 0.78 d.

During our SHERLOCK search, we found in the first run a signal of 18.75 d with an S/N = 14.9 and SDE = 14.1 in 9 of the 11 light curves processed. In the second run, we found a signal corresponding to TOI-2411 b with S/N = 8.4 and SDE = 16.7 in all the light curves. We executed the vetting module for the 18.75 d signal and found one red flag that might indicate a false positive: a momentum dump near the first transit. Momentum dumps might originate flux dimmings that can confuse the search algorithm; however, their shapes do

not necessarily resemble transit-like features. We carefully visually inspected the two events in the 120 s data, and they both look similar in shape, duration, and depth. Moreover, during the vetting, was also inspected the 600 s cadence, and the signal corresponding to the 18.75 d was also found. Hence, we concluded that the momentum dump, or any other systematic, was not the origin of the transit-like feature detected by SHERLOCK. Next, we conducted the statistical validation, which yielded values of FPP = 0.022 and NFPP = 0.0, falling in the likely planet area.

We then performed the fitting analysis and compared the two models as described for the previous candidates. We found $\Delta \ln Z = \Delta \ln Z_{\text{2-planet}} - \Delta \ln Z_{\text{1-planet}} \sim 66.1$, which highly supports the two-planet model. The candidate, CTOI-10837041.02, is found to be a mini-Neptune of 2.88 $R_\oplus$, whose transit depth is estimated to be $\sim$1.77 ppt (see Table 3), which can be detected using small- and medium-size telescopes. However, its moderately long orbital period of 18.75 d and estimated uncertainty in the transit time of $\Delta T \sim 6.7$ h make it very challenging to confirm its planetary nature utilizing ground-based facilities. Unfortunately, *TESS* will not revisit this star, and hence, as in the case of CTOI-420779000.02 (HAT-P-26 system), a filler program using *CHEOPS* would be a potential solution to recover and confirm the candidate as a genuine planet.

The two-planet model is displayed in Fig. 8, the posterior distribution is provided in Appendix C, Table C1, and the convergence of the model is visualized in Fig. C4.

4 THE HIDDEN GEMS PROJECT

Nearby and low-mass stars ($M < 0.6 \, M_\odot$) represent the best opportunity nowadays to study in great detail small transiting planets ($R_p < 4 \, R_\oplus$). Indeed, these stars ease the characterization of small exoplanets, not only in terms of radius and mass but also in terms of internal structure and atmospheric properties, which encodes information about physical and chemical properties, in turn providing insights into planets' formation and evolutionary history (see e.g. Pudritz, Cridland & Alessi 2018; Triaud 2021).

Summarizing, there are three critical points that support this reasoning. (1) their small sizes and lower masses compared to solar-like stars produce larger transit depths and induce larger radial velocities for a given planet, allowing for an accurate determination of planetary bulk properties, such as the radius and mass, with precisions better than $\sim$4 and $\sim$25 per cent, respectively. This allows us to derive the mean density and give a rough estimation of the bulk composition and internal structure (Otegi et al. 2020). This advantage has been the driver of many transits and Doppler surveys such as MEarth (Nutzman & Charbonneau 2008), SPECULOOS (Sebastian et al. 2021), CARMENES (Quirrenbach et al. 2014), and SPIRou (Artigau et al. 2014), to name a few. (2) Atmospheres of planets orbiting these small stars can be explored using transit spectroscopy, which relies on the chromatic signature of the chemical species present in the atmosphere, detectable with current and future space-based observatories such as *JWST* and *Ariel*. Transit spectroscopy allows three configurations to study planetary atmospheres: transmission spectroscopy, emission spectroscopy, and phase curves (Madhusudhan 2019). The possibility of atmospheric characterization of small planets with any of these three techniques has been the drive behind the detection of new planets around such stars, notably thanks to the *TESS* and *Kepler/K2* missions. It has even now become one of the main science goals of the community, with its importance recognized at the highest level with a recommendation by the *JWST* Exoplanet Working Group to dedicate over 500 h of

Director's Discretionary Time to this end.[11] (3) Low-mass stars are the most common in the solar neighbourhood (Winters et al. 2015), representing up to 70 per cent of the stellar content of our Galaxy. Moreover, they have a lower binary fraction than solar-like stars. In addition, computational simulations, as well as Doppler and transit surveys, have concluded that the fraction of stars hosting small planets increases with decreasing mass (see e.g. Bonfils et al. 2013; Dressing & Charbonneau 2015; Alibert & Benz 2017). By combining the high frequency of low-mass stars with the occurrence rate of small planets orbiting them, there might exist $\sim$10$\times$ more small planets orbiting these stars than around solar analogues. Hence, planets discovered orbiting low-mass stars are more representative of planet formation processes than planets orbiting single Sun-like stars (Triaud 2021).

Of particular interest are these systems hosted by a low-mass star where more than one transiting planet exists. In these cases, by studying different planets that were born in the same protoplanetary disc, we can get information not only about the planets but also about the formation and evolution of the whole planetary system. By aggregating information gleaned from the analysis of numerous multiplanetary systems, we will be able to formulate a robust planetary formation model that can resolve existing degeneracies and overcome the limitations currently constraining our understanding (see e.g. Raymond & Morbidelli 2022). This comprehensive model will eventually enable us to systematically decipher the underlying processes that govern planetary formation and evolution across different environments, unifying disparate observations and refining our predictions regarding the distribution and characteristics of exoplanets.

The full sample of transiting exoplanets around low-mass stars keeps on growing and now exceeds 1400[12] planets with 85 per cent of them being smaller than Neptune. As the race to detect new small planets amenable for detailed studies is still ongoing, we present the Hidden Gems project, which relies on the capabilities of SHERLOCK to discover extra planets in low-mass star systems hosting at least one transiting planet.

Our target list is constructed by filtering stars with $T_{\text{eff}} < 5300$ K to include K and M dwarf stars and select a volume-limited sample within a distance of 50 pc, which totals 167 transiting planets in 107 unique systems at the time of writing. However, this catalogue will be updated with the upcoming releases of newly discovered planetary systems. Occurrence rate studies show that M dwarfs host at least two planets on average (e.g. Dressing & Charbonneau 2015; Tuomi et al. 2019), this corresponds to twice the number reported by Kunimoto & Matthews (2020) for K dwarfs. It implies, on a statistical level, that there could still be additional planets in these systems. Moreover, a transiting system has a favourable orbital inclination that enhances the likelihood of finding extra planet candidates. *TESS* offers an exquisite opportunity, thanks to its long baselines of observations, to search for concealed low S/N signals that might have been unnoticed. The search will be conducted on each system's available *TESS* sectors and will be updated as soon as new data becomes public. This strategy increases the S/N of the putative detections, reaching smaller planetary radii and longer period candidates, as shown in Fig. 9.

---

[11]https://www.stsci.edu/files/live/sites/www/files/home/hst/about/space-telescope-users-committee/presentations-and-documentation/_documents/2023_nov/exoplanets-DDWG.pdf
[12]NASA Exoplanet Archive (2024 February 20) https://exoplanetarchive.ipac.caltech.edu/.
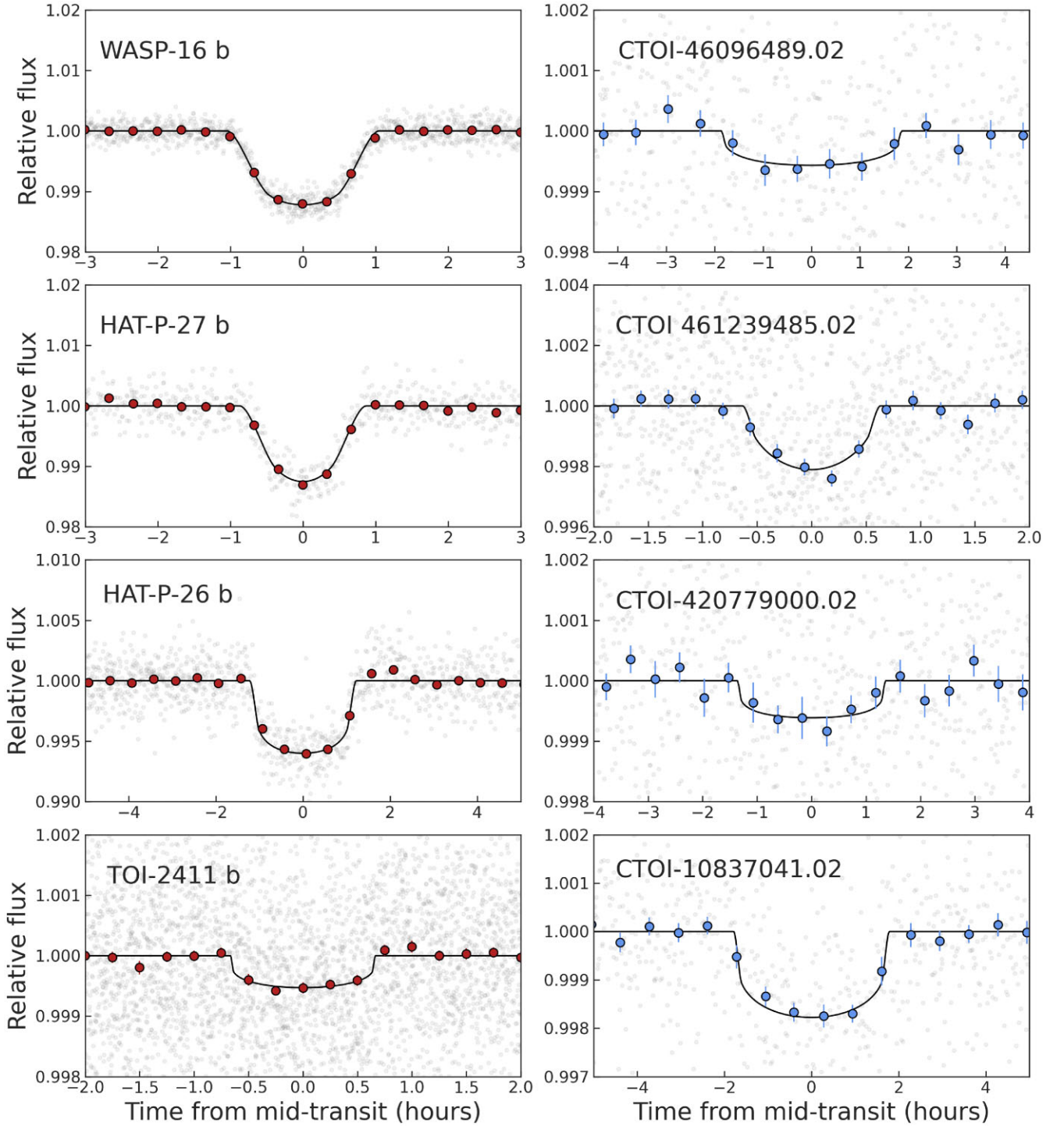
Figure 8. Detrended *TESS* photometry phase folded to the periods of the known planets (left column) and planetary candidates (right columns), along with the best-fitting transit models. The unbinned data points are depicted together to the circles with error bars corresponding to binned data adapted in each case to ease the visualization. From top to bottom, the WASP-16, the HAT-P-27, the HAT-P-26, and the TOI-2411 systems. The posterior distributions of these fits are provided in the Appendix C and the corner plots in Figs C1, C2, C3, and C4, respectively.

By developing this project, we will eventually be able to pinpoint the presence of small transiting planets in the habitable zones of their stars within the solar neighbourhood, setting the stage for detailed studies. Hence, the ultimate objective of this project is to compile an accurate census of nearby transiting potential habitable planets.

5 CONCLUSIONS

Providing open-sourced software packages is a valuable contribution to the scientific community, especially in cases where the documentation is abundant, detailed, and focused on hands-on examples. With SHERLOCK, we are not only doing that but also releasing a mature pipeline standardizing scientific procedures that increase
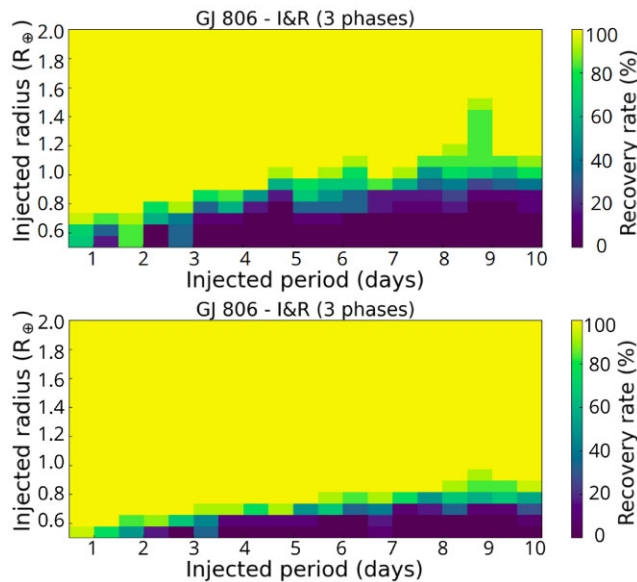
Figure 9. Improvements between the injection and recovery scenarios for a simple search on GJ 806 using different number of sectors. Top: Results of injection and recovery scenarios only using Sector 41. Bottom: The same retrieval scenarios using Sectors 41, 55, and 56.

the scientist's productivity, help the reproducibility of the end-to-end processes, and empower individual teams to conduct their own planetary searches relying on space-based missions.

SHERLOCK represents an advanced exoplanetary search pipeline that leverages a suite of robust and community-validated astrophysical packages including LIGHTKURVE, ELEANOR, TLS, TRICERATOPS, and ALLESFITTER. These are complemented by bespoke modules tailored to create a streamlined and integrated workflow. Initially, SHERLOCK systematically acquires and pre-processes available data, setting the stage for effective exoplanet detection. Subsequent phases involve rigorous vetting of candidates to identify and eliminate potential false positives through comprehensive analysis and statistical validation techniques. The pipeline further refines estimates of planetary characteristics through sophisticated data modelling, enhancing the precision of the derived orbital and physical parameters. SHERLOCK concludes the analyses by determining the best observational windows, crucial for planning ground-based confirmatory follow-ups. Each phase is encapsulated within dedicated modules, optimizing the accuracy and efficiency of the whole process of finding reliable planetary candidates. Still, we highlight that SHERLOCK may need manual fine-tuning to operate to its fullest potential, making it not an entire automated search pipeline, requiring visual inspection of detected signals before passing through all the modules.

Our performance test showed that SHERLOCK has an excellent recall of 0.98 when recovering SPOC official alerts confirmed as planets, which makes it a very reliable pipeline for the community. During this test, we not only recovered 98 per cent of the planets but also found some candidates that were confirmed planets but did not have any associated alerts, such as TOI-942 c and Kepler-89 c, strengthening the conclusion of being a reliable planetary searching tool. On top of that, we found four new planetary candidates unnoticed until now that fulfilled all our tests, only lacking ground-based observation to confirm their planetary nature fully. These four candidates were found to be in the planetary systems WASP-16 (CTOI-46096489.02), HAT-P-27 (CTOI-461239485.02), HAT-P-26 (CTOI-420779000.02), and TOI-2411 (CTOI-10837041.02).

The vetting implemented by SHERLOCK, based on the robust tests conducted by the SPOC pipeline, reports a variety of metrics that are marked as red flags whenever they are found to be above specific thresholds. However, low data quality sometimes produces inconclusive results for these metrics, especially on the centroids and optical ghost diagnoses. Hence, in these cases, the red flags have to be taken as warning bells instead of a robust indication of a false positive, which needs further exploration by the user. SHERLOCK's vetting will be further developed in coming releases towards a more robust data exploration, producing a more reliable set of metrics even for very low S/N signals. In particular, we are currently exploring implementing a deep learning neural network similar to ExoMiner (Valizadegan et al. 2022), allowing us to discern genuine planets from false positives unbiasedly.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Alibert Y., Benz W., 2017, A&A, 598, L5
Aller A., Lillo-Box J., Jones D., Miranda L. F., Barceló Forteza S., 2020, A&A, 635, A128
Artigau É. et al., 2014, in Ramsay S. K., McLean I. S., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V. SPIE, Bellingham, p. 914715
Astropy Collaboration, 2013, A&A, 558, A33
Astropy Collaboration, 2018, AJ, 156, 123
A-thano N. et al., 2023, AJ, 166, 223
Bakos G. Á., Hartman J. D., Torres G., Kovács G., Noyes R. W., Latham D. W., Sasselov D. D., Béky B., 2011, Eur. Phys. J. Web Conf., 11, 01002
Barceló Forteza S., Michel E., Roca Cortés T., García R. A., 2015, A&A, 579, A133
Barros S. C. C., Demangeon O., Deleuil M., 2016, A&A, 594, A100
Bashi D., Helled R., Zucker S., Mordasini C., 2017, A&A, 604, A83
Béky B. et al., 2011, ApJ, 734, 109
Benz W. et al., 2021, Exp. Astron., 51, 109
Bonfils X. et al., 2013, A&A, 549, A109
Borucki W. J. et al., 2003, in Blades J. C., Siegmund O. H. W., eds, Proc. SPIE Conf. Ser. Vol. 4854, Future EUV/UV and Visible Space Astrophysics Missions and Instrumentation. SPIE, Bellingham, p. 129
Borucki W. J. et al., 2010, Science, 327, 977
Bryson S. T. et al., 2013, PASP, 125, 889
Bryson S. T. et al., 2017, in Jenkins J. M., ed., Kepler Data Processing Handbook: KSCI-19081-002. NASA Ames Research Center, Moffett Field, CA, p. 35

Burdanov A., Delrez L., Gillon M., Jehin E., The SPECULOOS and TRAPPIST teams, 2018, in Deeg H., Belmonte J., eds, Handbook of Exoplanets. Springer, Cham, Switzerland, p. 1007

Burdanov A. Y. et al., 2022, PASP, 134, 105001

Caldwell D. A. et al., 2010, ApJ, 713, L92

Carleo I. et al., 2021, A&A, 645, A71

Carter J. A., Agol E., 2013, ApJ, 765, 132

Charpinet S. et al., 2010, A&A, 516, L6

Chen J., Kipping D., 2017, ApJ, 834, 17

Cincotta P., Simó C., 1999, Celest. Mech. Dyn. Astron., 73, 195

Cincotta P. M., Simó C., 2000, A&AS, 147, 205

Cincotta P. M., Giordano C. M., Simó C., 2003, Physica D: Nonlinear Phenomena, 182, 151

Cloutier R., 2019, AJ, 158, 81

Coughlin J. L. et al., 2016, ApJS, 224, 12

Cranmer M., Tamayo D., Rein H., Battaglia P., Hadden S., Armitage P. J., Ho S., Spergel D. N., 2021, Proc. Natl. Acad. Sci., 118, e2026053118

Crossfield I. J. M. et al., 2019, ApJ, 883, L16

Dawson R. I., Johnson J. A., 2018, ARA&A, 56, 175

Delrez L. et al., 2018, in Marshall H. K., Spyromilio J., eds, Proc. SPIE Conf. Ser. Vol. 10700, Ground-based and Airborne Telescopes VII. SPIE, Bellingham, p. 107001I

Delrez L. et al., 2021, Nat. Astron., 5, 775

Delrez L. et al., 2022, A&A, 667, A59

Demory B. O. et al., 2020, A&A, 642, A49

Díaz R. F., Almenara J. M., Santerne A., Moutou C., Lethuillier A., Deleuil M., 2014, MNRAS, 441, 983

Dransfield G. et al., 2024, MNRAS, 527, 35

Dreizler S. et al., 2024, A&A, 684, A117

Dressing C. D., Charbonneau D., 2015, ApJ, 807, 45

Espinoza N., Kossakowski D., Brahm R., 2019, MNRAS, 490, 2262

Feinstein A. D. et al., 2019, PASP, 131, 094502

Feliz D. L., Plavchan P., Bianco S. N., Jimenez M., Collins K. I., Villarreal Alvarado B., Stassun K. G., 2021, AJ, 161, 247

Foreman-Mackey D. et al., 2021, J. Open Source Softw., 6, 3285

Giacalone S. et al., 2021, AJ, 161, 24

Giacalone S. et al., 2022, AJ, 163, 99

Gillon M., 2018, Nat. Astron., 2, 344

Gillon M. et al., 2017, Nature, 542, 456

Gillon M. et al., 2020, BAAS, 52, 0208

Gillon M. et al., 2024, Nat. Astron., in press (arXiv:2406.00794)

Günther M. N., Daylan T., 2019, Astrophysics Source Code Library, record ascl:1903.003

Günther M. N., Daylan T., 2021, ApJS, 254, 13

Günther M. N. et al., 2019, Nat. Astron., 3, 1099

Hartman J. D. et al., 2011, ApJ, 728, 138

Hippke M., Heller R., 2019, A&A, 623, A39

Hippke M., David T. J., Mulders G. D., Heller R., 2019, AJ, 158, 143

Huang C. X. et al., 2020, Res. Notes Am. Astron. Soc., 4, 204

Jenkins J. M. et al., 2016, in Chiozzi G., Guzman J. C., eds, Proc. SPIE Conf. Ser. Vol. 9913, Software and Cyberinfrastructure for Astronomy IV. SPIE, Bellingham, p. 99133E

Jenkins J. S., Pozuelos F. J., Tuomi M., Berdiñas Z. M., Díaz M. R., Vines J. I., Suárez J. C., Peña Rojas P. A., 2019, MNRAS, 490, 5585

Jensen E., 2013, Astrophysics Source Code Library, record ascl:1306.007

Jones E. et al., 2001, SciPy: Open Source Scientific Tools for Python. Available at: http://www.scipy.org/

Kass R. E., Raftery A. E., 1995, J. Am. Stat. Assoc., 90, 773

Kostov V. B. et al., 2019a, AJ, 157, 124

Kostov V. B. et al., 2019b, AJ, 158, 32

Kovács G., Zucker S., Mazeh T., 2002, A&A, 391, 369

Kruse E., Agol E., Luger R., Foreman-Mackey D., 2019, ApJS, 244, 11

Kunimoto M., Matthews J. M., 2020, AJ, 159, 248

Lightkurve Collaboration, 2018, Astrophysics Source Code Library, record ascl:1812.013

Lissauer J. J. et al., 2012, ApJ, 750, 112

Lister T. A. et al., 2009, ApJ, 703, 752

Luger R., Agol E., Kruse E., Barnes R., Becker A., Foreman-Mackey D., Deming D., 2016, AJ, 152, 100

Luger R., Kruse E., Foreman-Mackey D., Agol E., Saunders N., 2018, AJ, 156, 99

Maciejewski G., Golonka J., Łoboda W., Ohlert J., Fernández M., Aceituno F., 2023, MNRAS, 525, L43

Madhusudhan N., 2019, ARA&A, 57, 617

Mallonn M. et al., 2019, A&A, 622, A81

Marcy G. W. et al., 2014, ApJS, 210, 20

Mayo A. W. et al., 2018, AJ, 155, 136

Melton E. J., Feigelson E. D., Montalto M., Caceres G. A., Rosenswie A. W., Abelson C. S., 2024, AJ, 167, 202

Montalto M., 2023, MNRAS, 518, L31

Montalto M. et al., 2020, MNRAS, 498, 1726

Morton T. D., 2012, ApJ, 761, 6

Morton T. D., Bryson S. T., Coughlin J. L., Rowe J. F., Ravichandran G., Petigura E. A., Haas M. R., Batalha N. M., 2016, ApJ, 822, 86

Murgas F. et al., 2023, A&A, 677, A182

Nutzman P., Charbonneau D., 2008, PASP, 120, 317

Ofir A., Dreizler S., 2013, A&A, 555, A58

Otegi J. F., Dorn C., Helled R., Bouchy F., Haldemann J., Alibert Y., 2020, A&A, 640, A135

Peterson M. S. et al., 2023, Nature, 617, 701

Pollacco D. L. et al., 2006, PASP, 118, 1407

Pozuelos F. J. et al., 2020, A&A, 641, A23

Pozuelos F. J. et al., 2023, A&A, 672, A70

Pudritz R. E., Cridland A. J., Alessi M., 2018, in Deeg H., Belmonte J., eds, Handbook of Exoplanets. Springer, Cham, Switzerland, p. 2475

Quinn S. N. et al., 2019, AJ, 158, 177

Quirrenbach A. et al., 2014, in Ramsay S. K., McLean I. S., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V. SPIE, Bellingham, p. 91471F

Raymond S. N., Morbidelli A., 2022, in Biazzo K., Bozza V., Mancini L., Sozzetti A., eds, Demographics of Exoplanetary Systems. Springer, Cham, Switzerland, p. 3

Rein H., Liu S. F., 2012, A&A, 537, A128

Rein H., Tamayo D., 2015, MNRAS, 452, 376

Ricker G. R. et al., 2016, in MacEwen H. A., Fazio G. G., Lystrup M., Batalha N., Siegler N., Tong E. C., eds, Proc. SPIE Conf. Ser. Vol. 9904, Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 99042B

Sahoo S. K. et al., 2020, MNRAS, 495, 2844

Savitzky A., Golay M. J. E., 1964, Anal. Chem., 36, 1627

Schanche N. et al., 2022, A&A, 657, A45

Sebastian D. et al., 2021, A&A, 645, A100

Thompson S. E. et al., 2018, ApJS, 235, 38

Thuillier A., Van Grootel V., Dévora-Pajares M., Pozuelos F. J., Charpinet S., Siess L., 2022, A&A, 664, A113

Torres G. et al., 2011, ApJ, 727, 24

Triaud A. H. M. J., 2021, in Madhusudhan N., ed., ExoFrontiers. IoP Publishing, Bristol, UK, p. 6–1

Tuomi M. et al., 2019, preprint (arXiv:1906.04644)

Twicken J. D. et al., 2018, PASP, 130, 064502

Valizadegan H. et al., 2022, ApJ, 926, 120

Vanderburg A., Johnson J. A., 2014, PASP, 126, 948

Van Grootel V. et al., 2021, A&A, 650, A205

von Essen C. et al., 2019, A&A, 628, A116

Weiss L. M. et al., 2013, ApJ, 768, 14

Wells R. D. et al., 2021, A&A, 653, A97

Winters J. G. et al., 2015, AJ, 149, 5

Wu D.-H., Rice M., Wang S., 2023, AJ, 165, 171

Zink J. K., Hardegree-Ullman K. K., Christiansen J. L., Dressing C. D., Crossfield I. J. M., Petigura E. A., Schlieder J. E., Ciardi D. R., 2020, AJ, 159, 154

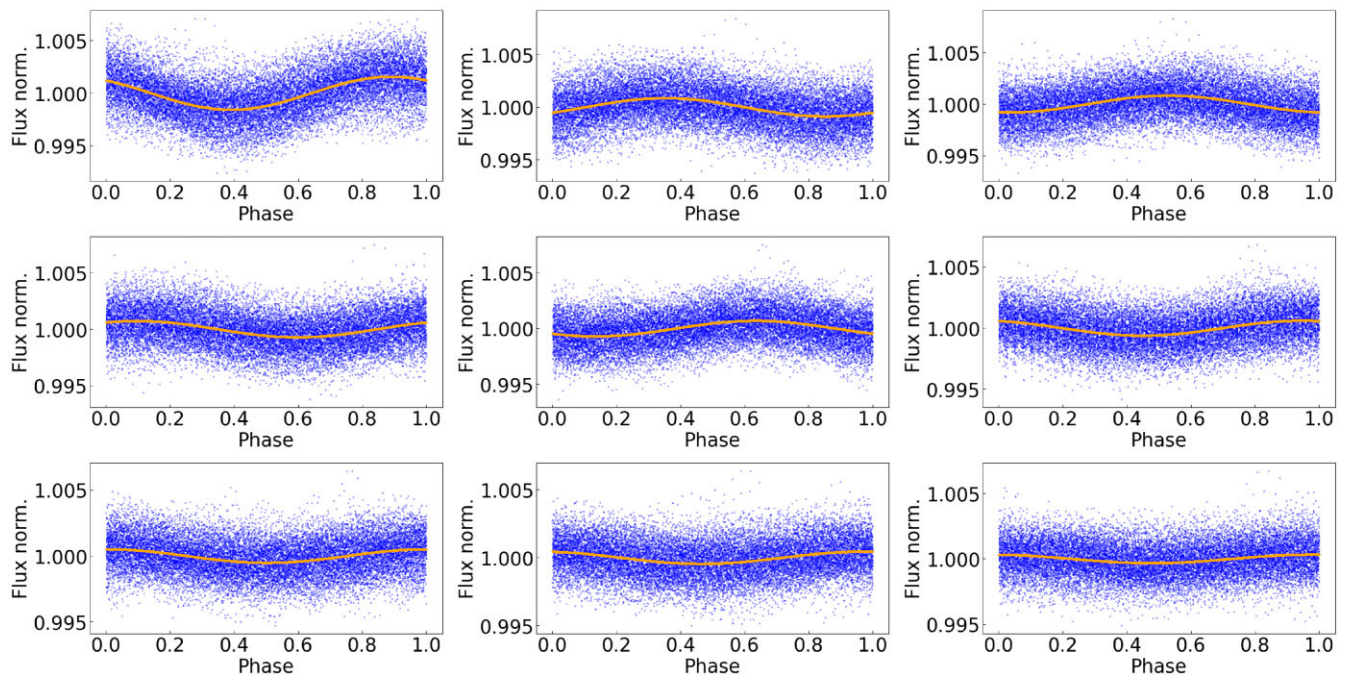Zong W., Charpinet S., Vauclair G., Giammichele N., Van Grootel V., 2016, A&A, 585, A22

APPENDIX A: TIC 169285097 sspe FIT SIGNALS

Figure A1. Output from SHERLOCKSSPE algorithm for Sector 2 data of TIC 169285097. From top-left to bottom-right, the first nine pulsations fits are represented after the previous ones have already been subtracted. The phased flux is represented with its best sinusoidal fit on each panel.

APPENDIX B: PERFORMANCE SURVEY CATALOGUE: TIC IDS

100566492, 102264230, 10837041, 111991770, 112099249, 120247528, 12862099, 136916387, 139375960, 142937186, 144700903, 146520535, 147797743, 157230659, 165297570, 168281028, 172598832, 175772482, 181949561, 188876052, 189380158, 189625051, 20096620, 201186294, 201604954, 206284139, 207077681, 21113347, 229047362, 237320326, 243921117, 245392284, 248853232, 250707118, 257567854, 262715204, 263179590, 269217040, 27039476, 273231214, 281408474, 283419843, 288246496, 291685334, 291751373, 292152376, 305048087, 32487566, 33595516, 336732616, 339522221, 347430350, 354619337, 368739874, 368805700, 369455629, 372172128, 374530847, 375942197, 37718056, 381856447, 391958006, 392476080, 398895470, 39926974, 399402994, 401604346, 420779000, 420814525, 425817867, 429302040, 431701493, 434226736, 436478932, 437704321, 441420236, 452808876, 453064665, 459730973, 460396820, 46096489, 461239485, 464646604, 466382581, 466390120, 467971286, 49043968, 51637609, 53728859, 54002556, 57984377, 6663331, 7088246, 90850770, 91051152, 9155187, 92449173, 9385460, 9443323, 97409519.

APPENDIX C: POSTERIOR DISTRIBUTIONS FROM THE FITS

Table C1. Posterior distributions for WASP-16, HAT-P$-$27, HAT-P$-$26, and TOI-2411 systems.

Parameter	Unit	WASP-16 b	CTOI-46096489.02
$R_p/R_\star$		$0.1150^{+0.0022}_{-0.0019}$	$0.0215^{+0.0017}_{-0.0022}$
$(R_\star + R_p)/a_p$		$0.1266^{+0.0036}_{-0.0040}$	$0.0493^{+0.0040}_{-0.0031}$
$\cos i_p$		$0.0921^{+0.0046}_{-0.0051}$	$0.016^{+0.013}_{-0.010}$
Mid-transit time, $T_{0;p}$	$\mathrm{BJD_{TDB}} - 245\,7000$	1613.66548 ± 0.00020	$1618.725^{+0.012}_{-0.014}$
Orbital period, P_p	d	3.118531 ± 0.000073	$10.457^{+0.018}_{-0.028}$
q_1			$0.60^{+0.23}_{-0.20}$
q_2			$0.35^{+0.28}_{-0.21}$
$\log \sigma_w$	log rel.flux.		-6.5560 ± 0.0055
GP: $\ln \sigma$			$-8.69^{+0.16}_{-0.15}$
GP: $\ln \rho$			$-0.75^{+0.34}_{-0.31}$

Parameter	Unit	HAT-P$-$27 b	CTOI-461239485.02
$R_p/R_\star$		$0.1209^{+0.0050}_{-0.0045}$	$0.0457^{+0.0042}_{-0.0040}$
$(R_\star + R_p)/a_p$		$0.1069^{+0.0059}_{-0.0063}$	$0.188^{+0.025}_{-0.017}$
$\cos i_p$		0.0767 ± 0.0076	$0.0130^{+0.033}_{-0.031}$
Mid-transit time, $T_{0;p}$	$\mathrm{BJD_{TDB}} - 245\,7000$	2705.8727 ± 0.0004	$2705.5090^{+0.0014}_{-0.0015}$
Orbital period, P_p	d	3.03908 ± 0.000019	$1.19949^{+0.00021}_{-0.00022}$
q_1			$0.75^{+0.16}_{-0.22}$
q_2			$0.67^{+0.21}_{-0.33}$
$\log \sigma_w$	log rel.flux.		-6.0260 ± 0.0073
GP: $\ln \sigma$			$-6.83^{+0.28}_{-0.24}$
GP: $\ln \rho$			0.52 ± 0.31

Parameter	Unit	HAT-P$-$26 b	CTOI-420779000.02
$R_p/R_\star$		$0.0722^{+0.0015}_{-0.0016}$	0.0230 ± 0.0024
$(R_\star + R_p)/a_p$		$0.0821^{+0.0014}_{-0.0013}$	0.0584 ± 0.0010
$\cos i_p$		$0.0331^{+0.0042}_{-0.0049}$	0.022 ± 0.014
Mid-transit time, $T_{0;p}$	$\mathrm{BJD_{TDB}} - 245\,7000$	2678.8923 ± 0.0005	$2683.1598^{+0.0010}_{-0.0014}$
Orbital period, P_p	d	4.23440 ± 0.00031	$6.5938^{+0.0087}_{-0.0069}$
q_1			$0.44^{+0.34}_{-0.23}$
q_2			$0.26^{+0.31}_{-0.16}$
$\log \sigma_w$	log rel.flux.		-6.3901 ± 0.0064
GP: $\ln \sigma$			$-8.78^{+0.19}_{-0.23}$
GP: $\ln \rho$			$-1.92^{+1.3}_{-0.75}$

| Parameter	Unit	TOI-2411 b	CTOI-10837041.02
$R_p/R_\star$		0.021 ± 0.001	$0.0390^{+0.0016}_{-0.0018}$
$(R_\star + R_p)/a_p$		$0.2230^{+0.0060}_{-0.0055}$	$0.02741^{+0.00094}_{-0.00081}$
$\cos i_p$		$0.0334^{+0.017}_{-0.020}$	$0.0115^{+0.0024}_{-0.0030}$
Mid-transit time, $T_{0;p}$	$\mathrm{BJD_{TDB}} - 245\,7000$	$2129.3198^{+0.0012}_{-0.0016}$	$2141.3147^{+0.0034}_{-0.0034}$
Orbital period, P_p	d	$0.78289^{+0.00014}_{-0.00018}$	$18.7496^{+0.0047}_{-0.0041}$
q_1			$0.41^{+0.34}_{-0.23}$
q_2			$0.41^{+0.32}_{-0.25}$
$\log \sigma_w$	log rel.flux.		-6.6475 ± 0.0045
GP: $\ln \sigma$			-8.961 ± 0.089
GP: $\ln \rho$			$-1.49^{+0.40}_{-0.35}$
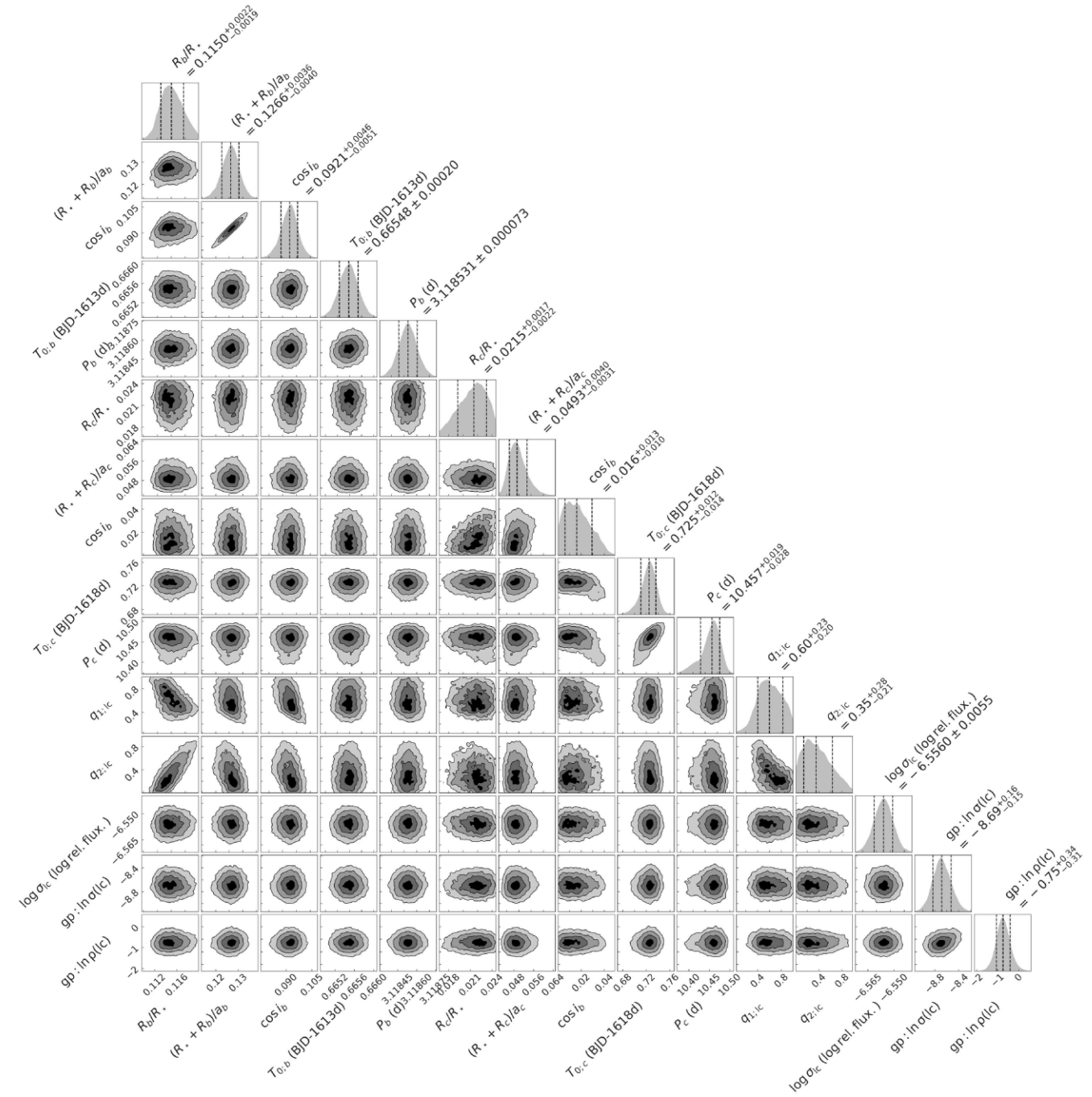
Figure C1. Posterior probability distributions for the fitting parameters used to model the planetary system WASP-16 as described in Section 2.6 considering the two-planet scenario from Section 3.1. The vertical dashed lines represent the median and the 68 per cent credible interval. The figure highlights the correlation (or absence thereof) between all the parameters. The values of each parameter can be consulted in Table C1.
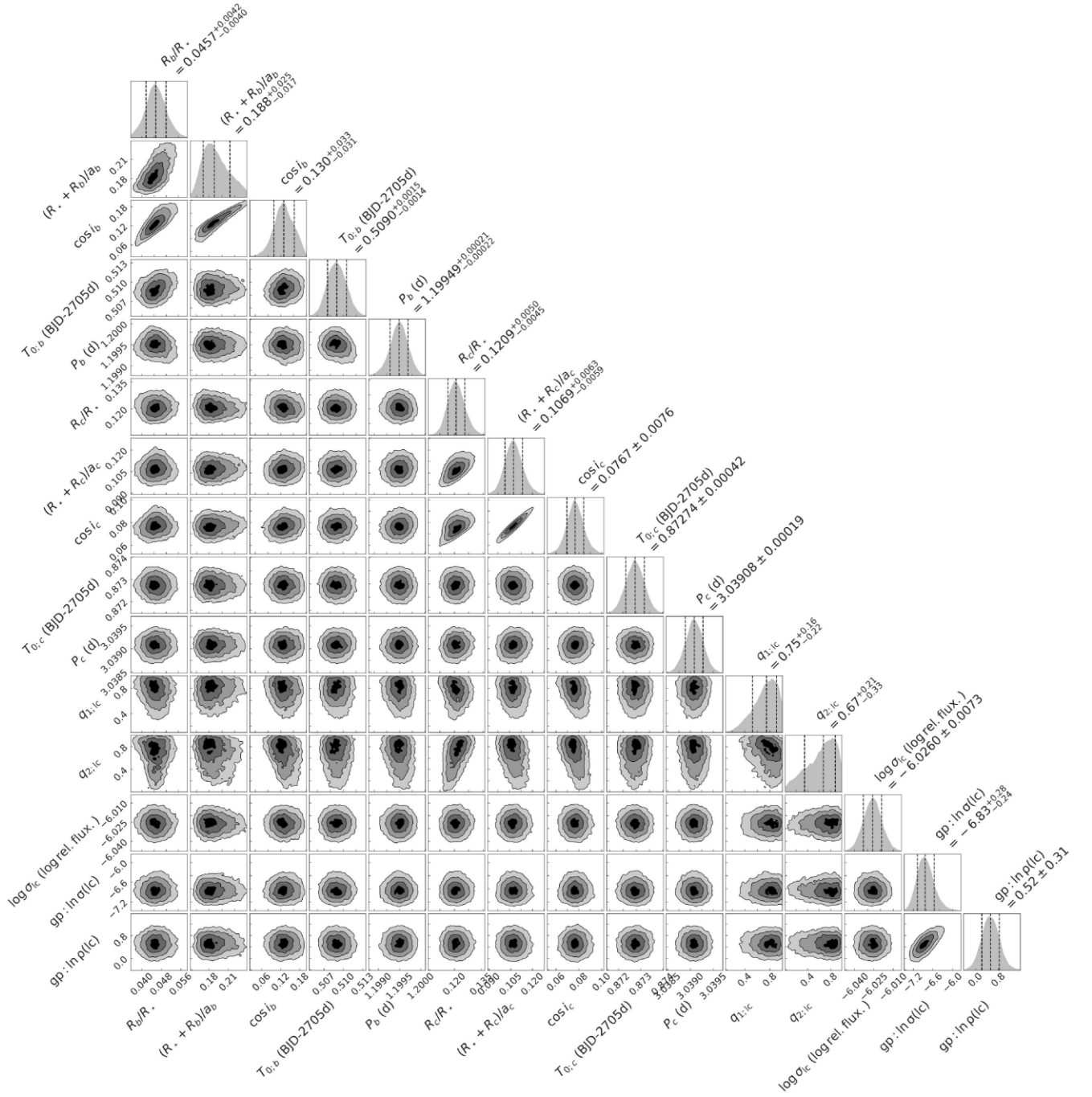
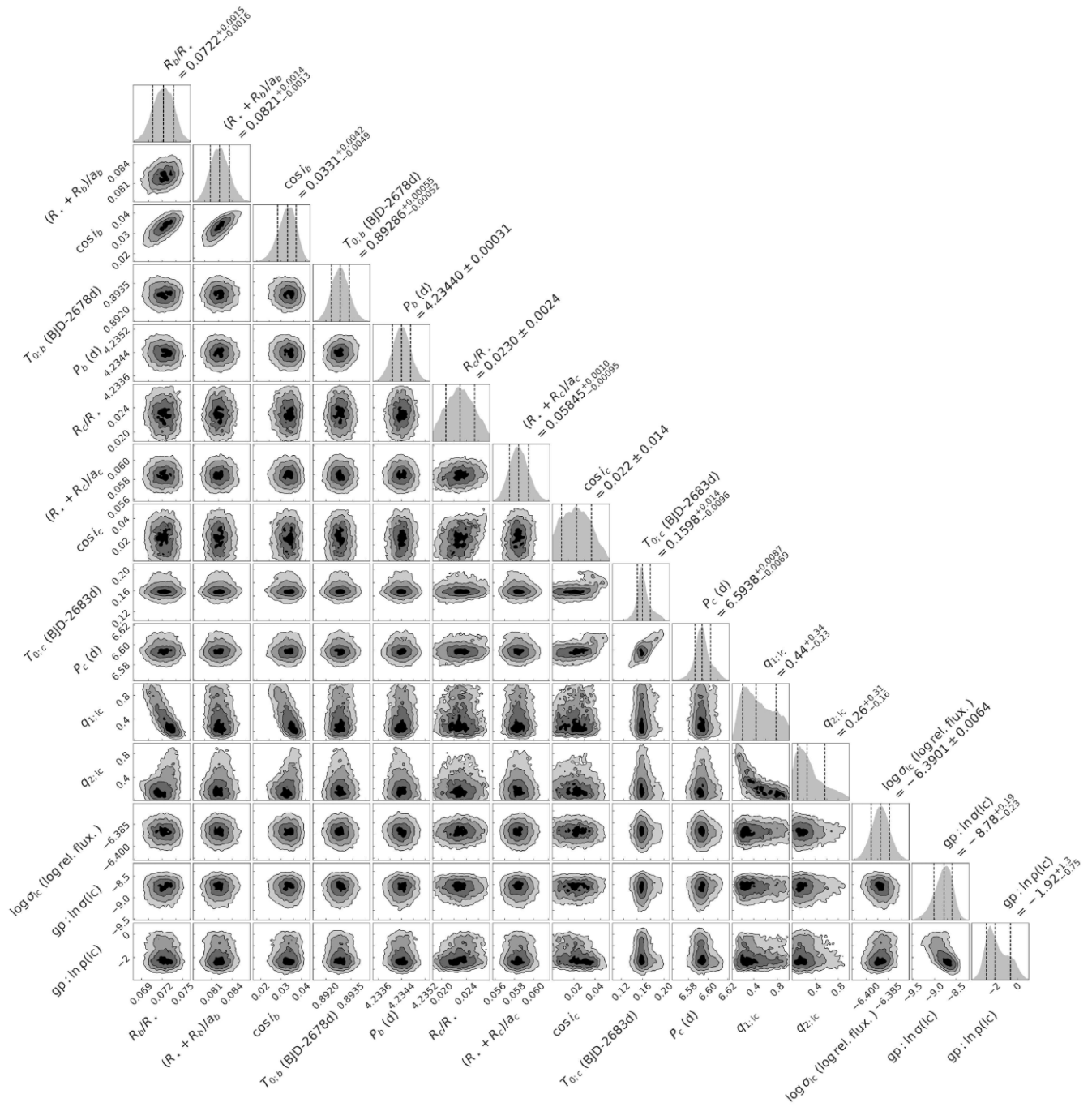**Figure C2.** Similar to Fig. C1, but for the HAT-P-27 system.

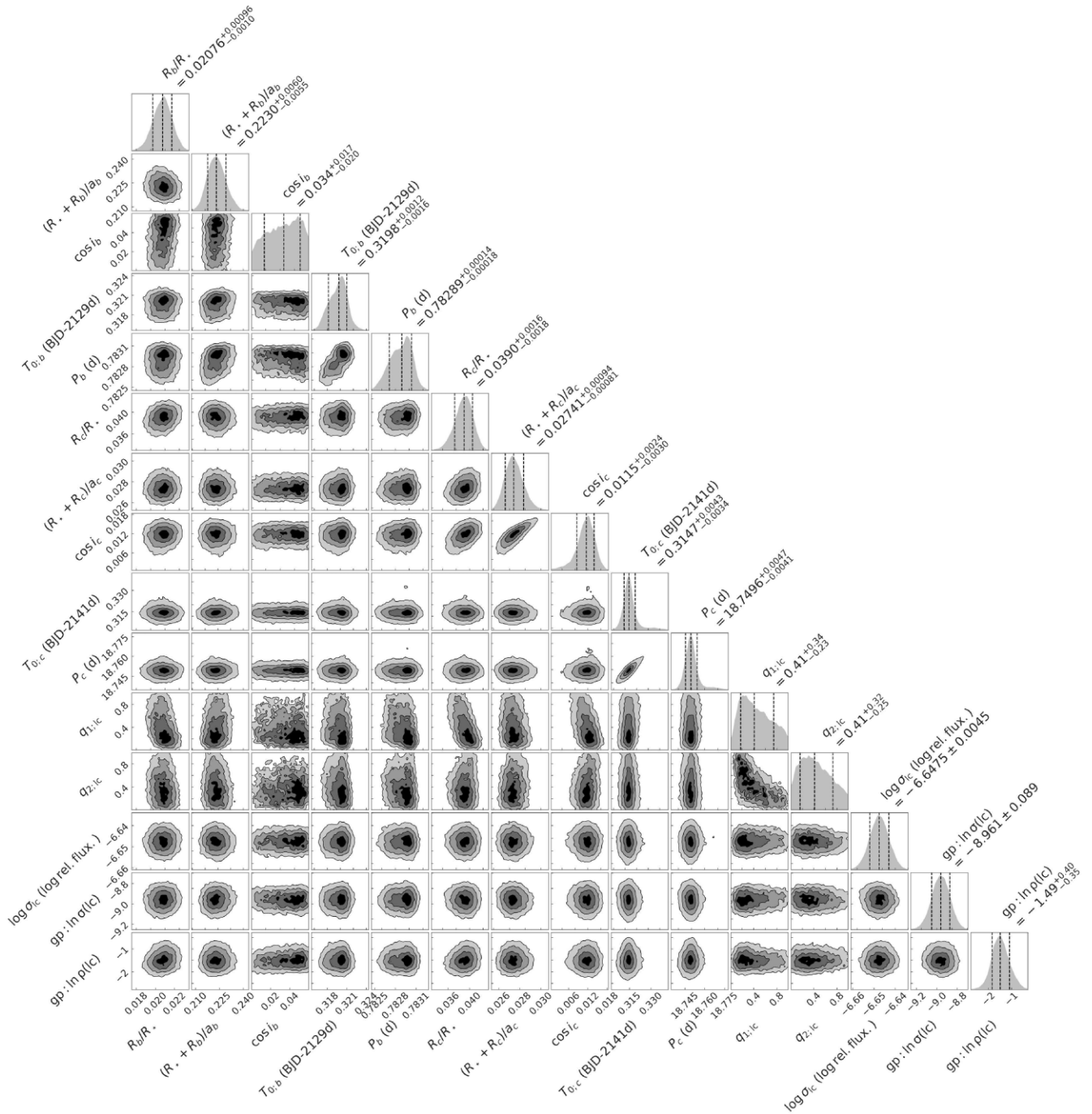**Figure C3.** Similar to Fig. C1, but for the HAT-P-26 system.

Figure C4. Similar to Fig. C1, but for the TOI-2441 system.

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