



HAL
open science

Overcoming the stellar activity impact on exoplanet detection and characterization: application to the follow-up of TESS candidates with the SOPHIE spectrograph

Neda Heidari

► **To cite this version:**

Neda Heidari. Overcoming the stellar activity impact on exoplanet detection and characterization : application to the follow-up of TESS candidates with the SOPHIE spectrograph. Astrophysics [astro-ph]. Université Côte d'Azur; Shahid Beheshti University (Tehran), 2022. English. NNT : 2022COAZ4110 . tel-04043297

HAL Id: tel-04043297

<https://theses.hal.science/tel-04043297>

Submitted on 23 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE DE DOCTORAT

Neda Heidari

Préparée au sein du :

Laboratoire J.-L. Lagrange (UMR 7293) & Shahid beheshti uiversity et Laboratoire d'Astrophysique de Marseille

sur le sujet :

**Surmonter l'impact de l'activité stellaire sur la detection et caractérisation d'exoplanètes :
application au suivi de candidats TESS avec le spectrographe SOPHIE**

Présentée en vue de l'obtention du grade de

docteur en

Sciences Fondamentales et Appliquées

d' université Côte d'Azur

et d' université Shahid Beheshti

Spécialité

Sciences de la planète et de l'univers

Dirigée par : Nematollah Riazi/ Farrokh Vakili

Co-encadrée par : Isabelle Boisse

Soutenu le : 29 novembre 2022

Composition du jury

Présidente du jury

Magali DELEUIL, PROFESSEUR DES UNIVERSITÉS, AMU

Rapporteurs

Anne-Marie LAGRANGE, DIRECTRICE DE RECHERCHE, LESIA

Claude CATALA, DIRECTEUR DE RECHERCHE, LESIA

Examineurs

Lionel BIGOT, CHARGÉ DE RECHERCHE, OCA

Sohrab RAHVAR, PROFESSEUR DES UNIVERSITÉS, SUT

Marzieh FARHANG, ASSISTANT PROFESSOR UNIVERSITÉ, SBU

Monika LENDL, ASSISTANT PROFESSOR UNIVERSITÉ, GENÈVE

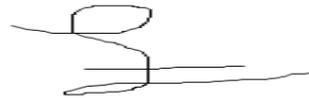
Sadollah NASIRI, PROFESSEUR DES UNIVERSITÉS, SUB

Affidavit

I, undersigned, Neda Heidari, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific direction of Isabelle Boisse, Nematollah Riazi and Farrokh Vakili in accordance with the principles of honesty, integrity and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with both the french/Iranian national charter for Research Integrity and the Cote d'Azur/ Shahid Beheshti University charter on the fight against plagiarism.

This work has not been submitted previously in these countries in the same or in a similar version to any other examination body.

Nice 9 August 2022



Cette œuvre est mise à disposition selon les termes de la [Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/).

List of publications

Publications :

1. **HD88986 : a multi-planet system with a temperate transiting sub-Neptune and a wide-orbit Jupiter mass planet**
[N. Heidari](#), I. Boisse and et al. (2022), **in prep**
2. **BEBOP III. Observations and an independent mass measurement of Kepler-16 (AB) b – the first circumbinary planet detected with radial velocities**
A. Triaud, M. Standing, [N. Heidari](#) and et al. (2022), MNRAS, 511(3), 3561-3570.
3. **HD207897 b : A dense sub-Neptune transiting a nearby and bright K-type star**
[N. Heidari](#) , I. Boisse and et al. (2022) A& A, , 658, A176.

SOPHIE related publications :

1. **The GAPS Programme at TNG XL : A puffy and warm Neptune-sized planet and an outer Neptune-mass candidate orbiting the solar-type star TOI-1422**
L. Naponiello, L. Mancini, ..., [N. Heidari](#) and et al. (2022), arXiv preprint, arXiv : 2207.03293v2, A& A, **Accepted**
2. **The HD 93963, A transiting system : a 1.04 d super-Earth and a 3.65 d sub-Neptune discovered by TESS and CHEOPS**
L. M. Serrano, D. Gandolf, ..., [N. Heidari](#) and et al. (2022), arXiv preprint, arXiv : 2207.13920, A& A, **Accepted**
3. **A warm super-Neptune around the G-dwarf star TOI-1710 revealed with TESS, SOPHIE and HARPS-N**
P.-C. König, M. Damasso, ..., [N. Heidari](#) and et al. (2022), arXiv preprint, arXiv : 2204.08984, A& A, **Accepted**.
4. **The SOPHIE search for northern extrasolar planets. XVII. A wealth of new objects : Six cool Jupiters, three brown dwarfs, and 16 low-mass binary stars**
S. Dalal, F. Kiefer, ..., [N. Heidari](#) and et al. (2021), A& A, 651, A11.
5. **The SOPHIE search for northern extrasolar planets. XVIII : Six new cold Jupiters, including one of the most eccentric exoplanet orbits**
O. D. S. Demangeon, S. Dalal, ..., [N. Heidari](#) and et al. (2021), A& A, 653, A78.

-
6. **TOI-1296b and TOI-1298b observed with TESS and SOPHIE : two hot Saturn-mass exoplanets with different densities around metal-rich stars**
C. Moutou, J. M. Almenara,....., [N. Heidari](#) and et al. (2021), A& A, 653, A147.
 7. **The EBLM project – VIII. First results for M-dwarf mass, radius and effective temperature measurements using CHEOPS light curves**
M. I. Swayne, P. F. L. Maxted,...., [N. Heidari](#) and et al. (2021), MNRAS, 506 (1), pp. 306-322.

Talks & Posters :

1. **cool stars 21, Toulouse (2022).**
Presented a poster on "Low-mass planet detection by the SOPHIE spectrograph"
2. **Iranian National Observatory : 3.4m telescope instrumentation and science cases, Tehran, Iran (2022).**
Oral presentation on the topic "HD207897 b : A dense sub-Neptune transiting a nearby and bright K-type star".
3. **TESS Science Conference II, Cambridge, MA (2021).**
Oral presentation on the topic "HD207897 b : A new transiting sub-Neptune with a high density".
4. **Semaine de l'Astrophysique Française (SF2A) (2021)**
Oral presentation on the topic "HD207897 b : A dense sub-Neptune transiting a nearby and bright K-type star"
5. **Semaine de l'Astrophysique Française (SF2A) (2021)**
Presented a poster on "Examining TESS light curve to sort out SOPHIE planet candidate"

Abstract

Low-mass planets, with a mass of lower than 30 Earth masses, have an extraordinarily diverse range in terms of compositions. Their bulk density, derived by combining radial velocity and photometric data, ranges from massive rocky super-Earth with iron-rich cores to low-density objects with deep gaseous atmospheres, so-called super-puffs. The extremely diverse composition of these planets raises several questions in theories of atmospheric loss and planetary formation. Our ability to answer these questions is currently limited by the low number of low-mass planets for which fundamental properties have been accurately characterized.

The goal of this thesis is to contribute to the study of the diversity of low-mass planets' composition, by increasing the sample of precisely characterized low-mass planets and making the link between the observations and the formation paths of these bodies. More specially, this thesis focuses on overcoming two big challenges 1) instrumental limitations, to improve radial velocity data reduction and processing to facilitate the detection of low-mass planets, and 2) detecting low-mass planets and well-characterizing them.

In radial velocity, a highly precise data reduction system, that can treat the possible systematic sources of errors, is required to reach the instrumental limit of detecting low-mass planets. Therefore, first I focused on troubleshooting and improving the SOPHIE Data Reduction Software (DRS). I have been updating the DRS to the latest python version and added several new features to it, including correcting the atmospheric dispersion effect, fixing the number of mask lines at different epochs, optimizing long-term variation from the zero-point, optimizing conditions for detecting solar contamination spectra, and correcting background contamination of calibration lamp. These attempts led to great improvements in RVs and FWHM activity indicators (see chapter 2 of this thesis).

Then, I faced the challenge of low-mass planet detection and accurately characterizing them. I carried out the photometry and radial velocity analysis of the SOPHIE SP1 targets, a dedicated program to probe low-mass planet detections. Through their radial velocity analysis, my main goal was to investigate for detecting the significant periodic signals in radial velocity data, and study a variety of stellar activity indicators, to help with confirmation of the nature of a signal. Through their photometry analysis, my investigation was with the aim of 1) trying to detect a new planet signal in TESS

data; II) investigating the likely activity-induced RV variations; III) evaluating the activity of host stars of SOPHIE planet candidates. These attempts resulted in the discovery and full characterization of two transiting low-mass planets and several other candidates, and successfully removed some false positives of SOPHIE planet candidates (see chapter 3 of this thesis).

Finally, I carried out a general study of the known low-mass planets as of 2022. I discussed some key exoplanet science questions, and how current and future instruments and methods will help to answer them (see chapter 4). The techniques developed and the results gathered in this thesis will greatly contribute to the accurate detection and well characterization of future planets detection by SOPHIE. Discovered low-mass planets will contribute to a deeper understanding of the low-mass planet populations.

Keywords: Radial velocity- photometry- low-mass planets- instrumental limitations- Precision Radial Velocity (PRV)

Résumé

Les exoplanètes de faible masse, c'est-à-dire celles ayant une masse inférieure à 30 masses terrestres, ont des compositions très différentes. Leur densité apparente, obtenue à partir de la vitesse radiale et des données photométriques, va des super-Terres rocheuses massives avec un noyau riche en fer aux "super-puffs", des objets de faible densité avec des atmosphères gazeuses épaisses. Les compositions très diversifiées de ces planètes soulèvent plusieurs questions pour les théories de perte atmosphérique et de formation planétaire. Notre capacité à y répondre est actuellement limitée par le nombre restreint de planètes de faible masse dont les propriétés fondamentales ont été précisément caractérisées.

L'objectif de cette thèse est d'étudier la diversité des compositions des planètes de faible masse, en augmentant l'échantillon de planètes précisément caractérisées et en faisant le lien entre les observations et les trajectoires de formation de ces corps. Plus particulièrement, dans cette thèse, je surmonte deux défis : 1. les limitations instrumentales, en améliorant la réduction et le traitement des données et (2) la détection de planètes de faible masse et leur caractérisation.

Lorsque l'on utilise la méthode des vitesses radiales, un système de réduction de données très précis et capable de traiter les éventuelles sources d'erreurs systématiques est nécessaire pour atteindre la limite instrumentale. Je me suis donc d'abord concentrée sur la résolution de problèmes et l'amélioration du logiciel de réduction de données SOPHIE (DRS). J'ai mis à jour le code de la DRS vers la dernière version du logiciel Python et y ai ajouté plusieurs nouvelles fonctionnalités, notamment corriger de l'effet de dispersion atmosphérique, fixer le nombre de lignes de masque à différentes époques, optimiser la variation à long terme à partir du point zéro, optimiser la détection de spectres contaminés par le spectre du Soleil et la corriger de la contamination de fond de la lampe d'étalonnage. Ces travaux ont mené à des améliorations considérables en terme de VRs et la dispersion de la mesure de la largeur à mi-hauteur des raies (voir le chapitre 2 de cette thèse).

Ensuite, j'ai entrepris la détection et la caractérisation de planètes de faible masse. D'une part, j'ai analysé les données photométriques disponibles et les vitesses radiales des cibles du sous-programme 1 de SOPHIE (SP1), un programme dédié à la détection de planètes de faible masse. Mon objectif principal, grâce à l'analyse des données spectroscopiques, était de détecter des signaux périodiques significatifs dans les

données de vitesse radiale et d'étudier une variété d'indicateurs d'activité stellaire, afin d'aider à confirmer la nature du signal. D'autre part, l'analyse photométrique m'a permis I) d'explorer la possible détection d'un nouveau signal planétaire dans les données TESS; II) d'étudier les variations probables de VR induites par l'activité; III) d'évaluer l'activité des étoiles hôtes des candidats planètes SOPHIE. Ces travaux ont conduit à la découverte de deux planètes en transit et de plusieurs candidats, et ont permis d'éliminer certains faux positifs (voir le chapitre 3 de cette thèse).

Enfin, j'ai fait une étude générale des planètes de faible masse. J'y discute certaines questions clés de la science des exoplanètes, et de la manière dont les instruments et méthodes actuels et futurs aideront à y répondre (voir chapitre 4). Les techniques développées et les résultats rassemblés dans cette thèse contribueront grandement à la détection précise et à la bonne caractérisation des futures planètes détectées par SOPHIE. Les futures découvertes de planètes de faible masse contribueront à une meilleure compréhension de cette population.

Mots clés: Vitesse radiale - photométrie - planètes de faible masse - limites instrumentales - vitesse radiale à haute-precision

Contents

Abstract	5
Résumé	7
Contents	9
1 Introduction	11
1.1 A brief history of exoplanet research	13
1.2 Searching for Exoplanets	14
1.2.1 Direct imaging	14
1.2.2 Microlensing	15
1.2.3 Radial velocity	15
1.2.4 Photometry	28
1.3 Characterising the exoplanets	39
1.3.1 Joint modeling of light and RV curves	39
1.3.2 Atmospheric characterization	39
1.3.3 Classification of Exoplanets	40
1.3.4 Internal planet structure	41
1.4 Thesis objective	42
2 Optimizing SOPHIE data reduction	45
2.1 The SOPHIE high precision spectrograph	47
2.2 The SOPHIE-red project and my contribution	50
2.3 Atmospheric dispersion effect correction	52
2.3.1 Method	53
2.3.2 Results	53
2.3.3 Conclusion	57
2.4 Solar contamination correction	58
2.4.1 Review on existing methods	58
2.4.2 Optimized conditions and results for simultaneous sky observa- tion mode	59
2.4.3 Optimizing conditions and results for simultaneous calibration mode	64
2.4.4 Conclusion	66
2.5 Correction of background contamination of the calibration lamp	67
2.6 Optimizing long-term variation from the zero-point	69
2.6.1 High resolution mode	69

2.6.2	High Efficiency mode	81
2.7	Fixing mask lines selection for spectra at different epochs	82
2.8	Conclusion	84
3	Searching for low-mass planets: a combination of radial velocity and photometric data	86
3.1	Searching for low-mass planets with the SOPHIE spectrograph (SP1)	88
3.1.1	Scientific objective	88
3.1.2	SP1 sample	88
3.1.3	Observation strategy	89
3.1.4	SP1 status in 2022	90
3.2	Objective of my work on SP1	90
3.3	Data processing and selections	91
3.4	Analysis of only RVs and activity indices	92
3.5	Stellar rotation period investigation	100
3.6	Discovery and characterization of new transiting planets	102
3.6.1	HD207897 b: A dense sub-Neptune transiting a nearby and bright K-type star	102
3.6.2	Two planets around the HD88986	120
3.6.3	A transiting planet candidate	132
3.7	Other contributions to SOPHIE RV programs	134
3.7.1	RV search around binary stars	135
3.7.2	High precision RV search around M dwarfs	137
3.7.3	High precision RV search for super-Earth in systems harboring giant planets	141
4	Conclusions and future perspectives	143
4.1	Overview of low-mass planet populations as of 2022	144
4.1.1	Low-mass planets radius gap	145
4.1.2	Radius-period diagram	146
4.1.3	Mass-radius diagram & internal composition model	147
4.2	Key questions in exoplanet science as 2022	148
4.3	What to expect from the next decades?	150
4.3.1	Future space telescopes	150
4.3.2	Future ground-based telescopes	154
4.4	My future role in answering exoplanet science key questions	155
5	Overview of the results	157
	Bibliography	159

Chapter 1

Introduction

Contents

1.1	A brief history of exoplanet research	13
1.2	Searching for Exoplanets	14
1.2.1	Direct imaging	14
1.2.2	Microlensing	15
1.2.3	Radial velocity	15
1.2.3.1	Orbital parameters	15
1.2.3.2	Doppler spectroscopy and measuring RVs	19
1.2.3.3	Limitations in the accuracy of the RV method	21
1.2.3.4	Spectroscopic surveys	26
1.2.4	Photometry	28
1.2.4.1	Limitation of photometry methods	33

- 1.2.4.2 Photometric surveys 34
- 1.3 Characterising the exoplanets 39
 - 1.3.1 Joint modeling of light and RV curves 39
 - 1.3.2 Atmospheric characterization 39
 - 1.3.3 Classification of Exoplanets 40
 - 1.3.4 Internal planet structure 41
- 1.4 Thesis objective 42

1.1 A brief history of exoplanet research

Are we alone? How many worlds like ours are in the universe? Such a curiosity about other worlds started in antiquity and led to the detection of the first exoplanet in 1992 via accurate timing of pulsars (Wolszczan et al., 1992). These exoplanets which are designated PSR1257b and PSR1257c were estimated to be more massive than the Earth, orbiting the pulsar PSR1257 in orbits comparable to the orbit of Mercury around the Sun. The first exoplanet discovered orbiting a sun-like star, using accurate RV measurements, was a giant planet in a 0.05 AU orbit around the main sequence star 51 Pegasi, with an orbital period of 4.23 days. (Mayor et al., 1995). This exoplanet is a gas giant with a surface temperature of about 1284 K. Discovery of a giant planet so near to a main sequence star raised fundamental questions regarding our understanding of the formation and evolution of planetary systems. In 1999, a new and significant step forward was done with the first transiting planet detection by Charbonneau et al., 1999. This planet orbits around the star HD 209458 every 3.52 days with a radius of 1.39 Jupiter radius. Two years later, in 2001, the first detection of an exoplanet's atmosphere was done by Charbonneau et al., 2002 with the detection of Sodium in the atmosphere of HD 209458b.

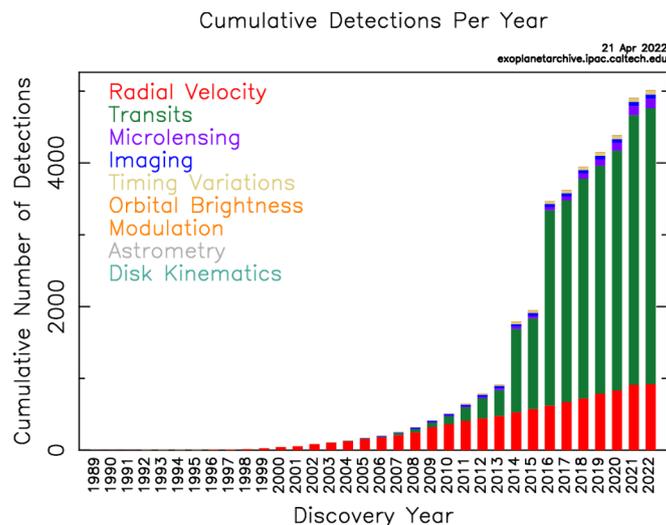


Figure 1.1 – Cumulative detection of exoplanets per year (from <http://exoplanetarchive.ipac.caltech.edu>).

After about three decades since the first exoplanet detection, the number of confirmed exoplanets registered in NASA Exoplanet Archive¹ exceeds 5000 at the time of this writing, with many more candidates awaiting confirmation. This is largely due to the progress of space and ground-based instruments. CoRoT (Baglin et al., 2006),

1. <http://exoplanetarchive.ipac.caltech.edu>

the first dedicated space mission to search for exoplanets around sun-like stars, was successfully launched in 2006. It has discovered dozen of exoplanets, including the first rocky exoplanet CoRoT-7b (Queloz et al., 2009; Léger et al., 2009). Two years later, in 2009, the Kepler mission (Borucki et al., 2010) started to measure the brightness of thousands of stars. It has detected 4767 planet candidates, of which 2711 have been confirmed², revolutionizing planetary science. Since then the number of detected exoplanets has been significantly increasing (see Fig. 1.1).

Today, the exciting field of exoplanet research continues to grow very fast. We proceed to improve our understanding of exoplanets, their formation, and evolution. Many spaces and ground-based surveys using novel methods have achieved better sensitivities for hunting Earth-size planets in the habitable zone around their host star. We are now taking a sure step toward detecting planets similar to Earth.

1.2 Searching for Exoplanets

To find the tiny signatures of exoplanets, scientists have been developing a variety of indirect and direct methods. Fig. 1.1, illustrates the various exoplanet detection methods and their contribution over the years. Among these methods, the transit and radial velocity (RV) methods are the most successful in discovering new planets, thanks to several ground instruments and space-based missions (e.g. CoRoT, Kepler, TESS, HARPS, SOPHIE). In the following, I will describe the four exoplanet detection methods that have been the most productive in discovering new planets, with a specific focus on RV and photometry, which form the basis of the current thesis.

1.2.1 Direct imaging

The direct imaging method is the only direct method for detecting exoplanets. In this method, a high-contrast image of the planet is obtained by blocking the glare of the host star and capturing the planet's reflected lights. These lights can be either the reflected light of the host star or the planet's thermal emissions. The direct imaging method is more efficient than RV and photometry in discovering outer planets. Through this method, one can obtain an estimate of projected separations of planet and star, planetary masses, period, effective temperatures, and gravities. Additionally, direct imaging of planets combining by the low, medium, or high-resolution spectrograph, allows us to explore the planet's atmosphere, and obtain necessary information about the planetary composition and physical properties (see Sect. 1.3.2). However,

2. https://exoplanetarchive.ipac.caltech.edu/docs/counts_detail.html

this method is not currently suitable for detecting small planets close to the host star. For further details about the method and its advantages see Lagrange, 2014. Also, for more information about several prospective futures of this method, see Chapter 4.

1.2.2 Microlensing

Microlensing is one of the indirect methods of exoplanet detection. In this method, when a massive foreground star passes in front of a distant background star, the foreground star acts as a gravitational lens and temporarily amplifies the light of the background star. If the foreground star hosts a planet, the planet also acts as a gravitational lens and induces a tiny magnification in light of the background star for a short time. Through this method, one can measure planet/star mass ratio and planet-star distance, and later by using the mass of the lens and the geometric configuration of the lens, one can obtain the mass and radius of the planet. The microlensing method is the most sensitive method to find wide-orbit low-mass planets or planets around very dim stars. However, detecting planets by this method is difficult due to the chance of occurring the event, the short time of its planetary light deviations, and not being re-observable. For more information about the method, see Gaudi, 2012.

1.2.3 Radial velocity

Radial velocity is one of the major indirect tools in detecting and characterizing exoplanets. The gravitational tug of the planet and parent star causes both orbits around the center of mass of the system with the same orbital period, although with a much lower speed for the host star. For example, Jupiter causes the Sun to move periodically with a velocity amplitude of 12.5 ms^{-1} . This movement will induce a Doppler shift of the spectra toward the blue when the star approaches, and toward the red when it recedes.

Here, I will describe how to determine a planet's orbital parameters through RV measurements and discuss the limitations of this method.

1.2.3.1 Orbital parameters

Although accurate extraction of the orbital parameters of exoplanets is a challenging task due to sparse sampling and noisy RVs (see Sect. 1.2.3.3), classical mechanics and some geometry is usually adequate in elementary modeling of the orbit of exoplanets and analyzing the available data. The RV of a planet in an elliptical orbit with

eccentricity e , period P , and semi-major axis a are given by (I mainly follow Danby, 1988 in this section):

$$v = v_0 + K_* [\cos(\omega_* + v_*) + e \cos \omega_*], \quad (1.1)$$

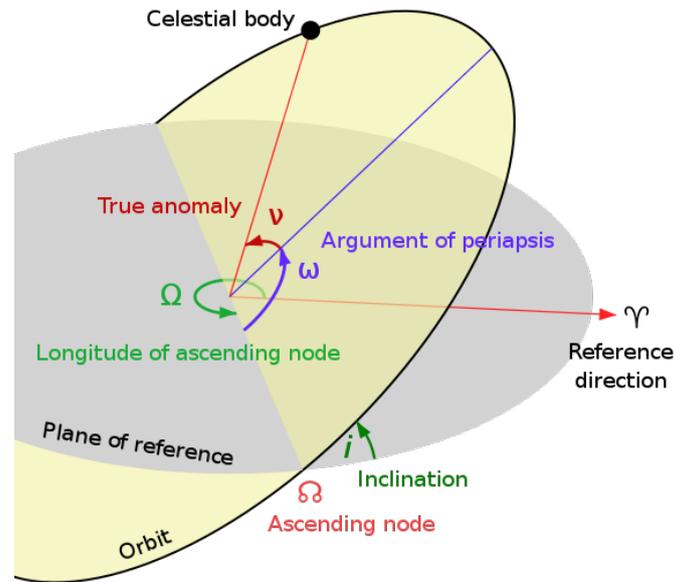


Figure 1.2 – Representation of Keplerian orbital period elements. Credited by the English Wikipedia.

in which v_0 is the system center of mass RV, K_* is the RV amplitude ($K_* = \frac{2\pi a \sin i}{P\sqrt{1-e^2}}$), ω_* is the argument of periastron, and v_* is the true anomaly. The conventions and terms are based on centuries of work on the solar system objects. True anomaly is calculated as follows:

$$M(t) \equiv \frac{2\pi}{P}(t - T), \quad (1.2)$$

where P is the orbital period, and T is the time of periastron passage. Eccentric anomaly E is related to M via the Kepler equation:

$$M = E - e \sin(E). \quad (1.3)$$

This nonlinear equation is solved for E , numerically. After determining E for a planet, the true anomaly v is easily calculated using

$$\cos v = \frac{\cos(E) - e}{1 - e \cos(E)}. \quad (1.4)$$

Since we are primarily interested in the planet mass M_p , eccentricity e , and semi-major axis a , we have to use the mass equation

$$K_* = \frac{M_p \sin(i)}{(M_* + M_p)^{2/3}} \left(\frac{2\pi G}{P} \right)^{1/3} \frac{1}{\sqrt{1 - e^2}}, \quad (1.5)$$

in which i is the inclination angle of the orbit (the angle between the normal to the orbit and the line of sight), and M_* is the mass of the host star. Normally, $M_p \ll M_*$ and equation (1.5) can be simplified accordingly. The mass of the star can be estimated using the spectroscopic data (through spectral type and luminosity class, Santos et al., 2013; Sousa et al., 2018), while K_* and P are determined directly from the velocity curve. Moreover, the eccentricity e and argument of pericentre ω are defined using the shape of the velocity curve. Nearly circular orbits ($e \ll 1$) lead to a more or less sinusoidal shape of the velocity curve, while for some reason like prior planet–planet scatterings (Chatterjee et al., 2008; Jurić et al., 2008), transfer of angular momentum due to other planets in the system, wide-binary companion (Kaib et al., 2013; Malmberg et al., 2007) or the close passage of a stellar binary companion (Halbwachs et al., 2005; Kane et al., 2014), can cause highly eccentric orbits, that show spiky velocity variations. Since the inclination angle can not be determined from the RV curve, it can only provide a lower limit to the mass of the planet:

$$M_p \geq (M_*)^{2/3} K_* \left(\frac{P}{2\pi G} \right)^{1/3} \sqrt{1 - e^2}. \quad (1.6)$$

According to equation 1.5, massive planets, with a shorter orbital period, orbiting low-mass stars create a larger RV semi-amplitude. Thus, they are more easily detectable by this method. RV detection of a planet depends on the amplitude of the signal K , the RV uncertainty σ_{RV} and the number of RV measurements N (assuming a circular orbit):

$$S/N_{RV} \approx \sqrt{N} \frac{K}{\sigma_{RV}} \quad (1.7)$$

To determine the most significant periodicity(es) in the data, a generalized Lomb-Scargle (also known as generalized L-S) periodogram is a commonly used method. This method which is based on least square analysis was first introduced by Lomb,

1976 and Scargle, 1982 and later several extensions of this periodogram like generalized Lomb-Scargle periodogram (Zechmeister et al., 2009), floating means periodogram (Cumming et al., 1999; VanderPlas et al., 2015) and the Date-Compensated Discrete Fourier Transform (Ferraz-Mello, 1981) were proposed. The main difference of these extensions with the Lomb-Scargle periodogram is adding an offset to the model at equation 1.8 (in L-S, the mean of data was subtracted. It was supposed the mean of data ($y=0$) is equal to the Sine function model).

In generalized L-S, assuming that the data is in the form of a time series (y_i, t_i, σ_i) with σ_i the RV's uncertainty, one uses the periodic fit (including a constant as an offset):

$$y_i \equiv A \cos(\omega t_i) + B \sin(\omega t_i) + c \quad (1.8)$$

For constructing a periodogram, minimization of the squared differences between data and models divided by data errors should be performed (Zechmeister et al., 2009). If this minimum value is shown with χ^2 , Zechmeister et al., 2009 showed that periodogram power can be introduced with the following equation:

$$p(\omega) \equiv \frac{\chi_0^2 - \chi(\omega)^2}{\chi_0^2}. \quad (1.9)$$

where χ_0^2 is the non-varying part of (1.8) (for more detail of analytical calculations see Zechmeister et al., 2009). Note that $0 \leq p(\omega) \leq 1$ with the maximum $p(\omega) = 1$ corresponds to a perfect fit. In other words, the periodogram represents the power spectrum of the time series data. The main period of the system should appear as the highest peak in the periodogram.

False peaks usually appear due to sampling of observation or statistical fluctuations which appear as noise in the data which makes the identification of real peaks difficult. In order to evaluate the validity of each peak, one uses the so-called false-alarm probability (FAP) or P-value which is a measure of how much is the probability of finding a peak with a specific magnitude in a random background (Gaussian random number).

There are several methods for computing the FAP, from analytical methods (e.g. Scargle, 1982; Baluev, 2008; Horne et al., 1986) to Monte-Carlo simulations (Süveges, 2014)³ that some of them have been contested (Koen, 1990). One of the commonly

3. AstroPY has provided a package for computing FAP with different method <https://docs.astropy.org/en/stable/timeseries/lombscargle.html>

used analytical methods was introduced by Baluev in a series of works (Baluev, 2008; Baluev, 2009; Baluev, 2013; Baluev, 2015). This method is based on Rice method derived from the extreme value theory. In this series of works they applied this model in both linear periodogram (like L-S periodogram, Baluev, 2008) and non linear periodogram (like Keplerian periodogram, Baluev, 2015). This method has shown its efficiency in computing FAP with considering white noise. Recently, Delisle et al., 2020 extended this method with considering correlated noise which is more realistic in most of astronomical data sets.

Another method for computing FAP is bootstrap re-sampling (VanderPlas, 2018) which repeatedly produces many random number of data with Gaussian (normal) distribution via Mont Carlo simulation. Then, periodograms are computed for each simulated data set. For each periodogram, the highest peak is selected and a distribution of them is constructed. As an example, the 1% FAP is related to the power which is only reached 1% of the time.

1.2.3.2 Doppler spectroscopy and measuring RVs

Figure (1.3), illustrates RVs from observations to planet detection. I explained how exoplanets are detected through RVs and the orbital parameters are extracted (Sect. 1.2.3.1). But how do we measure the stellar RV? The planet and its host star orbit around their common center of mass. These movements will induce a Doppler shift of the spectra toward the blue when the star approaches, and towards the red when it recedes. In the standard rest frame and neglecting relativistic effects, the wavelength shift will be given by:

$$\frac{\Delta\lambda}{\lambda_0} = \frac{\lambda_{obs} - \lambda_0}{\lambda_0} = \frac{v_r}{c} \quad (1.10)$$

where λ_{obs} is measured wavelength, λ_0 is the wavelength at the source and c is the light speed. Measuring the wavelength shift from a single spectral line is not very accurate. Instead, one can consider all spectral lines and measure the wavelength shifts by using the cross-correlating function (CCF) method and increase the signal to noise (S/N). To do this, one needs to cross-correlate the numerical mask (binary mask) with the observed spectra (Queloz, 1995; Baranne et al., 1996, see Fig 1.4 for schematic procedure). The numerical masks are the box-shaped emission lines placed in the most prominent lines (Pepe et al., 2002).

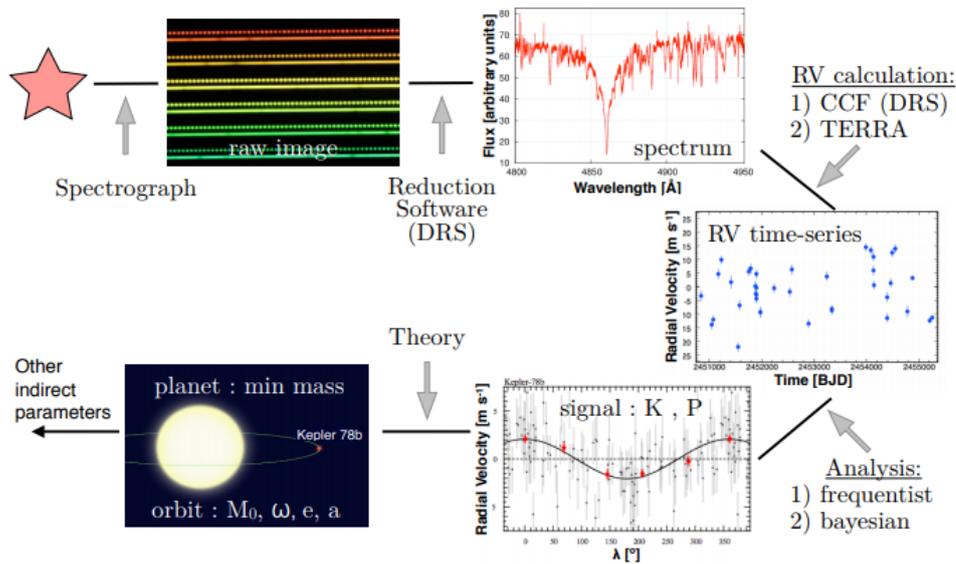


Figure 1.3 – RVs from observation to planet detection. Artistic impression by David A. Aguilar

The nearly Gaussian shape of the CCF (see Fig. 1.4) comes from the approximate shape of spectral lines. Its full width at half maximum (FWHM) depends on spectral type, metallicity, and rotation of star. A Gaussian is fitted on the CCF and RV is determined according to the extremum.

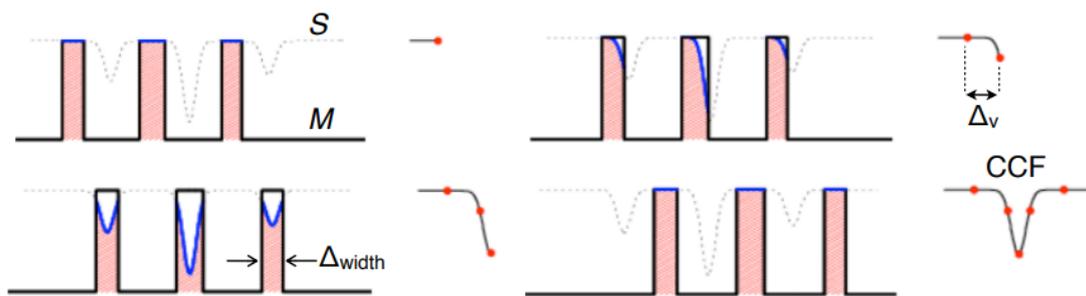


Figure 1.4 – Cross-correlation method. The good performance of cross correlation function relies on appropriate choice of Δ_{width} and Δ_v as well as creating an accurate mask. Adopted from (Eggenberger et al., 2010).

This method is not only very useful for computing the RVs but also helps to monitor the stellar activity. Star spots and other star activities distort the CCF symmetries. Thus, one can measure bisector span (BISS), FWHM, and CCF contrast as a stellar activity indicator (see Fig. 1.5). The CCF contrast is the depth of Gaussian fit on

CCF. To calculate BISS (Povich et al. 2001), as illustrated in Figure 1.5, the CCF is subdivided horizontally and the midpoints of the line segments are calculated. In the following, two different vertical regions have been defined in the CCF and a mean velocity has been calculated for each region: v_1 for $0.3 \leq CCF \leq 0.45$ (top) and v_2 for $0.6 \leq CCF \leq 0.9$ (bottom). Then BISS is derived by $v_2 - v_1$.

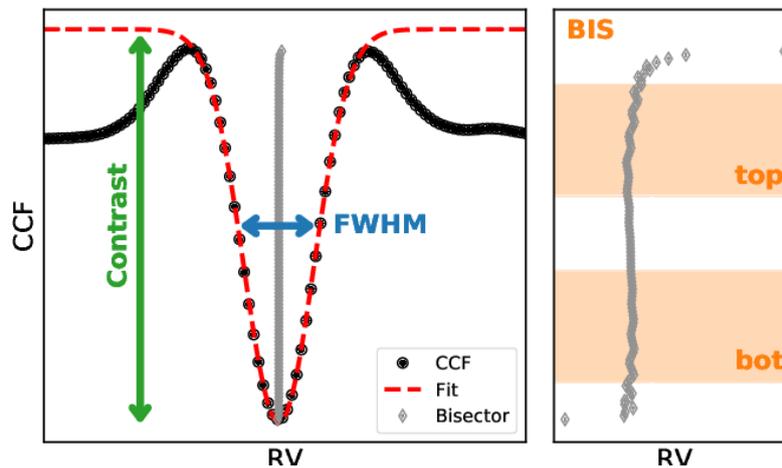


Figure 1.5 – The CCF activity indicators. The black dots are showing CCF and the best Gaussian fit is illustrated by the red dashed line. FWHM and CCF contrast are derived from the fit. The bisector span is shown with grey diamonds with a zoom to the central region (right). The indicated top and bottom regions are used to compute the BIS. One can see the magnified horizontal line that is distorted by stellar activity. Picture from Lafarga et al., 2020.

Measuring RVs by CCF method is currently used in the “official” pipeline of many spectrographs such as HARPS (ESO) and SOPHIE (OHP).

1.2.3.3 Limitations in the accuracy of the RV method

Instrumental limitations

There are several technical issues in the spectroscopic studies of exoplanets (Desidera et al., 2004). Old fashion spectra used to have uncertainties in the RVs of the order of a few $km s^{-1}$. This uncertainty is far worse than the precision needed to detect exoplanets that have amplitudes of the order of a few meters per second or even less. Employing more advanced technology and innovative methods reduced uncertainties to about $10 m s^{-1}$ in 1995 and to about $3 m s^{-1}$ in 1998. Nowadays, the accuracy of a few tens of centimeters per second has been reached.

A first, and fundamental, limit to RV precision is related to the photon noise. As the arrival of photons are independent of each other, one can expect that it follows

a Poisson distribution. Therefore, a count of N number of photons has a standard deviation of \sqrt{N} . Bouchy et al., 2001 studied the RV uncertainty and introduce the following formula as a fundamental noise limitation in velocity measurements:

$$\delta V_{RMS} \equiv \frac{c}{Q\sqrt{N_{e^-}}} \quad (1.11)$$

where c is the speed of light, Q is a quality factor that depends on stellar spectral type, stellar $v \sin i$ and spectral resolution, and N_{e^-} is the total number of counted photo-electrons. The latter also depends on exposure time, stellar magnitude, and the total efficiency of a telescope.

Among the other important instrumental limitations, one can mention thermal, mechanical, and pressure stability, detector effects, and wavelength calibration. Because a large part of this thesis is dedicated to the instrumental improvement of the SOPHIE spectrograph, I will explain them in more detail in Sect. 2.

Non instrumental

One of the biggest obstacles in RVs comes from the star itself. Stars are not a quiescent light source but exhibit several physical phenomena which can interfere with spectroscopic observations of exoplanets, leading to false detection or missing signals from the RV (Saar et al., 1997; Huélamo et al., 2008). For example α Cen Bb was claimed in 2012 by a group working with HARPS, to be an Earth-mass planet orbiting every 3.2 days around the star α Cen (Dumusque et al., 2012). α Cen B is a K-type star at a distance of 4.73 light-years from us. This could be a potentially important discovery since this star is the nearest star with a spectral type close to the spectral type of the Sun. However, a team of researchers at the University of Oxford showed that it can be a false detection caused by stellar activity (Rajpaul et al., 2015b). Various types of stellar activity can lead to such false detections (Da Silva et al., 2012; Meunier et al., 2013). It is therefore quite important to identify and characterize various types of stellar activity to distinguish them -as accurately as possible- from the variations caused by exoplanets.

I hereby explain the most important types of stellar activity (see Fig. 1.6) which can influence the spectroscopic and photometric observations (in particular those of low mass planets), as the followings:

- Oscillations:

Oscillations of outer layers of the main sequence and sub-giant stars are driven by underlying convection currents which cause pressure waves in the stellar

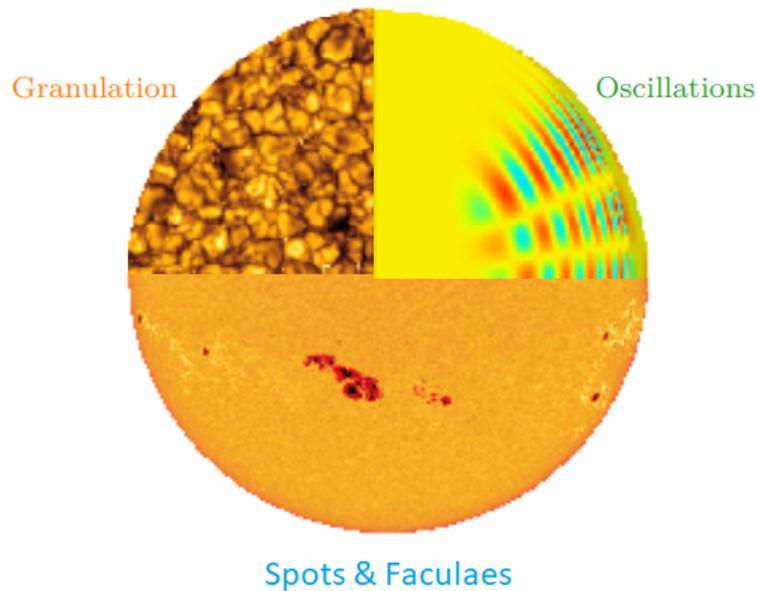


Figure 1.6 – Illustration of different stellar activities. Picture from Faria, 2018.

atmosphere (see Fig. 1.6). These oscillations have periods of the order of several minutes (for the sun this is around 5-15 min) and increase with a stellar mass between the main sequence. Theory suggest oscillations period increase with the square root of the star mean density (Christensen-Dalsgaard, 2004). The RV signature of oscillations is between the order of 0.1 to 4 m.s^{-1} (Dumusque et al., 2011a; Queloz et al., 2001), depending on the spectral type of stars and an evolutionary stage (Christensen-Dalsgaard, 2004; Schrijver et al., 2000).

- Granulation, mesogranulation, and supergranulation:
Also caused by the underlying convection currents in the stellar atmosphere, producing a convection cell from rising hot plasma at the center which is surrounded by colder descending ones at the outer edge (see Fig. 1.6). These effects last for shorter than 25 min for granulation (Tarbell et al., 1989; Del Moro, 2004) and up to timescales of 33 h for mesogranulation, and supergranulation (Del Moro et al., 2004). Because the rising (thus blue shifted) materials are hotter, therefore they contribute to more light than colder descending (thus red-shifted) plasma. Hence, it causes a shift towards blue on the spectrum known as "convective blueshift" (Dravins et al., 1981; Dravins, 1982). The effect of granulation, mesogranulation and supergranulation on RV is of the order of meters per second level (Dumusque et al., 2011b).
- Rotational modulation:
This effect is associated with the magnetic field. The magnetic field in the area which is stronger (of the order of 1 to 2 kG) blocks the convection and creates

a colder region in the photosphere so cold dark spot (see Fig. 1.6). The size of spots can vary up to a few tens of thousands of km, and their lifetime is from days to weeks (Foukal et al., 1986). The faculae are also photospheric events caused by the emergence of small magnetic flux tubes. Their size is smaller than star spots (less than a few hundred km) and the temperature is less than spots (Schrijver, 2002; Johnson et al., 2021). They can create an inhomogeneous flux deficit (for spots) or excess (for faculae) on the surface of the star (flux effect). Because the spots (or faculae) rotate together with the star when the star is approaching (or receding) they block half of the hemisphere from blue-shifting (or red-shifting). The blue shift and red shift no longer cancel each other and lead to asymmetries in the red-shifted and blue-shifted spectra (Aigrain et al., 2012). Note that the faculae contribute to smaller flux effects than spots, due to their lower temperature (Meunier et al., 2010a). In addition to the flux effects of spots and faculae, the magnetic field can inhibit convective motions (Tarbell et al., 1987; Hanslmeier et al., 1991) and lead to spots and faculae red-shifted. The amplitude of the resulting perturbations in RV measurements depends on the declination of the dark and bright spots on the stellar surface with our line of sight, and the magnitude of the stellar activity and can be of the order of several meters per second (Meunier et al., 2010b; Dumusque et al., 2014). These effects can appear as a periodic signal in periodogram which is comparable to P_{rot} or its harmonics $P_{rot}/2$, $P_{rot}/3$ and etc.

— long magnetic cycles:

Although individual spots/faculae have a lifetime ranging from a few days to weeks and the total number of spots varies periodically with a period of the years (Saar et al., 1999). As an example, our Sun has a magnetic cycle of 11 years (Maunder, 1904). This effect can be comparable with planet signals with longer periods.

Many stars are more active than the Sun. Reinhold et al., 2020 studied the photometric variation of 369 solar-like stars by combining photometric observations from Kepler with astrometric data from the Gaia spacecraft. They found out that most solar-like stars show higher photometric variability than the Sun and accordingly they are more active than the Sun. Stellar activity depends on mass, spectral type, rotation, and the age of the star. Younger stars rotate more rapidly and are therefore usually more active. The effect of activity in the RV also depends on stellar rotation: when the star rotates faster, the deformation of the lines is more important. For a more detailed discussion about the dependence of stellar activity on mass, age, and spectral type refer to Fischer et al., 2015 and Macintosh et al., 2015.

How can we disentangle the effects of stellar activity from a real planetary signal? Since stellar activity is a disturbing agent in the detection and accurate determination of exoplanet parameters, one should seek ways to identify and disentangle their effects.

Here I explain a few of the methods which have been developed for dealing with stellar activities, as the following :

— Observational strategies:

The existing observational strategies have been improved to an extent that can tackle major disturbances due to oscillations and granulation. The oscillation effect can be reduced with an observation strategy of 15 min duration exposure and 3 observations of 10 min duration per night, separated by 2h (Dumusque et al., 2011a) to remove granulation. However, uncertainties due to rotational modulation phenomena such as spots and plages are still disturbing, since signals due to these effects can hide or mimic those of exoplanets within this period range (Bonfils et al., 2007; Huélamo et al., 2008).

— Line-profile and chromospheric activity indicators:

One technique for the investigation of the stellar activities is to consider some indicators from spectral lines, including asymmetry of the CCF (e.g. bisector span, FWHM, CCF contrast. See 1.2.3.2 for more explanations), $\log(R'_{HK})$ index, and $H\alpha$ index. The $H\alpha$ absorption line at 656.281 nm is well known to indicate stellar activities. Depending on the activity level and spectral type of the star, this line becomes deeper in the absorption line or increases a brightness of the line and even it can convert to emission lines (Cram et al., 1979; Pasquini et al., 1991). To extract the $H\alpha$ index, for example, one can follow Boisse et al., 2009 or Da Silva et al., 2011 recipe, which is a relative flux at the core of $H\alpha$ divided by the nearby continuum region as a reference. The CaII H and K emission lines centered at 396.9 nm and 393.4 nm respectively, are also common to track the star activity level (see e.g. Noyes et al., 1984a). The $\log(R'_{HK})$ index is defined by the sum of flux at CaII H and K divided by two nearby continuum regions as reference (see e.g. Boisse et al., 2009 and references therein). The value of $\log(R'_{HK}) \leq -4.8$ can be interpreted as a quiet star and the star may be active otherwise. This indicator can also be used for computing the stellar rotational period (Noyes et al., 1984a; Astudillo-Defru et al., 2017a). Additionally, for M dwarf stars Na I D1 and D2 lines can also be a signature of stellar activity (Da Silva et al., 2012).

— modelling activity-induced RV variations:

Modeling stars with various activities are also possible, thanks to the underlying physics and detailed computer codes. Some of these codes include complex physical activities which take place in the inner parts and outer atmosphere of the star, and some use statistical methods to removed correlated noise. As an example, Rajpaul et al., 2015a introduced a Gaussian process (GP) framework for statistically accounting for the quasi-periodic (QP) behavior of stellar activity signals. The activity signals are quasi-periodic as the active regions evolve and decay and also on the star surface which rotates, they come in and out of our view. In GPs, one can model the stellar activity with a quasi-periodic kernel

(note that there are also several other kernels):

$$\kappa_{QP}(t, t') = \exp \left[\frac{-\sin^2\left(\frac{\pi(t-t')}{P}\right)}{2\lambda_p^2} - \frac{(t-t')^2}{2\lambda_e^2} \right]; \quad (1.12)$$

where P and λ_p are related to the period and length-scale (decaying) of the periodic component of stellar activities, and λ_e is related to the evolutionary timescale of active regions. While λ_e has units of time, λ_p is dimensionless because it is relative to P . A prior can be set for these parameters on a given star. For example, if we can estimate the rotation period of the star (e.g. from photometric observations), one can set a prior on P . Then a Markov chain Monte Carlo (MCMC) method will be applied to find the best-fit of parameters on equation 1.12. I note that this method is widely used in the literature (e.g., Bluhm et al., 2021; Gan et al., 2021), with promising results.

Boisse et al., 2012a also developed a tool for fast simulation of starspot and plage effects on the photometric and RV data, called Spot Oscillation And Planet (SOAP) code. The tool was later improved by Oshagh et al., 2013 and Dumusque et al., 2014. The software which is publicly available⁴ can be utilized to estimate the unwanted effects of starspots and oscillations from the light and RV curves.

— Long-term and continuous photometry of stars:

Long-term and continuous photometry of stars (with a ground-based telescope or spacecraft like TESS, Kepler, etc) can provide us with more information about the activities and rotational periods of the target stars.

1.2.3.4 Spectroscopic surveys

Detecting exoplanets using the RV method needs continuous observations of a sample of stars. To do this job several spectrographs are used. Here, I highlighted some of the main surveys in visible and infrared:

— Visible:

- HARPS:

The HARPS (High Accuracy Radial Velocity Planet Searcher) spectrograph which is mounted on the ESO's 3.9 telescopes in Chile, covering the 380-690 nm wavelength with a resolution of 115 000. HARPS can reach an RV precision of about 0.8 m/s for S/N=200 at 550 nm (Fischer et al., 2016).

- HARPS-N:

4. <http://www.astro.up.pt/soap2>

HARPS-N (Cosentino et al., 2012) is the northern version of HARPS, installed on the Galileo 3.58m telescope in La Palma, Spain. The wavelength coverage is 378-691 nm and it has a resolution power of 115,000, reaching to a 0.8 m/s RV precision for S/N= 200 at 550 nm (Fischer et al., 2016; Dumusque et al., 2015b).

- **HIRES:**
The HIRES (High-Resolution Echelle Spectrograph, Vogt, 1994) installed on the 10m Keck telescope in Hawaii is another important facility that contributes to accurate RV measurements needed for exoplanet research. This spectrograph has a resolution of 55000 and covers a 364-800 nm wavelength range, reaching a 1.5 m/s precision for S/N=200 at 550 nm (Fischer et al., 2016; Spronck et al., 2015).
 - **ESPRESSO:**
ESPRESSO (Echelle Spectrograph for Rocky Exoplanet- and Stable Spectroscopic Observations) is a third-generation echelle spectrograph mounted on an 8.2 m Very Large Telescope (VLT). It covers the wavelength range from 380-686 nm with a resolution of 200,000, reaching a RV precision of 50 cm/s for S/N ~ 115 at 550 nm (Pepe et al., 2021).
 - **SOPHIE:**
SOPHIE (Spectrograph pour l'Observation des Phénomènes des Intérieurs et des Exoplanètes, meaning spectrograph for the observation of the phenomena of the stellar interiors and the exoplanets) is another facility that is mounted on the 1.93m reflector at the Haute-Provence Observatory located in southeast France (Bouchy et al., 2009a). This spectrograph takes data of exoplanets in the wavelength range 387-694 nm and has a resolution power of $\approx 75\,000$. After undergoing several improvements, SOPHIE with new name of SOPHIE+ has reached a precision of 1-2 $m s^{-1}$, capable of detecting lower-mass exoplanets (Bouchy et al., 2013; Perruchot et al., 2011). See Sect. 2.1 for more information about this spectrograph. Note that a major part of this thesis is dedicated to the results acquired by the SOPHIE data.
- near infrared:
- **CARMENES:**
CARMENES (Calar Alto high-Resolution search for M dwarfs with Exoearths with Near-infrared and optical Échelle Spectrographs, Quirrenbach et al., 2014) is mounted in the 3.5 m telescope at Calar Alto (Spain). It has two arms and can cover two wavelength ranges: visible (520-960 nm) and near-infrared (960-1710 nm), with a resolution power of 94,600 in visible and 80,400 in near-infrared. CARMENES can reach a RV precision of 1 m/s (Quirrenbach et al., 2018).
 - **SPIRou:**

The SPIRou is mounted in a 3.6 m Canada-France-Hawaii telescope and covers a wavelength range of 980-2350 nm. This spectrograph has a resolution power of 70,000 (Artigau et al., 2014), reaching to a precision of 1 m/s.

- NIRPS:
NIRPS (Near Infra-Red Planet Searcher, Bouchy et al., 2017) is also planned to be used with the same telescope (3.9 m) of HARPS with a resolution of $\approx 100\,000$. This instrument will search for exoplanets in the infrared region of the spectrum with wavelength range of 950-1800 nm, and aiming to reach 1 m/s RV precision.

1.2.4 Photometry

The first report of using transits to detect exoplanets appeared in Rosenblatt, 1971 and to date, 3850 transiting planets are detected. Details of this method for exoplanet detection have now developed considerably, although the principles remain the same. Transits lead to a periodic drop in the measured flux of the star as the planet moves across the apparent disk of the star (see Fig. 1.7).

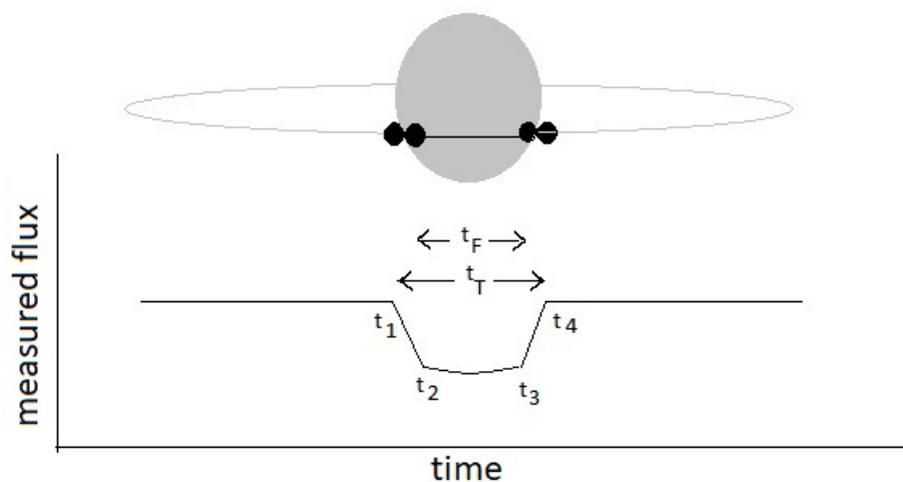


Figure 1.7 – Light variation due to the transit of an exoplanet. t_F is the time between the start and end of inner contacts, while t_T is the corresponding time interval for outer contacts. Also, t_1 is the time of first outer contract, t_2 is time of first inner contract, t_3 is time of second inner contract and t_4 is time of second outer contract of apparent disks.

The synthesis and analysis of eclipsing binary light curves is thoroughly studied in classical studies (see Kopal, 1955; Wilson et al., 1971), which are a supportive background for the exoplanet light curve analysis. Seager et al., 2003 used a few simplifying assumptions to derive R_p/R_* (planet-to-star ratio of radii), a/R_* (orbital radius to the radius of the parent star), and the inclination angle i (i.e. the angle between the normal to the orbit and the line of sight). The orbital period P , depth of transit δ , duration of transit t_T , and duration of the flat bottom t_F are more directly determined from the light curve. The simplifying assumptions are as follows (Seager et al., 2003):

- The planetary orbit is circular (this assumption is better realized for short orbit, tidally affected planets).
- There is no limb darkening (i.e. one assumes that the stellar disk is uniformly bright in all wavelengths).
- The disk of the planet is completely dark.
- There is no third light from any other nearby object (e.g. a companion star).
- The validity of the canonical stellar mass-radius relation.

Using the above assumptions, it is easy to derive the following useful relation:

$$\delta \equiv \frac{F_* - F_{tr}}{F_*} = \left(\frac{R_p}{R_*} \right)^2. \quad (1.13)$$

For the hypothetical transit of Jupiter across the Sun, as viewed by a distant observer, $\delta \sim 0.01$, while for the Earth-Sun $\delta \approx 0.0008$, corresponding to a change in the apparent magnitude of about $\delta m_J \sim 0.025 \text{ mag}$ and $\delta m_E \sim 0.002 \text{ mag}$, respectively. The simple estimate goes as follows: $m = -2.5 \log L + m_0$, thus $\delta m = -2.5 \frac{\delta L}{L} = 2.5 \frac{\delta I}{I} = 2.5 \delta$. These values show that the observation of exoplanetary transits requires high precision photometry at the $m \text{ mag}$ level or better.

The time interval of full planetary transit time t_T obeys the relation

$$\left[\frac{\sin(\pi t_F/P)}{\sin(\pi t_T/P)} \right]^2 = \frac{[1 - R_p/R_*]^2 - [a/R_* \cos(i)]^2}{[1 + R_p/R_*]^2 - [a/R_* \cos(i)]^2}. \quad (1.14)$$

In the large orbit limit ($a \gg R_*$) and $\frac{\pi t_T}{P} \ll 1$, we obtain the following simplified

relation

$$\left(\frac{t_F}{t_T}\right)^2 \simeq \frac{[1 - R_p/R_*]^2 - [a/R_* \cos(i)]^2}{[1 + R_p/R_*]^2 - [a/R_* \cos(i)]^2}. \quad (1.15)$$

Furthermore, since for large orbits, transits occur only if i is nearly 90 degrees ($\cos(i) \ll 1$)⁵, we obtain

$$t_T \simeq \frac{P R_*}{\pi a} \sqrt{\left[1 + \frac{R_p}{R_*}\right]^2 - \left[\frac{a}{R_* \cos(i)}\right]^2}. \quad (1.16)$$

The parameter b which gives the ratio of minimum projected planet-orbital radius (i.e. $a \cos i$) to the radius of the star (R_*) is given by

$$b \equiv \frac{a \cos i}{R_*} = \sqrt{\frac{(1 - \sqrt{\delta})^2 - [\sin^2(\pi t_F/P) / \sin^2(\pi t_T/P)](1 + \sqrt{\delta})^2}{1 - \sin^2(\pi t_F/P) / \sin^2(\pi t_T/P)}}, \quad (1.17)$$

which in the limit $\pi t_T/P \ll 1$ and $a \gg R_*$, leads to the simplified equation

$$b \simeq \sqrt{\frac{(1 - \sqrt{\delta})^2 - [(t_F/t_T)^2](1 + \sqrt{\delta})^2}{1 - (t_F/t_T)^2}}. \quad (1.18)$$

Another important quantity which can be easily deduced from the observables δ , P , t_F , and t_T is

$$\frac{a}{R_*} = \sqrt{\frac{(1 + \sqrt{\delta})^2 - b^2[1 - \sin^2(\pi t_T/P)]}{\sin^2(\pi t_T/P)}}. \quad (1.19)$$

Again, using the approximation $a \gg R_*$, we obtain

$$\frac{a}{R_*} \simeq \frac{2P}{\pi} \frac{\delta^{1/4}}{\sqrt{t_T^2 - t_F^2}}. \quad (1.20)$$

5. For the transit to occur, we should have $\cos i < \frac{R_* + R_p}{a}$. For example, to a distant observer, Jupiter shows transit if $i > 89^\circ.9$.

Using the Kepler's third law $P^2/(4\pi^2) = a^3/G(M_* + M_p)$ in which G is the Newton's gravitational constant, and the approximation $M_p \ll M_*$, we are led to

$$a = \left(\frac{P^2 G M_*}{4\pi^2} \right)^{1/3}. \quad (1.21)$$

The orbital inclination angle i is then approximately given by

$$i \simeq \cos^{-1} \left(\frac{bR_*}{a} \right). \quad (1.22)$$

The mean density of the parent star ρ_* in terms of the mean density of the Sun ρ_\odot is obtained as

$$\frac{\rho_*}{\rho_\odot} = \frac{M_*/M_\odot}{R_*^3/R_\odot^3} = \frac{4\pi^2}{GP^2} \left(\frac{a}{R_*} \right)^3 \simeq \frac{4\pi^2}{GP^2} \left[\frac{(1 + \sqrt{\delta})^2 - b^2 [1 - \sin^2 \frac{\pi t_T}{P}]}{\sin^2(\pi t_T/P)} \right]^{3/2} \quad (1.23)$$

which for $a \gg R_*$ and $M_p \ll M_*$ reduces to

$$\frac{\rho_*}{\rho_\odot} \simeq \frac{32P}{\pi G} \frac{\delta^{3/4}}{(t_T^2 - t_F^2)^{3/2}}. \quad (1.24)$$

One should keep in mind that the above simplified analysis of the transit light curve rests upon the approximations of circular orbit, uniform stellar disk brightness and the absence of anomalies like star spots, star variability, third light etc. Since the number of discovered exoplanets is rapidly growing, such a simple analysis is very helpful in obtaining the parameters for a large number of transit light curves and doing statistical studies of exoplanet physical parameters. On the other hand, accurate and more detailed investigation of individual transit light curves requires a more realistic model, beyond these simplified modelling.

For a more accurate analysis, one can follow Winn, [2010b](#). Using rectangular coordinates with the X-axis in the direction of descending node,

$$X = -r \cos(\omega + f) \quad (1.25)$$

$$Y = -r \sin(\omega + f) \sin i \quad (1.26)$$

$$Z = r \sin(\omega + f) \cos i, \quad (1.27)$$

in which (as before) $r = \frac{a(1-e^2)}{1+e\cos f}$, f being the true anomaly and i the orbital inclination. Since

$$r_{sky} \equiv \sqrt{X^2 + Y^2} = \frac{a(1-e^2)}{1+e\cos f} \sqrt{1 - \sin^2(\omega + f) \sin^2 i}, \quad (1.28)$$

transits and occultations (if any) occur at $f_{transit} = +\frac{\pi}{2} - \omega$ and $f_{occultation} = -\frac{\pi}{2} - \omega$, respectively. We therefore have the following expressions for the impact parameter of transits and occultations:

$$b_{transit} = \frac{a \cos i}{R_*} \left(\frac{1-e^2}{1+e \sin \omega} \right), \quad (1.29)$$

$$b_{occultation} = \frac{a \cos i}{R_*} \left(\frac{1-e^2}{1-e \sin \omega} \right). \quad (1.30)$$

The probability of transits for an exoplanet characterized by (R_*, R_p, a, e) , randomly observed by the observer is given by:

$$P_{transit} = \left(\frac{R_* \pm R_p}{a} \right) \left(\frac{1+e \sin \omega}{1-e^2} \right), \quad (1.31)$$

$$P_{occultation} = \left(\frac{R_* \pm R_p}{a} \right) \left(\frac{1-e \sin \omega}{1-e^2} \right). \quad (1.32)$$

These expressions are obtained, by calculating the ratio of the total solid angle corresponding to all orientations leading to transit/occultation, to the complete solid angle 4π . Note that +/- signs in equations (1.31) and (1.32) correspond to excluding/including partial transits/occultations. In the limit $R_p \ll R_*$ and $e \rightarrow 0$, we obtain

$$P_{transit} \approx P_{occultation} \approx \frac{R_*}{a} \approx 0.005 \left(\frac{R_*}{R_\odot} \right) \left(\frac{a}{1AU} \right)^{-1}. \quad (1.33)$$

The time interval between any two observable events (for example between the first and second contacts which is called ingress), is given by:

$$\tau_2 - \tau_1 = \frac{P}{2\pi\sqrt{1-e^2}} \int_{f_1}^{f_2} \left(\frac{r(f)}{a} \right)^2 df \quad (1.34)$$

For circular orbits ($e=0$), we obtain:

$$\tau_{total} \equiv t_{IV} - t_1 = \frac{P}{\pi} \sin^{-1} \left(\frac{R_*}{a} \frac{\sqrt{(1+k)^2 - b^2}}{\sin i} \right) \quad (1.35)$$

and

$$\tau_{full} \equiv t_{III} - t_{II} = \frac{P}{\pi} \sin^{-1} \left(\frac{R_*}{a} \frac{\sqrt{(1-k)^2 - b^2}}{\sin i} \right) \quad (1.36)$$

in which $k = \frac{R_p}{R_*}$ and b is the impact parameter. Measuring τ_{total} and τ_{full} , and P , therefore, leads to relational constraints between the system parameters. For example, from 1.35 and 1.36, one obtain:

$$\frac{(1+k)^2 - b^2}{(1-k)^2 - b^2} = \frac{\sin^2 \left(\frac{\pi \tau_{total}}{P} \right)}{\sin^2 \left(\frac{\pi \tau_{full}}{P} \right)}, \quad (1.37)$$

which sets a relation between k and b . Also, we have, approximately,

$$\frac{\tau_{egress} - \tau_{ingress}}{\tau_{egress} + \tau_{ingress}} \approx e \cos \omega \left(\frac{R_*}{a} \right)^3 (1 - b^2)^{3/2}. \quad (1.38)$$

1.2.4.1 Limitation of photometry methods

There are several limitations and false positive scenarios on detecting planets by transit method which can mimic or hide exoplanet transits. Here, I mention the most important false positive scenarios and ways to resolve them:

- Brown dwarf or low-mass star: since dwarf stars/Brown dwarfs and giant planets have approximately the same size, the transit of dwarfs across the disk of a much larger star can lead to almost the same light drop as the exoplanet transits (see Fig. 1.8 b). In order to distinguish dwarfs from exoplanets, one can use RV measurements, since dwarf stars are more massive than giant planets and cause higher amplitude velocity curves of the host star. As an example, the OGLE-TR-122b light curve resembles a planetary transit. However, RV curve proves that the companion is a very low mass ($\sim 0.08 M_\odot$), main-sequence star (Pont et al., 2005).

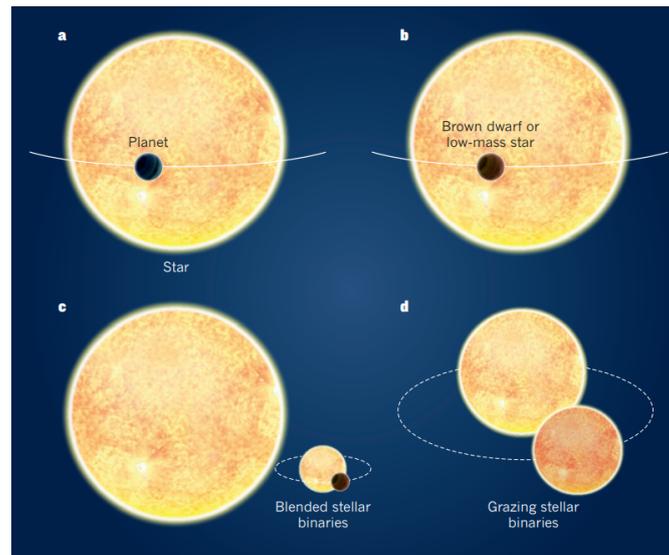


Figure 1.8 – False positive scenarios in planet detection by the transit method (see the text). Image credit Cameron, 2012.

- Blended stellar binaries: a third body light leads to a shallower relative light minimum of the eclipses and thus resembling an exoplanet transit (see Fig. 1.8 c). Such triple stellar systems can be distinguished from a planetary system, since they may show a double peak in the cross correlation function of the spectrum with the spectrum of the third star. Another way to distinguish these two from each other is to compare the stellar density deduced from the light curve with the one deduced from the spectrum. If the two average densities are too different, it could be a sign that there is a third body which causes decrease in the depth of the observed light minimum (Seager et al., 2003).
- Grazing stellar binaries: partially eclipsing binary systems have shallow minima which can be confused with exoplanet transits (see Fig. 1.8 d). Transits can be distinguished from partial eclipses via the shape of the minima. Transits produce U-shape minima while partial eclipses are V-shaped.

In addition, the activity of the host star can hide or make it difficult detecting of the shallowest transit events. To detect planets around very active stars, one widely used method is modeling the activity using the GPs (see Sect. 2.5).

1.2.4.2 Photometric surveys

Detecting the planets using the transit method has been done through both ground and space observations. Among the ground-based photometric observations, one

can mention HAT (Hungarian-made Automated Telescope⁶) and WASP (Wide Angle Search for Planets⁷) projects, which together have discovered several planets. Although many transit light curves are obtained by ground-based telescopes, space observatories are far more promising. They have many merits over ground-based observations, including:

- Avoiding atmospheric extinction and turbulence, light pollution, etc.
- Being able to keep observing up to 24 hours a day.

Here, I highlight a few last and current space missions that are dedicated to detecting planets with one focus on TESS data which I used its data.

- CoRoT:
CoRoT (Convection, Rotation et Transits planétaires, Baglin et al., 2006) was a satellite doing observations between 2006 to 2013. It had an observation strategy consistent with staring at a given star field for a duration of between 21 to 152 days (Deleuil et al., 2021). In November 2011, about 600 exoplanet candidates were screened by the CoRoT team. Among them, 34 planets were discovered and fully characterized by ground-based observations (Deleuil et al., 2021). A particularly notable discovery from this mission is the detection of the first rocky exoplanet – CoRoT-7b – as it was confirmed by RV follow-up (Queloz et al., 2009; Léger et al., 2009). However, the stars that CoRoT observed have a magnitude of roughly between 11 to 16 which makes it very difficult to have a RV follow-up to confirm and characterize the candidates.
- Kepler:
The Kepler mission (Borucki et al., 2010) did its main observations between 2009 to 2013, covering 115 square degrees, containing continuous observation of about 100,000 stars with an unprecedented precision of order 20 ppm for star magnitude $V=12$ for a 6.5-hour transit. One of the main advantages of Kepler was its observation strategy which was staring at stars for a long time. This strategy helped to detect planets at the longer orbital periods as well as to detect enough number of shallower transit events to increase the S/N of their detection. However, like CoRoT, Kepler observed the faint star typically between $V \approx 13-16$ magnitude which makes it again difficult to do RV follow-up observations. After 4 years, due to the failure of its second reaction wheel, NASA had to change the mission. From May 2014, Kepler continued with the new name K2 resuming its observation (Howell et al., 2014). Unfortunately, K2 had less precision compared to Kepler. Finally, it retired on October 30, 2018. During 9.6 years in orbit, Kepler sent the data to Earth in short cadence (1 minute),

6. HAT Exoplanet Surveys: <https://hatsurveys.org/>

7. WASP: <https://wasp-planets.net/>

long cadence (30 minutes), and quarterly full-frame images (30 min). Kepler's observations led to the discovery of more than 2,600 planets with different diversities from Earth-size planets to Jupiter size.

- TESS:

The Transiting Exoplanet Survey Satellite (TESS, Ricker et al., 2015) is a NASA space telescope was launched in 2018 and has been designed to observe nearly the whole sky with photometric precision from 60 ppm to 3 percent (Oelkers et al., 2018). TESS initially had planned to observe the sky for 2 years and 26 sectors. Later, because of the well-health of the telescope, its mission was extended to 4 years.

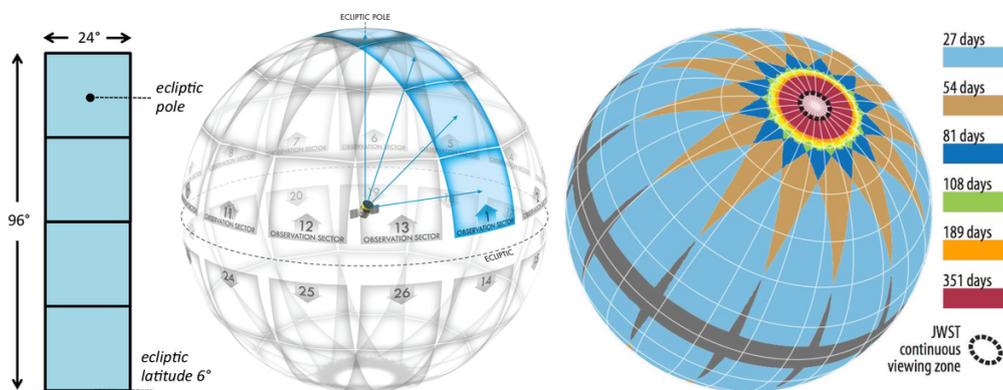


Figure 1.9 – *Left*: TESS cameras are aligned and cover 24*96 degrees. *Center*: TESS field of view (FOV) on each sector in the first 2-years of its mission and their overlaps. *Right*: TESS observation strategy for all the sky. Image credit Ricker et al., 2014.

TESS has a highly elliptical orbit in which at perigee, its distance from the Earth is $17 R_{Earth}$ while at apogee, this distance is $59 R_{Earth}$. TESS has 4 cameras and each of them has a 10.5 cm entrance pupil diameter, with a CCD containing 4096×4096 pixels, and $24^\circ \times 24^\circ$ field of view. The combined field of view for the all four TESS camera is $24^\circ \times 96^\circ$ (see Fig. 1.9 left). Every two years, TESS observes 85 % of the sky in 26 sectors, divided into 13 sectors for each hemisphere. Each TESS sector is 27.4 days (see Fig. 1.9 center), and each sector has 2 orbits (13.7 days) with a 2:1 resonance with the Moon. Each TESS sector has a slight overlap, leading to longer observation for some parts of the sky (see Fig. 1.9 right). As an example, at ecliptic poles duration of observations is 351 days. This is not accidental. The reason is these regions correspond to the continuous observation zone of the James Webb Space Telescope (JWST). Therefore, TESS plays an important role in providing targets for atmospheric follow-up by JWST.

TESS CCD are sensitive to a range of wavelengths from 600 to 1000 nm. In comparison to Kepler, the bandpass is extended further to red wavelengths, offering

the observation of a larger sample of M dwarfs (see Fig. 1.10). Another notable advantage of this mission is that TESS mostly observes stars that are brighter than magnitude 12, allowing follow-up studies to confirm and characterize the candidates.

TESS main scientific goals⁸ are:

1- Observing 200,000 pre-selected bright stars to detect planets with periods <10 days and radii <2.5 R_E in the solar neighborhood, as well as discovering planets with radii > 2.5 R_E distributed across the celestial sphere.

2- Discovering long-period planets, up to 120 days, in the ecliptic pole where TESS continues viewing zone is placed (see Fig. 1.9).

3- Measuring the mass of at least 50 transiting planets with a radius smaller than 4 R_E .

To date (May 2022), 217 TESS exoplanets have been confirmed, with an additional 5725 candidates awaiting confirmation⁹. Out of these 5725 TESS candidates, TESS discovered more than 1180 candidates with radii smaller than 4 R_E , more than 80 of which have been confirmed and have a mass measurement¹⁰.

TESS provides two formats of scientific data:

1. In 2-minute (short) cadence for 200,000 pre-selected stars from the TESS Input Catalog (TIC).

2. In 30-minutes with the full-frame image (FFI) for all other stars.

The Science Processing Operations Center (SPOC, Jenkins et al., 2016) pipeline, which is based on the Kepler Science Operations Center (KSOC) pipeline, reduces the short cadence of TESS data. This reduction for both spacecraft includes pixel calibration, removing systematic errors, calculating flux, and extracting raw light curves. The light curve obtained from simple aperture photometry (SAP) flux measurements (Twicken et al., 2010; Morris et al., 2017) does not include correction for common trends and instrumental effects, leaving them for users to remove. While flux measurements based on pre-search data conditioned simple aperture photometry (PDCSAP, Smith et al., 2012) are based on more processed data. This pipeline also produces a calibrated full-frame images (FFIs). Furthermore, the pipeline searches for periodic signals that could be caused by planets. The TESS vetting process then determines whether this is a false positive (FP) or a planet candidate. If this is a planet candidate,

8. <https://heasarc.gsfc.nasa.gov/docs/tess/primary-science.html>

9. <https://tess.mit.edu/publications/>

10. <https://tess.mit.edu/publications/>

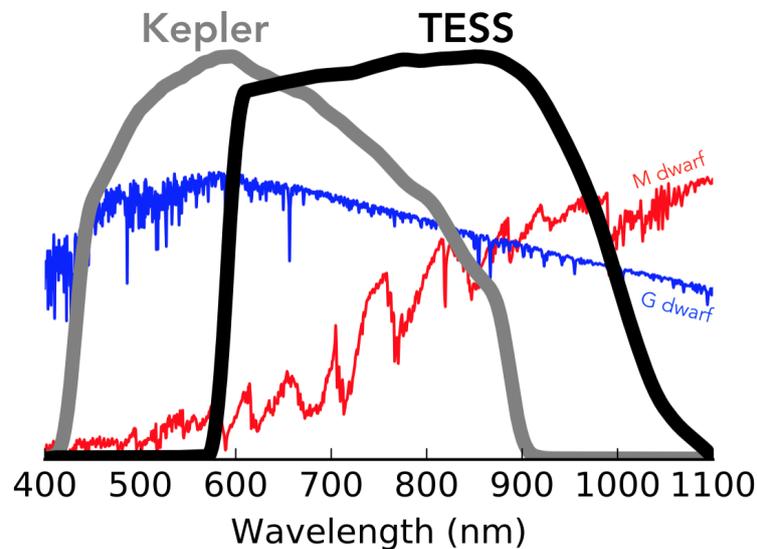


Figure 1.10 – TESS bandpass in comparison with Kepler. TESS bandpass is extended further to red wavelengths, offering the observation of a larger sample of M dwarfs.

known as TESS Object of Interest (TOI), it will be assigned by a planet number¹¹ and announced to the community for confirmation. Note that TESS data is publicly available on the Mikulski Archive for Space Telescopes (MAST)¹².

- CHEOPS:

CHaracterising ExOPlanet Satellite (CHEOPS, Fortier et al., 2014) is an European space telescope that was launched in 2019. It covers a wavelength range of 400-1100 nm and has a high precision of 20 ppm in 6 hours of integration time for a star with magnitude of $6 \leq V \leq 9$ (Benz et al., 2021). CHEOPS is the first dedicated exoplanet mission with the primary goal of following up on previously discovered exoplanets, improving (when possible) the radius measurement of a planet or providing a radius on planets that already have a mass measurement using other methods. Knowing both the mass and the radius of a planet, allow one to access its density and model the planet's internal structure. Note that the highly precise radius measurement of CHEOPS can greatly help to break the degeneracies in modeling planet interior structure, leading to improve our understanding of the planetary formation. CHEOPS can also study the atmosphere of planets using phase curves or the secondary eclipse method (Lendl et al., 2020, see Sect. 1.3.2 for more information about the methods). Additionally, the high precision of CHEOPS allows one to study

11. List of TOI can be found on <https://exofop.ipac.caltech.edu/tess/index.php> and <https://exofop.ipac.caltech.edu/cfop.php>, respectively

12. <http://archive.stsci.edu/kepler>

the stellar physics variability and deformation of exoplanets. For instance, CHEOPS's observations led to the discovery of WASP-103b, a rugby ball-shaped exoplanet caused by tidal interaction between the planet and the host star (Barros et al., 2022). For more details about the technical and scientific goals of the mission see Benz et al., 2021.

1.3 Characterising the exoplanets

1.3.1 Joint modeling of light and RV curves

Among ~ 5000 ¹³ discovered planets (and many more candidates), transiting planets (~ 3780) have a considerable impact on improving our understanding of the formation and evolution of planetary systems. Such planets when orbiting a bright host star, allow RV follow-up. Through the combination of both photometry and RV method, a planet can be accurately characterized in terms of fundamental parameters such as mass and radius and thus their density, allowing modeling of their internal structure and composition. Determining the mean density is very important because it is a criterion to show the type of a planet. As an example, the mean density of the Earth is 5.5 gr cm^{-3} and we know that it is rocky and Jupiter has a mean density of 1.33 gr cm^{-3} and we know it is a gas giant planet.

1.3.2 Atmospheric characterization

Studying the atmosphere of a planet is one of the most exciting areas of exoplanet characterization. Indeed, it is a great window to access planetary formation, evolution, composition, and also their habitability. Atmospheric study through breaking the degeneracies in planetary interior structure models can greatly help to study the diversity of planets and provide additional constraints on planetary formation models (see Sect. 1.3.4).

To study exoplanet atmosphere, for only transiting planets, one can use transmission spectroscopy, secondary eclipse, or phase curves method. When the planet passes in front of its host star, a small fraction of the host star's light will be absorbed by the planet's atmosphere. In this method, so-called transmission spectroscopy, through recording the planet's spectrum, one can infer the chemical composition of the planet's atmosphere or surface. In this method the predicted S/N can be calculated

13. <http://exoplanet.eu>

following the below formula (Kempton et al., 2018):

$$S/N \propto \frac{R_P^3 \times T_{eq}}{M_P \times R_*^2} \times 10^{-\frac{m_J}{5}} \quad (1.39)$$

where R_P and R_* are the radius of the planet and the star, respectively. M_P is mass of the planet, T_{eq} is the planet's equilibrium temperature, and m_J is the apparent magnitude of the star in J-band. Note that photometric bands differ depending on the instruments. Also, the planet is assumed to be a black body object with an albedo of zero when calculating its equilibrium temperature. In the secondary eclipse method, one probes the planetary atmosphere by measuring the difference between the emitted flux of the summed planet and star before an eclipse and only the star itself during the eclipse. One can also continue to measure the planet's emission variation during the complete planet's orbit and obtain a phase-curve resolved from planet-emitted light (Crossfield et al., 2020). The study of the planet's emitted light can be done through both photometry and spectroscopy methods. Massive short-period planets, having larger emission atmospheric signals, are more suitable for the study of their atmosphere through secondary eclipse and phase curves methods.

For non-transiting and transiting planets, one can also study the planet's atmosphere through direct imaging. In this method, by using low, medium, or high-resolution spectra, the planet is resolved from the star. Then, the reflected light of the star from the planet is directly measured. This method is suitable for the giant planets ($\geq 1 M_{Jup}$) which are at large separation (tens of AU) from their host star (Cheetham et al., 2019).

1.3.3 Classification of Exoplanets

As the number of detected planets increases, their wide diversity was also revealed. Astronomers classify planets based on planetary mass and radius. Classification by size is as follows (Borucki et al., 2011):

- Earth-size ($<1.25 R_E$)
- Super-Earths ($1.25 - 1.75 R_E$)
- Sub-Neptunes ($1.75 - 3.5 R_E$)
- Neptune-size ($3.5 - 6 R_E$)
- Jupiter-size ($6 - 15 R_E$)

It is worth mentioning that different authors may use different radius boundaries to classify planets (e.g. Morbidelli et al., 2016). Regarding planet mass, one can also

classify planets as follows (Stevens et al., 2013):

- Earth planets ($0.1 - 2 M_E$)
- Super-Earths ($2 - 10 M_E$)
- Neptunes ($10 - 100 M_E$)
- Jupiters ($100 - 10^3 M_E$)
- Super-Jupiters ($10^3 M_E - 13 M_J$)
- Brown Dwarfs ($13 M_J - 0.07 M_{sun}$)
- Stellar Companions ($0.07 M_{sun} - 1 M_{sun}$)

It is also possible to regroup some of these boundaries. For instance, planets with a mass under 30 Earth masses are known as 'low-mass planets' (Fortney et al., 2013; Jontof-Hutter, 2019). I note that classifications based on compositions of planets are also commonly used, which I will explain in details in the following section 1.3.4. For the rest of this thesis, I use the mentioned characterizations.

1.3.4 Internal planet structure

Precise mass and radius measurements of planets, allow us to model the planetary bulk composition. The most important classifications based on planetary composition are as following¹⁴:

- Giant planets: these planets are similar to our own gas giant planets in the solar system. The giant planets have a planet core which is covered by a thick layer of hydrogen and/or helium.
- Neptunes: these planets typically have a core composed of rock and heavier metals. This core is covered by a mantle of water, ammonia, and methane (referred as ice because they are usually found frozen in the outer solar system) with an envelope mostly composed from hydrogen and helium.
- Super-Earth: this class of planets that do not exist in our solar system, have a composition made of gas, rock, or a combination of both.
- Rocky (or terrestrial): they are mostly consistent with a composition of the rock, silicate, water, and/or carbon. Such planets may possess a thin layer of atmosphere.

Of course, there are more types or sub-types of planetary composition, particularly

14. This category is based on Nasa exoplanet exploration <https://exoplanets.nasa.gov/what-is-an-exoplanet/planet-types/overview/>

toward the low-mass planet regimes. For example, intermediate-size planets, those between super-Earths and sub-Neptunes could be made of 50% of rock and 50 % of ice, so-called 'water world' (Zeng et al., 2019), e.g. GJ 1132 b (Southworth et al., 2017). Another example is planets with a similar mass to super-Earth $\leq 10 M_E$ and mostly close to their host star (0.1-0.3 AU), but with a large radius of $\geq 4 R_E$, so-called 'super-puff' (e.g. Kepler-51 system (Masuda, 2014)). The composition of such planets with extremely low density ($\leq 0.3 \text{ g cm}^{-3}$) is consistent with a significant contribution of gas mass fraction of $> 10 \%$ (Lopez et al., 2014), exceptionally greater than other super-Earth populations. The extremely low density of such planets is a puzzle from the point of view of planetary formation scenarios (Lee et al., 2016b; Piro et al., 2020; Wang et al., 2019).

1.4 Thesis objective

Fig. 1.11 top, shows a radius-mass diagram of known low-mass planets from the NASA Exoplanet Data Archive (July 12, 2022) with 50 % precision on mass and radius, overplotted by different density lines. This figure clearly illustrates that the low-mass planets have an extreme diversity in their density, and consequently in their bulk composition. Planets with a few Earth masses can have a size between super-Earth or sub-Saturn (see Fig. 1.11 bottom), and planets with a few Earth radii can have a wide range in size between 2–15 Earth masses (Jontof-Hutter, 2019). This wide range in density and composition, even within a system (e.g. Kepler-36 b (Carter et al., 2012)), raises several questions regarding theories of planet formation and atmospheric loss. It is currently under debate which possible scenario(s) mainly shape the different diversity of low-mass planets' composition. These scenarios include: dynamical evolution over the age of system (Pu et al., 2015), existence of a giant planet (Huang et al., 2017), host star obliquity (Spalding et al., 2016), resonant chains which remain stable or may cause an interruption in postmigration (Izidoro et al., 2017), multiple formation channels (Lee et al., 2016a), atmospheric loss (e.g. Inamdar et al., 2015), or variation in formation conditions (Moriarty et al., 2016). Addressing our questions regarding the diversity of low-mass planets' composition is currently mostly limited by the number of well-characterized planets.

The goal of this thesis is to contribute to the study of the diversity of low-mass planets composition, by increasing the number of precisely characterized low-mass planets, making the link between the observations and the formation paths of these bodies, and providing essential constraints on scenarios of low-mass planet formation. To achieve this goal, three years ago in 2019, I set two main steps as follows:

- 1) Overcoming the obstacle of instrumental limitation: to improve RV data reduction and processing which facilitate the detection of low-mass planets (see Chapter 2

for details of method).

2) Overcoming the challenge of low-mass planets detection and accurately characterizing them: to extend the sample of well-characterized low-mass planets in the different environments to examine the different formation theories (see chapter 3 for details of method).

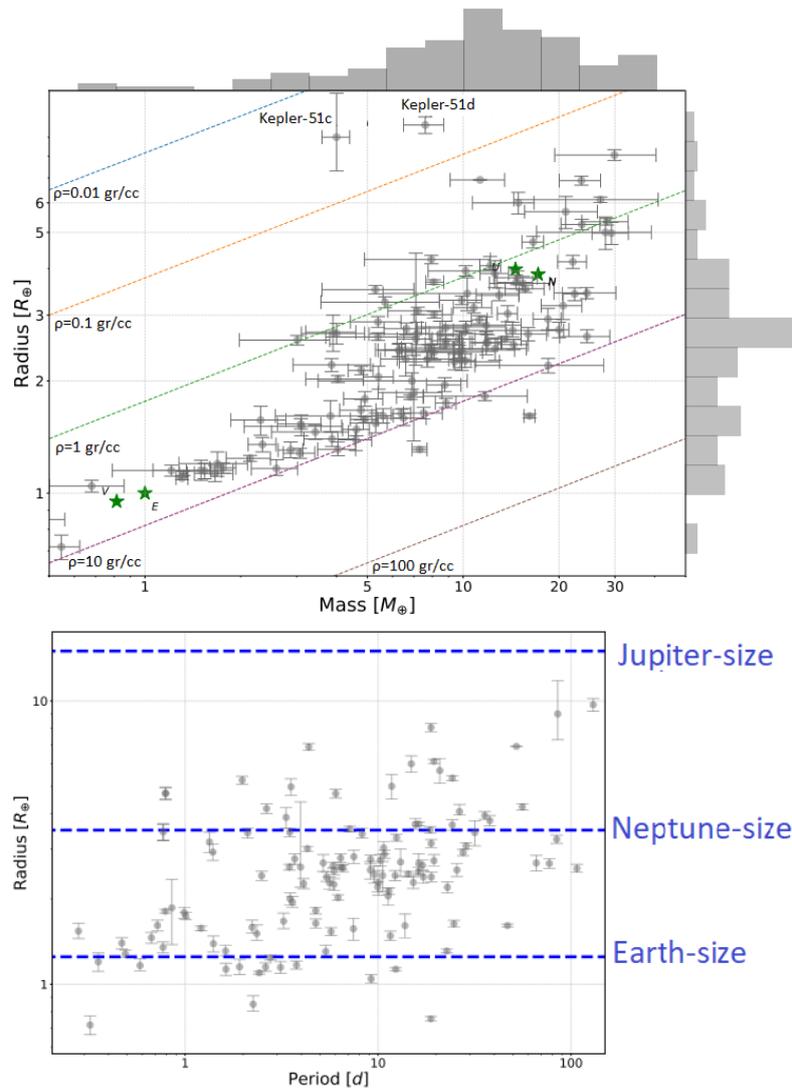


Figure 1.11 – *Top*: Radius-mass diagram of the known low-mass planets from the NASA Exoplanet Data Archive (July 12, 2022) with 50 % precision on mass and radius. The plot is overplotted by different density lines, with point out the Kepler-51 system as an example of 'supper puff' planets. *Bottom*: Radius-period diagram of the same planets, showing that low-mass planets can have a size between Earth to sub-Saturn. These two figures clearly show the high diversity of low-mass planets.

To do so, I used both photometric and RV data, which combining them for transiting planets, allows one to measure the fundamental parameters such as mass and radius, leading to inferring their density and bulk composition. For photometric data, I used mostly TESS data which at the start of my Ph.D., 2019, began to observe the northern hemisphere. For RV data, I mostly used SOPHIE high-precision RV data as part of its consortium.

By the end of my Ph.D., I carried out a general study of the known low-mass planets as of 2022. I reviewed the key exoplanet science questions remaining to be answered in the future. I also investigate how these questions can be answered by current and future instruments which gives me a way forward for my future research projects (see [chapter 4](#)).

Chapter 2

Optimizing SOPHIE data reduction

Contents

2.1	The SOPHIE high precision spectrograph	47
2.2	The SOPHIE-red project and my contribution	50
2.3	Atmospheric dispersion effect correction	52
2.3.1	Method	53
2.3.2	Results	53
2.3.2.1	Improvements of mean RVs RMS	53
2.3.2.2	Improvements of mean FWHM RMS	54
2.3.2.3	Detecting new activity signals in FWHM	56
2.3.3	Conclusion	57
2.4	Solar contamination correction	58
2.4.1	Review on existing methods	58

2.4.2	Optimized conditions and results for simultaneous sky observation mode	59
2.4.3	Optimizing conditions and results for simultaneous calibration mode	64
2.4.4	Conclusion	66
2.5	Correction of background contamination of the calibration lamp	67
2.6	Optimizing long-term variation from the zero-point	69
2.6.1	High resolution mode	69
2.6.1.1	Building a new flux corrected master constant time series	71
2.6.1.2	HD185144 activity correction	71
2.6.1.3	Building a FWHM master constant correction	78
2.6.2	High Efficiency mode	81
2.7	Fixing mask lines selection for spectra at different epochs	82
2.8	Conclusion	84

2.1 The SOPHIE high precision spectrograph

SOPHIE is a high-resolution spectrograph mounted on the 1.93 m telescope at Observatoire de Haute-Provence (OHP), France (telescope shown in Fig. 2.1, left). It is a fiber-fed echelle spectrograph in a temperature and pressure-stabilized environment that has been in operation since 2006 (Perruchot et al., 2008; Bouchy et al., 2013). For stabilizing the temperature, three levels of temperature control have been designed: 1) the spectrograph is located in a thermally insulated chamber; 2) this chamber is placed in an air-conditioned room; and 3) the room has an insulated wall. To ensure constant pressure, the spectrograph is supported by the telescope pier structure and positioned on shock absorbers. Furthermore, all of its dispersive components are located in a constant-volume vessel filled with Nitrogen. The spectrograph has a $2k \times 4k$ CCD with $15 \mu\text{m}$ pixel size and an overall CCD efficiency of 80%, cooled to -100°C . Figure 2.1 (right) shows a picture of SOPHIE, and Figure 2.2 shows a schematic of the spectrograph with details on its different parts.

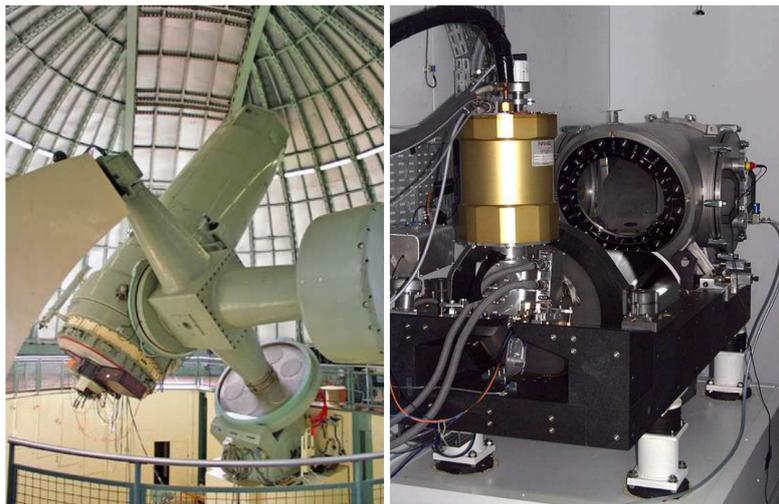


Figure 2.1 – The SOPHIE spectrograph (*right*) mounted on the 1.93 m telescope (*left*) at OHP.

The spectrograph has 39 spectral orders and covers the wavelength range from 387 nm to 694 nm. Each order has a spectral range of about 100 \AA and overlaps with adjacent orders. It operates in two modes:

- High resolution (HR) mode with a resolution power of 75, 000 and a precision of $1\text{-}2 \text{ m.s}^{-1}$ (Bouchy et al., 2013). This mode is used for the brightest stars ($V < 12$).
- High efficiency (HE) mode with a resolution power of 39, 000 and a precision of $3\text{-}4 \text{ m.s}^{-1}$ (Bouchy et al., 2013). This mode is selected for the observation of

fainter stars ($V > 12$).

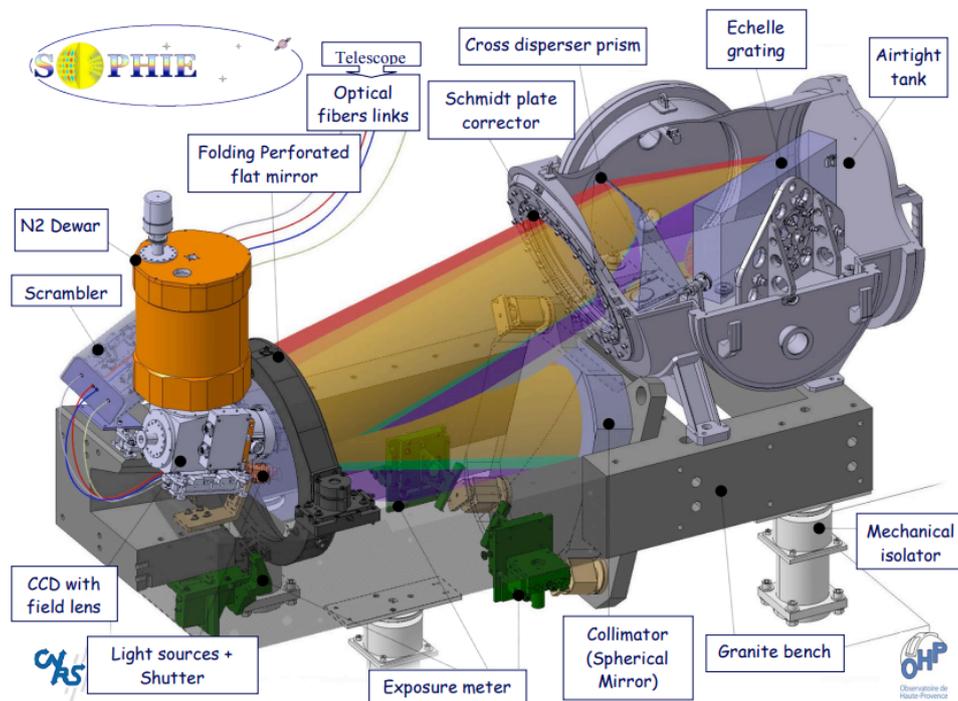


Figure 2.2 – Schematic of the SOPHIE spectrograph. Image credit: Perruchot et al., 2008.

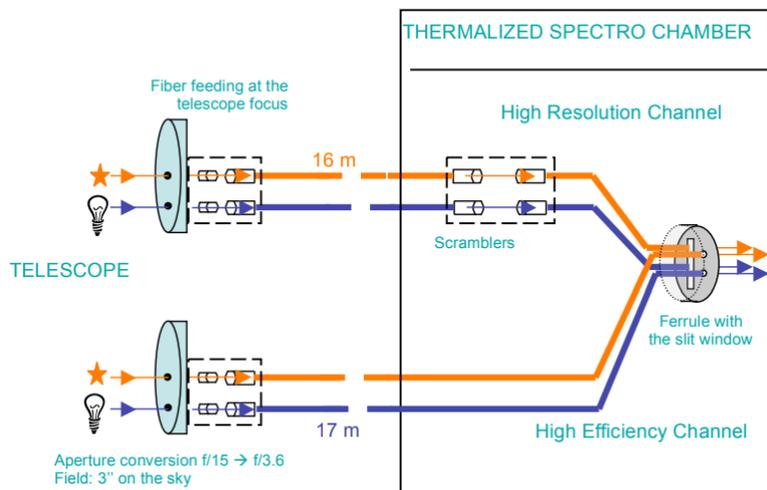


Figure 2.3 – Schematic picture of two observation modes of HR and HE in the SOPHIE spectrograph. Image credit: Perruchot et al., 2008.

The differences between these the two modes are due to the two extra features of the HR mode: 1) a $40.5 \mu\text{m}$ slit on the fiber at the entrance to the spectrograph (Fig. 2.3)

which allows a higher resolution, and 2) a scrambler to increase homogenization and stabilize the illumination which improves the RVs' precision. Despite the advantages of these two extra features, they cause a loss of flux. Therefore, the HR mode is only suitable for sufficiently bright targets ($V < 12$).

Each mode uses two optical fibers (Fig. 2.3), separated by 1.86 arc-min, so-called 'fiber A' and 'fiber B'. Fiber A, also known as a science fiber, is dedicated to star observations, and fiber B can be used for either simultaneous sky observations or calibration lamp observations. The option to be used depends on the brightness of the target and the goal of the observation. For the brighter targets (typically magnitude < 9), the RV photon noise is compatible with the accuracy of the instrument. In this case, to obtain the best accuracy, fiber B will be used for simultaneous calibration lamp observations to track the small drift of the telescope. However, not performing sky observations comes with the risk of not being able to correct the Moon contamination. For the fainter stars (typically magnitude > 9), the RV photon noise is higher than the accuracy of the instrument. Therefore, fiber B will be used for simultaneous sky observations to record the sky background contamination.

SOPHIE has an automatic Data Reduction Software (DRS) which was initially adopted from the HARPS spectrograph. Each observation night, the DRS begins its operation with afternoon calibrations, which include creating dark current, flat field maps, wavelength solutions from a Thorium-Argon (Th-Ar) spectrum, and setting a zero-point with Fabry-Perot etalon (FP) for measuring drift over the night. Then, during the night, it performs the following operations a few seconds after the end of the observation:

- Corrects the data from cosmic rays, one-dimensional flat-field, and CTI (charge transfer inefficiency) effect. The latter is caused by the imperfect transfer of charge between two adjacent pixels in the CCD (Bouchy et al., 2009b).
- Calculates the Barycentric Julian Date (BJD), and the Earth's barycentric velocity (BERV) which is the Earth's motion relative to the Solar System barycenter, projected in the direction of the line of sight.
- Measures the drift of the spectrograph for simultaneous calibration lamp observation. This is done through the Th-Ar (up to semester 2017A) or the FP (since semester 2017B) with respect to the zero-point from the afternoon calibration. To measure small wavelength shifts, the spectrograph drift must be precisely measured. To do so, the FP provides better precision on the wavelength solution compared to the Th-Ar lamp. This is due to the fact that it presents several lines that are uniformly spaced and cover the entire wavelength range of the spectrograph (Das et al., 2018).

- Separates the data in the 39 orders and stores it together with other information (e.g. BJD) in a two-dimensional FITS file. Part of a raw image of SOPHIE spectral orders, with a star on fiber A and the FP or Th-Ar on fiber B, is shown in Fig. 2.4.
- Measures the RV and some activity indicators (for the detailed method, see Sect. 1.2.3.2) and adds this information as keywords in the headers of FITS files.
- Removes the BERV effect, merges the spectral orders, and calculates a one-dimensional spectrum. Then it stores them in a one-dimensional FITS file.

The SOPHIE Exoplanets consortium has different planet research programs, which I will describe in chapter 3 of this thesis. For a complete description of the optical design of the instrument, see Perruchot et al., 2008.

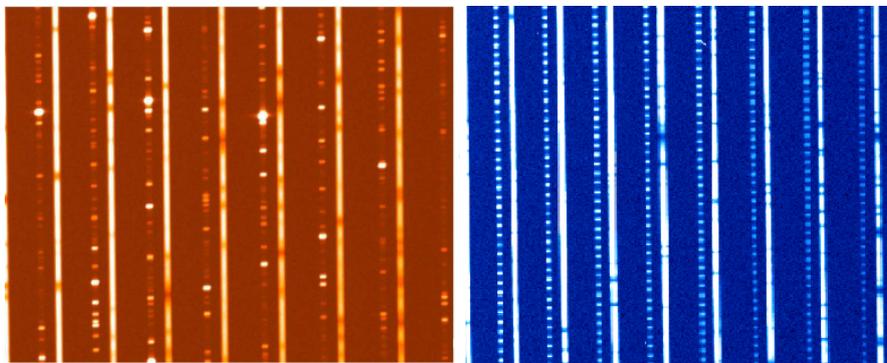


Figure 2.4 – Part of a raw image of SOPHIE spectral orders, with a star on fiber A and Th-Ar (*left*) or the FP (*right*) on fiber B. Around 8 spectral orders can be seen in the figure. The distance between the orders varies from ~ 33 pixels in the red to 70 pixels in the blue part of the spectrum. The absorption lines of the star spectrum, and the emission lines of the calibration lamp, are shown in the figure. The distance between each fiber in the same order is about 17 pixels.

2.2 The SOPHIE-red project and my contribution

In the early years of SOPHIE operation, the RV precision on stable stars was limited to 5-6 m/s (Boisse et al., 2009; Diaz et al., 2012), which was below the accuracy required to detect low-mass planets. In 2013, a significant upgrade was accomplished by utilizing a new fiber link that included octagonal-section fibers (Bouchy et al., 2013). This enables RV precision to reach 1-2 m/s level, making it suitable for discovering low-mass planets with masses down to 5-10 M_E . SOPHIE's enhancement is continuing with a new project so-called 'SOPHIE-red'. In this project, several improvements are

planned, including changing the camera, changing the fiber train, improving the DRS, and automatically monitoring the instrument.

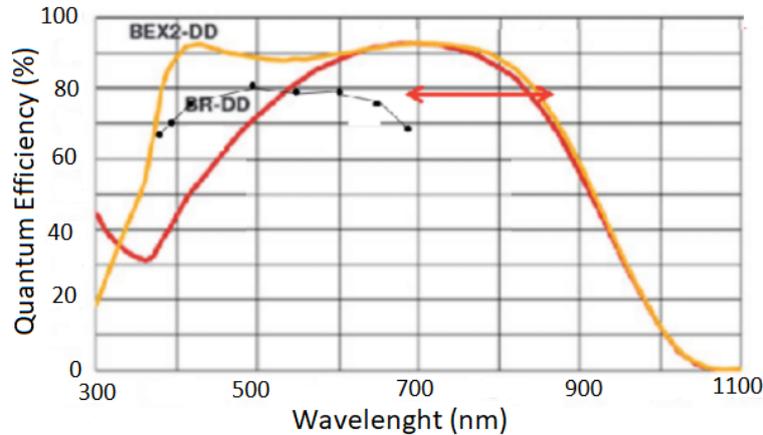


Figure 2.5 – The difference between the quantum efficiency of the current (black) and planned (yellow) CCD in the different wavelengths. The red color is related to a different proposed CCD with a different design that was not selected in the end.

To do this, the current CCD will be replaced by a new $4K \times 4K$ CCD (instead of $2k \times 4k$ for the current one), with a $15 \mu\text{m}$ pixel size, and an overall efficiency of 90 % (vs 80 % for the current camera). Additionally, with this new CCD, the SOPHIE wavelength will be extended toward the red (around 5 spectral orders more in red wavelength), hence the name 'SOPHIE-red'. Fig. 2.5 illustrates the difference between the quantum efficiency of the current (black) and planned (yellow) CCDs in the different wavelengths. It clearly shows the improvements in the red domain. Additionally, one of the major issues with the current camera is that the CTI effect is quite large. The new camera is expected to have a much lower CTI effect. The new camera will be installed in spring 2023, therefore SOPHIE-red will be available in semester 2023B. With this upgrade, SOPHIE will stay a key instrument in exoplanet detection.

A highly precise data reduction system can carefully treat potential systematic error sources and increase data precision. Therefore, in addition to this instrumental improvement, an enhancement of the SOPHIE DRS was performed, which I worked on. The first part of this work was the update of the DRS from python2 (Fortran) to python3.7 for computational improvements. This allowed me to acquire a detailed understanding of the software. Secondly, I have developed and tested methods to improve the DRS. In the following sections, I present my work on optimizing the DRS.

2.3 Atmospheric dispersion effect correction

One of the problems for any astronomical ground-based observation is the atmospheric dispersion effect. The atmosphere, due to its changing refraction, acts like a prism and disperses light as it passes through. This wavelength-dependent phenomenon influences the distribution of the stellar flux.

In a high-precision spectrograph that aims for very precise velocities, correcting this effect should be considered. Indeed, the atmospheric dispersion can introduce a slope to the continuum (see Fig. 2.6) and thus change the measured mean RV of the targets (Pepe et al., 2008). This effect is more serious for blue wavelengths with higher refraction, where the activity indicator of Ca II is located. This activity indicator is important for detecting stellar activities.

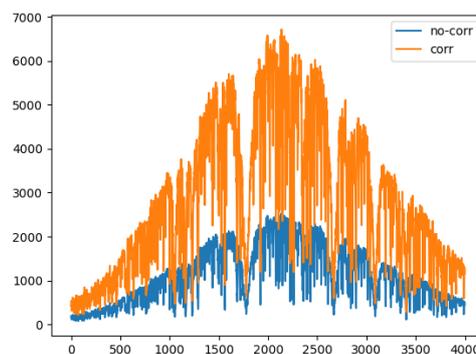


Figure 2.6 – Before and after flux correction of one of the SOPHIE stars in the spectral order of 3.

The SOPHIE spectrograph has an Atmospheric Dispersion Compensator. However, it still leaves residuals of atmospheric dispersion in the data, which should be taken into account. Here, I investigate how to correct this effect from SOPHIE spectra. To do so, I first describe the method (Sect. 2.3.1). Then, I show the results of tests of this method on SOPHIE Sub-programme 1 (SP1) targets (Sect. 2.3.2). SP1 is a SOPHIE high-precision program dedicated to the discovery of super-Earths and Neptunes (for details on the SP1 targets see Chapter. 3).

2.3.1 Method

To correct this effect, we first observed a list of standard stars with different spectral types (F, G, and K), in high-resolution mode with a high signal-to-noise ratio (S/N) and low air mass. Then, by normalizing the sum of each spectrum order's flux to the flux at an arbitrary wavelength (e.g. 550 nm), we built several templates for different stellar spectral types (Modigliani et al., 2019). These templates are saved as a reference for the expected flux distribution for each spectral type.

Then, I adopted the flux correction function which was inactive in the current DRS. Based on the target spectral type, this function first finds the best available flux template in the SOPHIE archive. Then, to correct for the atmospheric dispersion, the target spectrum is scaled by multiplying it by the flux ratio of the target spectrum and the corresponding template (Wehbe et al., 2020). In fact, it corrects the spectrum's flux distribution with respect to the reference templates, which, because of their low air masses, represent our anticipated flux distribution. If the flux ratio of the spectrum and the corresponding template is too low (<0.25) or too high (>3), due to a mismatch with the flux template, the function is automatically turned off. I note that these values are adopted from the HARPS DRS.

2.3.2 Results

I applied the flux correction method on all the SOPHIE SP1 stars and derived RVs and associated activity indicators. Then, I investigated the impact of flux correction on the data. The results are presented below. All of the tests in this chapter are performed on SP1 stars with more than ten observations, for a total of 96 targets.

2.3.2.1 Improvements of mean RVs RMS

Here, as a case study, I first describe the results of applying the flux correction method to one of the SOPHIE constant stars, HD89269A. Then, I show the results of using this method on all of the SP1 targets as well as the other SOPHIE constant stars.

HD89269A is a G4 spectral type star. Because of its low RV dispersion, it is used as a SOPHIE constant star. By applying the flux correction method on this star, I reduced the RVs RMS by 45 cm/s, which is 38 % of the star mean error bars 1.2 m/s (see Table 2.1). Fig. 2.7 shows the RV time series of the star with and without flux correction. Following the correction, a number of data points exhibit improved consistency with other data. In fact, these data points have a higher air mass (>1.4), and thus they are

Table 2.1 – Results of the template correction method on constant stars

Target	No. observations	RMS-corrected (m/s)	RMS-uncorrected (m/s)	Mean error bars (m/s)	Air-mass ranges
HD185144	410	1.98	2.00	1.09	1.1- <1.6
HD9407	185	2.30	2.34	1.66	1.1- <1.5
HD89269A	391	2.78	3.23	1.21	1- <1.6
HD221354	32	2.34	2.32	1.37	1- <1.6
Mean		2.35	2.47	1.33	

more polluted by atmospheric dispersion; this method successfully corrects them.

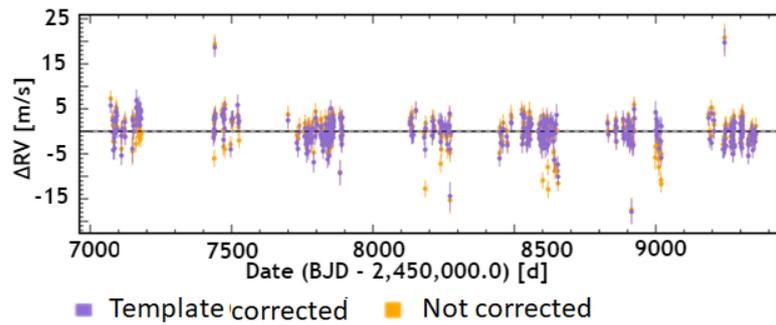


Figure 2.7 – RVs of HD89269A, a SOPHIE constant star, with and without flux correction. The RVs RMS is reduced by 45 cm/s, or 38% of the star mean error bars 1.2 m/s. Several data points after the correction show better consistency with other data.

I applied the flux correction method on other SOPHIE constant stars (see sect. 2.6 for more explanation of the targets). The results are presented in Table 2.1, which shows a 12 cm/s improvement in mean RVs RMS. This improvement is 9 % of the mean error bar of 1.33 m/s (last column). Also, by applying the flux correction on all SP1 targets, I improved their mean RVs RMS by 8 cm/s, from an initial value of 7.86 m/s to the final value of 7.78 m/s. This improvement represents 7 % of the mean error bars of stars (1.2 m/s).

2.3.2.2 Improvements of mean FWHM RMS

The Full-Width Half of Maximum (FWHM) of the RV CCF is one of the important proxies to search for stellar activities. FWHM has been demonstrated to be a reliable tracer of stellar activity on other high-precision spectrographs (Astudillo-Defru et al., 2017b; König et al., 2022). Until now, it was not the case for the SOPHIE data, due to this parameter being too noisy (Boisse et al., 2012b). In this section, I show the substantial improvement obtained by applying the flux correction method, first on the HD89269A FWHM, then on all SP1 targets.

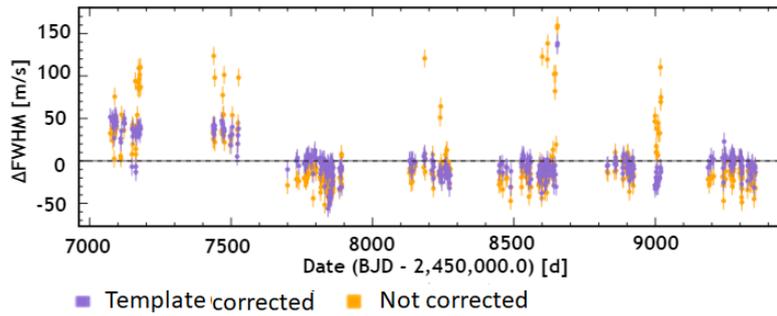


Figure 2.8 – The orange points show the FWHM time series of HD89269A, which includes several outliers, before applying the flux correction. By applying the flux correction, these outliers in the FWHM time series become consistent with the rest of the data (purple points), and the time series improves greatly in stability.

HD89269A: I applied the flux correction method to the SOPHIE constant star HD8926A. The orange points in Fig. 2.8, represent the FWHM time series before applying the flux correction; it includes several outliers. These outliers, as I previously mentioned, are linked to dates with higher air masses (> 1.4), which led to higher levels of atmospheric dispersion pollution. The flux correction eliminated the outliers in the FWHM time series (purple points), making them consistent with other points, and the time series became significantly more stable. Indeed, the mean FWHM RMS was reduced from 29 m/s to 21 m/s. This 8 m/s improvement is three times greater than the 2.8 m/s error bar.

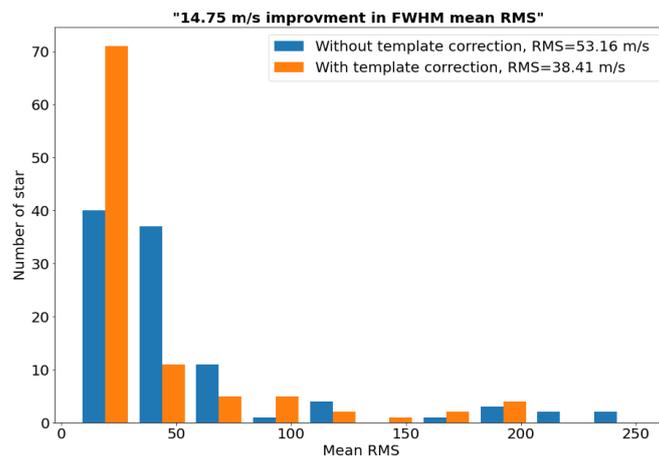


Figure 2.9 – Distribution of the FWHM means RMS of SOPHIE SP1 targets, with/without performing the flux correction method. Applying the flux correction led to a 14.75 m/s improvement in the FWHM mean RMS.

All SP1 targets: I used the flux correction method on 96 SP1 stars, which resulted in a mean FWHM RMS reduction from 53.16 m/s to 38.41 m/s, a gain of 14.75 m/s. This gain is 5 times bigger than the 2.8 m/s mean FWHM error bars of these stars (see Fig. 2.9).

2.3.2.3 Detecting new activity signals in FWHM

FWHM activity indicator can help in confirming the nature of a signal and correctly interpreting data (Santos et al., 2014; Lovis et al., 2011). I present a case study here to demonstrate that the SOPHIE FWHM data after flux correction can achieve this objective.

Fig. 2.10 (top), shows the RV periodogram of one of the SOPHIE stars with a signal at 5.9 d with a FAP of below 1 %. The periodogram of the bisector span also displays the same signal with a FAP below 10 % (*second* panel). Thus, it is likely that the signal at RVs has a stellar activity origin. When the flux correction was applied to this star's data, there were no differences in the RV and bisector periodograms, but there were in the FWHM. FWHM periodograms without and with flux correction are displayed, respectively, in the *third* and *fourth* panels. A peak with a close period to the RV signal appears in the periodogram of the corrected FWHM activity indicator, which was not present before the correction. This demonstrates that FWHM is correctly tracing an activity signal after flux correction. Therefore, the improvement in FWHM following the flux correction can greatly help in distinguishing between planet signals and false positives.

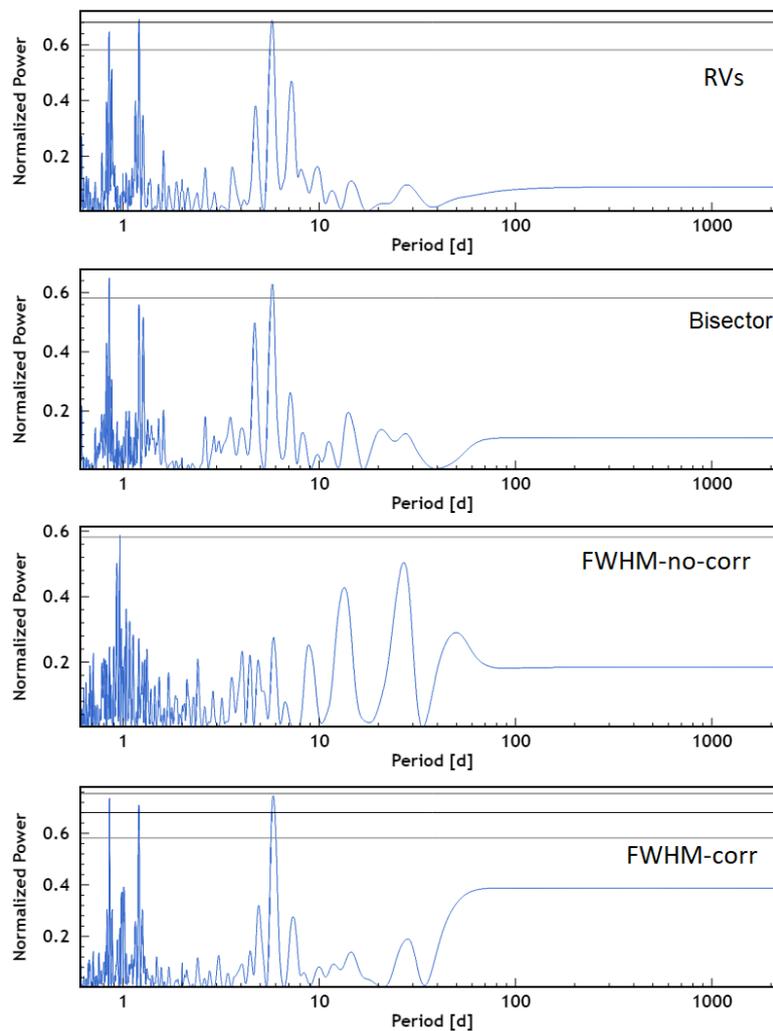


Figure 2.10 – *First-panel*: RVs periodogram of one of the SOPHIE stars, with a signal at 5.9 d and its aliases. *Second-panel*: Bisector periodogram of this star which displays the same signal as RVs. I note that I found no differences between the bisector and RVs periodograms before and after applying the flux correction method. *Third-panel*: FWHM periodogram before performing the flux correction. *Fourth-panel*: FWHM periodogram after applying the flux correction. The same signal in RVs appears in the FWHM corrected periodogram.

2.3.3 Conclusion

In this section, I presented the correction of the atmosphere dispersion effect by applying the flux correction method. To correct for this effect, first, we built several reference templates by observing a list of constant stars with different spectral types

at low air mass. Then, using these templates, I corrected the impact of atmosphere dispersion on the spectra.

I implemented and tested this method on 96 stars from the SP1 sample. I found a great improvement in both the mean RVs RMS (7 % of the mean error bars) and the mean FWHM RMS (5 times greater than the mean error bars). The FWHM is a good proxy for discriminating between false positives and real planet signals. As a result of this attempt, the FWHM has become significantly more stable and can now be used to track stellar activity.

2.4 Solar contamination correction

As a part of the SOPHIE data reduction optimization, here I investigate spectral contamination caused by background sky brightness, which is one of the systematic noise sources on RVs. Background sky brightness, significantly due to the Sun's light reflected by the Moon, pollutes the spectral lines, distorting the shape of the CCF and shifting the extracted RVs. This effect should be considered in order to obtain an accurate data analysis.

2.4.1 Review on existing methods

In SOPHIE, to evaluate the sky background brightness, mostly for the fainter stars ($V > 9$), fiber B is used to monitor the sky simultaneously with the target observation on fiber A. Then, for correcting this contamination, one proposed method is the procedure of Pollacco et al., 2008. It includes the following steps: 1) scaling the calculated CCF on fibers A and B by the throughput of the two fibers; and 2) subtracting these two scaled CCF from each other in order to obtain an expected clean stellar CCF. This method has been successfully performed on several SOPHIE RV measurements (e.g. Hébrard et al., 2008, Santerne et al., 2011). Another method for correcting the polluted spectra is sky model subtraction (Roy et al., 2020). To do this, first, the CCF of the sky fiber is computed. Then, its CCF continuum level is used to scale a sky model, which is a synthetic or high-S/N solar spectrum. After that, the scaled sky model is subtracted from the star fiber spectrum. This method is particularly useful when sky and star fibers have different geometry or resolution. However, it is not as simple as direct CCF subtraction, and the extra steps (e.g., choosing a true sky spectrum) may cause more errors. As a third approach, one can mention the method used in the ESPRESSO spectrograph. In this method, after scaling the two spectra by the throughput of the two fibers, the sky spectrum is re-binned to the same wavelength scale as fiber A. This resampling is required because the center of each pixel is not

exactly the same in the fiber A and fiber B wavelengths. Then, the two spectra are subtracted, either pixel by pixel or after smoothing the sky spectrum.

Systematically subtracting the sky spectrum from that of the star (Moon correction) would lead to introducing extra noise when the moonlight contamination is negligible; applying a condition to detect the Moon-polluted spectra before applying the correction prevents this. Diaz et al., 2012 considered a data point to be contaminated if it fulfills two conditions:

1. The moonlight should contribute significantly to the star fiber. They translate this condition as having a S/N of star fiber > 3 , and a sky CCF-contrast $> 1\%$, i.e. Moon CCF is detected in the sky spectrum rather than noise.
2. The star and Moon RV should be close. They refer this condition to $|RV_{target} - BERV| < 7 \times FWHM_{target}$. In this condition, BERV represents the Earth's motion relative to the solar system's barycenter, projected in the direction of the line of sight. BERV in the time of observation and in the direction of the target is within ~ 1 km/s of Moon RVs (Diaz et al., 2012). Consequently, it is fairly reasonable to consider BERV to be a Moon RV.

For the observations obtained with a simultaneous calibration lamp observation, for which no sky monitoring is possible, the authors considered a spectrum to be Moon-polluted when the Moon phase is more than 40 % visible in the sky. Note that the Moon phase is the Moon's apparent shape in the sky (e.g., new moon or full moon). Besides this, the same criterion based on the BERV was applied, i.e. $|RV_{target} - BERV| < 7 \times FWHM$.

To achieve the desired precision in the RVs analysis, in the subsequent sections, I attempted to optimize the conditions for detecting Moon-contaminated spectra for both simultaneous sky and calibration lamp observation modes.

2.4.2 Optimized conditions and results for simultaneous sky observation mode

If the Moon RV is sufficiently separated from the stellar RV, its influence on the stellar RV is negligible and no correction is required. Here, to introduce an optimized condition, I verify how far this separation goes. To do this, I used two tests on the Kepler16AB spectra and investigated the effect of the sky CCF on the star CCF. Kepler16AB is a binary star, hosting a planet with an orbital period of 229 d (Triaud et al., 2022). SOPHIE observed this star simultaneously with the sky background from 2016

to 2021. These tests are (see Fig. 2.11):

1. Examine the observations in which the sky CCF and the star CCF are at various separations from each other.

2. Examine the effect of the two CCF on each other for different amounts of sky and star CCFs overlap. This is tested by gradually widening the RV steps for the computing star CCF.

The full results of these tests and computed RVs are presented in Table 2.2. They show: 1. the effect of the sky-CCF on the star-CCF with and without overlaps is much smaller than the amount of RV correction provided by the Moon correction method (last column). For instance, in Fig. 2.11 (top), the narrowest star CCF width (dark blue) does not overlap with the sky CCF (cyan), and the RV is 20.474 km/s. When the CCF of the sky spectrum and the star are fully overlapped, for the largest star CCF width (light green), the RV changes to 20.477 km/s (i.e. $\Delta R = 3$ m/s), while the correction by the Moon correction method is $\Delta R = 26$ m/s. Therefore, the SOPHIE Moon correction routine is likely to overestimate the corrections. 2. A comparison of RV variations for different amounts of CCFs overlap, for the dates 2018-07-29 and 2019-06-15, where the sky CCF is sufficiently separated from the star CCF (> 2 FWHM, FWHM ~ 9 km/s), shows that the sky CCF has no significant effect on the star CCF.

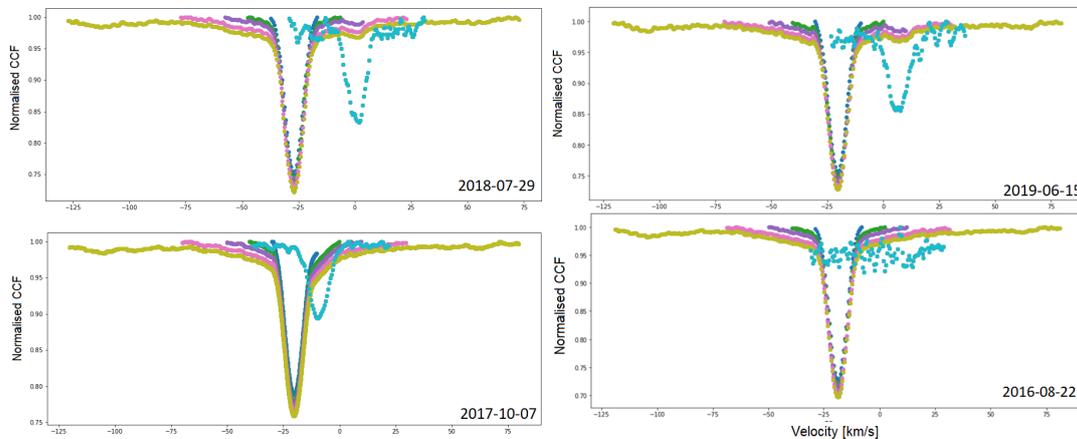


Figure 2.11 – The sky and the star CCFs of Kepler16AB. The cyan color represents the sky CCF, and other colors correspond to the star CCFs that were computed by the various RV steps. With this test for each data point, I investigate the effect of the sky-CCF on the star-CCF.

Given the small effect of the sky CCF on the star CCF when they are far enough apart ($> 2 \times \text{FWHM}$), and the significant change of RVs after the Moon correction process, I suggest applying the Moon correction to the data when the separation between the sky-CCF and the star-CCF is $< 2 * \text{FWHM}$.

Accordingly, I consider a spectrum to be contaminated if it fulfills three sets of conditions. Condition 1:

- $5 \text{ km/s} < \text{CCF FWHM of the sky spectrum} < 15 \text{ km/s}$
- $\text{CCF contrast of the sky spectrum} > 1 \%$

These conditions ensure that the CCF fitted on the sky fiber corresponds to the sky background and not spurious noises. To demonstrate them, I plotted one of the simultaneous sky observations in Fig. 2.12; star and sky CCF (top) and a zoomed-in sky CCF (bottom). As shown in the figure, the sky CCF is consistent with noise. By fitting a Gaussian to it, the CCF contrast and the FWHM have small values, of 0.8 % and 2 km/s, respectively. Therefore, using condition 1 to distinguish the sky background from the noise is reasonable. I note that these values are adopted from the work of A. Santerne (personal communication). Condition 2:

- $|\text{RV of sky spectrum} - \text{BERV}| < 2.0 \text{ km/s}$

Because the Moon RV is close to the BERV ($\sim 1 \text{ km/s}$) in the line of sight and at the time of observation, this condition confirms that the RV on the sky CCF is related to the Moon. If the spectrum fulfills the first two conditions, the third condition is examined. Condition 3:

- $|\text{star RV} - \text{sky RV}| < 2 \times \text{star FWHM}$

This condition, which I verified above, is to ensure that the sky and star RV have an effect on each other.

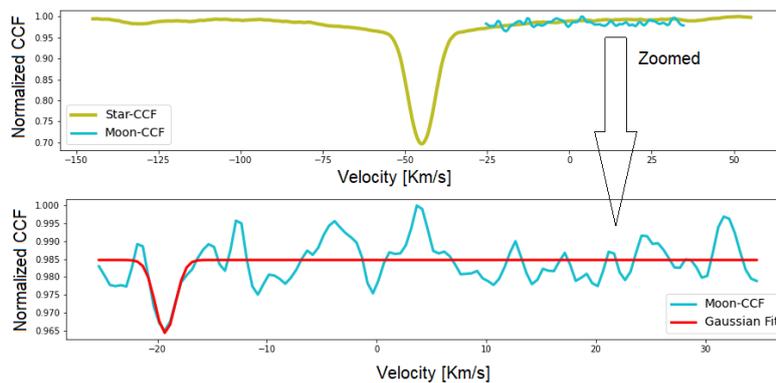


Figure 2.12 – *Top*: The star and the sky CCF of one of the SOPHIE observations, *Bottom*: zoomed on the sky CCF. The sky CCF, as shown in the figure, is consistent with noise. By fitting a Gaussian to it, the CCF contrast and FWHM have small values of 0.8 % and 2 km/s, respectively.

Table 2.2 – RVs of Kepler16AB, derived by different RV steps for the computing CCF (CCF-width). Note that all values are in km/s.

Day	RV _{fiberB}	$ RV_{fiberA} - RV_{fiberB} $	CCF-width= fwhm	CCF-width= 2*fwhm	CCF-width= 3*fwhm	CCF-width= 5*fwhm	CCF-width= 100	Moon-correction
2019-06-15	RV=6.524	~ 3*FWHM	RV=-20.474	RV=-20.474	RV=-20.476	RV=-20.476	RV=-20.477	0.026
2018-07-29	RV=0.675	~ 3*FWHM	RV=-27.152	RV=-27.152	RV=-27.154	RV=-27.151	RV=-27.151	0.013
2016-08-22	RV=-7.226	~ 1.3* FWHM	RV=-20.385	RV=-20.386	RV=-20.385	RV=-20.386	RV=-20.388	No Moon
2017-10-07	RV=-9.405	~ 1.2* FWHM	RV=-20.289	RV=-20.276	RV=-20.264	RV=-20.249	RV=-20.244	0.177
2018-10-18	RV=-9.127	~ 1.8* FWHM	RV=-25.628	RV=-25.627	RV=-25.625	RV=-25.627	RV=-25.631	0.002
2018-11-17	RV=-8.540	~ 1.7* FWHM	RV=-23.968	RV=-23.972	RV=-23.966	RV=-23.964	RV=-23.965	0.008

Results: I extracted 160 Kepler16’s data points with and without applying the optimized Moon conditions (see Fig. 2.13 top). Before applying the optimized conditions, there were 31 data points that had been corrected by the Moon correction method; after that, there were 9 data points. As shown in Fig. 2.13 (top), several data points have less dispersion after using the optimized Moon correction conditions. In fact, after removing the binary’s keplerian orbit, the RVs RMS decrease from 13.84 m/s to 13.34 m/s, resulting in a gain of 0.5 m/s. This gain is small in comparison to the mean RV error bars of this star, which is 12 m/s and dominated by photon noise for this faint star ($G_{mag}=11.7$). However, as shown in Fig. 2.13, it slightly improved the detection of planet signals at 228 d, reaching a FAP of 0.01%.

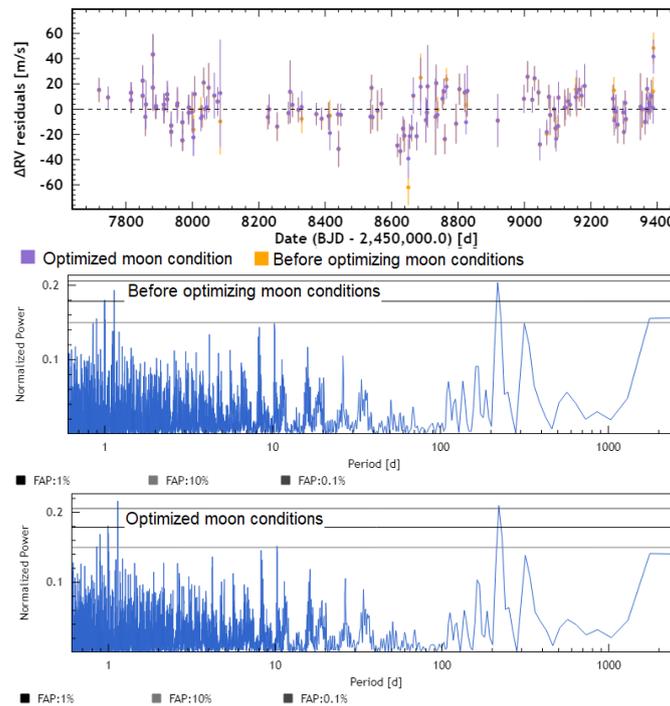


Figure 2.13 – *Top*: RVs before and after applying optimized conditions, after removing the binary’s keplerian orbit. *Middle*: RVs periodogram before applying optimized conditions. *Bottom*: RVs periodogram after applying optimized conditions. The planet signal in 228 d slightly improved, reaching a FAP of 0.01 %.

This optimized criterion was also applied to 80 stars of one of the SOPHIE programs, called BEBOP. BEBOP is a program dedicated to finding circumbinary planets (see Sect. 3.7.1, for more explanations). BEBOP observations are mostly done simultaneously with sky observation, which is why I tested this new condition on this program. As a result, through optimizing the Moon-contaminated condition, I improved the mean RV RMS of the BEBOP data by 16 cm/s. This improvement is small in comparison to the mean error bars of the stars, which is 35 m/s. However, together with the other

improvements in this chapter, it will push SOPHIE’s detection limit.

2.4.3 Optimizing conditions and results for simultaneous calibration mode

To detect low-mass planets using precise radial velocity (PRV), systemic noises should be carefully controlled to ensure that no outlier remains in the data set. Outliers can, in fact, hide the presence of a planet signal (e.g. Cloutier et al., 2019). Therefore, it is necessary to detect the Moon polluted spectra, as one of the outlier sources, and exclude them from our analysis. In the simultaneous calibration observation mode, fiber A monitors the target while fiber B measures the drift of the spectrograph. Thus, we do not have the opportunity to measure the sky’s background light and investigate whether there is a background contamination effect or not. Hence, having optimized conditions for detecting and flagging Moon-contaminated data is essential to achieving the goal of precise RV. Here, I investigate a new condition for detecting Moon-contaminated data.

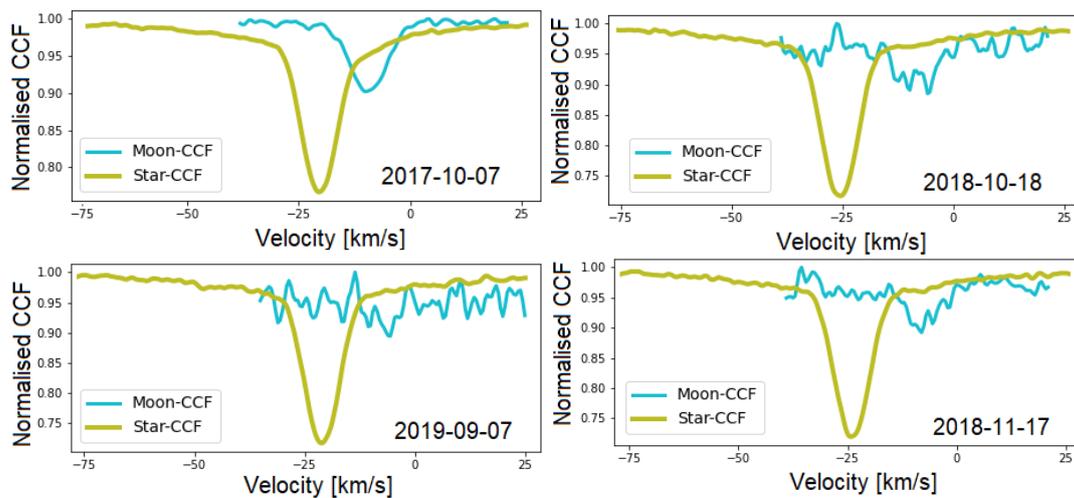


Figure 2.14 – Extracted star and Moon CCF for the different observations of Kepler16AB.

I performed a variety of tests on the Kepler16 AB target, which is monitored simultaneously with sky observation. Since the recorded spectra on fiber B were available, I detected Moon-contaminated data points using the conditions described in sect. 2.4.2. Then, I investigated in properties of these data points, including the Moon phase at the time of observations, the Moon-target separation, and the sky background level. Note that the sky level is a criterion for estimating the sky background light. It is calculated by collecting the CCD pixels of the main target that are above a certain threshold and

Table 2.3 – Moon information of Kepler16AB observations.

Day	Moon-phase	Moon-target separation (degree)	sky level [ADU* pixel ⁻¹]
Contaminated dates:			
2019-06-15	98%	79	12
2018-07-29	96%	74	15
2018-10-18	69%	73	13.9
2018-11-17	70 %	79	0.48
2017-10-07	94 %	95	27.4
Not contaminated dates:			
2019-09-07	68 %	75	6.3
2016-08-22	75%	90	4.3

* ADU stands for Analog-Digital Units. The number of photons detected in each pixel is converted into a digital number, which is the number of ADUs. This is done using an analog-to-digital converter (ADC).

treating the rest as the sky level. The sky level is calculated by the SOPHIE DRS and it is available in the header of each FITS spectrum file. The results of this test are presented in Table 2.3, and CCFs are plotted in Fig. 2.14 and Fig. 2.11. According to this test, I suggest a spectrum to be considered Moon-contaminated if it fulfills the three sets of conditions listed below. Condition 1:

- At the time of observation, the Moon must be visible above the horizon from the observatory.

If Condition 1 is fulfilled, Condition 2 should be checked. Condition 2:

- Either the Moon phase is more than 68% in the sky at the time of observation, and the sky level is higher than the average of all target observations (e.g., in the case of Kepler16, the average sky level is 10);
- or the Moon-target separation is less than 30 degrees. I note that this value has already been confirmed, and it is being considered in the observation strategies (SOPHIE communication).

In the above conditions, I use the *ephem* python program to calculate the Moon phase and position at the time of observation. If the first two conditions are fulfilled, the third condition is then checked. Condition 3:

- $|RV_{target} - BERV| < 2 * FWHM$

This condition is adopted from the simultaneous sky observation condition, presented

Table 2.4 – Summary of applying Moon conditions for simultaneous calibration mode.

Target	NO. data	NO. contaminated	NO. correctly detected	NO. not-correctly detected
Star 1	11	2	2	2
Star 2	36	5	3	1
Star 3	12	6	3	–

in Sect 2.4.2. As I mentioned before, it is based on the fact that at the time of observation, the Moon RV and BERV are less than ~ 1 km/s apart in the direction of the target. Therefore, it is reasonable to consider BERV as the RV for the Moon. Using these conditions, the detected Moon-contaminated spectra in simultaneous calibration and sky observation mode are the same, except on the date 2018-11-17. While this data point is detected as Moon contaminated by simultaneous sky observation conditions and using the recorded sky background, the contamination could not be resolved by the simultaneous calibration conditions. The reason for this is that, for unknown reasons, the sky level on this date has the lowest value among the 122 data points of Kepler16AB.

In Table. 2.4, I show the results of applying this condition to 3 stars from the BEBOP program, which were observed using simultaneous sky observations. For each star, the total number of Moon-contaminated data using the recorded sky spectrum is shown in the third column. First, I assumed there was no recorded sky spectrum for these stars and used the simultaneous calibration conditions to find Moon-contaminated spectra. The detected contaminated spectra were then validated using information from the recorded spectrum on fiber B (fourth and fifth columns). As a result, the detected contaminated dates generally have a good agreement with reality when using the registered sky observation (compare third and the fourth column), but sometimes with a few data points difference (extra or less). For example, in the case of star1, the condition detected all contaminated data. However, it incorrectly flagged two extra data points. Therefore, conservatively, I suggest removing the detected contaminated dates, if they are above 5σ from the mean RVs.

2.4.4 Conclusion

The RVs can suffer from systematic errors of up to several m/s caused by moonlight, in particular, for the fainter star. Hence, Moon contamination is a serious problem when aiming for very precise velocities and subsequently finding low-mass planets. To correct this, for simultaneous sky observation mode, a condition to determine the contaminated spectrum prior to correction is necessary. This is due to the fact that when moonlight is negligible, subtracting the sky and star spectrum (Moon correction)

would result in additional noise. Therefore, I provide optimized conditions for finding Moon-polluted spectra. For simultaneous calibration observation mode where no sky spectrum is available, I also propose a condition for detecting possible contaminated spectra and removing them from the analysis, if they are above 5 sigmas.

Despite the fact that these conditions and subsequent data correction result in improved data (see Sect. 2.4.2), the best way to reduce the contamination is to incorporate it into the observation strategy. According to SOPHIE's current observation strategy, the observer should not observe a star if the Moon and the star are closer than 30 deg, and $|RV_{Moon} - BERV| < 15 \text{ km/s}$. The conditions developed in this thesis can also be applied to the observation strategy. The combination of using these developed conditions to account for the Moon at the time of observation, as well as to detect polluted spectra for correction, will have a significant impact on RVs.

2.5 Correction of background contamination of the calibration lamp

The combined analysis of RVs and activity indicators lets us identify the origin of a signal. Therefore, accurate activity indicators are essential for data interpretation. As shown in Fig. 2.4, in a CCD image, the recorded light from fibers A and B in a spectral order are close (~ 17 pixels) to each other. It leads to a small but non-negligible amount of light diffusion due to the calibration lamp from fiber B to fiber A (Lovis et al., 2011). Therefore, the key step before deriving the $\log(R'_{HK})$ and $H\alpha$ activity indexes is to subtract the background light from the star spectrum caused by the diffuse light from the Th-Ar or FP calibration lamp.

To correct this contamination, the SOPHIE routine employs different methods, depending on the calibration lamps. For the Th-Ar lamp, a background is estimated from the flux of fiber B in the same spectral order by fitting a polynomial function on local minima. On the other hand, for the FP lamp which is installed since semester 2017B, this background is directly measured through a Dark-FP frame, i.e. no light on fiber A, and FP calibration lamp illumination on fiber B (Hobson, 2019). However, after more years of observations, I clearly saw a jump between the data taken by the two calibration lamps (see Fig. 2.16).

To correct this jump, I directly measured this contamination by observing a Dark-Th-Ar frame, i.e. no light illumination on fiber A; Th-Ar illuminating on fiber B. I repeated this observation with the different exposure times (300 s, 600 s, 900 s, 1800 s). They showed that the flux of calibration lamps is slightly increasing depending on the exposure time (Fig. 2.15 left, dashed line). Consequently, the measured back-

grounds are also changing with exposure time, but in a non-linear way depending on the spectral order (Fig. 2.15 left, star-line). I used a Dark-Th-Ar frame with an exposure time of the 1800s as a reference because it has the highest flux. Then, I constructed a weighted 'master background' by multiplying it by: 1. the Th-Ar flux of the simultaneous calibration lamp for each observation, and 2. a scaling factor per order based on the difference in behavior between the measured backgrounds and the flux of the calibration lamp with exposure time (Fig. 2.15 left). The calibration lamp contamination is then corrected by subtracting the weighted master background from the target spectrum.

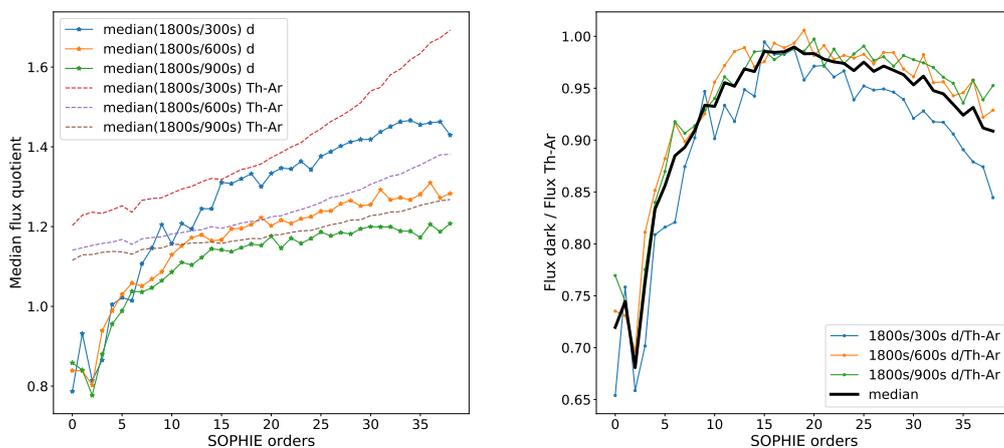


Figure 2.15 – *Left*: Median flux quotient per order for backgrounds (or darks, star-line) and Th-Ars (dashed line). Each exposure time is presented with a different color. The exposure of the 1800s is the reference point. *Right*: Quotient of Dark and Th-Ar calibration per order for each exposure time (dotted lines) and their median (solid black line). The median line is used to scale the master background.

I tested this new background correction on one of the SOPHIE constant stars, HD89269A (see Fig. 2.16). This star was chosen because its FWHM and bisector indicate that there is no evidence of long-term activity. The test results are summarized in Table 2.5. The median of $H\alpha$ index after direct background measurement for simultaneous Th-Ar exposures, shows much better consistency with simultaneous FP observations. Fig. 2.16 clearly shows that the jump between the two data sets is successfully removed. Accordingly, because direct measurement of this contamination for both lamps results in much better data consistency, it is likely that the DRS's routine for simultaneous Th-Ar background correction was overestimated. Therefore, I suggest using direct measurement to correct the background light of the Th-Ar calibration lamp.

Table 2.5 – Median of $H\alpha$ index for different background correction methods, for the constant star HD89269A. Th-Ar and FP simult indicate observations with simultaneous Th-Ar and FP calibration lamps, respectively.

Correction	Median $H\alpha$
Th-Ar simult- old method	0.1073
Th-Ar simult- direct measurement	0.1124
FP simult- direct measurement	0.1136

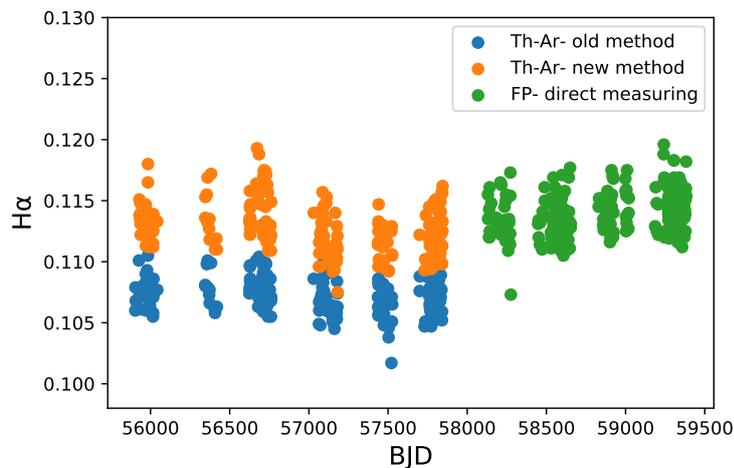


Figure 2.16 – $H\alpha$ activity indicator of SOPHIE constant star HD89269A, for different background correction methods. The direct measurement of background for simultaneous Th-Ar exposures (orange points) shows much better consistency with simultaneous FP observations (green points).

2.6 Optimizing long-term variation from the zero-point

2.6.1 High resolution mode

Courcol et al., 2015 showed that the SOPHIE spectrograph suffers from a long-term variation from the zero-point due to instrumental offsets. This long-term variation is monitored by systematically observing a set of so-called "constant stars" each night. The constant stars are chosen to be 4 super-constant stars: HD185144, HD9407, HD221354, and HD89269A. These stars were initially monitored for several years by the HIRES spectrograph, and showed an RV dispersion of 2.0, 1.7, 1.9, and 2.0 m/s

respectively for HD185144, HD9407, HD221354, and HD89269A. Because of their low RV dispersion, they have been used as constant stars for SOPHIE since 2011.

To reach the precision over a long time span, this long-term variation must be taken into account. To correct this, I adopted a code developed by B. Courcol-Rozès (Courcol et al., 2015) and later improved by J. Rodrigues and N. Hara (personal communication). The corrections have the following steps:

- Building a master constant time series from the 4 super constant stars: First, the median RVs (V_0) of each star is subtracted from the RVs (V), i.e. $\Delta V = V - V_0$. Then, the RV variations bigger than 15 m/s are flagged as outliers and removed. This is due to the fact that the constant stars did not show big RV variations. Finally, all RVs (ΔV) of constant star observations are combined, and a sliding median of the data is computed by grouping the data in each jump. To identify data jumps, close data and their error bars are compared, and a threshold as a significance level is defined.
- Residuals optimization: Zero-point drift is a long-term variation. Therefore, the mean velocity of stars can be affected by the seasons in which they are observed. This means that the mean of a star's RV residuals after master constant correction $offset_0$ is non-zero. To account for this instrumental effect, it is necessary to repeat the first step in the next iteration by adding an offset to the data, i.e. $\Delta V = V - V_0 + offset_0$. If, after this new correction and the next algorithm iteration, the mean of RV residuals after master constant correction $offset_1$ is again non-zero, the correction is still not satisfactory. It is necessary to start over by $offset_0 + offset_1$. This process is usually repeated two or three times until $offset_n = 0$.
- New constant star selections: Once the master constant time series is created, it is subtracted from the RVs of all SP1 targets. If the corrected star RVs show a low RV dispersion of < 3.5 m/s and no periodic signal in the periodogram with a FAP of below 10%, the star is chosen as a new constant star. This star will be used in the next iteration of the master constant's construction.
- Loop back to the first step: Steps 1–3 are repeated until no new constant stars are added on step 3.

The master constant correction is a powerful method for removing instrumental offsets and more accurately analyzing the data. However, there is still the possibility of leaving out some instrumental variations. As an example, in most of the SOPHIE targets, I found a signal at 365 d that could be attributed to instrumental factors (e.g. Dumusque et al., 2015a). Therefore, I planned for improvements in master constant corrections, as one of the most effective steps in RV precision.

To do so, I examined every step of the construction of master constant correction, including re-checking the stability of super-constant stars by searching for any correlation between RVs and their activity indicators. Furthermore, because I improved the data as described above, it was worthwhile to re-calculate the master constant time series.

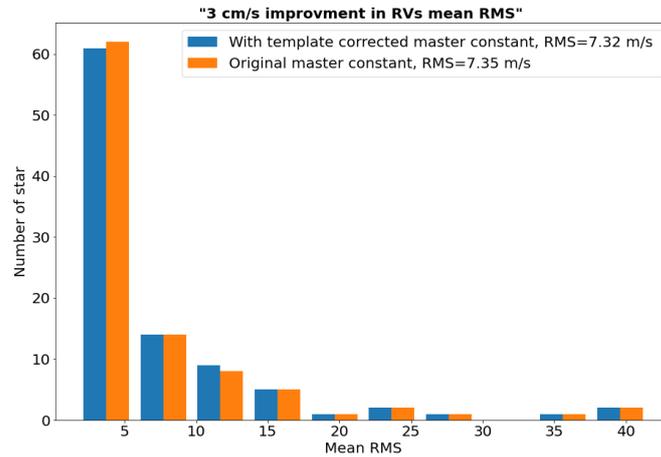


Figure 2.17 – Histogram of mean RVs RMS of 96 SP1 stars for two different data sets: 1) the data and applied master constant correction are flux corrected, 2) the data and applied master constant correction are not flux corrected. The figure shows the mean RV RMS of stars is improved by 3 cm/s.

2.6.1.1 Building a new flux corrected master constant time series

I extracted all RV data of SOPHIE SP1 stars by applying the flux correction method (see Sect. 2.3). Then, using this new data set, I created a new master constant time series and subtracted it from the 96 SP1 stars (see Fig. 2.17). This attempt led to a 3 cm/s improvement in the mean RVs RMS of stars, which is 3 % of their mean error bars of 1.2 m/s. This improvement is not significant, however, every small improvement towards more precise radial velocity is important when searching for low-mass planets.

2.6.1.2 HD185144 activity correction

By using the new background correction of the calibration lamp on activity indicators (see Sect. 2.5), I found a signature of long-term activity in the $H\alpha$ and $\log(R'_{HK})$ of one of the super-constant stars 'HD185144' (see Fig. 2.18). This star is one of the most observed stars by SOPHIE, with more than 1000 data points over the course of 10 years. This star's activity cycle was already known thanks to the HIRES data (Isaacson et al.,

2010). Later, it was confirmed by SOPHIE when the first master constant correction was built (Courcol-Rozès, 2016). At that moment, in 2016, the amplitude of the signal was estimated to be less than 1.5 m/s, which was negligible given the spectrograph's precision. Thanks to the SOPHIE improvements, the RV precision has improved, and the impact of the stellar activity on HD185144 RVs needs to be corrected. Therefore, I searched for HD185144's activity effect on the RVs and its corresponding effect on the master constant time series.

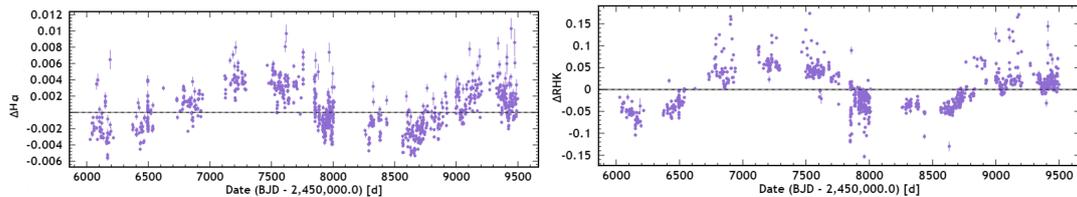


Figure 2.18 – $H\alpha$ and $\log(R'_{HK})$ of one of the super-constant stars 'HD185144'.

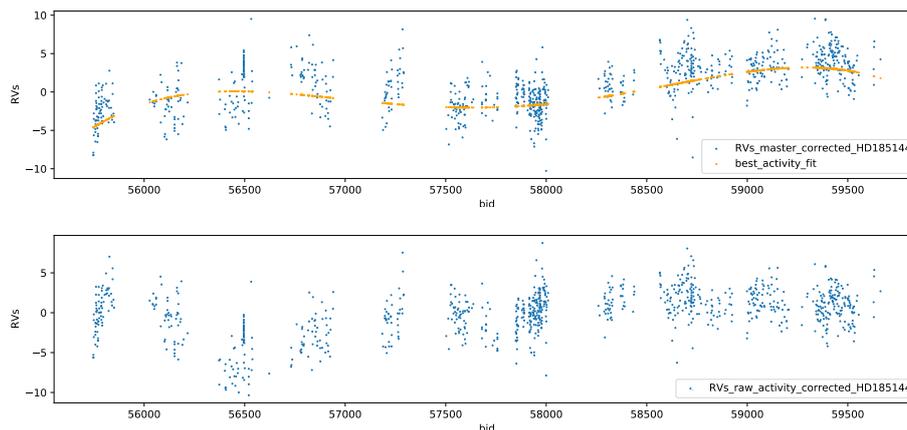


Figure 2.19 – *Top*: Blue points represent RVs of HD185144, master constant corrected by 24 other stars. The overplotted curve is the best-fitted polynomial curve to remove the star activity. *Bottom*: raw RVs of HD185144 corrected by its activity.

To do so, by excluding the HD185144 star, I built a master constant time series. Then, using this time series, I corrected the SOPHIE long-term variations of HD185144 RVs (Fig. 2.19 top) and searched for the activity signature. As shown in Fig. 2.19, the RVs of HD185144 are affected by a long-term activity period with an amplitude of ~ 2.5 m/s. To correct this activity, I used the best fit of a simple polynomial curve on the RVs (Fig. 2.19 top orange line). The same phase was then kept and subtracted from the raw data before the master constant correction (see Fig. 2.19 bottom). In Fig. 2.20 and Fig. 2.21, I followed the same procedure to correct the activity of the bisector and FWHM time series of HD185144.

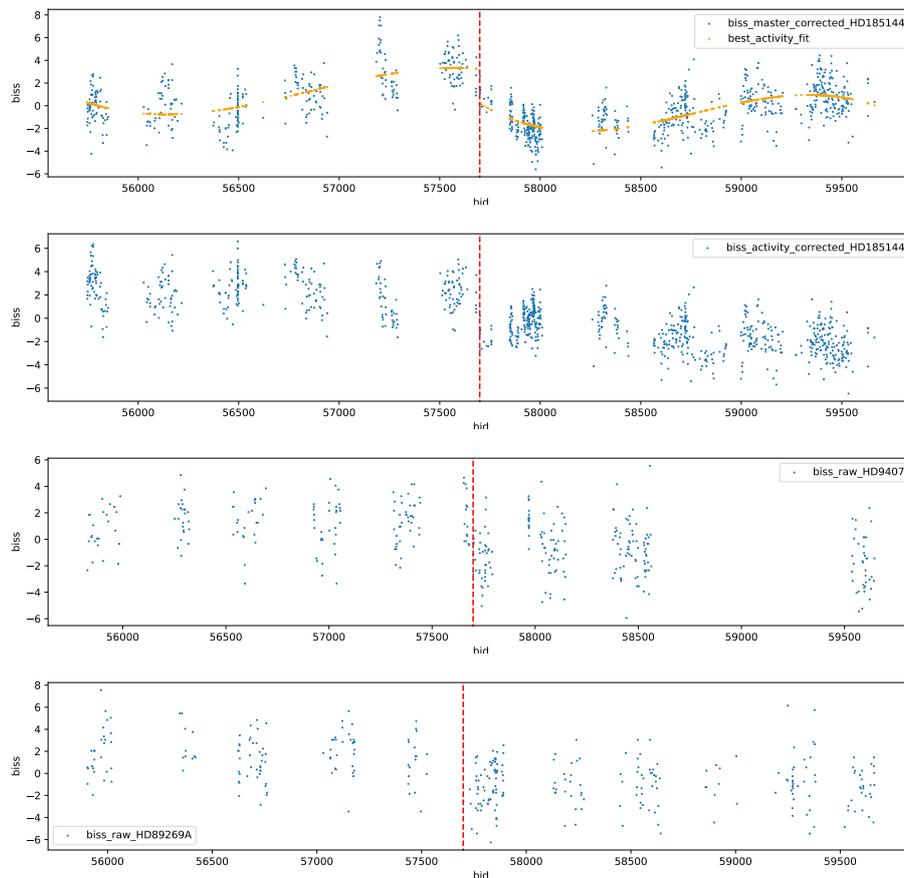


Figure 2.20 – *First*: HD185144 bisectors, master constant corrected by 24 other stars, overplotted by the best-fitted polynomial curve to remove the activity. *Second*: raw bisectors of HD185144 corrected by its activity. *Third and fourth*: the bisector time series of two other super-constant stars demonstrates the good activity correction of HD185144. The bisector time series of HD185144 after activity correction are compatible with the bisector of other super-constant stars.

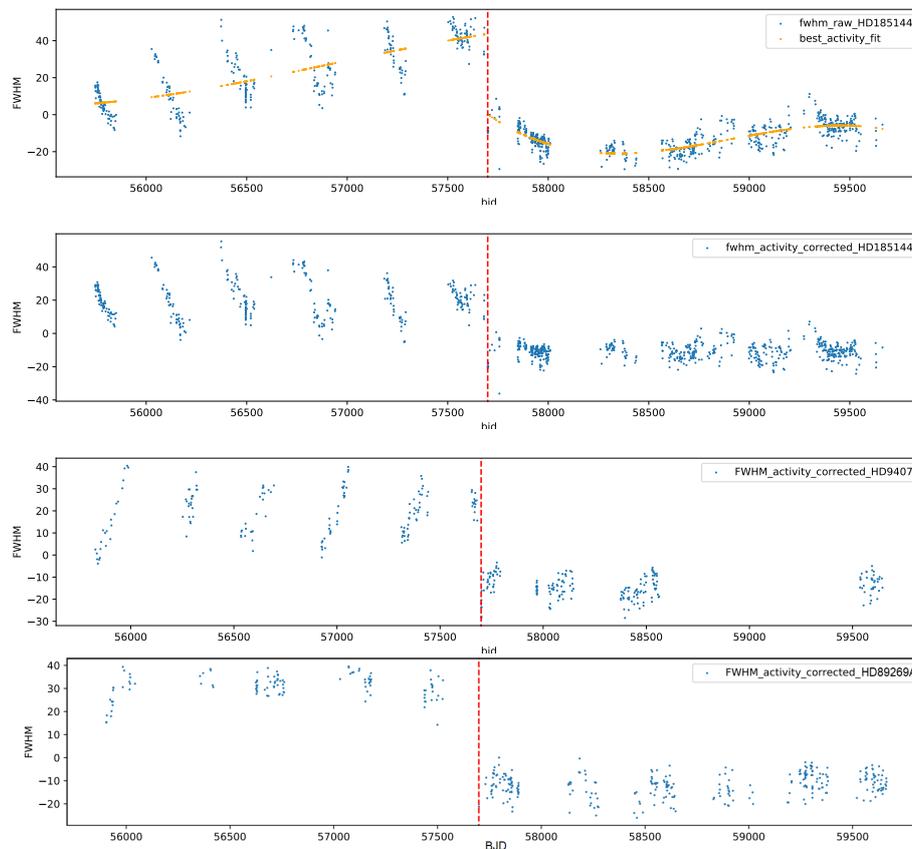


Figure 2.21 – *First*: FWHM of HD185144 star overplotted by the best-fitted polynomial curve. *Second*: Raw FWHM of HD185144 corrected by its activity. *Third* and *fourth*: the FWHM time series of two additional super-constant stars to verify good activity correction of HD185144. Indeed, after activity correction, the FWHM time series of HD185144 is compatible with the FWHM time series of other super-constant stars.

Then, I created a new bisector and RVs master correction by adding the activity-corrected HD185144 data into the process for building the master constant correction. Figures 2.23 and 2.22 show these new optimized master constant time series, which exhibit a long-term variation with several sharp jumps. Some of the jumps are attributed to the instrumental offsets, which I marked with vertical lines in RVs and listed in Table 2.6.

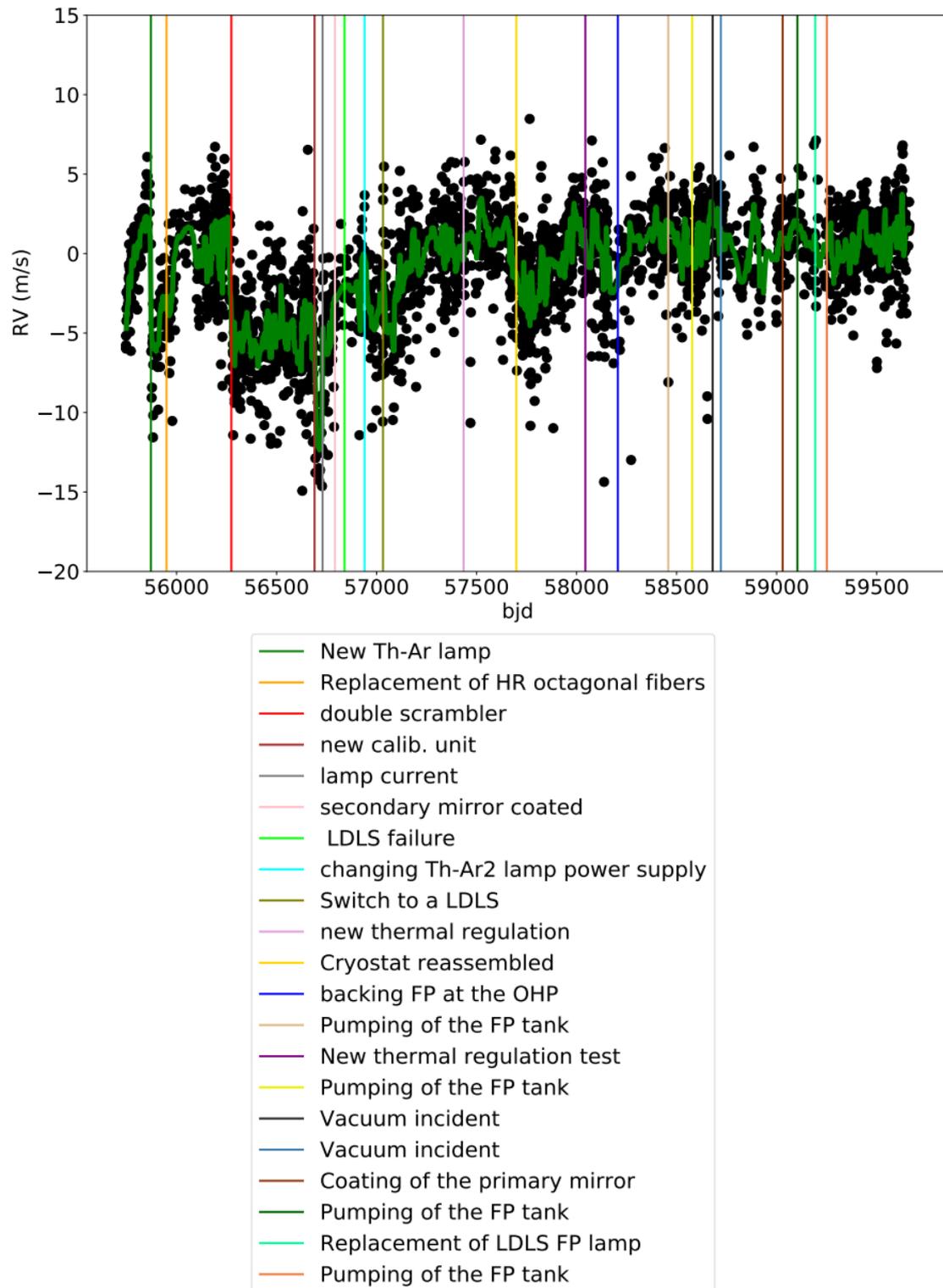


Figure 2.22 – Optimized master constant time series on RVs. The instrumental offsets are overplotted by vertical lines.

Table 2.6 – Summary of known date and cause of the SOPHIE RV long-term drift.

BJD	Cause
55872	New Th-Ar lamp
55950	Replacement of HR octagonal fibers
56274	Integration of new HR octagonal fibers after the data scrambler
56690	A Th-Ar lamp was changed
56730	A new calibration unit was installed
56792	The secondary mirror of the telescope was coated
56840	LDLS failure, switch to a tungsten lamp
56940	Modification of the Th-Ar2 lamp power supply; Switch to a LDLS
57032	Switch to an LDLS
57700	disassembled/ reassembled the Cryostat; detector was realigned.
58044-46	New thermal regulation test.
58206	FP is back at the OHP
58458	Pumping of the FP tank
58578	Pumping of the FP tank
58680	Incident in vacuum; pumping of the SOPHIE Cryostat
58711-21	Incident in vacuum; increasing temperature of the SOPHIE camera
59030	Coating of the primary mirror
59104	Pumping of the FP tank
59193	Replacement of LDLS FP lamp
59251	Pumping of the FP tank

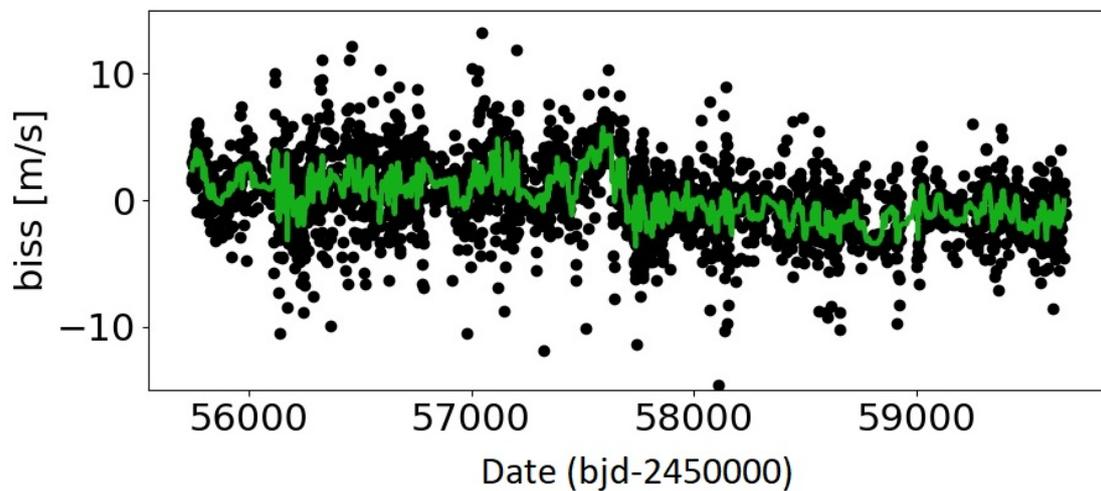


Figure 2.23 – Optimized master constant time series on bisector.

Results: I applied this new master constant correction to 96 SP1 targets. Fig. 2.24, shows the distribution of the mean RV RMS of these targets corrected by the master constant time series, both with and without the activity correction of HD185144. By removing the activity of HD185144, I reduced the mean RVs RMS by 1 cm/s. As expected, this gain is insignificant when compared to the mean error bars of the SP1 sample, which is 1.2 m/s.

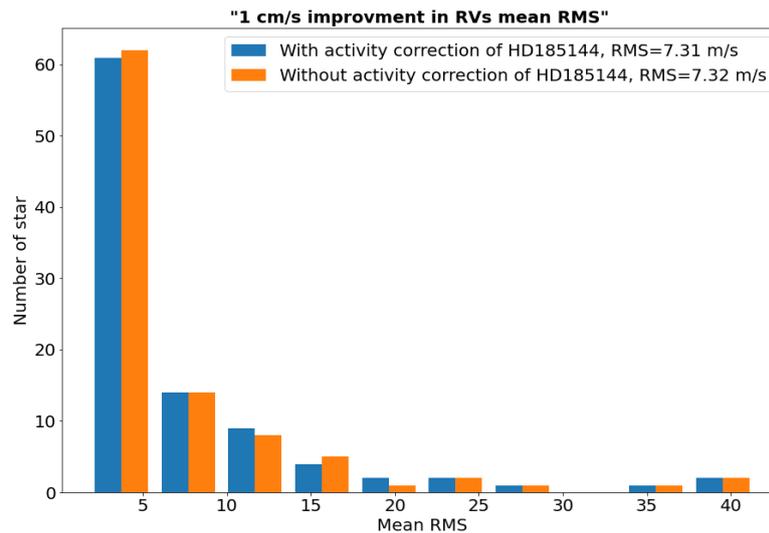


Figure 2.24 – Distributing of the mean RVs RMS of SP1 targets, corrected by the master constant time series, both with and without the activity correction of HD185144.

A better indicator of the gain, is the disappearance of the yearly peak at 360 d for most of the targets. This peak is likely the result of instrumental factors. Furthermore, the signal of candidates becomes stronger. To demonstrate this, I plotted the periodograms of one of the SP1 stars with two planetary candidates at 39.8 d and 14 d in Fig. 2.25, by removing subsequently the highest peak in each periodogram. The *left* panels are periodograms of the data after applying the master constant correction, which is corrected from the HD185144 activity. The *right* panels show the same periodogram without correcting the activity of HD185144. As this plot illustrates, after correcting the activity of HD185144, the planet signals at 39.8 d and 14 d appear stronger (first and second rows), and the peak at 360 d (Fig. 2.25 third row) is removed.

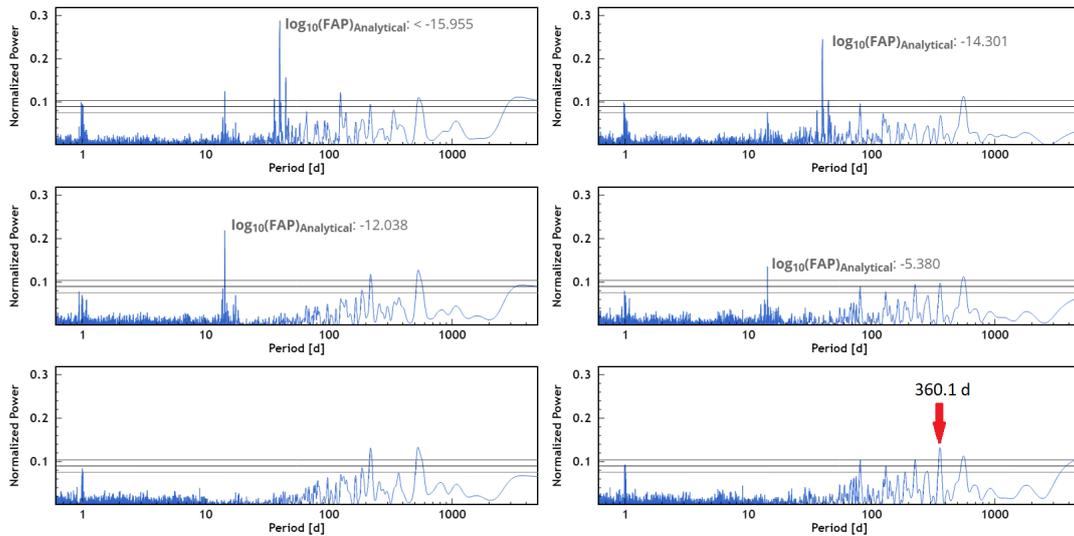


Figure 2.25 – Periodograms of one of the SP1 stars, by removing subsequently the highest peak in each periodogram. The *left* panels show periodograms of the data after applying the master constant, which is corrected by the activity of HD185144. The *right* panels show the same periodogram without removing the activity of HD185144. As this plot illustrates, after correcting the activity of HD185144, planet signals are stronger (first and second rows), and the peak at 360 d (third row) is removed.

2.6.1.3 Building a FWHM master constant correction

In non-well-stabilized spectrographs, the FWHM time series is affected by the instrumental profile. Hence, it is very important to track how the FWHM evolves depending on the instrument. As I previously demonstrated, the SOPHIE FWHM time series after the atmospheric dispersion correction is now sufficiently stable (see sect. 2.3.2.2). This enabled me to track instrumental instabilities in the FWHM data.

To do this, I first applied the template correction method to all SOPHIE SP1 targets, which led to a great improvement in the stability of the FWHM time series. Then, I built a master constant time series with 10 constant stars and in total ~ 2500 data points.

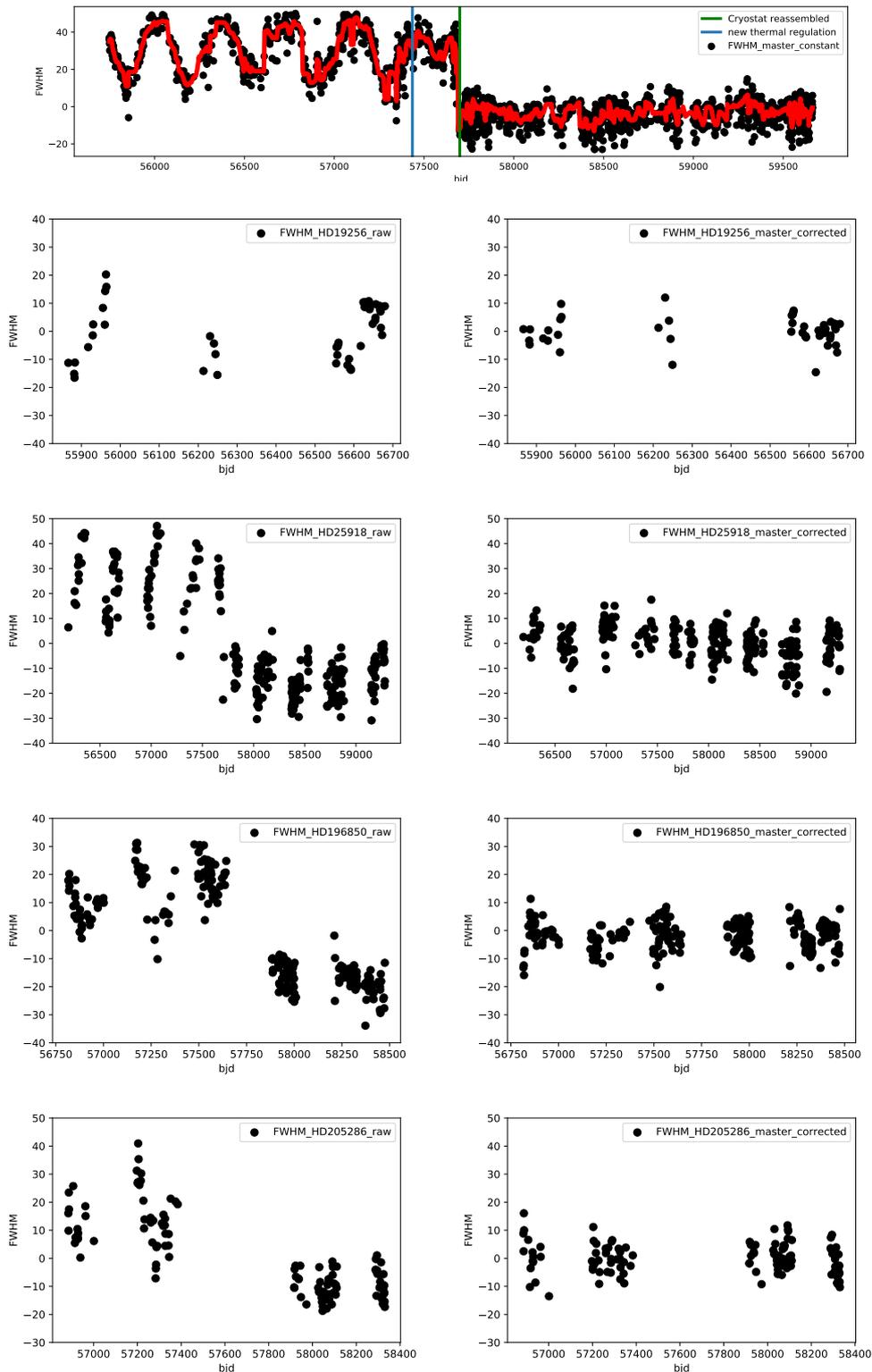


Figure 2.26 – *Top*: FWHM master constant time series. (*Rest of the panels*): the FWHM time series of stars before (*left*) and after (*right*) applying the FWHM master correction. The figure clearly shows the jump, and yearly variations are removed.

To create this time series, I removed the data with $S/N < 100$ to choose the best-quality data points. Then, I followed the same procedure of constructing the RV master constant (Sect. 2.6), except for the star selection. Indeed, due to the large data jump as well as the high variability of the FWHM time series, I used constant stars that have a longer time span. The FWHM master constant time series is shown in Fig. 2.26. The sharp jump at 57699 is due to the time that the cryostat was disassembled and reassembled and the detector was realigned. In the first part of the data sets, before $\text{bjd}=57435$, the data sets are dominated by an annual cycle, with an amplitude of ~ 20 m/s. After this date, due to the new thermal regulation of SOPHIE, the yearly variations are greatly reduced.

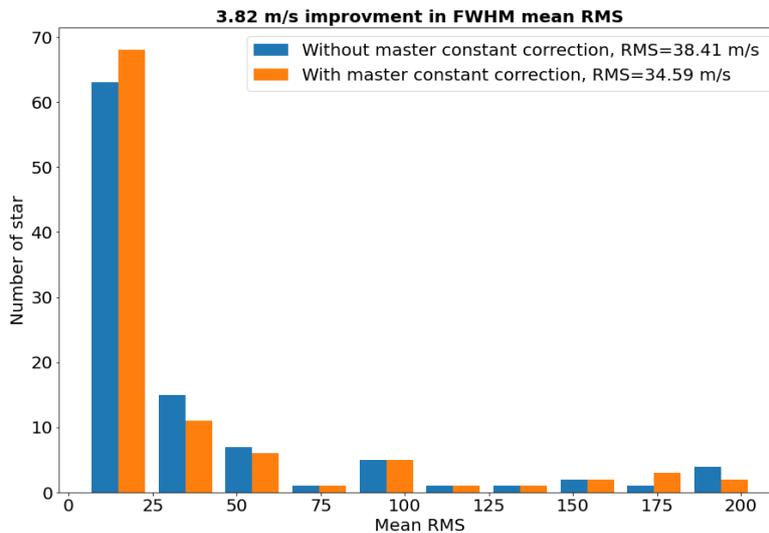


Figure 2.27 – Distributing of mean FWHM RMS of SP1 targets with/without applying the FWHM master constant time series.

Results: In Fig. 2.26, I show the FWHM time series of a few stars before and after applying the FWHM master correction. The figure clearly illustrates that after applying FWHM master correction, the jump and yearly effects on the data disappear, and the FWHM time series becomes significantly more stable. I also applied the FWHM master constant correction on all SP1 targets. As a result, I successfully improved the mean FWHM RMS by 3.82 m/s (see Fig. 2.27). This value is 1.4 times greater than the mean error bars of 96 SP1 stars 2.8 m/s. By combining this work with the atmospheric dispersion correction, the data are significantly improved, allowing it to be used as an important probe in the search for stellar variability.

2.6.2 High Efficiency mode

I also built a master constant correction on the data taken by High Efficiency (HE) mode. To do this, I used the RVs of 3 super-constant stars (HD185144, HD89269A, HD9407). I removed the low S/N data (below 70-85 depending on the star), and the nightly drift of the spectrograph as measured by the calibration lamp. The CTI corrections are also applied following Santerne et al., 2012.

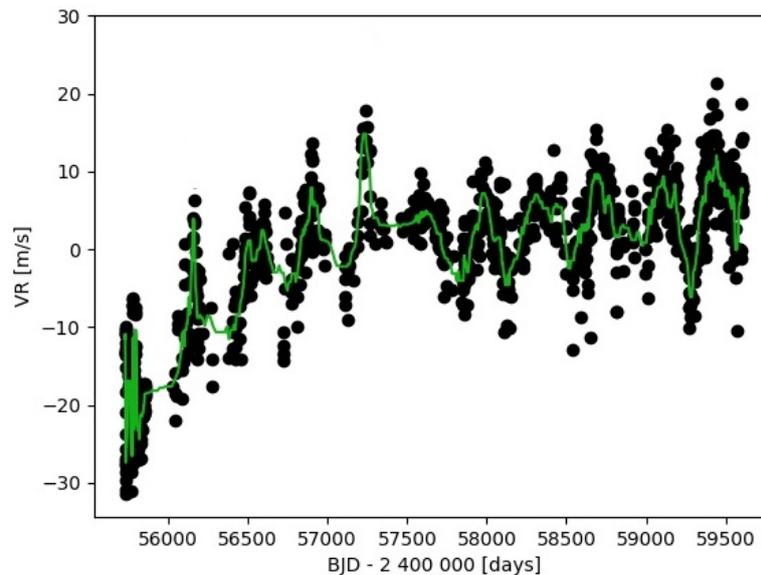


Figure 2.28 – Master constant time series for HE mode.

To correct the activity cycle of HD185144, I used the phase of activity from HR mode and subtracted it from the RVs of the star in HE mode. The reason for this is that the data in HR mode are more precise, and thus the phase of activity in this mode is more accurate. Then, I followed the same procedure of master constant correction on HR mode (see Sect. 2.6.1.3). Fig. 2.28 shows the master constant time series in HE mode. This time series shows a long-term drift with the yearly cycles. The cause of the observed drift and the yearly cycles, which can be seen in each of the three constant stars individually, is unknown.

2.7 Fixing mask lines selection for spectra at different epochs

As I explained in detail in the Sect. 1, the CCF method is a standard approach for extracting RVs. In this method, a spectrum is cross-correlated by a binary mask which is consistent with the box-shaped theoretical positions and widths of the absorption lines (Pepe et al., 2002).

In addition to the Doppler motion induced by the planet and stellar activity, the spectra are shifted by barycentric Earth radial velocity (BERV). This shift depends on the position of the star in the sky and varies by ± 30 km/s throughout the year. Accordingly, the lines on the CCD's edge are not always visible. This effect can cause a yearly signal in the data. Despite the fact that this optimized master constant correction successfully removed the yearly signal from the most of the SOPHIE targets, the signal still remained in some of the stars. Here, I attempted to fix the number of mask lines at different epochs with the motivation of removing the yearly signal and also having a stable RV computation. Indeed, by fixing mask line selection, we can ensure that the number of lines used to calculate RVs remains constant.

To fix the number of mask lines in the different epochs, I applied a cut-off on the mask lines for each spectral order. A similar cut-off has been used in the HARPS, CARMENES (Lafarga et al., 2020), and one version of SOPHIE data reduction that is no longer in use. This cut-off is introduced as follows in each spectral order:

$$\lambda_{min} = \lambda_1(1 + BERV_{max}/c)/(1 + v_{min}/c) \quad (2.1)$$

$$\lambda_{max} = \lambda_n(1 - BERV_{max}/c)/(1 + v_{max}/c) \quad (2.2)$$

Where λ_1 and λ_n represent the maximum and minimum wavelengths for each spectral order, $BERV_{max}$ is the maximum of barycentric shift, v_{min} and v_{max} are the minimum and maximum RVs of a target, respectively, and c is the speed of light.

Results: I implemented this cut-off to the SOPHIE data reduction system and tested it on HD88986 star (see sect. 3.6.2 for more information about the star). Before applying the method, the number of lines used to calculate RV CCF was 6862 at a minimum BERV of -27 km/s and 6997 at a maximum BERV of 30 km/s. After using this method, the number of lines for each observation epoch is set to 6927. In addition, the mean RV RMS of the data was reduced from 7.96 m/s to 7.89 m/s, resulting in a 7 cm/s gain. This improvement is small when compared to the mean RV error bars of 1.6 m/s. However, it might help in planet detection. In Fig. 2.29, I plotted the RV periodograms

for this star before and after implementing the method. The figure shows that the signal at 29 d is slightly stronger, reaching a FAP level below 1 %. The signal at 29 d is the estimated stellar rotational period.

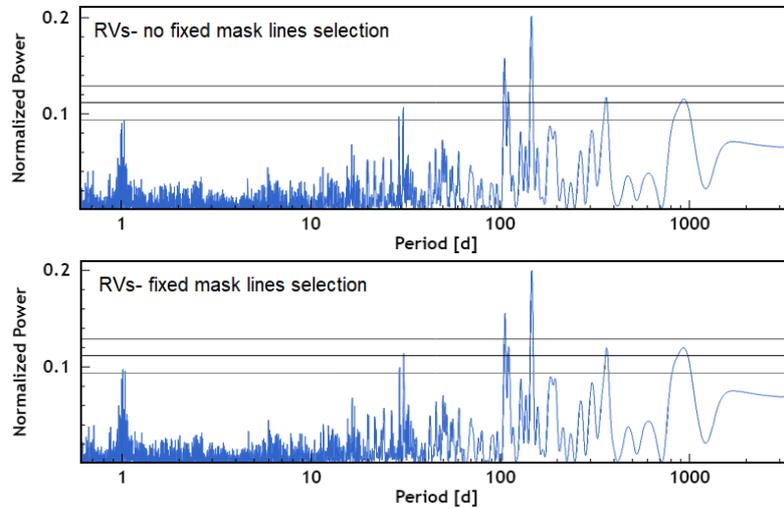


Figure 2.29 – RVs periodogram of HD88986, before and after fixing the mask lines selection.

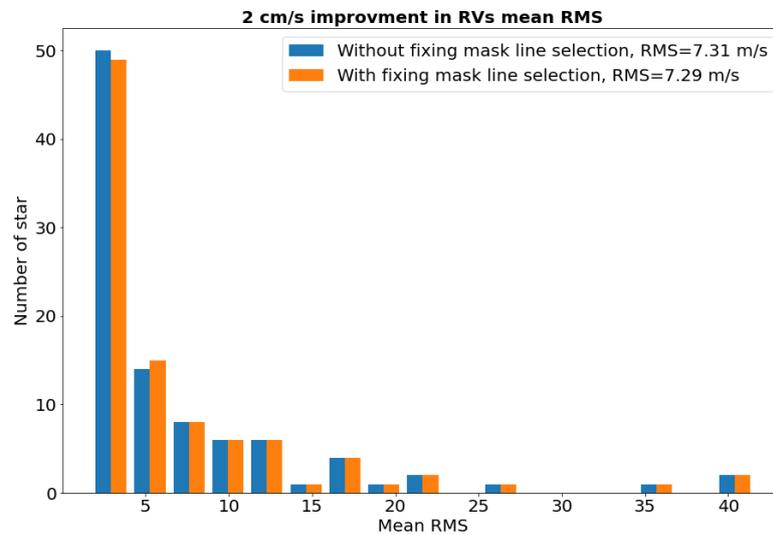


Figure 2.30 – The distribution of mean RVs RMS of SP1 targets derived with and without fixing the line masks at different epochs.

I also extracted the RVs of 96 SOPHIE SP1 targets using this method. Fig. 2.30, presents the distribution of mean RVs RMS of these stars derived with and without fixing the mask lines at different epochs. As shown in the figure, by fixing the number of mask lines, I improved the mean RV RMS of 96 SOPHIE SP1 targets by 2 cm/s. This

value is insignificant when compared to the mean error bars of stars 1.2 m/s. However, as I demonstrated in the case of HD88986 for the peak of 29 d, it strengthens the signals.

I note that even after fixing the mask lines, a few targets still suffer from the yearly signals. The other possible reasons should be investigated in the future (e.g. testing the stitching effect of CCD (Dumusque et al., 2015a)).

2.8 Conclusion

In this chapter, I presented my work on facing the challenge of instrumental limitations, which is one of the biggest obstacles in low-mass planets detection. This attempt led to removing several instrumental effects on RVs, FWHM and activity indicators. Particularly, applying the flux correction method as well as constructing a master constant correction on FWHM, resulted in a significant improvement of 18.57 m/s in the mean RMS of this activity indicator (see Fig. 2.31 left). In fact, thanks to these efforts, SOPHIE's FWHM can now be used as a reliable tracer of stellar activity. In RVs, by applying a flux correction, fixing the number of mask lines at different epochs, and removing the effect of HD185144's activity cycle from the master constant time series, I improved the mean RVs RMS by 6 cm/s (see Fig. 2.31 right). These successful attempts will have a notable impact on future planet detections by SOPHIE in terms of the desired precision on planet parameters, and distinguishing between planet and activity signals.

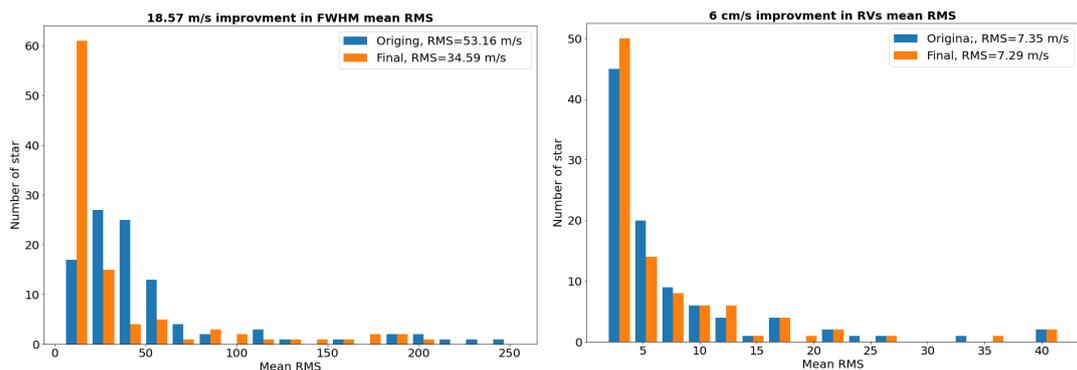


Figure 2.31 – Distribution of RVs and FWHM mean RMS for 96 SP1 targets with/without applied improvements during my Ph.D.

Additionally, I investigated the effect of moonlight pollution as a source of noise in RV measurements. For simultaneous calibration mode, I presented a set of conditions for finding and flagging the possible Moon-contaminated spectrum without access

to the recorded sky observation. For simultaneous sky observation mode, I also presented an optimized condition for detecting Moon-contaminated spectra that will later be corrected by the Moon correction method.

To further improve the accuracy of the SOPHIE spectrograph, there are some steps that should be tested or implemented in the SOPHIE DRS in the future. For example: applying the blaze correction to the spectrum before calculating the CCF for extracting the RVs. This will help to remove any continuum slope around the spectral lines (Modigliani et al., 2019), improving the Moon correction method. For instance, instead of subtracting the CCF of the two fibers (star and sky background) from one another, the correction can be done at the level of the spectrum and through subtracting the spectra of the two fibers (Modigliani et al., 2019); testing and implementing the GP method for constructing the master constant correction to provide errors on it; searching for the cause of yearly signal in the data, e.g. testing stitching effect of CCD (Dumusque et al., 2015a); building a reference template to correct atmospheric dispersion effects for HE modes.

Chapter 3

Searching for low-mass planets: a combination of radial velocity and photometric data

Contents

3.1	Searching for low-mass planets with the SOPHIE spectrograph (SP1)	88
3.1.1	Scientific objective	88
3.1.2	SP1 sample	88
3.1.3	Observation strategy	89
3.1.4	SP1 status in 2022	90
3.2	Objective of my work on SP1	90
3.3	Data processing and selections	91
3.4	Analysis of only RVs and activity indices	92
3.5	Stellar rotation period investigation	100

3.6	Discovery and characterization of new transiting planets	102
3.6.1	HD207897 b: A dense sub-Neptune transiting a nearby and bright K-type star	102
3.6.2	Two planets around the HD88986	120
3.6.3	A transiting planet candidate	132
3.7	Other contributions to SOPHIE RV programs	134
3.7.1	RV search around binary stars	135
3.7.2	High precision RV search around M dwarfs	137
3.7.3	High precision RV search for super-Earth in systems harboring giant planets	141

3.1 Searching for low-mass planets with the SOPHIE spectrograph (SP1)

3.1.1 Scientific objective

The SOPHIE exoplanet consortium has been leading several long-term planet observations (Bouchy et al., 2009a). Among them, the SOPHIE sub-program 1, known as SP1, is dedicated to the search for low-mass planets, between Super-Earths (2-10 M_{earth}) and Neptunes (10-30 M_{earth}), around G and K spectral type stars. The program's objectives are:

- To study the statistical properties of planets, measure the occurrence rate of low-mass planets, and constrain the models of formation and evolution of low-mass planets by expanding the sample;
- To detect habitable super-Earths and Neptunes;
- To study the relationship between exoplanets and the chemical composition of their host stars.

3.1.2 SP1 sample

The initial target list of SP1 was selected in July 2012, by taking into account several criteria:

- Visible from the northern hemisphere;
- Spectral types G and K;
- A color cut of $0.6 < B-V < 1.4$, i.e. dwarf stars;
- In the solar neighborhood, at a distance of less than 35 pc, according to the Hipparcos catalog;
- Not very active ($V_{sini} < 4.5$ km/s, $\text{Log}(R'HK) < -4.8$);
- Not already known to have a stellar or planetary companion;
- Not observed in HARPS-N GTO.

The Final SP1 sample includes 190 stars, of which 71 stars are of spectral type G and 119 are of spectral type K (see Fig. 3.1, first row). Their magnitude goes from 5.3 to 9.7 (see Fig. 3.1, second row right) with a B-V color between 0.6-1.4 (see Fig. 3.1, second row left). These criteria make SP1 stars suitable for discovering low-mass planets (see Courcol-Rozès, 2016 for more information about the SP1 program).

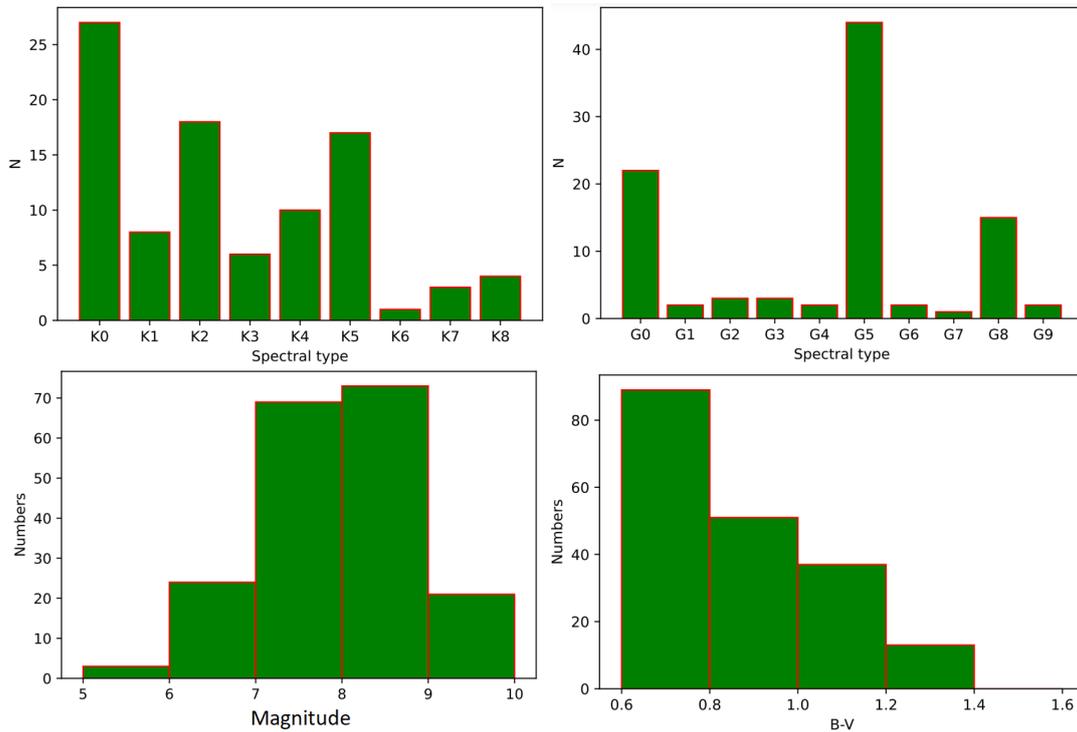


Figure 3.1 – Histogram of spectral types (*first row*), Magnitudes (*second row left*) and B-V color (*second row right*) of the SP1 sample.

3.1.3 Observation strategy

To achieve the goals of the SP1 program, the highest precision of the spectrograph is needed, requiring addressing instrumental and stellar limitations. To do so, we use SOPHIE's high-resolution mode (resolution power of $\lambda/\Delta\lambda \approx 75000$) with a simultaneous calibration lamp observation. The latter allows us to monitor the instrumental drift during the observation to increase the precision of the data. The exposure time is set to 600 to 1200 s, depending on the stellar magnitude. This helps to: 1. achieve an S/N greater than 100; 2. average the stellar oscillations; 3. obtain a photon noise of less than 2 m/s. Additionally, the observer should avoid: observing at airmass > 1.6 to help minimize the effects of absorption and emission of the atmosphere; and observing when the Moon and the star are closer than 30 deg and $|RV_{Moon} - BERV| > 15 \text{ km/s}$. The observer should also increase the exposure time when the seeing is important ($> 4''$). This setup increases the precision of the data and the possibility of discovering low-mass planets.

3.1.4 SP1 status in 2022

The allocated time for SP1 observations by SOPHIE is around 15 nights per semester. From July 2012 to the time of writing of this thesis in 2022, 96 stars have been observed with at least 10 measurements each, with a precision typically between 2-4 m/s (see Fig. 3.2 left).

As shown in Fig. 3.2 right, most of the targets were observed less than 50 times, with two peaks between 30-40 measurements. This is because after taking typically 30-40 measurements, a first analysis is performed, leading to deciding whether to continue or stop the observation of this star (e.g., due to stellar activity). The stars with more than 50 measurements may show a signature of one or even several planet candidates. The number of observations needed to detect a low-mass planet with very low amplitude (1-2m/s) is typically greater than 100 measurements, depending on the quality of the data and period of the planet. As an example, 290 RV measurements of star HD 158259 have been taken by SOPHIE from 2012 to 2019, and 5 planets and one strong candidate have been revealed around this star (Hara et al., 2020).

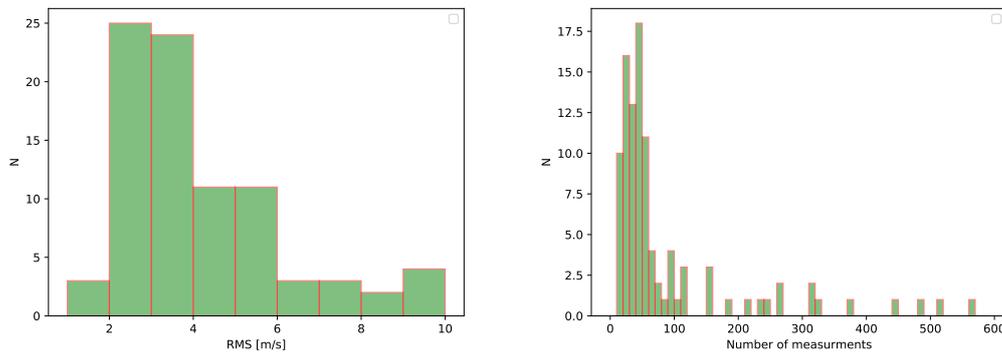


Figure 3.2 – Histogram of RV RMS (left) and the number of SOPHIE observation points (right) for the SP1 stars.

3.2 Objective of my work on SP1

As detailed in chapter 1, the presence of stellar activity can interfere with spectroscopic and photometric observations of exoplanets, leading to false detection or missing signals from the RVs. Stellar activity can also influence the accurate determination of exoplanet parameters. These effects pose a serious challenge in the detection of low-mass planets, which is the subject of the present work.

Using simultaneous RV and photometric measurements delivers a wealth of information about the stellar activity, a key to evaluating the activity of the host star and performing accurate data analysis. As an example, for those SP1 stars with less than 50 measurements, two possibilities arise: either the RVs periodogram did not show any periodic signal, or the RV signal has a corresponding signal in one of the activity indicators. In the latter case, there is always the possibility that the RV signal is a planet and the stellar rotational signal (or its harmonic $P_{rotation}/2$ and $P_{rotation}/3$) is close to it. As an example, observations of HD207897 were stopped after 44 RV measurements, because the planet signal at 16.2 days was near half of the stellar rotational peaks, and it was likely that the RV signal was due to stellar activity. But later, we discovered that the origin of the signal is planetary by using TESS photometric observations (see Sect. 3.6.1).

Additionally, a combination of both RV and photometric measurements for transiting planets, allows us to measure fundamental parameters like mass and radius, leading to measuring the mean density. With this information in hand, we can distinguish between the type of planets (rocky or gas), model the planetary atmosphere, and study the diversity of planetary composition.

Accordingly, to seek these purposes, I analyzed both RV and photometric measurements (in case of availability) of SP1 stars. This work particularly focused on:

- Evaluating the activity of the host star of the SOPHIE planet candidates;
- Investigating the likeliness of activity-induced RV variations;
- Trying to detect new planet signals in RVs and/or photometric data, and well characterize them.

3.3 Data processing and selections

For detecting a low-mass planet with a small Keplerian amplitude (1-2 m/s), careful data processing is required. For SP1 targets, I carry out its data processing, by developing an automatic pipeline that has the following steps:

- Extracting the RV data using improved SOPHIE data reduction systems including the flux corrections of spectra, fixing mask lines selection at different epochs, and CTI corrections (see Chapter 2 for the detail of my work on the DRS improvements).
- Subtract the drift of the spectrograph from RVs. The nightly drift of the spectrograph is tracked and measured by simultaneous calibration lamp observations (see Sect. 2.1 for more explanation of calibration lamp).

- Removing the bad quality data with:
 - Lower S/N: the threshold is typically $S/N=50$. However, depending on the magnitude of the stars, the low S/N data might be removed up to $S/N=80$ in the data analysis.
 - High drift of the spectrograph: usually, the measured drift is around 3 m/s at maximum. If it exceeds 5 m/s, the observer should repeat the wavelength solution to reset the zero point. However, due to different reasons, it still happens that the measured drift for some nights exceeds more than 15 m/s. It shows that the instrument was unstable. I remove them from data analysis which would be detrimental to high precision.
 - High/low calibration lamp flux: the flux of the calibration lamp is usually between 0.2 and 1, otherwise, it would not be validated. Therefore, I remove such points.
 - Moonlight affected data (see Sect. 2.4 for more details).
- Performing the optimized master constant correction on RVs, bisector, and FWHM time series (see Sect. 2.6).
- Extracting the activity indicators, like $\log(R'_{HK})$ and H α index with new background correction of calibration lamps (see Sect. 2.5).
- Saving the final reduced RVs in the ASCII files.

3.4 Analysis of only RVs and activity indices

In each semester, the last data sets of SP1 stars are analyzed, and the target observations for the following semester are prioritized based on their results. Here, I present an analysis of RV data as well as activity indicators for recently observed SP1 stars. This subset of stars, which was intensively observed by SOPHIE during my Ph.D., had previously displayed several interesting signals in their RV analysis over the previous few years. For computing the periodograms, I used the website of the Data and Analysis Center for Exoplanets (DACE, Delisle et al., 2016)¹. The results of my analysis are summarized in Table 3.1, and each target is discussed in the following.

1. Available at <https://dace.unige.ch>

Table 3.1 – Summary of SP1 RVs analysis.

star	N_{obs}	$P_{candidate}$ (d)	$P_{H\alpha}$ (d)	$P_{\log(R'_{HK})}$ (d)	P_{Biss} (d)
star1	142	360*	—	> 300*	237*
star2	144	> 200*, 20 [×]	> 164*	> 127*	21 [×] , 51 [†] , 613 [†]
star3	484	12.93*, 6*, 10*, 374*, > 2000*	> 1400*, 259.7 [†]	> 160*	—
star4	313	> 2000*, 653*, 6.5*, 107.4 [†]	> 2000*, 360*	> 2000*, 450*, 180*	326 [†]
star5	234	10.12*	> 2000*, 380*, 32 [†]	1400*, 360*	580 [†] , 44 [†] , 31 [†]
star6	161	25.8*, 55.9*	26*, 109*	25.7*	—
star7	314	14.25*, 39.8*, 528.8*	183*, > 200*	> 200*	—
star8	218	2441*, 31.8*, 16.4*, 1.1 [†] ,	> 2000*	> 100*, 31.5 [×] , 37.1 [†]	> 2000*
star9	267	25.5*, 5.5*	122*, 183*, 400*	> 900*, 360*	171 [×]

Note: *, [†], [×] denote a periodic signal below FAP 0.1%, 1%, 10%, respectively. Because the data for these stars are not public, the names of the stars are not provided.

Star 1- activity (Fig. 3.3): This star is a G-type star with a magnitude of $V=7.0$. We gathered 142 SOPHIE observations from October 2014 to March 2020, with an RV RMS of 2.84 m/s, mean error bars of 1.2 m/s, and $\langle \log(R'_{HK}) \rangle = -4.9$. I removed 3 data points due to their low $S/N < 50$. The RVs present a long-term drift that has a corresponding signal at the long period in the periodogram of $\log(R'_{HK})$. There is also a strong correlation between the RVs and the $\log(R'_{HK})$ ($R=0.6$). Therefore, it is likely that this drift is caused by stellar activity. After removing the drift by a polynomial line, there is a signal at 360 d with a FAP of below 0.1 %. The origin of this peak is uncertain. However, it is compatible with a yearly signal, which is most likely the result of instrumental factors. Additional DRS improvements might help to understand the origin of this signal (e.g. testing the stitching effect of CCD (Dumusque et al., 2015a)). The residuals show no interesting signal.

Star 2- activity+hint (Fig. 3.4): This star is a G2V star with a magnitude of $V=6.76$. From September 2014 to November 2020, 144 data points were collected with an RV RMS of 11.6 m/s and mean error bars of 1.1 m/s. 3 data points are discarded due to the low $S/N < 50$. The RVs periodogram shows a long-term signal that has a corresponding signal in the activity indicator of $H\alpha$ and $\log(R'_{HK})$. There is also a strong correlation between the RVs and the $\log(R'_{HK})$ ($R=-0.6$). Therefore, it is likely that this signal has an activity origin. Furthermore, this star has a signal at 20 days with a FAP of below 10 %. More observations are required to determine the origin of this signal and get a sufficient level of detection for publication.

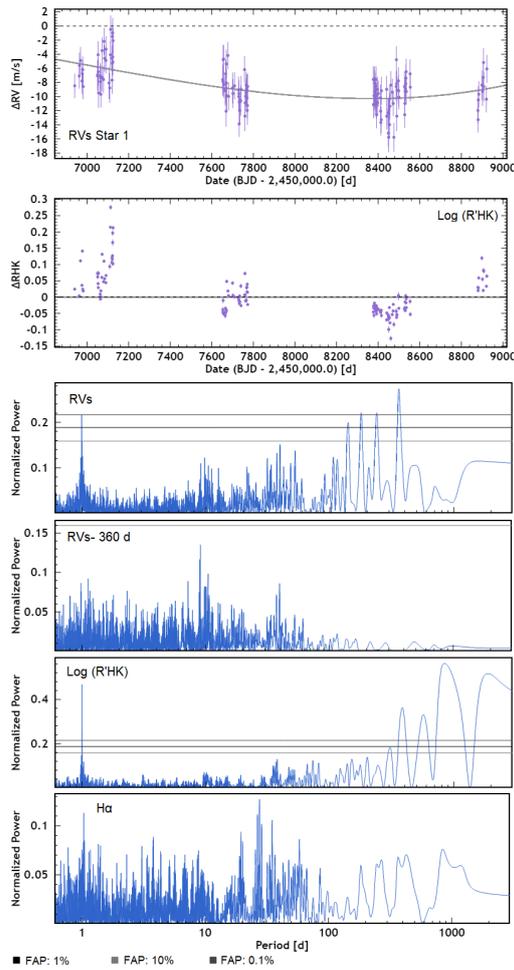


Figure 3.3 – RVs, $\log(R'_{HK})$ time series, and periodograms for star1.

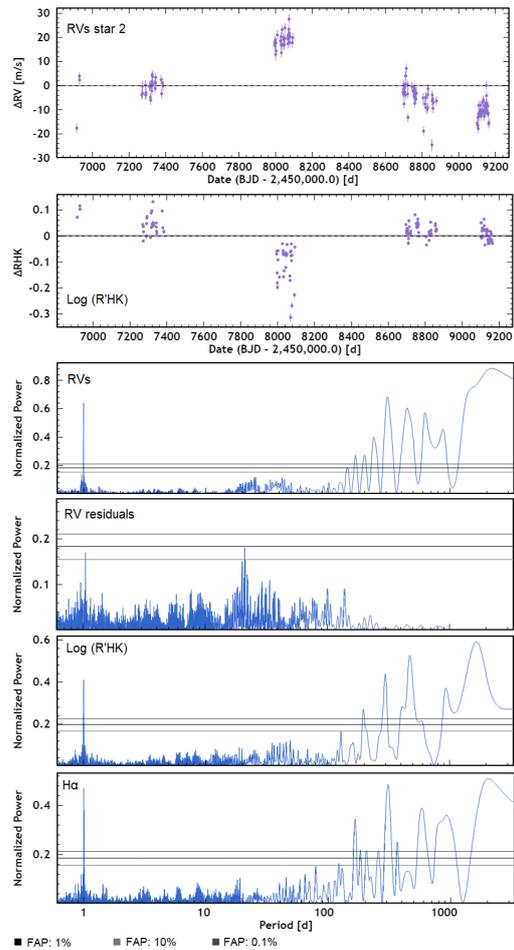


Figure 3.4 – RVs, $\log(R'_{HK})$ time series, and periodograms for star2.

Star 3- activity+ 3 detections (Fig. 3.5): This star is a G9V star with a magnitude of $V=7.82$. We have 484 data which are taken from September 2012 to January 2022 with an RV RMS of 3.7 m/s and mean RV error bars of 1.6 m/s. In total 28 data points are removed: 7 data points with a low S/N <50 , 19 data points due to the use of the different mask for calculating data, and 2 data points being outliers. I subsequently fitted a Keplerian model to the highest peak at each RVs periodogram, revealing 5 signals with a FAP level below 0.01 % at 374 d, 12.9 d, 10 d, 6 d, and a long period of >2000 . Except for the long-period signal at >2000 d, none of these signals has a corresponding peak in the periodogram of activity indicators. Therefore, it is likely that the signals at 12.9 d, 10 d, and 6 d have a planetary origin. Because the long-period signal has a corresponding peak in the $H\alpha$ activity indicator, it is likely that its origin is stellar. Furthermore, the peak at 374 d is consistent with the yearly signal and is likely

Table 3.2 – Results of Keplerian fits on star 3

Period	K (m/s)	M sin i (M_E)
12.9	1.9	6.3
10	1.7	5.6
6	1.5	4.0

due to instrumental factors. As I previously mentioned, improvement in DRS might help to identify the source of this signal. The results of fitting Keplerian models are presented in Table. 3.2.

Star 4- activity+ 2 detections+hint (Fig. 3.6): This star is a G5 star with a $V=7.22$ magnitude. The observations of the star were performed from February 2012 to March 2022 and 313 RV measurements were collected. The RV RMS and the mean RV error bars are 11.6 m/s and 1.3 m/s, respectively. 3 data points are removed because of the low $S/N < 50$. The first peak in the periodogram is around 3800 d, with a FAP of less than 0.1 %. A similar signal can be found in the $H\alpha$ and $\log(R'_{HK})$ activity indicators, indicating that it is likely caused by stellar activity. After removing this peak, there is a signal in 633 d with a FAP below 0.1 % that reveals a signal at 6.5 d after being fitted by the Keplerian model. Both signals at 633 d and 6.5 d do not have a corresponding signal in the periodogram of activity indicators. Hence, it is likely that the planetary is the origin of these signals. The Keplerian model on both planets shows that the planet at 633 d has a Keplerian amplitude of 5.8 m/s and a minimum mass of 70.80 M_E ; the planet at 6.5 d has a Keplerian amplitude of 2.6 m/s and a minimum mass of 7.5 M_E . In the RV residuals, there are a bulk of signals around 100 d that are required additional observations to determine their origin and statistically become stronger.

Star 5- drift+1 detection (Fig. 3.7): This star is a G2 spectral type star with a magnitude of $V=5.3$. We have 234 measurements from November 2015 to March 2021 with an RV RMS of 5.2 m/s and mean error bars of 1 m/s. 2 data points are removed due to the low $S/N < 50$. The RVs of this target presents a drift. Given that the periodogram of the $H\alpha$ activity indicator shows a long period signal, it is likely that the stellar activity is the origin of this drift. After removing this drift with a linear line, the periodogram shows a strong signal at 10.1 d with a FAP below 0.1 %. There is no corresponding peak in the activity indicators for this signal. Therefore, it is likely that its origin is planetary. I note that excluding or including the removal of drift in the analysis, had no effect on the signal detection. A Keplerian model fitted to the signal yields a Keplerian amplitude of 4.5 m/s and a minimum mass of 15.2 M_E .

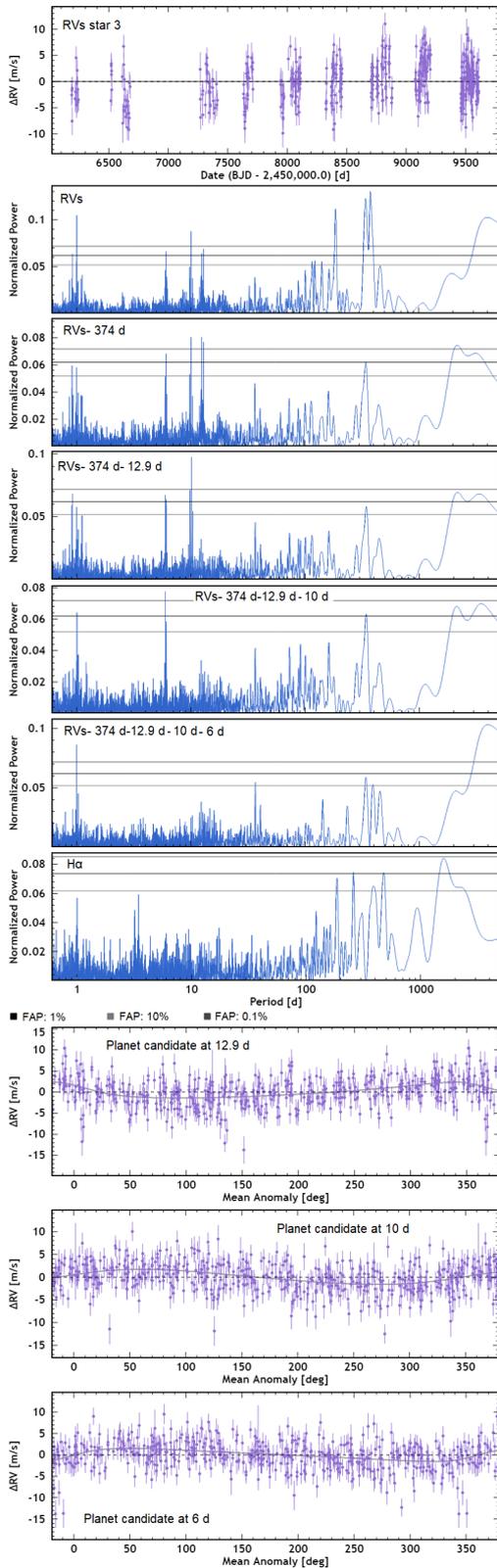


Figure 3.5 – RVs time series, periodograms, and phase folded RVs at the planetary candidates period for star 3.

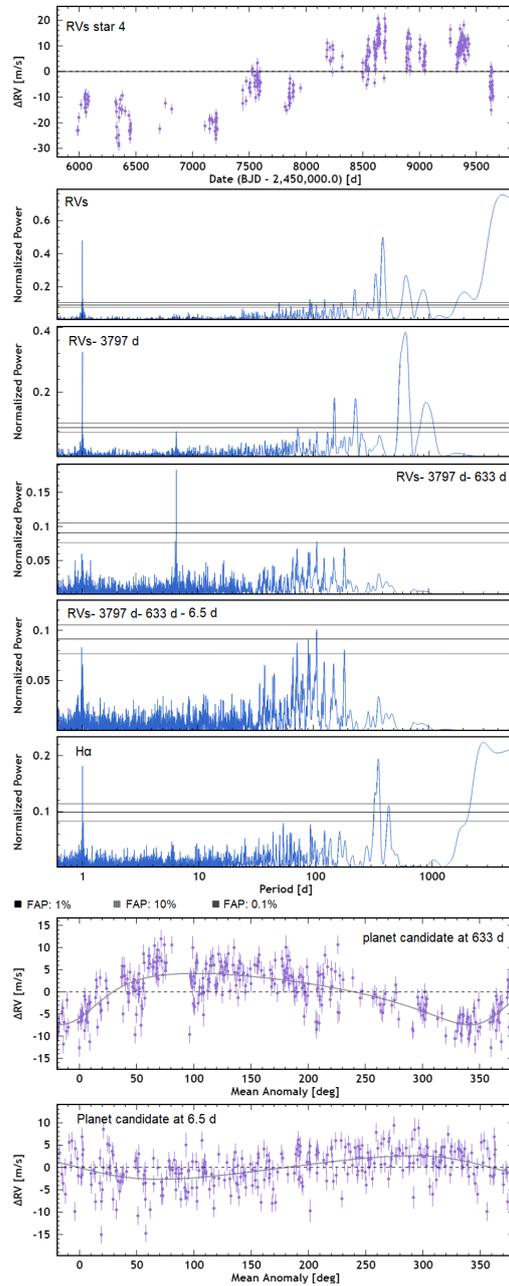


Figure 3.6 – RVs time series, periodograms, and phase folded RVs at the planetary candidates period for star 4.

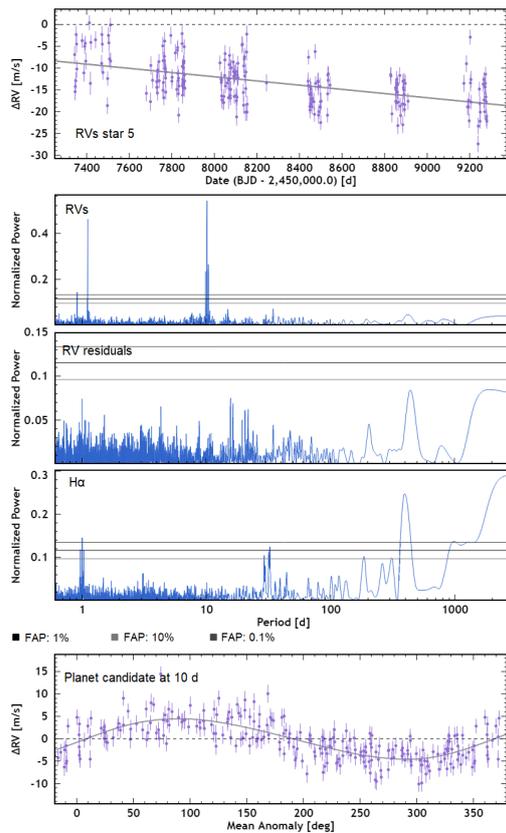


Figure 3.7 – RVs time series, periodograms, and phase folded RVs at the planetary candidate period for star 5.

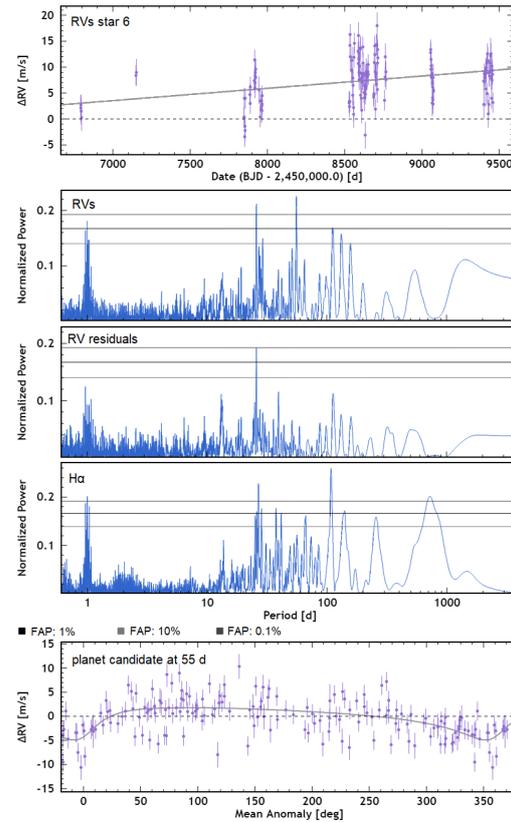


Figure 3.8 – RVs time series, periodograms, and phase folded RVs at the planetary candidate period for star 6.

Star 6- drift+ 1 detection (Fig. 3.8): This star is a K5 star with $V=8.55$ magnitude. The observations of this star were done from May 2014 to August 2021 with an RV RMS of 4 m/s and mean error bars of 1.5 m/s. 6 and 4 data points are discarded because of low $S/N < 50$ and outliers, respectively. After fitting a drift on the RVs, the periodogram shows a signal at 55.9 d with a FAP of less than 0.1% and no corresponding activity signal. Therefore, it is likely that its origin is planetary. The Keplerian fit on this signal gives a Keplerian amplitude of 3.4 m/s and a minimum mass of $16.4 M_E$. This star is also affected by the activity of the host star, and the RV signal in 25.8 d has a corresponding peak in the $H\alpha$ activity indicator. Therefore, the origin of this signal is likely due to the activity of its host stars. Furthermore, the cause of long-term drift is not clear. However, because this star is active $\langle \log(R'_{HK}) \rangle = -4.6$, the cause of this drift is likely to be stellar activity. More observations are required to investigate it.

Star 7- 2 detections+hint (Fig. 3.9): This is a G2V star with a magnitude of $V=7.07$. We have 314 data from May 2012 to July 2021 with an RV RMS of 3.4 m/s and mean error bars of 1.7 m/s. 4 data points are removed due to the low $S/N < 50$ and 2 more are detected as outliers. The first RVs periodogram displays a clear signal at 39.8 d with a FAP below 0.1 %. After fitting the Keplerian model to this signal, the RVs periodogram shows a 14 d signal with a FAP below 0.1 %. Since these peaks have no corresponding signals in the periodogram of activity indicators, it is likely that they are caused by planets. A Keplerian fit to both signals reveals that the planet at 39 d has a Keplerian amplitude of 2.7 m/s and a minimum mass of $14.6 M_E$, whereas the planet at 14 d has a Keplerian amplitude of 2.1 m/s and a minimum mass of $7.8 M_E$. The RV residuals show a signal at 528.8 d (and its alias at 218 d) with a FAP below 0.1 %. This star requires further observations because the periodograms of activity indexes show several peaks with periods of $> 200d$, and it is unclear whether the signal at 528.8 d is due to stellar activity or a planet.

Star 8- activity+hint (Fig. 3.10): This star is a K0 star with a magnitude of $V=7.54$. The observations of this star are done from May 2013 to November 2021, gathering 218 data with an RV RMS of 7.4 m/s and mean error bars of 1.3 m/s. In total 4 data points are removed: 2 with low S/N and 2 being outliers. I estimated the star rotational period 37_{-6}^{+3} d, following the method of Noyes et al., 1984b. The periodogram of $\log(R'_{HK})$ shows a bulk of peaks in this area. The first RVs periodogram displays a long period peak, which has a corresponding signal in the $H\alpha$ and $\log(R'_{HK})$ activity indicators. This star is an active star with $\langle \log(R'_{HK}) \rangle = -4.7$ and its RVs and $\log(R'_{HK})$ activity indexes are highly correlated ($R= 0.8$). Therefore, it is likely that this peak has an activity origin. After removing the long-term peak, the periodogram shows a signal at 31.8 d with a FAP below 0.1 %. The signal is likely caused by the activity because it is compatible with the estimated star rotational period, and also has a corresponding peak at 31.5 d with a FAP of 10 % in the periodogram of $\log(R'_{HK})$. In addition, after removing the signal at 31.8 d, the periodogram reveals another signal at 16.4 d with a FAP of below 0.1 %. Because this signal is also close to half of the signal at 31.5 d in the periodogram of $\log(R'_{HK})$, determining its origin is very difficult. Furthermore, the target has a signal at 1.1 d and 27 d in the RV residuals with a FAP of less than 1 %. These signals have no corresponding peaks in the periodogram of activity indicators. More observations are required to verify the origin of all these signals, and the signals at RV residuals statistically become stronger.

Star 9- 2 detections (Fig. 3.11): This star is a G5 star with a magnitude of $V=7.81$. We have 267 data from July 2012 to October 2020 with an RV RMS of 4.8 m/s and mean RV error bars of 1.3 m/s. 13 data points were discarded due to their low $S/N < 50$, and 1 because of being an outlier. RVs periodogram of this star shows a peak at 25.5 d with a FAP of below 0.1 %. After fitting it, the periodogram reveals another signal at

5.5 d with a FAP of below 0.1 %. These peaks have no corresponding signals in the periodogram of activity indicators. Therefore, it is likely that they have a planetary origin. Applying the Keplerian model to them shows a Keplerian amplitude of 4.1 m/s with a minimum mass of 19 M_E for the 25.5 d signal, and a Keplerian amplitude of 3.2 m/s with a minimum mass of 8.6 M_E for the 5.5 d signal.

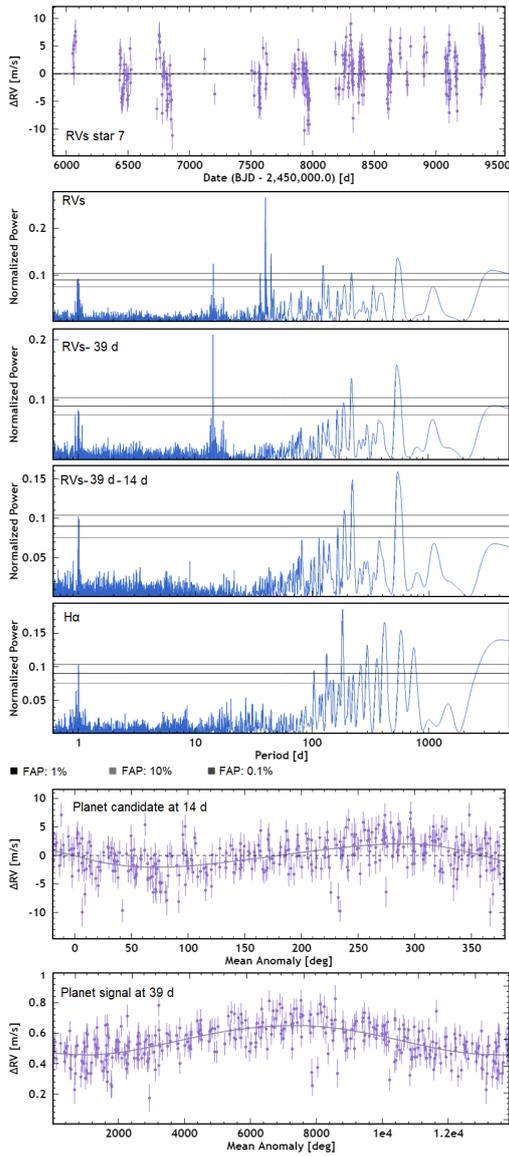


Figure 3.9 – RVs time series, periodograms, and phase folded RVs at the planetary candidates period for star 7.

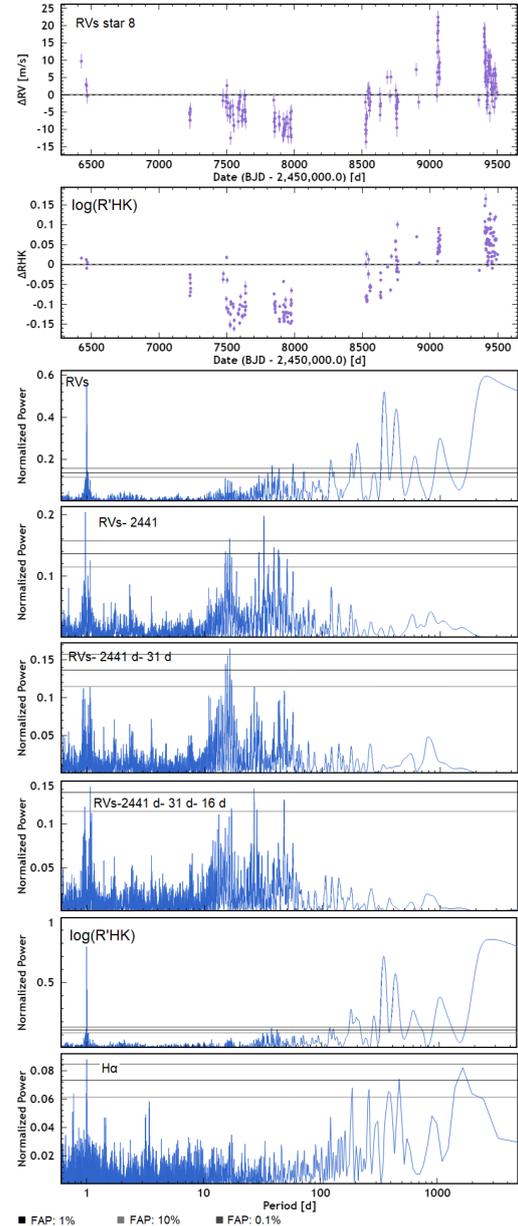


Figure 3.10 – RVs, $\log(R'_{HK})$ time series, and periodograms for star 8.

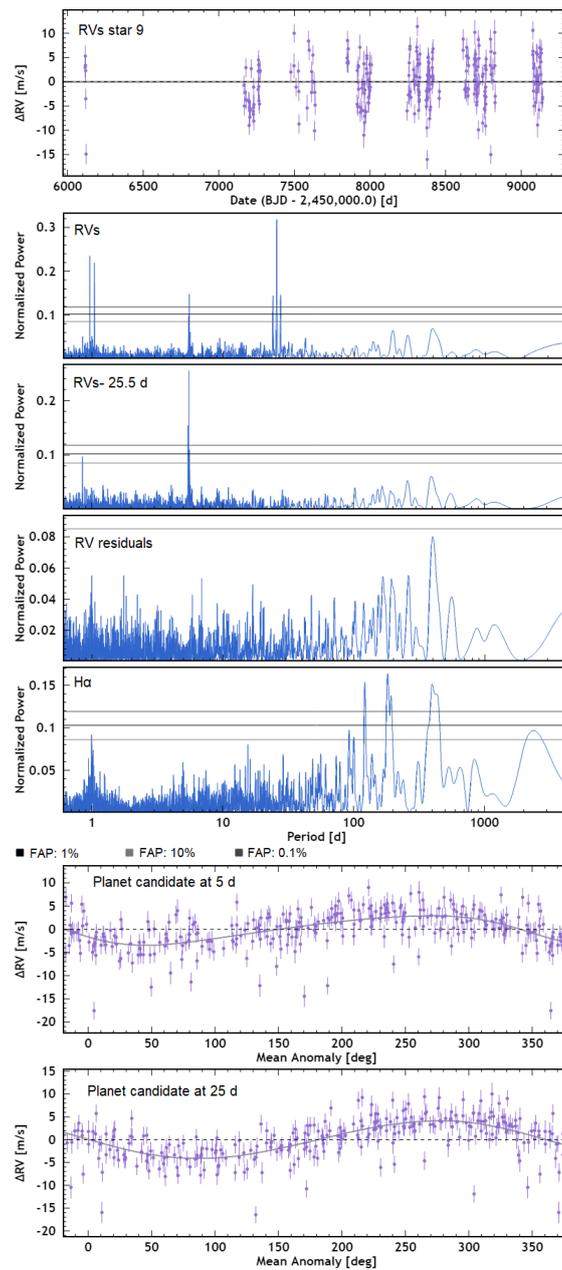


Figure 3.11 – RVs time series, periodograms, and phase folded RVs at the planetary candidates period for star 9.

3.5 Stellar rotation period investigation

By incorporating TESS photometry into the analysis of RV periodic signals, some false positive RV candidates can be eliminated, indicating that the RV variation could

have been caused by stellar rotation rather than an exoplanet. With this objective in mind, I investigated the stellar rotational signature in TESS photometric data for about 40 SP1 stars at the beginning of my Ph.D. I could not derive the stellar rotation period for all of the stars in the sample because some of them lacked photometric data. Another reason is that the TESS observation time span is short (~ 27 d), and cannot cover one phase of the stellar rotation for some stars. Finally, not all stars exhibit the activity signature in their photometric data due to the lack of active regions. For instance, of the sample of nine stars mentioned above, 4 stars were not observed by TESS between sectors 13-26 while it was observing the northern hemisphere (beginning of my Ph.D.). Following Noyes et al., 1984b recipe, 3 stars are estimated to have rotational periods greater than the TESS time span observation. Therefore, searching for the stellar rotational period was limited to two stars.

Here, I present one example, HD82106, where I successfully removed a false positive SOPHIE planet candidate by analyzing the SOPHIE RV measurements and the TESS photometric data. HD82106 is a K3 spectral-type star with a magnitude of $V = 7.1$. The star was monitored by SOPHIE between February to April 2017, gathering 22 RV measurements with a dispersion of 6.7 m/s. The star is active with $\langle \log(R'_{HK}) \rangle = -4.4$. Its RV periodogram (Fig. 3.12 right panels, top) illustrates a peak at 5.9 d with a FAP of below 1%. The same peak was found in the periodogram of the bisector span (see Fig. 3.12 right panels, bottom).

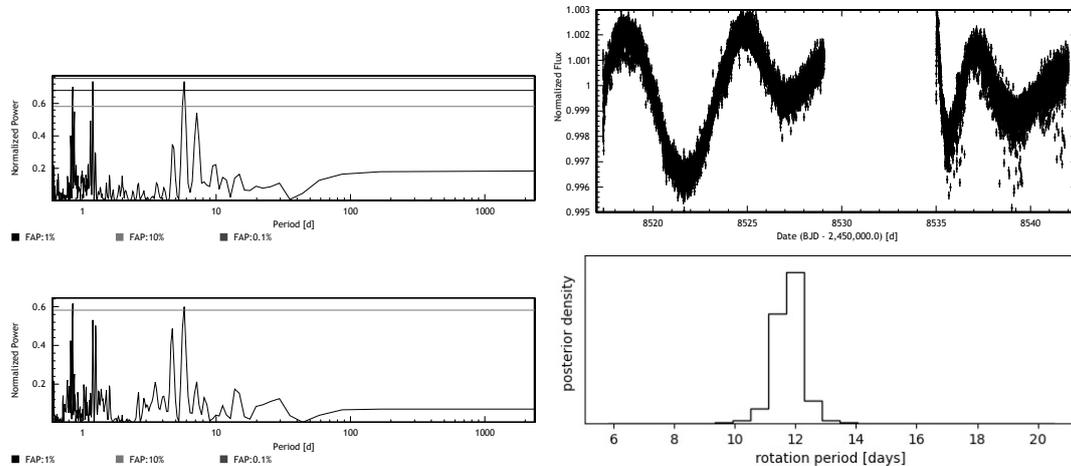


Figure 3.12 – *Right panels*: Periodogram of RVs (*top*) and bisector span (*bottom*) of HD82106. *Left panels*: TESS light curve of the target (*top*) and the posterior distribution of the stellar rotational period (*bottom*).

The star was also observed by TESS in sector 8 from 2nd to 28th February 2019. The star’s TESS light curve provides additional support for the fact that the star is very active (see Fig. 3.12 left panels, top). To find the star rotational period, I used the EXOPLANET package (Foreman-Mackey et al., 2021) and performed a Gaussian

Processes (GPs) model on the TESS data. The posterior distribution of the stellar rotational period shows a peak at 12 d (see Fig. 3.12 left panels, bottom). Given the star is very active and there is a corresponding peak in the bisector activity indicators, the RVs signal at 5.9 d is likely half of the rotational period peak at 12 d.

3.6 Discovery and characterization of new transiting planets

3.6.1 HD207897 b: A dense sub-Neptune transiting a nearby and bright K-type star

HD207897 b is a transiting sub-Neptune that orbits the nearby (28 pc) and bright ($V=8.37$) K0V star HD207897 (TOI-1611) every 16.20 days. This discovery is based on TESS photometric data as well as RV measurements from the SOPHIE, Automated Planet Finder (APF), and HIRES high-precision spectrographs. The results are published in Heidari et al., 2022, for which I gathered 96 specialists from around the world and led the publication. I summarise the main properties of the planet here and give more details in the subsequent pages.

This planet has a radius of $2.50 \pm 0.08 R_E$ and a mass of either $14.4 \pm 1.6 M_E$ or $15.9 \pm 1.6 M_E$ with nearly equal probability. The two solutions correspond to two possibilities for the stellar activity period. The density accordingly is either $5.1 \pm 0.7 \text{ g cm}^{-3}$ or $5.5^{+0.8}_{-0.7} \text{ g cm}^{-3}$, making it one of the relatively rare dense rocky sub-Neptunes. The existence of this dense planet at only 0.12 AU from its host star is unusual in the currently observed sub-Neptune population. We investigated different scenarios. One possibility is that the planet has lost most of its volatile elements by evaporation, but for the case of HD 207897 b with an orbital period of 16.20 d and receiving an incident flux of $F = 26.3 F_E$, this is not a satisfactory answer. Even when we consider an extreme evaporation process (Des Etangs, 2007), the mass loss of the planet would be just $0.1 M_E$ during the entire lifetime of the star, which cannot account for its high density. The most likely scenario is that this planet has migrated to its current position.

HD 207897 b: A dense sub-Neptune transiting a nearby and bright K-type star

N. Heidari^{1,2,3,*}, I. Boisse^{3,*}, J. Orell-Miquel^{4,5}, G. Hébrard^{6,7}, L. Acuña³, N. C. Hara⁸, J. Lillo-Box⁹, J. D. Eastman¹⁰, L. Arnold^{6,11}, N. Astudillo-Defru¹², V. Adibekyan^{13,14}, A. Bieryla¹⁰, X. Bonfils¹⁵, F. Bouchy⁸, T. Barclay^{17,16}, C. E. Brasseur¹⁸, S. Borgniet⁴⁰, V. Bourrier⁸, L. Buchhave⁴⁴, A. Behrard^{19,**}, C. Beard²⁰, N. M. Batalha²¹, B. Courcol⁸, P. Cortés-Zuleta³, K. Collins¹⁰, A. Carmona¹⁵, I. J. M. Crossfield²², A. Chontos^{23,**}, X. Delfosse¹⁵, S. Dalal⁷, M. Deleuil³, O. D. S. Demangeon^{13,14}, R. F. Díaz²⁴, X. Dumusque⁸, T. Daylan^{25,***}, D. Dragomir²⁶, E. Delgado Mena¹³, C. Dressing²⁷, F. Dai²⁸, P. A. Dalba^{29,****}, D. Ehrenreich⁸, T. Forveille¹⁵, B. Fulton³⁰, T. Fetherolf²⁹, G. Gaisné¹⁵, S. Giacalone²⁷, N. Riazi¹, S. Hoyer³, M. J. Hobson^{31,32}, A. W. Howard³³, D. Huber²³, M. L. Hill²⁹, L. A. Hirsch³⁴, H. Isaacson^{27,35}, J. Jenkins⁴⁶, S. R. Kane²⁹, F. Kiefer^{7,40}, R. Luque^{4,5}, D. W. Latham¹⁰, J. Lubin²⁰, T. Lopez³, O. Mousis³, C. Moutou⁴¹, G. Montagnier⁶, L. Mignon¹⁵, A. Mayo²⁷, T. Močnik³⁹, J. M. A. Murphy^{21,**}, E. Palle^{4,5}, F. Pepe⁸, E. A. Petigura³⁶, J. Rey³⁷, G. Ricker²⁵, P. Robertson²⁰, A. Roy^{18,38}, R. A. Rubenzahl^{33,**}, L. J. Rosenthal³³, A. Santerne³, N. C. Santos^{13,14}, S. G. Sousa¹³, K. G. Stassun⁴², M. Stalport⁸, N. Scarsdale²¹, P. A. Strøm^{7,47,48}, S. Seager^{25,49,50}, D. Segransan⁸, P. Tenenbaum⁴³, R. Trongsgaard⁴⁴, S. Udry⁸, R. Vanderspek²⁵, F. Vakili², J. Winn⁴⁵, and L. M. Weiss²³

(Affiliations can be found after the references)

Received 29 May 2021 / Accepted 8 September 2021

ABSTRACT

We present the discovery and characterization of a transiting sub-Neptune that orbits the nearby (28 pc) and bright ($V = 8.37$) K0V star HD 207897 (TOI-1611) with a 16.20-day period. This discovery is based on photometric measurements from the Transiting Exoplanet Survey Satellite mission and radial velocity (RV) observations from the SOPHIE, Automated Planet Finder, and HIRES high-precision spectrographs. We used EXOFASTv2 to model the parameters of the planet and its host star simultaneously, combining photometric and RV data to determine the planetary system parameters. We show that the planet has a radius of $2.50 \pm 0.08 R_E$ and a mass of either $14.4 \pm 1.6 M_E$ or $15.9 \pm 1.6 M_E$ with nearly equal probability. The two solutions correspond to two possibilities for the stellar activity period. The density accordingly is either $5.1 \pm 0.7 \text{ g cm}^{-3}$ or $5.5^{+0.8}_{-0.7} \text{ g cm}^{-3}$, making it one of the relatively rare dense sub-Neptunes. The existence of this dense planet at only 0.12 AU from its host star is unusual in the currently observed sub-Neptune ($2 < R_E < 4$) population. The most likely scenario is that this planet has migrated to its current position.

Key words. planets and satellites: detection – stars: activity – planets and satellites: individual: HD 207897 – planets and satellites: individual: TOI-1611 – planets and satellites: individual: TIC ID 264678534

1. Introduction

The brightness of more than 200 000 stars has been monitored by Transiting Exoplanet Survey Satellite (TESS) (Ricker et al. 2015) with two-minute cadence during its two-year primary mission. The observed stars are closer and brighter (typically 30–100 times brighter) than the stars *Kepler* surveyed. This offers us a unique opportunity for furthering our knowledge in planetary science with follow-up observations: those from ground-based high-precision spectrographs to confirm the planetary nature and mass measurement, which together with the radius allows us to determine the bulk composition of planets; and ground- and space-based observations to provide atmospheric characterization, for instance, with the upcoming *James Webb* Space Telescope (Gardner et al. 2006).

The NASA *Kepler* mission (Borucki et al. 2010) has discovered a large number of planets of intermediate size, with radii

between those of Earth and Neptune. They are also known as sub-Neptunes. Because the size of planets is directly dependent on the physical mechanisms in their formation and evolution, the absence of sub-Neptunes in our Solar System and their abundance among the exoplanet population has raised numerous fundamental questions. Many theoretical and statistical studies have been made of such planets. Rogers (2015) showed that most planets with $R > 1.6 R_E$ have a low density and are inconsistent with a purely rocky composition. Fulton et al. (2017) demonstrated that the sub-Neptunes' distribution of radii is bimodal, with two peaks centered at $1.2 R_E$ and $2.4 R_E$. This reveals a gap in planet radii between 1.5 – $2.0 R_E$. While this bimodal distribution can be explained by photoevaporation (Owen & Wu 2013, 2017; Lopez & Fortney 2014) and core-powered mass loss (Ginzburg et al. 2016, 2018), the composition and origin of close-in sub-Neptunes are not yet clear. Additional detections and precise characterizations are the keys to progress in answering our questions about their nature. One of the primary goals of TESS is to measure the mass of at least 50 transiting planets with a radius smaller than 4 Earth radii (Ricker et al. 2015). As of 22 April 2021, TESS has found more than

* Corresponding authors; N. Heidari (email: neda.heidari@lam.fr), I. Boisse (email: isabelle.boisse@lam.fr).

** NSF Graduate Research Fellow.

*** Kavli Fellow.

**** NSF Astronomy and Astrophysics Postdoctoral Fellow.

750 candidates¹ (so-called TESS objects of interest, or TOI) with radii smaller than $4 R_E$. To date, 122 of them have been confirmed and have a mass measurement (e.g., Dragomir et al. 2019; Günther et al. 2019; Nielsen et al. 2020).

Here, we announce the detection and characterization of a sub-Neptune orbiting a bright ($V = 8.4$) K0 star using TESS photometric data and SOPHIE, Automated Planet Finder (APF), and HIRES RVs. In Sect. 2 we describe the variety of observations that were used to characterize the HD 207897 (TOI-1611) system, including photometric, spectroscopic, and high-resolution imaging data. In Sect. 3 we analyze the data and present the characterization of the host star and planet, combining models on RVs and transit data. Finally, we present our discussion and conclude in Sect. 4.

2. Observations

In this section, we provide a summary of all observations of HD 207897 taken by TESS and ground-based facilities.

2.1. TESS photometry

TESS planned to observe 80% of the sky in 26 sectors for 2 yr. Each sector lasts for approximately 27 days, and the sectors partially overlap². Because HD 207897 is located near the north ecliptic pole where sectors overlap, it appears in many sectors. As reported in the Web TESS Viewing Tool (WTV)³, HD 207897 observations were taken in five sectors divided into two continuous periods from sectors 18–20 and 25–26, with a total time span of 131 days. After observations of sector 18 were completed, the MIT Quick Look pipeline (QLP; Huang et al. 2020a,b) detected the signature of two transits of HD 207897 b at a period of 16.20 d, and an alert was issued on 19 December 2019 by the TESS Science Office. No transits occurred during sector 19 as the sole transit fell in the data gap between the two orbits. After observations of sector 20 were downlinked, the Science Processing Operations Center (SPOC; Jenkins et al. 2016) at NASA Ames Research Center conducted a transit search (Jenkins 2002; Jenkins et al. 2010) and detected two transits of HD 207897 b at a signal-to-noise ratio (S/N) of 23.3. A limb-darkened transit model was fitted to the transit signature (Li et al. 2019), which passed all the diagnostic tests presented in the Data Validation report (Twicken et al. 2018), including the odd-even transit depth test, the ghost diagnostic test, and the difference-image centroiding test, which located the source of the transit signatures within 0.42 ± 2.5 arcsec of the target star HD 207897. Two additional transits were observed and detected in sector 25, and one more transit occurred in sector 26. A multisector search of sectors 18–26 by the SPOC detected 7 transit events of HD 207897 b in total at an S/N of 34.3 and an average depth of 913.7 ± 21.1 ppm. No additional transiting-planet signatures were detected in any of the SPOC or QLP runs.

We used the short-cadence (2 min) observations of TESS data from these five sectors in our photometric analysis, which we present in Sect 3.2. These data were reduced by the SPOC Pipeline and are publicly available on Mikulski Archive for Space Telescopes (MAST)⁴.

¹ <https://tess.mit.edu/publications/>

² <https://tess.mit.edu/observations>

³ <https://heasarc.gsfc.nasa.gov/cgi-bin/tess/webtess/wtv.py>

⁴ <https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html>

2.2. High-resolution spectroscopy

2.2.1. SOPHIE

Before TESS observations, HD 207897 had been monitored by the SOPHIE spectrograph (Perruchot et al. 2008; Bouchy et al. 2013) between 2012 and 2015. The star measurements were conducted as a part of the Recherche de Planètes Extrasolaires (RPE) subprogram dedicated to detecting Neptunes and Super-Earths orbiting nearby and bright solar-type stars (Bouchy et al. 2009; Courcol et al. 2015; Hara et al. 2020). The observations were performed using high-resolution mode (resolution power of $\lambda/\Delta\lambda \approx 75\,000$) with a simultaneous thorium-argon (Th-Ar) calibration lamp, allowing us to monitor the instrumental drift. We collected 44 high-resolution spectra with an RV root mean square (RMS) of 6.9 m s^{-1} (see Sect. 3.3.1 for more details about the derivation of the RVs). A first analysis of the data in 2015 showed a periodicity close to 16.3 d. However, our preliminary analysis based on an erroneous estimate of the rotational period of the star caused us to attribute the 16.3 d signal to stellar spot modulation at $P_{\text{rot}}/2$, and observations of the star were stopped. After the release of the TESS data, we resumed SOPHIE observations to gather more data to better separate the activity signal and planetary RV variation, and also to search for other possible planets. The star was observed again in 2020 with a simultaneous Fabry-Perot (FP) reference spectrum. Twenty-four additional high-resolution spectra with an RV RMS of 4.43 m s^{-1} were gathered.

Our final SOPHIE dataset includes 68 spectra. With an exposure time ranging from 1000 to 1500 s, we achieved a median S/N of 97.7 per pixel at 550 nm. The mean RV uncertainty, computed using the quadratic sum of photon noise and wavelength calibration error, is 1.7 m s^{-1} . The RV data set are presented in Table A.1 after the corrections described in Sect. 3.3.1 were applied.

2.2.2. HIRES

From the summit of Maunakea, we first observed HD 207897 using the Keck I telescope and HIRES spectrometer (Vogt 1994) from 2003 July 7 as part of the California Planet Search, and additional RVs were collected for 11 months beginning on 2020 January 21. Thirty-seven RVs were collected using the B5 decker ($0.87'' \times 5.0''$), resulting in a resolution of 50 000. The median exposure time was 231 s, the average S/N per pixel was 220, and the internal uncertainty was 1.04 m s^{-1} . The Doppler pipeline and observing setup follows the California Planet Search procedure outlined in Howard et al. (2010). The RV data set are presented in Table A.2. We note that data before 2004 July 9 (three data points) used a different CCD detector. For the sake of simplicity, we only used them in our RV periodogram by applying an offset term (Fig. 1).

2.2.3. APF

From the summit of Mt. Hamilton at the Lick Observatory, we collected 23 RVs of HD 207897 using the Levy Spectrograph (Burt et al. 2014) on the APF from 2020 June 2 until 2021 February 28. Twenty-three RVs were collected. Their median exposure time was 1200s, the S/N per pixel was 86, and the internal uncertainty was 3.0 m s^{-1} . With a resolution of 100 000, this slit-fed spectrograph uses the iodine-cell technique to calculate RVs according to Butler et al. (1996). For additional details on the instrumental setup and data reduction, see Fulton et al. (2015). The full list of corrected RVs can be found in Table A.3.

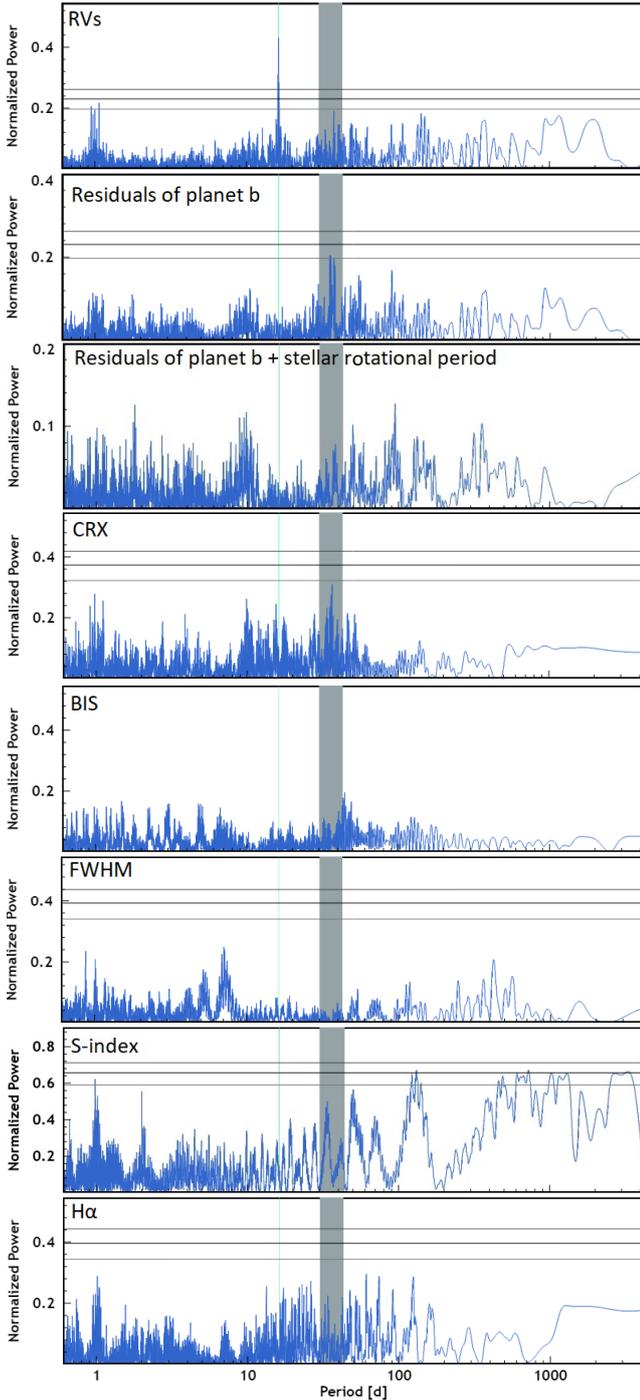


Fig. 1. Normalized periodograms of RVs and activity indexes for HD 207897. *From top to bottom:* RVs, residual of RVs after a Keplerian fit on 16.20 d, residuals of the fit on the planet and stellar rotation signal, CRX, bisector, FWHM, and $H\alpha$ index. The vertical cyan line marks the position of the highest peak in the RV periodogram at 16.20 d and shows no corresponding peak in the stellar activity periodograms. Horizontal lines indicate the 0.1, 1, and 10% FAP level from top to bottom. The vertical gray strip highlights the position of the rotational stellar period as estimated in Sect. 3.3.2.

2.2.4. FIES

From the Roque de los Muchachos Observatory in La Palma, Spain, we observed with the FIBre-fed Échelle Spectrograph (FIES; [Telting et al. 2014](#)) at the 2.56 m Nordic Optical Telescope and obtained four spectra between 2019 December 27 and

2020 January 15. We used the high-resolution fibre ($R \sim 67\,000$) and extracted the spectra following [Buchhave et al. \(2010\)](#). The S/N per resolution element at 550 nm ranges between 62 and 127. We used the FIES data for stellar classification (see Sect. 3.1) and refrained from including them in the RV analysis because all the observations were acquired near phases 0.25 and 0.75.

2.2.5. TRES

From the Fred Lawrence Whipple Observatory (FLWO) at the top of Mt. Hopkins, Arizona, USA, we observed with the 1.5 m Tillinghast Reflector Telescope using the Tillinghast Reflector Echelle Spectrograph (TRES; [Fűrész 2008](#); [Mink 2011](#)) to obtain one spectrum ($S/N \sim 36$) on UT 31 December 2019. TRES is a fiber-fed optical spectrograph with a resolving power of $R \sim 44\,000$. The spectrum was extracted following the procedures outlined in [Buchhave et al. \(2010\)](#) and was used to derive stellar parameters as described in Sect. 3.1.

2.3. High spatial resolution imaging

We used the AstraLux high spatial resolution camera ([Hormuth et al. 2008](#)) at the 2.2 m telescope of the Calar Alto Observatory (CAHA, Almeria, Spain) to unveil possible close companions to HD 207897. This instrument applies the lucky-imaging technique to retrieve diffraction-limited images from point source objects by acquiring thousands of short-exposure frames that freeze atmospheric variations and thus produce diffraction-limited images. We observed this target on 2020 February 26 under good weather conditions and a mean seeing of 0.9 arcsec. The brightness of this target allowed us to obtain 11 000 frames with an individual exposure time of 10 milliseconds. We used the Sloan Digital Sky Survey z filter (SDSSz), which is best suited for obtaining the highest possible resolution with AstraLux. We also used a field of view windowed to 6×6 arcsec. All frames were reduced by the instrument pipeline ([Hormuth et al. 2008](#)), which performs the basic reduction (bias subtraction and flat field correction), aligns all frames, and measures the Strehl ratio ([Strehl 1902](#)) of the individual images. This metric was then used to select the best images to perform the final stacking. We used the best 10% of the frames to produce the final high-resolution image. The final image does not show signs of close companions. We performed a dedicated search by removing the instrumental point spread function scaled to the target peak flux. No additional sources were present in the image. We then followed the procedures described in [Lillo-Box et al. \(2012, 2014\)](#) by using our own developed astrasens package⁵ to obtain the sensitivity of our image by performing an injection or retrieval of artificial sources. The sensitivity curve is shown in Fig. 2.

By using this sensitivity curve, we additionally estimated the probability of an undetected blended source in our high-spatial resolution image (BSC, see the procedure described in [Lillo-Box et al. 2014](#)). We use a python implementation of this approach (bsc, by J. Lillo-Box), which uses the TRILEGAL⁶ galactic model (v1.6 [Girardi et al. 2012](#)) to retrieve a simulated source population of the region around the corresponding target⁷. This simulated population is used to compute the density of stars around the target position (radius $r = 1^\circ$) and to derive the probability of chance alignment at a given contrast

⁵ <https://github.com/jlillo/astrasens>

⁶ <http://stev.oapd.inaf.it/cgi-bin/trilegal>

⁷ This is done in python by using the [Bhatti et al. \(2020\)](#) implementation by [Bhatti et al. \(2020\)](#).

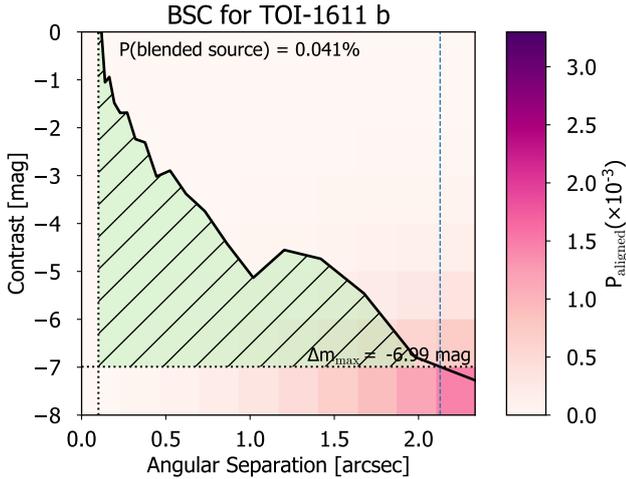


Fig. 2. Blended source confidence curve from the AstraLux SDSSz image (solid black line). The color in each angular separation and contrast bin represents the probability of a chance-aligned source with these properties at the location of the target, based on the TRILEGAL model (see Sect. 2.3). The maximum contrast of a blended binary capable of mimicking the planet transit depth is shown as a dotted horizontal line. The green shaded region represents the unexplored regime in the high-spatial resolution image. The BSC corresponds to the integration of P_{aligned} over this shaded region.

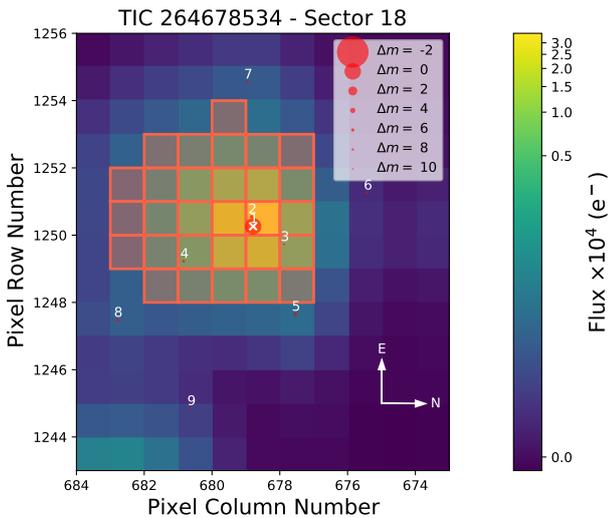


Fig. 3. Target pixel file of TESS for HD 207897 corresponding to sector 18, created with `tpfplotter` (Aller et al. 2020). The red square is related to the SPOC aperture mask, and red circles indicate the *Gaia* DR2 magnitude of nearby stars with the size depending on their brightness. HD 207897, number 1, has a *Gaia*(mag) = 8.1, while the nearest star, TIC 264678535, number 2, has a *Gaia*(mag) = 13.8.

magnitude and separation. In this case, given the AstraLux high-resolution image, we obtained a probability of an undetected blended source of 0.041%. We consider this probability low and the odds that such an undetected source might be an appropriate binary capable of mimicking the transit signal to be negligible. In addition, by using `tpfplotter` code⁸ (Aller et al. 2020), we plotted the target pixel file from TESS for HD 207897 b (see Fig. 3). We searched for possible light contamination by considering the nearby stars TIC 264678535 ($T_{\text{mag}} = 13.3$), TIC 264678538 ($T_{\text{mag}} = 15.69$) and TIC 264678529 ($T_{\text{mag}} = 14.39$),

⁸ <https://github.com/jlillo/tpfplotter>

which all lie inside the selected apertures analyzed by PDC-SAP. Because the *Gaia* *G* bandpass is quite similar to the TESS bandpass, we used the *Gaia* fluxes of these stars and estimated the level of contamination. The total flux due to nearby stars was only 0.8% of the HD 207897 flux, which is automatically corrected by SPOC. Because the probability of undetected blended sources from the AstraLux high-resolution image is low, the amount of light contamination due to nearby stars is negligible, and because similar Keplerian amplitudes were derived with different masks (see Sect. 3.3.1), we conclude that the transit and RV variations originate from the main target within the TESS TPF (HD 207897) and are of planetary origin.

3. Analysis and results

3.1. Stellar parameters

To obtain the stellar atmospheric parameters, we first summed the 66 spectra of SOPHIE after correcting for RV variations of the star, for barycentric Earth radial velocity, and correcting for the background of the calibration lamps (Hobson 2019). This resulted in a high S/N per pixel spectrum of 772 at 550 nm.

The T_{eff} and [Fe/H] were calculated following the procedure described in Santos et al. (2013) and Sousa et al. (2018). The procedure is based on the equivalent width of the Fe I and Fe II lines and assumes a balance in excitation and ionization of iron lines in local thermal equilibrium. The abundances of Mg and Si were derived by closely following the curve-of-growth analysis methods described in our previous works (Adibekyan et al. 2012, 2015). The abundances of C and O are very difficult to determine for stars cooler than about 5200 K (Delgado Mena et al. 2010; Bertran de Lis et al. 2015). We estimated the abundances of these elements empirically by using a machine-learning algorithm (we used the estimator `RandomForestRegressor`) from the Python Scikit-learn package (Pedregosa et al. 2011). The estimation of C and O was based on the abundance of Mg and Fe (Delgado Mena et al. 2017). Our initial sample was based on the HARPS sample, for which abundances of Mg and Fe are available (Adibekyan et al. 2012). Then we derived the O abundance for 535 stars and the C abundance for 758 stars following the method described in our previous works (Delgado Mena et al. 2010; Bertran de Lis et al. 2015). These samples were used as training and test datasets. The resulting abundances and other stellar parameters of HD 207897 are presented in Table 1.

As an independent stellar classification analysis, we used the FIES and TRES data following (SPC; Buchhave et al. 2012, 2014). The FIES analysis uses five spectral orders spanning a wavelength range from 5065 to 5320 Å, and the TRES analysis uses three spectral orders spanning the range from 5060 to 5300 Å. The spectra were compared to a library of synthetic templates to measure the effective temperature T_{eff} , surface gravity $\log g$, projected rotational velocity $v \sin i$, and metallicity [m/H] (a solar mix of metals). We analyzed the spectra individually and calculated a weighted average of each parameter. The results were $T_{\text{eff}} = 5085 \pm 50\text{K}$, $\log g = 4.48 \pm 0.10$, $v \sin i < 2\text{ km s}^{-1}$, and [m/H] = -0.23 ± 0.08 , which agrees with the results of the other methods.

We also performed an analysis of the broadband spectral energy distribution (SED) of the star together with the *Gaia* EDR3 parallax (with no systematic offset applied; see, e.g., Stassun & Torres 2021), in order to determine an empirical measurement of the stellar radius, following the procedures described in Stassun & Torres (2016); Stassun et al. (2017a,b). We pulled the $B_T V_T$ magnitudes from *Tycho-2*, the *JHK_S*

Table 1. Stellar properties of HD 207897.

Other identifiers		
	TIC 264678534	
	HD 207897	
	BD +83 617	
	HIP 107038	
	<i>Gaia</i> DR2 2300641567596591488	
	2MASS J21404490+8420005	
Parameter	HD 207897	References
Astrometric properties		
Parallax (mas)	35.3446 ± 0.0468	<i>Gaia</i> DR2 ^(*)
	35.3581 ± 0.0159	<i>Gaia</i> EDR3
Distance	28.25 ± 0.03	<i>Gaia</i> DR2
α (hms)	21 : 40 : 44.78	<i>Gaia</i> DR2
δ (dms)	+84 : 20 : 00.56	<i>Gaia</i> DR2
Photometric properties		
<i>B</i> – <i>V</i>	0.86 ± 0.02	HIP
<i>V</i> (mag)	8.37 ± 0.0015	HIP
<i>Gaia</i> (mag)	8.1304 ± 0.0004	<i>Gaia</i> DR2
<i>Gaia</i> _{BP} (mag)	8.6051 ± 0.0025	<i>Gaia</i> DR2
<i>Gaia</i> _{RP} (mag)	7.5402 ± 0.0014	<i>Gaia</i> DR2
TESS(mag)	7.58 ± 0.006	TESS
<i>J</i> (mag)	6.830 ± 0.023	2MASS
<i>H</i> (mag)	6.391 ± 0.034	2MASS
<i>K</i> _s (mag)	6.312 ± 0.026	2MASS
<i>W</i> ₁ (mag)	6.262 ± 0.088	WISE
<i>W</i> ₂ (mag)	6.219 ± 0.025	WISE
<i>W</i> ₃ (mag)	6.275 ± 0.015	WISE
<i>W</i> ₄ (mag)	6.233 ± 0.045	WISE
Spectroscopic properties		
Spectral type	K0V	HIP
ξ_t (kms ⁻¹)	0.53 ± 0.10	Sect. 3.1
$\log(R'_{\text{HK}})$	-4.83 ± 0.10	Sect. 3.3.2
$v \sin i$ (km s ⁻¹)	< 2	Sect. 3.3.2
[Fe/H] dex	-0.21 ± 0.02	Sect. 3.1
[C/H] dex	-0.23 ± 0.07	Sect. 3.3.2
[O/H] dex	-0.11 ± 0.08	Sect. 3.3.2
[Mg/H] dex	-0.17 ± 0.06	Sect. 3.3.2
[Si/H] dex	-0.19 ± 0.05	Sect. 3.3.2
Bulk properties		
Mass (M_{sun})	$0.80^{+0.036}_{-0.030}$	Sect. 3.4
	0.84 ± 0.05	Sect. 3.1
Radius(R_{sun})	$0.779^{0.019}_{-0.018}$	Sect. 3.4
	0.785 ± 0.014	Sect. 3.1
$\log g$ (cgs)	$4.559^{+0.026}_{-0.025}$	Sect. 3.4
$L_s(L_{\text{sun}})$	$0.360^{+0.019}_{-0.014}$	Sect. 3.4
T_{eff} (K)	5070^{+60}_{-57}	Sect. 3.4
P_{rot} (days)	37 ± 7	Sect. 3.3.2

Notes. ^(*)We applied the offset correction as prescribed in [Lindgren et al. \(2018\)](#).

magnitudes from 2MASS, the *W*₁–*W*₄ magnitudes from WISE, the $G_{\text{BP}}G_{\text{RP}}$ magnitudes from *Gaia*, and the NUV magnitude from GALEX. Together, the available photometry spans the full stellar SED over the wavelength range 0.2–22 μm (see Fig. 4).

We performed a fit using Kurucz stellar atmosphere models, with the effective temperature (T_{eff}), metallicity ([Fe/H]), and

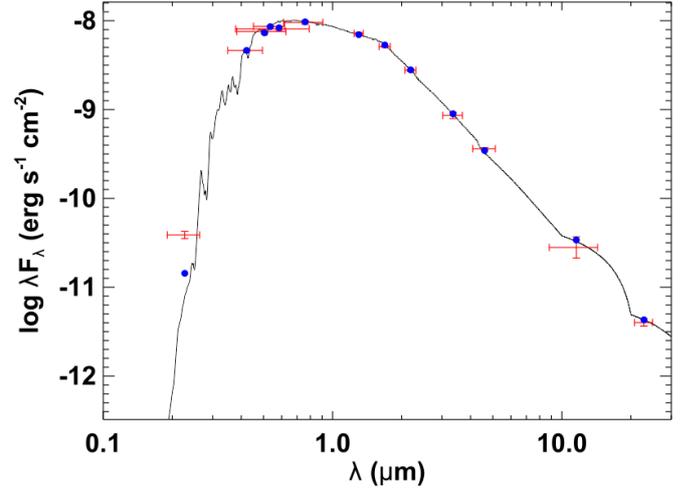


Fig. 4. Spectral energy distribution of HD 207897. Red symbols represent the observed photometric measurements, where the horizontal bars represent the effective width of the passband. Blue symbols are the model fluxes from the best-fit Kurucz atmosphere model (black).

surface gravity ($\log g$) adopted from the spectroscopic analysis. We also fixed the extinction $A_V \equiv 0$ based on the proximity of the stars (see Table 1). The resulting fit (Fig. 4) has a reduced χ^2 of 1.1, excluding the GALEX NUV flux, which indicates a moderate level of activity. When the (unreddened) model SED is integrated, the bolometric flux at Earth is $F_{\text{bol}} = 1.401 \pm 0.016 \times 10^{-8} \text{ erg s}^{-1} \text{ cm}^{-2}$. The F_{bol} and T_{eff} together with the *Gaia* parallax results in the stellar radius, $R_{\star} = 0.785 \pm 0.014 R_{\odot}$. In addition, we can estimate the stellar mass from the empirical relations of [Torres et al. \(2010\)](#), giving $M_{\star} = 0.84 \pm 0.05 M_{\odot}$. Additionally, we cross-checked these values with a different method by EXOFASTv2 in Sect 3.4.

3.2. Photometry data analysis

The photometry was extracted with the pre-search data-conditioned simple aperture photometry (PDC-SAP) pipeline ([Stumpe et al. 2012](#); [Stumpe et al. 2014](#); [Smith et al. 2012](#)) provided by the TESS team. We removed nans and flagged low-quality data and 5σ outliers. We then normalized and detrended the light curve with a spline-robust iterative sigma-clipping method ([Schoenberg 1946](#)) using the Wotan package ([Hippke et al. 2019](#))⁹. This method detrends the light curve by fitting the spline through minimizing the sum of squared residuals along with iteratively sigma-clipping. This step is one of the most important steps in the photometric analysis as it reduces the number of false-positive signals by removing instrumental and stellar noise. On the other hand, these detrending algorithms always include the risk of changing the transit depth or even fully remove shallow transits. For this reason, we tested different values of knots and chose the value of 0.7, which was the highest value that appeared suitable for removing the light-curve variabilities. The resulting light curve is shown in Fig. 5, and the transit events are marked with red triangles.

To search for any periodic signals in the data, we used the transit least-squares (TLS) algorithm ([Hippke & Heller 2019](#)), a method for investigating planetary transits that takes the stellar limb-darkening ([Mandel & Agol 2002](#)) and the effects of

⁹ <https://github.com/hippke/wotan>

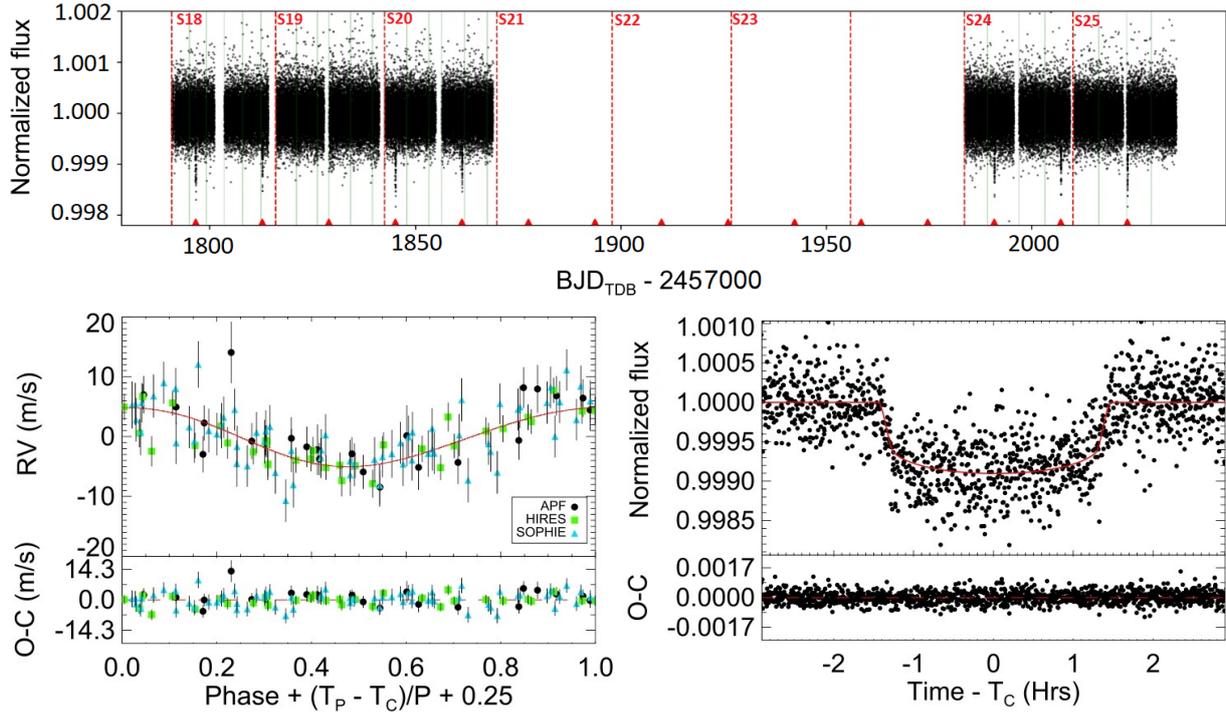


Fig. 5. TESS light curves and EXOFASTv2 best-fit models for RVs and photometric data. *Top panel:* full PDC-SAP (two-minute) TESS light curve after detrending, taken from sectors 18, 19, 20, 25, and 26. The red triangles indicate the transit events, and the green vertical lines indicate the momentum dumps of the spacecraft, which occur every 2.5 days. None of the transit events occurred during momentum dumps. *Bottom left:* phase-folded SOPHIE, APF, and HIRES RVs of HD 207897 b at the period of 16.20 d. *Bottom right:* TESS phase-folded light curve.

planetary ingress and egress into account. This method is publicly available¹⁰ and has been optimized for the signal detection efficiency (SDE) of small planets. We searched for periodic signals in the range 0.6–122 d and sampled 28607 periods within this range. The result showed a prominent periodic signal that occurred every 16.20 d with an S/N of 43, an SDE of 76.2 and a false-alarm probability (FAP) lower than 0.01% (see Fig. B.1). After masking the signal at 16.20 d, we ran the TLS again and did not see any periodic signals. A search for a long period beyond 122 d did not show any significant signals either.

3.3. Radial velocity data analysis

3.3.1. RV data reduction on SOPHIE

The SOPHIE data were reduced with the SOPHIE data reduction pipeline (DRS, Bouchy et al. 2009), which extracts the RV by cross-correlating spectrum with a binary mask and then fits a Gaussian of the cross-correlation function (CCF) (Pepe et al. 2002). We tested different masks, including G2, K0, and K5. Because they showed similar Keplerian amplitude variations, it is unlikely that these variations are produced by blend scenarios composed of stars of different spectral types (Bouchy et al. 2008). We finally adopted the RV data derived with the K5 mask because it presented a smaller RV dispersion.

We subsequently excluded six RV points that did not reach the required quality: three points with a lower S/N than required $S/N_{550} > 50$, two points with moonlight pollution, and one point that was an outlier and resulted from an observation performed without simultaneous calibration. We corrected the remaining points for the charge transfer inefficiency (CTI) effect (Santerne et al. 2012). This correction ranged from 1.4 to 6 m s⁻¹ with a

mean RV correction of 2.1 m s⁻¹. We then removed the nightly drift of the telescope measured from the simultaneous Th-Ar or FP reference spectra. This correction also falls within the 0.1 to 9 m s⁻¹ range, with a mean RV of 2.1 m s⁻¹.

The next step was to remove a long-term drift of the zero-point due to the instrumental effect identified in SOPHIE RV data (Courcol et al. 2015; Hobson et al. 2018). To track this offset, so-called ‘constant’ stars were monitored each night. We combined these observations to build a time series of the RV master constant (see Courcol et al. 2015 for more details of the method). We then estimated the long-term zeropoint drift as a function of time for each observation night and subtracted it from the HD 207897 RV data. This correction is in the 0.2–6 m s⁻¹ range with a mean value of 3.3 m s⁻¹. Because we saw the effects of the long-term drift of the zeropoints on the bisector span of constant stars, we applied the same processes to the bisector of HD 207897 as well. The correction for bisector was between 4 and 11 m s⁻¹ with a mean value of 2 m s⁻¹. After these corrections were applied, the HD 207897 RVs were reduced from an original RMS of 6.2 m s⁻¹ to a final RMS of 4.9 m s⁻¹.

3.3.2. Stellar rotation and activity

To investigate the activity of the star, we used several indicators such as bisector span, CCF FWHM, chromatic RV index (CRX), $H\alpha$, and $\log(R'_{HK})$ from SOPHIE spectra. We also used the S-index from the HIRES spectra.

We obtained the bisector span and CCF FWHM from the SOPHIE data reduction system. The CRX was extracted using the SERVAL (Zechmeister et al. 2018) code. To compute the $H\alpha$ index, which measures the flux in the $H\alpha$ line, we followed the definitions of Boisse et al. (2011). To do this, we also applied

¹⁰ <http://github.com/hippke/tls>

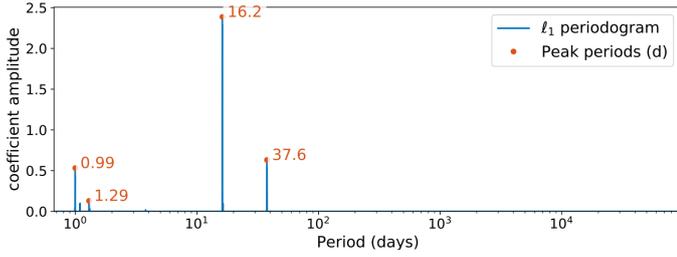


Fig. 6. ℓ_1 periodogram of SOPHIE RVs for HD 207897 b following Hara et al. (2017). The main peaks of the periodogram are indicated with orange circles.

background correction to the data as appropriate for the two different calibration lamps (for further details, see Hobson 2019).

We derived the S-index from HIRES spectra following Baliunas et al. (1995) and Paulson et al. (2002). We also computed the $\log(R'_{\text{HK}})$ index on SOPHIE spectra following the method of Noyes et al. (1984), but the S/N is low in all first (bluest) orders, where Ca H&K lines are located, with one exception in 2020 ($S/N = 33$). We decided to sum the bluest order of spectra with $S/N > 20$ to reach better S/Ns and then calculated $\log(R'_{\text{HK}})$. We first did this for the 2012–2015 spectra, of which 22 had an $S/N > 20$, giving a value of $\log(R'_{\text{HK}}) = -4.83 \pm 0.10$. Only one 2020 spectrum presents an $S/N > 20$, so that no summation was possible. This spectrum yields a value of $\log(R'_{\text{HK}}) = -4.78 \pm 0.10$, indicating that the star may have a similar level of activity as during the 2012–2015 period. Our value of $\log(R'_{\text{HK}})$ for 2012–2015 is consistent with $\log(R'_{\text{HK}}) = -4.86$ reported by Brewer et al. (2016). This value indicates a modest activity level of HD 207897.

To confirm our $\log(R'_{\text{HK}})$ value, we used the GALEX near-ultraviolet (NUV) flux excess of the star (Fig. 4). The observed NUV excess implies a chromospheric activity of $\log R'_{\text{HK}} = -4.82 \pm 0.05$ through the empirical relations of Findeisen et al. (2011), consistent with the value obtained spectroscopically. Moreover, the NUV-estimated activity implies an age of $\tau_{\star} = 4.4 \pm 0.9$ Gyr according to the empirical relations of Mamajek & Hillenbrand (2008).

We estimated a stellar rotation period of 37 ± 7 d with the method proposed in Noyes et al. (1984), which agrees well with an estimated period of 36^{+5}_{-4} d following Mamajek & Hillenbrand (2008). These estimates are consistent with the value of 38 days from Isaacson & Fischer (2010).

We sought to constrain the P_{rot} of the star with RVs and activity indicators. We subtracted a linear drift in the CCF FWHM, CRX, and RVs time series, and a cubic drift in the S-index. Moreover, an offset term was fit on $H\alpha$ data between the two different calibration lamp background corrections (before and after BJD = 57284.4185). Then, we computed the periodogram of RVs, the RV-residuals of Keplerian models, and the activity indicators (Fig. 1). The RV residuals of planet b show two peaks at 35.9 and 37.6 days near the 10% FAP.

We ran an ℓ_1 periodogram on the RVs (see Fig. 6) for comparison. This was obtained with the same procedure as in Hara et al. (2020): we considered alternative noise models characterized by an autocovariance function that is a sum of white, correlated, and quasi-periodic components with different amplitudes and timescales and ranked them with cross-validation. The correlated component is a Gaussian kernel characterized by its timescale and amplitude, and the quasi-periodic component is a Gaussian kernel as in Haywood et al. (2014), characterized by its decay timescale, period, and amplitude. The

amplitudes of the white, correlated, and quasi-periodic terms were taken on a grid (0 to 3 m s^{-1} with a 0.5 m s^{-1} step), the timescales 0, 3, and 6 days for the correlated term, 30, 60, and 90 days for the quasi-periodic term, and 37.6 days for the period of the quasi-periodic component. We also included one offset per instrument in the base model. Figure 6 corresponds to the highest-ranked noise model. We found a significant signal at the planet period (FAP of 2×10^{-9}) and found a signal at 37.6 d with an FAP of 0.5. The signal at 37.6 d very likely corresponds to a stellar signal.

While the periodogram of the CCF FWHM and $H\alpha$ activity indicators does not exhibit any significant signals, CRX shows a peak near to 10% FAP at 36.5. The S-index periodogram shows several peaks at the long periods, but no correlation ($R = 0.16$) is found between the S-index and HIRES RVs. We note that given the relatively low value of $v \sin i$ ($2 \pm 1 \text{ km s}^{-1}$), we did not detect any correlations between the RVs and its residuals and bisector.

We searched the SAP and PDC-SAP light curves for a signal of the rational period of the star in the photometry. For this, we applied the systematics-insensitive periodogram (SIP) method¹¹ (Angus et al. 2016; Hedges et al. 2020) for SAP and the Gaussian process (GP) model on PDC-SAP light curves. The SIP method detrends the SAP light curve (see Fig. B.2 bottom) from TESS instrument systematics and also calculates the Lomb-Scargle periodogram and without requiring a predetrending of the light curves as other methods (e.g., the autocorrelation function, ACF, by McQuillan et al. 2013). This method was initially used for the *Kepler* mission and has recently been successfully applied for TESS data such as TOI-1259A (Martin et al. 2021) and TOI-700 (Hedges et al. 2020). The SIP periodogram (see Fig. B.2 top) does not exhibit any significant signal for HD 207897 b. We also applied a GP model on PDC-SAP light curves, but it did not display any convincing signal either.

To examine the origin of the 37.6 d signal further, following Hara et al. (2022), we investigated the phase and amplitude consistency of the 37.6-day signal in the SOPHIE RVs. We used the statistic defined in Eq. (14) of Hara et al. (2022). We fit an offset and a linear trend as well as a sinusoidal model for the 16.2-day planet. Adopting a value of the RV jitter of 3.16 m s^{-1} , we find that the hypothesis that the phase and amplitude of the 37.6-day signal is constant is rejected at 2σ . This further supports the hypothesis that this signal is due to activity.

Based on the estimated stellar rotational period value and the signal at the CRX activity indicator, it is likely that the signals at 35.9 and 37.6 days are due to the stellar rotation period. We take these peaks in our joint model analysis with EXOFASTv2 in Sect. 3.4 into account.

3.3.3. Radial velocity results

The combined RV data of HD 207897 are plotted in Fig. 7 after fitting a zeropoint offset for each RV dataset (see Table 2). We removed a linear trend of $-0.28 \pm 0.11 \text{ m s}^{-1} \text{ yr}^{-1}$ from RVs. We note that no notable differences were found in our results when this linear drift was included or excluded. However, because we saw a clear linear drift in the activity indicators such as the CCF FWHM and CRX, we decided to keep the drift.

We investigated the footprints of the planet in RVs by searching for periodic signals. To do this, we used the website of the Data and Analysis Center for Exoplanets (DACE, Delisle et al. 2016)¹² and computed periodograms for the RVs (Fig. 1).

¹¹ <https://github.com/christinahedges/TESS-SIP>

¹² Available at <https://dace.unige.ch>

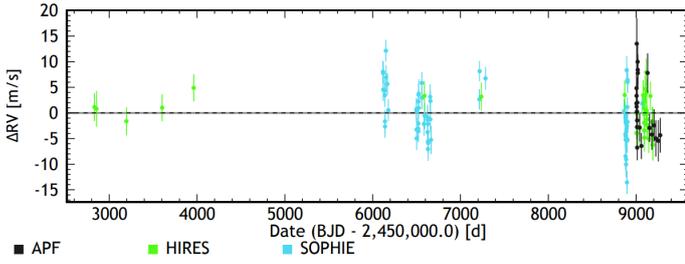


Fig. 7. Radial velocities of HD 207897 from SOPHIE, APF, and HIRES before correction for the linear trend.

The RV periodogram displays a clear peak at 16.20 d with a value lower than the 0.1% FAP (Baluev 2008). Moreover, the periodogram of activity indicators shows no corresponding peak. This shows that this periodic signal is likely due to a planet and not to stellar activity. See Fig. 1.

Figure 6 shows that the ℓ_1 periodogram confirms that the most prominent signal appears at 16.20 days with an FAP of 2×10^{-9} . The value of the second-highest peak at 37.7 days agrees with the estimated rotational period of the star (see Sect. 3.3.2).

3.4. Joint modeling of RV and photometry

To explore all the parameters of the system, including the host star and the planet, we simultaneously and self-consistently modeled the photometric observations of the star from five sectors of TESS and the RV observations from SOPHIE, HIRES, and APF using the fast exoplanetary fitting package (EXOFASTv2¹³, Eastman et al. 2013, 2019; Eastman 2017). This global modeling software uses a differential evolution Markov chain coupled with a Metropolis-Hastings Monte Carlo sampler that uses error scaling to explore the system parameters.

EXOFASTv2 fits a total of 33 free parameters for the HD 207897 system, which can be divided into the following categories:

- Twelve parameters related to the planet and activity period: the mid-transit time T_C , planet orbital period P , the ratio of the planet to star radius R_p/R_{star} (only for the planet), the orbital inclination i (only for the planet), the RV semi-amplitude K , and two more free parameters related to the eccentricity e .

- Two parameters for each RV instrument: instrumental offset and jitters. One free parameter is also fit for drift on RVs.

- Two limb-darkening coefficients for TESS photometric bandpasses, along with baseline flux and variance are fit for the transit light curve.

- Eleven stellar parameters: Stellar mass M_* , stellar radius R_* , and effective temperature T_{eff} by the MIST model, stellar radius $R_{*,\text{SED}}$ and effective temperature $T_{\text{eff,SED}}$ by the SED model, observed metallicity $[\text{Fe}/\text{H}]$, theoretical metallicity at the star's birth $[\text{Fe}/\text{H}]_0$ by the MIST model, age, equivalent evolutionary point EPP, V -band extinction A_v , and distance d .

Before running EXOFASTv2, we set a Gaussian prior on T_{eff} and $[\text{Fe}/\text{H}]$ from our spectral analysis results, presented in Sect. 3.1. We also imposed a prior on the *Gaia* DR2 parallax after applying the offset correction as described in Lindgren et al. (2018). The broadband photometry presented in Table 1 is

¹³ EXOFASTv2 is available at <https://github.com/jdeast/EXOFASTv2>

also included. We enforced the upper limit for the V -band extinction (A_v) from Schlegel et al. (1998) and Schlafly & Finkbeiner (2011). The broadband photometry, the *Gaia* parallax, and A_v allowed us to model the stellar spectral energy distribution, which is key to constraining the stellar radius. EXOFASTv2 interpolates a precompiled 4D grid of bolometric corrections ($\log g$, T_{eff} , $[\text{Fe}/\text{H}]$, and reddening) to directly compute the broadband photometry flux. We did not set the limb darkening to let EXOFASTv2 constrain the best parameters of a quadratic limb darkening through the Claret (2017) tables for the TESS bands and the stellar atmosphere parameters (T_{eff} , $[\text{Fe}/\text{H}]$ index, and $\log g_*$). We used the Mesa isochrones and stellar tracks evolutionary model (MIST, Dotter 2016; Choi et al. 2016) to derive the full stellar parameters by combining our data. We also allowed the fitting of a linear slope on the original RVs as part of our joint analysis.

EXOFASTv2 considers the chains to be well mixed when the Gelman-Rubin statistic (Gelman et al. 2004, 1992; Ford 2006) decreases below 1.01. The Gelman-Rubin statistic describes how similar the chains are; a value under 1.01 shows that the chains are well mixed. The list of our priors and final median values of the posterior distributions together with their 1σ confidence intervals of the full system parameters are reported in Table 2.

As we showed in Sect. 3.3.2, HD 207897 has a moderate stellar activity and the signals at 35.9 and 37.6 d are likely to be due to the stellar rotation periods. Because these signals affect the mass estimate of the planet slightly, we considered them as an additional Keplerian fit in our global analysis. A more ideal solution to take activity into account is detrending the RVs using activity indicators (e.g., the S-index or the FWHM), but this is currently not possible with EXOFASTv2. Furthermore, we did not have the same activity indicators for all RVs for a detrending. As an independent analysis, we used only SOPHIE data and trained GP against CRX. The results agreed completely with the EXOFASTv2 global modeling. When we used two Keplerian models in EXOFASTv2, we did not fix the period and let EXOFASTv2 find the best activity period between the 35.9 and 37.6 d signals. After EXOFASTv2 converged, we saw a bimodality in the posterior distribution for the stellar activity periods (see Fig. 8, top). We therefore present the final median posterior distribution values of the two most probable solutions in Table 2. We also report their calculated probabilities based on the area of the posterior distributions. The probability of the most likely values is 54%, and that of the less likely values is 46%.

HD 207897 is a main-sequence K0 dwarf star. We found that its most likely values for mass is $0.800^{+0.036}_{-0.030} M_{\odot}$, for the radius, this value is $0.779^{+0.019}_{-0.018} R_{\odot}$, and for T_{eff} it is 5070^{+60}_{-57} K. These values agree well with the result of our stellar analysis in Sect. 3.1 and also with the stellar parameters from *Gaia* Data Release 2, such as $T_{\text{eff}} = 5052^{+114}_{-78}$ K and a radius of $0.79^{+0.80}_{-0.75} R_{\odot}$ (Brown et al. 2018).

We also show that the planet has a period of 16.20 d, a radius of $2.5 \pm 0.08 R_{\text{E}}$ and a mass of either $14.4 \pm 1.6 M_{\text{E}}$ or $15.9 \pm 1.6 M_{\text{E}}$ with nearly equal probability. These two solutions correspond to two possibilities for the stellar activity period. Hence, the density is either $5.1 \pm 0.7 \text{ g cm}^{-3}$ or $5.5^{+0.8}_{-0.7} \text{ g cm}^{-3}$. The two values agree with each other inside their error bars.

The Keplerian solution and transit model for HD 207897 b based on the most likely values are shown in Fig. 5 (bottom). The Keplerian model based on the most likely activity period of 37.6 d is illustrated in Fig. 8.

Table 2. Median values and 68% confidence interval for HD 207897 b and its host star.

Parameter	Units	Model priors	Activity on 35.9 d (Prob. = 46%)	Activity on 37.6 d (Prob. = 54%)
Stellar parameters:				
M_*	Mass (M_\odot)	–	$0.801^{+0.036}_{-0.031}$	$0.800^{+0.036}_{-0.030}$
R_*	Radius (R_\odot)	–	$0.779^{+0.019}_{-0.018}$	$0.779^{+0.019}_{-0.018}$
$R_{*,\text{SED}}$	Radius ⁽¹⁾ (R_\odot)	–	$0.7691^{+0.011}_{-0.0098}$	$0.7692^{+0.011}_{-0.010}$
L_*	Luminosity (L_\odot)	–	$0.360^{+0.019}_{-0.014}$	$0.360^{+0.019}_{-0.014}$
F_{Bol}	Bolometric Flux (cgs)	–	$0.00000001440^{+0.00000000076}_{-0.00000000056}$	$0.00000001439^{+0.00000000077}_{-0.00000000055}$
ρ_*	Density (cgs)	–	$2.39^{+0.19}_{-0.18}$	$2.39^{+0.19}_{-0.18}$
$\log g$	Surface gravity (cgs)	–	$4.558^{+0.027}_{-0.025}$	$4.559^{+0.026}_{-0.025}$
T_{eff}	Effective Temperature (K)	$\mathcal{N}[5012,80]$	5071^{+59}_{-57}	5070^{+60}_{-57}
$T_{\text{eff,SED}}$	Effective Temperature ⁽¹⁾ (K)	–	5096^{+78}_{-54}	5096^{+78}_{-54}
[Fe/H]	Metallicity (dex)	$\mathcal{N}[-0.21, 0.08]$	$-0.040^{+0.037}_{-0.040}$	$-0.042^{+0.038}_{-0.041}$
[Fe/H] ₀	Initial Metallicity ⁽²⁾	–	$-0.028^{+0.050}_{-0.052}$	$-0.030^{+0.050}_{-0.053}$
Age	Age (Gyr)	–	$7.1^{+4.5}_{-4.4}$	7.1 ± 4.5
EEP	Equal Evolutionary Phase ⁽³⁾	–	343^{+17}_{-26}	343^{+17}_{-27}
A_V	V-band extinction (mag)	$\mathcal{U}[0,0.3534]$	$0.074^{+0.081}_{-0.052}$	$0.074^{+0.081}_{-0.052}$
σ_{SED}	SED photometry error scaling	–	$2.10^{+0.77}_{-0.47}$	$2.10^{+0.77}_{-0.48}$
ϖ	Parallax (mas)	$\mathcal{N}[35.345, 0.047]$	$35.345^{+0.048}_{-0.047}$	35.345 ± 0.047
d	Distance (pc)	–	28.292 ± 0.038	28.293 ± 0.038
$\dot{\gamma}$	RV slope ⁽⁴⁾ ($\text{m s}^{-1}/\text{day}$)	–	-0.00078 ± 0.00030	$-0.00077^{+0.00027}_{-0.00028}$
Planetary parameters:				
P	Period (days)	–	16.202157 ± 0.000085	$16.202159^{+0.000085}_{-0.000083}$
R_P	Radius (R_E)	–	$2.505^{+0.081}_{-0.077}$	$2.501^{+0.082}_{-0.078}$
M_P	Mass (M_E)	–	15.9 ± 1.6	14.4 ± 1.6
T_C	Time of conjunction ⁽⁵⁾ (BJD _{TDB})	–	$2458926.10942^{+0.00047}_{-0.00049}$	$2458926.10942^{+0.00047}_{-0.00048}$
T_T	Time of minimum projected separation ⁽⁶⁾ (BJD _{TDB})	–	$2458926.10944^{+0.00047}_{-0.00049}$	$2458926.10943^{+0.00047}_{-0.00048}$
T_0	Optimal conjunction Time ⁽⁷⁾ (BJD _{TDB})	–	$2458926.10942^{+0.00047}_{-0.00049}$	$2458926.10942^{+0.00047}_{-0.00048}$
a	Semi-major axis (AU)	–	$0.1164^{+0.0017}_{-0.0015}$	$0.1163^{+0.0017}_{-0.0015}$
i	Inclination (Degrees)	–	$88.755^{+0.067}_{-0.066}$	88.757 ± 0.067
e	Eccentricity	–	$0.047^{+0.053}_{-0.033}$	$0.047^{+0.057}_{-0.033}$
ω_*	Argument of Periastron (Degrees)	–	-125^{+68}_{-95}	-100 ± 110
T_{eq}	Equilibrium temperature ⁽⁸⁾ (K)	–	$632.3^{+8.2}_{-7.0}$	$632.2^{+8.1}_{-7.0}$
τ_{circ}	Tidal circularization timescale (Gyr)	–	38700^{+8900}_{-7900}	35300^{+8300}_{-7300}
K	RV semi-amplitude (m s^{-1})	–	4.65^{±0.45}	4.24^{±0.45}
R_P/R_*	Radius of planet in stellar radii	–	$0.02946^{+0.00048}_{-0.00044}$	$0.02943^{+0.00049}_{-0.00046}$
a/R_*	Semi-major axis in stellar radii	–	$32.12^{+0.85}_{-0.82}$	$32.13^{+0.85}_{-0.82}$
δ	Transit depth (fraction)	–	$0.000868^{+0.000029}_{-0.000026}$	$0.000866^{+0.000029}_{-0.000027}$
Depth	Flux decrement at mid transit	–	$0.000868^{+0.000029}_{-0.000026}$	$0.000866^{+0.000029}_{-0.000027}$
τ	Ingress/egress transit duration (days)	–	$0.00684^{+0.0010}_{-0.00072}$	$0.00678^{+0.0011}_{-0.00077}$
T_{14}	Total transit duration (days)	–	$0.1214^{+0.0014}_{-0.0013}$	$0.1213^{+0.0014}_{-0.0013}$
T_{FWHM}	FWHM transit duration (days)	–	$0.1145^{+0.0012}_{-0.0011}$	$0.1145^{+0.0012}_{-0.0011}$
b	Transit Impact parameter	–	$0.712^{+0.041}_{-0.039}$	$0.709^{+0.043}_{-0.042}$
b_S	Eclipse impact parameter	–	$0.687^{+0.035}_{-0.047}$	$0.689^{+0.036}_{-0.048}$
τ_S	Ingress/egress eclipse duration (days)	–	$0.00640^{+0.00066}_{-0.00069}$	$0.00644^{+0.00070}_{-0.00071}$
$T_{S,14}$	Total eclipse duration (days)	–	$0.1208^{+0.0016}_{-0.0019}$	$0.1209^{+0.0018}_{-0.0019}$
$T_{S,\text{FWHM}}$	FWHM eclipse duration (days)	–	0.1143 ± 0.0015	0.1143 ± 0.0016
$\delta_{S,2.5\mu\text{m}}$	Blackbody eclipse depth at 2.5 μm (ppm)	–	$0.205^{+0.026}_{-0.021}$	$0.204^{+0.026}_{-0.021}$
$\delta_{S,5.0\mu\text{m}}$	Blackbody eclipse depth at 5.0 μm (ppm)	–	$7.09^{+0.50}_{-0.45}$	$7.07^{+0.50}_{-0.45}$
$\delta_{S,7.5\mu\text{m}}$	Blackbody eclipse depth at 7.5 μm (ppm)	–	$20.2^{+1.1}_{-1.0}$	$20.2^{+1.1}_{-1.0}$
ρ_P	Density (cgs)	–	$5.52^{+0.82}_{-0.73}$	$5.05^{+0.77}_{-0.69}$
$\log g_P$	Surface gravity	–	$3.393^{+0.051}_{-0.053}$	$3.353^{+0.053}_{-0.056}$
Θ	Safronov Number	–	0.0647^{±0.0067}	0.0589^{±0.0066}
$\langle F \rangle$	Incident Flux ($10^9 \text{ erg s}^{-1} \text{ cm}^{-2}$)	–	$0.0361^{+0.0019}_{-0.0016}$	$0.0361^{+0.0019}_{-0.0016}$
T_P	Time of Periastron (BJD _{TDB})	–	$2458916.2^{+3.4}_{-4.3}$	$2458917.0^{+4.5}_{-5.4}$
T_S	Time of eclipse (BJD _{TDB})	–	$2458934.08^{+0.25}_{-0.40}$	$2458934.18^{+0.30}_{-0.38}$
T_A	Time of Ascending Node (BJD _{TDB})	–	$2458921.93^{+0.22}_{-0.42}$	$2458922.00^{+0.25}_{-0.42}$
T_D	Time of Descending Node (BJD _{TDB})	–	$2458930.16^{+0.31}_{-0.26}$	$2458930.18^{+0.36}_{-0.27}$
V_c/V_e	–	–	$1.010^{+0.067}_{-0.034}$	$1.007^{+0.069}_{-0.039}$
$e \cos \omega_*$	–	–	$-0.012^{+0.024}_{-0.039}$	$-0.003^{+0.029}_{-0.037}$
$e \sin \omega_*$	–	–	$-0.011^{+0.034}_{-0.065}$	$-0.007^{+0.039}_{-0.066}$
$M_P \sin i$	Minimum mass (M_E)	–	15.8 ± 1.6	14.4 ± 1.6
M_P/M_*	Mass ratio	–	0.0000594 ± 0.0000058	0.0000540 ± 0.0000058

Notes. Gaussian priors are presented by $\mathcal{N}(a, b)$, where a and b are the mean and width, respectively. Likewise, the uniform prior is denoted by $\mathcal{U}(c, d)$, and c and d present bounds on the parameter. The two solutions correspond to two possibilities for the stellar activity period. The highlighted parameters present a difference exceeding 0.5σ between the two solutions. See Table 3 in Eastman et al. (2019) for a detailed description of all parameters. ⁽¹⁾This value ignores the systematic error and is for reference only. ⁽²⁾The metallicity of the star at birth. ⁽³⁾Corresponds to static points in a star’s evolutionary history. See Sect. 2 in Dotter (2016). ⁽⁴⁾Reference epoch = 2456438.359500. ⁽⁵⁾Time of conjunction is commonly reported as the “transit time”. ⁽⁶⁾Time of minimum projected separation is a more correct “transit time”. ⁽⁷⁾Optimal time of conjunction minimizes the covariance between T_C and Period. ⁽⁸⁾Assumes no albedo and perfect redistribution.

Table 2. continued.

Parameter	Units	Model priors	Activity on 35.9 d (Prob. = 46%)	Activity on 37.6 d (Prob. = 54%)
d/R_*	Separation at mid transit	–	$32.6^{+2.1}_{-1.6}$	$32.5^{+2.2}_{-1.7}$
P_T	A priori non-grazing transit prob	–	$0.0298^{+0.0015}_{-0.0018}$	$0.0299^{+0.0017}_{-0.0019}$
$P_{T,G}$	A priori transit prob	–	$0.0316^{+0.0016}_{-0.0019}$	$0.0317^{+0.0018}_{-0.0020}$
P_S	A priori non-grazing eclipse prob	–	$0.0307^{+0.0021}_{-0.0015}$	$0.0306^{+0.0022}_{-0.0016}$
$P_{S,G}$	A priori eclipse prob	–	$0.0326^{+0.0023}_{-0.0016}$	$0.0325^{+0.0023}_{-0.0017}$
Stellar activity parameters:				
$P_{Activity}$	Period (days)	–	$35.923^{+0.069}_{-0.067}$	$37.627^{+0.55}_{-0.076}$
$T_{C,activity}$	Time of conjunction ⁽⁵⁾ (BJD _{TDB})	–	2458917.1 ± 2.2	$2458910.7^{+2.3}_{-2.6}$
$e_{Activity}$	Eccentricity	–	$0.14^{+0.13}_{-0.10}$	$0.15^{+0.13}_{-0.11}$
$\omega_{*,activity}$	Argument of Periastron (Degrees)	–	-127^{+98}_{-96}	33^{+94}_{-92}
$K_{Activity}$	RV semi-amplitude (m s ⁻¹)	–	$2.38^{+0.51}_{-0.52}$	$2.25^{+0.50}_{-0.49}$
Telescope parameters for SOPHIE:				
γ_{rel}	Relative RV Offset ⁽⁴⁾ (m s ⁻¹)	–	-6327.36 ± 0.56	-6327.53 ± 0.52
σ_J	RV Jitter (m s ⁻¹)	–	$3.16^{+0.42}_{-0.37}$	$3.05^{+0.41}_{-0.37}$
σ_J^2	RV Jitter Variance	–	$10.0^{+2.8}_{-2.2}$	$9.3^{+2.7}_{-2.1}$
Telescope parameters for APF:				
γ_{rel}	Relative RV Offset ⁽⁴⁾ (m s ⁻¹)	–	3.1 ± 1.1	$2.9^{+1.1}_{-1.0}$
σ_J	RV Jitter (m s ⁻¹)	–	$1.6^{+1.2}_{-1.6}$	$1.4^{+1.3}_{-1.4}$
σ_J^2	RV Jitter Variance	–	$2.7^{+5.5}_{-3.4}$	$2.0^{+5.4}_{-3.2}$
Telescope parameters for HIRES:				
γ_{rel}	Relative RV Offset ⁽⁴⁾ (m s ⁻¹)	–	$1.61^{+0.83}_{-0.86}$	$1.81^{+0.81}_{-0.76}$
σ_J	RV Jitter (m s ⁻¹)	–	$2.15^{+0.46}_{-0.38}$	$2.42^{+0.49}_{-0.40}$
σ_J^2	RV Jitter Variance	–	$4.6^{+2.2}_{-1.5}$	$5.9^{+2.6}_{-1.8}$
Wavelength parameters:				
			TESS	TESS
u_1	linear limb-darkening coeff	–	0.386 ± 0.046	$0.385^{+0.046}_{-0.047}$
u_2	quadratic limb-darkening coeff	–	0.188 ± 0.048	$0.189^{+0.047}_{-0.048}$
Transit parameters:			TESS UT 2019-10-10 (TESS)	TESS UT 2019-10-10 (TESS)
σ^2	Added Variance	–	$0.0000000244 \pm 0.0000000011$	$0.0000000244 \pm 0.0000000011$
F_0	Baseline flux	–	$0.9999796^{+0.0000022}_{-0.0000023}$	$0.9999796^{+0.0000022}_{-0.0000023}$

4. Internal structure

In order to characterize the internal structure of HD 207897 b, we performed a Markov chain Monte Carlo (MCMC) Bayesian analysis (Dorn et al. 2015) using the interior composition model introduced in Brugger et al. (2017), Mousis et al. (2020), and Acuña et al. (2021), which comprises three layers: a Fe-rich core, a silicate-rich mantle, and a water layer. With an input equilibrium temperature of 637 K, assuming an albedo of zero, the irradiance that HD 207897 b received is enough to present vapor and supercritical phases if water is found on its surface. Therefore we coupled an atmosphere-interior model that calculates the surface conditions and the contribution of the atmosphere to the total radius.

We considered two scenarios to obtain the interior structure of HD 207897 b: scenario 1, in which only the mass and radius of the planet are considered as inputs to the MCMC analysis (shown in Table 2), and scenario 2, in which the planetary mass and radius and the stellar Fe/Si and Mg/Si mole ratios (see Table 1) are the input data. To compute the Fe/Si and Mg/Si mole ratios with the stellar abundances, we followed the approach depicted in Brugger et al. (2017) and Sotin et al. (2007). We obtained Fe/Si = 0.74 ± 0.09 and Mg/Si = 1.11 ± 0.20 . The outputs of the MCMC analysis are the posterior distributions functions (PDF) of the core mass fraction (CMF), the water mass fraction (WMF), and the atmospheric parameters, which are the temperature at 300 bar, the planetary albedo, and the atmospheric thickness from transit pressure to 300 bar. We assume a

water-rich atmosphere. Table 3 shows the 1D, 1 σ confidence intervals of the MCMC output parameters.

In the most general case (scenario 1), up to 31% of the mass of HD 207897 b can be in the form of a hydrosphere. It reaches supercritical or superionic phases at its base (Mazevet et al. 2019) with the most likely value of the mass. This value increases slightly with the less likely value, although both cases are consistent with a water-rich planet with a water mass fraction of 20–30%. The 1 σ confidence interval limits the maximum CMF to 0.50 (see Fig. 9). In addition, a Fe-depleted planet (CMF=0) is possible but unlikely. Assuming a pure silicate interior, the WMF is found to be 2.2×10^{-5} in HD 207897 b, which corresponds to a pressure at the base of the hydrosphere of approximately 300 bar.

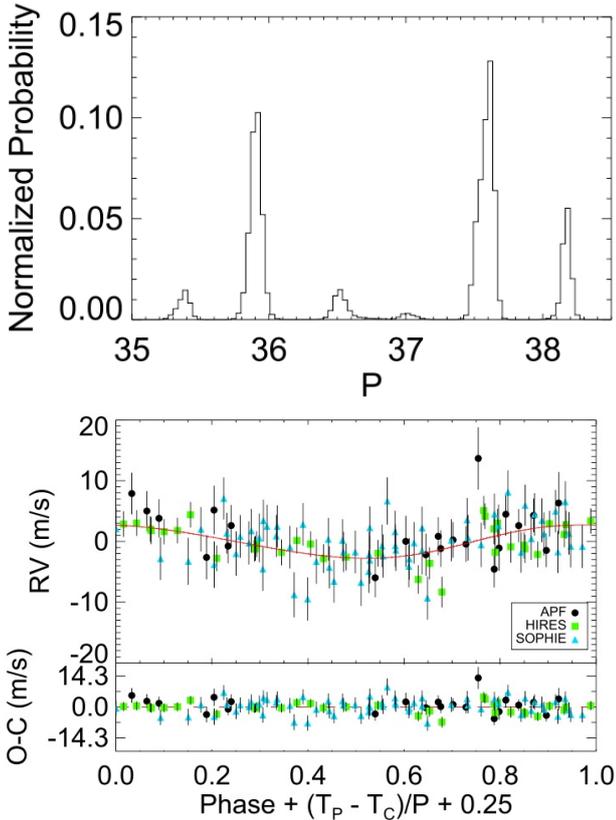
When we take the stellar abundances into account to constrain the planetary Fe/Si and Mg/Si mole ratios, the CMF is calculated to be 0.21, which is lower than the Earth's CMF (0.32). In this scenario, the WMF would be below 0.16, implying that HD 207897 b can be considered a water-rich planet. If the temperature and pressure of the hydrosphere are high enough to sustain a supercritical regime, then the atmosphere is extended and constitutes approximately 20% of the total planetary radius.

5. Discussion and summary

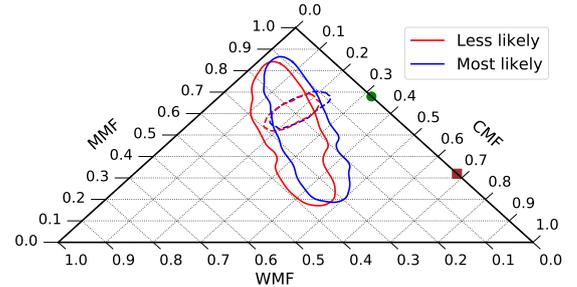
We detected and characterized a sub-Neptune orbiting HD 207897 with a period of 16.202161 ± 0.000083 d. We used TESS photometry data along with SOPHIE, HIRES, and

Table 3. 1σ confidence intervals of the interior and atmosphere MCMC output parameters in the two different compositional scenarios (see text).

Parameter	Activity on 37.6 d		Activity on 35.9 d	
	Scenario 1	Scenario 2	Scenario 1	Scenario 2
Core mass fraction, CMF	0.26 ± 0.18	0.19 ± 0.03	0.23 ± 0.18	0.19 ± 0.03
Water mass fraction, WMF	0.22 ± 0.09	0.19 ± 0.07	0.25 ± 0.10	0.21 ± 0.07
Temperature at 300 bar, T_{300} [K]	2715 ± 22		2742 ± 24	2737 ± 24
Thickness at 300 bar, z_{300} [km]	234 ± 22		265 ± 26	257 ± 25
Albedo, a_p	0.26 ± 0.01			
Core+Mantle radius, [R_p units]	0.72 ± 0.08	0.75 ± 0.05	0.70 ± 0.08	0.74 ± 0.05


Fig. 8. Bimodality in the posterior distribution for the stellar activity periods (*top*) and two-Keplerian model solution on the activity signal at 37.6 d with the initial period at 16.2 d (*bottom*).

APF RVs observations. We found that the planet has a radius of $2.5 \pm 0.08 R_E$ and a mass of either $14.4 \pm 1.6 M_E$ with a probability of 56% or $15.9 \pm 1.6 M_E$ with a probability of 46% based on bimodal results of the stellar activity period. The density accordingly translates into a high density of either $5.1 \pm 0.7 \text{ g cm}^{-3}$ or $5.5^{+0.8}_{-0.7} \text{ g cm}^{-3}$. We used the same mass and radii bounds as Otegi et al. (2020) from the NASA Exoplanet Data Archive¹⁴ (December 5, 2020) and plot all the sub-Neptune-sized planets ($2 < R/R_E < 4$) with a determined semimajor axis and luminosity in Fig. 10. This plot shows that HD 207897 b joins the group of dense sub-Neptunes such as HD 119130 b ($\rho_p = 7.4^{+1.6}_{-1.5} \text{ g cm}^{-3}$; Luque et al. 2019), GJ143 b ($\rho_p = 7^{+1.6}_{-1.3} \text{ g cm}^{-3}$; Dragomir et al. 2019), Kepler-10 c ($\rho_p = 7.1 \pm 1 \text{ g cm}^{-3}$; Dumusque et al. 2014), TOI-849 b ($\rho_p = 5.2^{+0.7}_{-0.8} \text{ g cm}^{-3}$; Armstrong et al. 2020), Kepler-538 b ($\rho_p = 5.4 \pm 1.3 \text{ g cm}^{-3}$; Mayo et al. 2019), Kepler-411 b

¹⁴ <https://exoplanetarchive.ipac.caltech.edu/>

Fig. 9. 2D 1σ confidence regions of HD 207897 b for the mass estimation with activity period on 37.6 d (blue) and with activity period on 35.9 d (red). Solid lines indicate the confidence intervals of scenario 1, and dashed lines correspond to scenario 2, in which the stellar abundances are also included as input data in the MCMC interior structure analysis. The MMF is defined as $\text{MMF} = 1 - \text{CMF} - \text{WMF}$. The green dot and brown square indicate the position of Earth and Mercury in the ternary diagram, respectively.

($\rho_p = 9.9 \pm 1.3 \text{ g cm}^{-3}$; Sun et al. 2019), K2-110 b ($\rho_p = 5.2 \pm 1.2 \text{ g cm}^{-3}$; Osborn et al. 2017), and K2-263 b ($\rho_p = 5.7^{+1.6}_{-1.4} \text{ g cm}^{-3}$; Mortier et al. 2018). These planets are relatively close to their host stars and may have a similar formation history.

The question now is how HD 207897 b and other similar planets with such a high density can exist at a close distance to their host star. One possibility is that the planet has lost most of its volatile elements by evaporation, but for the case of HD 207897 b with an orbital period of 16.20 d and receiving an incident flux of $F = 26.3 F_E$, this is not a satisfactory answer. Even when we consider an extreme evaporation process (Lecavelier Des Étang 2007), the mass loss of the planet would be just $0.1 M_E$ during the entire lifetime of the star, which cannot account for its high density. HD 207897 b is unlikely to have formed in situ. According to Schlichting (2014), the maximum isolation mass that can form at a distance of $a = 0.12 \text{ AU}$ is only $\sim 0.06 M_E$ assuming a minimum mass solar nebula (MMSN). A disk surface density ~ 41 times higher than that of the solar nebula would be required to form a planet as massive as HD 207897 b at this distance. Two possible scenarios can be considered that are both consistent with the MMSN, however: Either the material from the outer region migrated and formed the planet HD 207897 b (e.g., Chatterjee & Tan 2013), or the formation of HD 207897 b occurred far away from the disk and the planet subsequently migrated to its current location (e.g., McNeil & Nelson 2010; Kley & Nelson 2012). The second scenario could have been triggered by another planet in this system. The hint of a long-term trend on RVs allows for the presence of another planet. The high occurrence rate of long-period giant planets (mass $> 0.3 M_J$) in systems harboring small planets

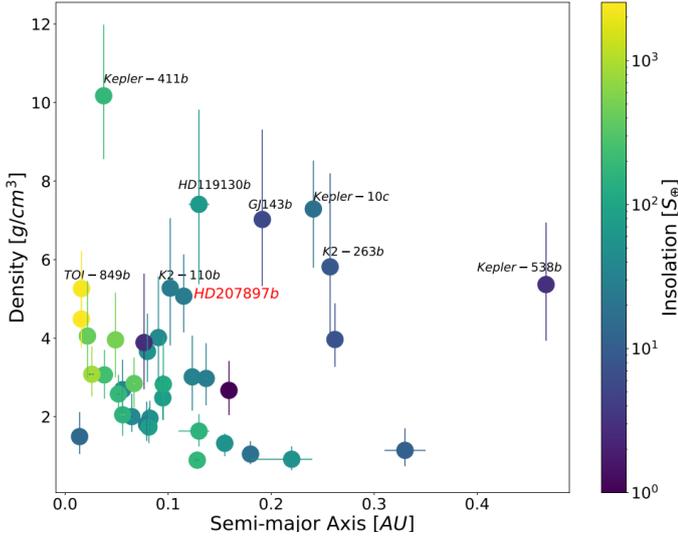


Fig. 10. Density-semimajor-axis diagram for HD 207897 b and other sub-Neptune-sized planets ($2 < R_E < 4$) with known semimajor axes, luminosity, and accurate mass and radius (Otegi et al. 2020). Dots are colored with planet insolation in Earth units ($S_p/S_{\oplus} = (L_*/L_{\odot}) \times (AU/a_p)^2$). The mentioned planets have a density higher than 5 g cm^{-3} .

(planets with masses or radii between those of Earth and Neptune) (Zhu & Wu 2018; Schlecker et al. 2021) can also support this scenario. In the case of HD 119130 b, Luque et al. (2019) suggested a migration scenario triggered by other planets to explain the linear drift in their RV data. However, more photometry and radial velocity observations are still needed to understand the planetary system HD 207897.

Furthermore, the brightness ($K = 6.3$ mag) of HD 207897, its relatively small stellar radius, and its quite nearby distance (28 pc) would make HD 207897 b a good target for atmospheric characterization. Similarly, the close distance and high brightness for a transiting planet host, make HD 207897 b an excellent target for studying the architecture of the system through ground-based observations. For example, measuring the host star spin-orbit alignment (obliquity) using the Rossiter-McLaughlin (RM, Rossiter 1924; McLaughlin 1924) anomaly can provide us with important information about planetary migration and evolution.

Acknowledgements. We warmly thank the OHP staff for their support on the observations. X.B., I.B. and T.F. received funding from the French Programme National de Physique Stellaire (PNPS) and the Programme National de Planétologie (PNP) of CNRS (INSU). We also acknowledge the financial support of French embassy in Tehran. N.H. acknowledges F. Vakili for his constant academic and administrative support. N.H. warmly thanks Ph. Stee, J. L. Beuzit, E. T. Givenchy for all their help. N.H. also thanks Jason D. Eastman and Michael Hippke who have written EXOFASTv2 and TLS package, respectively, for their guides. This publication makes use of The Data & Analysis Center for Exoplanets (DACE), which is a facility based at the University of Geneva (CH) dedicated to extrasolar planets data visualisation, exchange and analysis. DACE is a platform of the Swiss National Centre of Competence in Research (NCCR) PlanetS, federating the Swiss expertise in Exoplanet research. The DACE platform is available at <https://dace.unige.ch>. J.L.B. acknowledges financial support received from “la Caixa” Foundation (ID 100010434) and from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 847648, with fellowship code LCF/BQ/PI20/11760023. This work has been carried out in the frame of the National Centre for Competence in Research “Planets” supported by the Swiss National Science Foundation (SNSF). This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (project SPICE DUNE, grant agreement No 947634). This work was supported by FCT – Fundação para a Ciência

e Tecnologia (FCT) through national funds and by FEDER through COMPETE2020 – Programa Operacional Competitividade e Internacionalização by these grants: UID/FIS/04434/2019; UIDB/04434/2020; UIDP/04434/2020; PTDC/FIS-AST/32113/2017 and POCI-01-0145-FEDER-032113; PTDC/FIS-AST/28953/2017 and POCI-01-0145-FEDER-028953. V.A., E.D.M., N.C.S., and S.G.S. also acknowledge the support from FCT through Investigador FCT contracts nr. IF/00650/2015/CP1273/CT0001, IF/00849/2015/CP1273/CT0003, IF/00169/2012/CP0150/CT0002, and IF/00028/2014/CP1215/CT0002, respectively, and POPH/FSE (EC) by FEDER funding through the program “Programa Operacional de Factores de Competitividade – COMPETE”. O.D.S.D. is supported in the form of work contract (DL 57/2016/CP1364/CT0004) funded by FCT. A.C. et P.C. acknowledge funding from the French National Research Agency (ANR) under contract number ANR-18-CE31-0019 (SPLASH). N.A.D. acknowledges the support of FONDECYT project 3180063. S.H. acknowledges CNES funding through the grant 837319. X.D. and G.G. acknowledge funding in the framework of the Investissements d’Avenir program (ANR-15-IDEX-02), through the funding of the “Origin of Life” project of the Univ. Grenoble-Alpes. This work was supported by Fundação para a Ciência e Tecnologia (FCT) and Fundo Europeu de Desenvolvimento Regional (FEDER) via COMPETE2020 through the research grants UIDB/04434/2020, UIDP/04434/2020, PTDC/FIS-AST/32113/2017 and POCI-01-0145-FEDER-032113, PTDC/FIS-AST/28953/2017 and POCI-01-0145-FEDER-028953. O.D.S.D. is supported in the form of work contract (DL 57/2016/CP1364/CT0004) funded by FCT. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center for the production of the SPOC data products. Funding for the TESS mission is provided by NASA’s Science Mission directorate. We acknowledge the use of public TESS Alert data from pipelines at the TESS Science Office and at the TESS Science Processing Operations Center. This research has made use of the Exoplanet Follow-up Observation Program website, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. This paper includes data collected by the TESS mission that are publicly available from the Mikulski Archive for Space Telescopes (MAST). D.D. acknowledges support from the TESS Guest Investigator Program grant 80NSSC19K1727 and NASA Exoplanet Research Program grant 18-2XRP182-0136. T.D. acknowledges support from MIT’s Kavli Institute as a Kavli postdoctoral fellow. E.A.P. acknowledges the support of the Alfred P. Sloan Foundation. L.M.W. is supported by the Beatrice Watson Parrent Fellowship and NASA ADAP Grant 80NSSC19K0597. A.C. is supported by the NSF Graduate Research Fellowship, grant No. DGE 1842402. D.H. acknowledges support from the Alfred P. Sloan Foundation, the National Aeronautics and Space Administration (80NSSC19K0379), and the National Science Foundation (AST-1717000). I.J.M.C. acknowledges support from the NSF through grant AST-1824644. P.D. acknowledges support from a National Science Foundation Astronomy and Astrophysics Postdoctoral Fellowship under award AST-1903811. A.B. is supported by the NSF Graduate Research Fellowship, grant No. DGE 1745301. R.A.R. is supported by the NSF Graduate Research Fellowship, grant No. DGE 1745301. C.D.D. acknowledges the support of the Hellman Family Faculty Fund, the Alfred P. Sloan Foundation, the David and Lucile Packard Foundation, and the National Aeronautics and Space Administration via the TESS Guest Investigator Program (80NSSC18K1583). J.M.A.M. is supported by the NSF Graduate Research Fellowship, grant No. DGE-1842400. J.M.A.M. and P.C. acknowledge the LSSTC Data Science Fellowship Program, which is funded by LSSTC, NSF Cybertraining Grant No. 1829740, the Brinson Foundation, and the Moore Foundation; their participation in the program has benefited this work. This paper is partially based on observations made with the Nordic Optical Telescope, operated by the Nordic Optical Telescope Scientific Association at the Observatorio del Roque de los Muchachos, La Palma, Spain, of the Instituto de Astrofísica de Canarias.

References

- Acuña, L., Deleuil, M., Mousis, O., et al. 2021, *A&A*, **647**, A53
 Adibekyan, V. Z., Sousa, S. G., Santos, N. C., et al. 2012, *A&A*, **545**, A32
 Adibekyan, V., Figueira, P., Santos, N. C., et al. 2015, *A&A*, **583**, A94
 Aller, A., Lillo-Box, J., Jones, D., Miranda, L. F., & Forteza, S. B. 2020, *A&A*, **635**, A128
 Angus, R., Foreman-Mackey, D., & Johnson, J. A. 2016, *ApJ*, **818**, 109
 Armstrong, D. J., Lopez, T. A., Adibekyan, V., et al. 2020, *Nature*, **583**, 39
 Balunas, S. á., Donahue, R., Soon, W., et al. 1995, *ApJ*, **438**, 269
 Baluev, R. V. 2008, *MNRAS*, **385**, 1279
 Bertran de Lis, S., Delgado Mena, E., Adibekyan, V. Z., Santos, N. C., & Sousa, S. G. 2015, *A&A*, **576**, A89
 Bhatti, W., Bouma, L., Joshua, John, & Price-Whelan, A. 2020, waqasbhatti/astrobases: astrobases v0.5.0
 Boisse, I., Bouchy, F., Hébrard, G., et al. 2011, *A&A*, **528**, A4

- Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Science*, **327**, 977
- Bouchy, F., Moutou, C., Queloz, D., et al. 2008, *Proc. Int. Astron. Union*, **4**, 129
- Bouchy, F., Hébrard, G., Udry, S., et al. 2009, *A&A*, **505**, 853
- Bouchy, F., Díaz, R., Hébrard, G., et al. 2013, *A&A*, **549**, A49
- Brewer, J. M., Fischer, D. A., Valenti, J. A., & Piskunov, N. 2016, *ApJS*, **225**, 32
- Brown, A., Vallenari, A., Prusti, T., et al. 2018, *A&A*, **616**, A1
- Brugger, B., Mousis, O., Deleuil, M., & Deschamps, F. 2017, *ApJ*, **850**, 93
- Buchhave, L. A., Bakos, G. Á., Hartman, J. D., et al. 2010, *ApJ*, **720**, 1118
- Buchhave, L. A., Latham, D. W., Johansen, A., et al. 2012, *Nature*, **486**, 375
- Buchhave, L. A., Bizzarro, M., Latham, D. W., et al. 2014, *Nature*, **509**, 593
- Burt, J., Hanson, R., Rivera, E., et al. 2014, *SPiE*, **9152**, 915211
- Butler, R. P., Marcy, G. W., Williams, E., et al. 1996, *PASP*, **108**, 500
- Chatterjee, S., & Tan, J. C. 2013, *ApJ*, **780**, 53
- Choi, J., Dotter, A., Conroy, C., et al. 2016, *ApJ*, **823**, 102
- Claret, A. 2017, *A&A*, **600**, A30
- Courcol, B., Bouchy, F., Pepe, F., et al. 2015, *A&A*, **581**, A38
- Delgado Mena, E., Israelian, G., González Hernández, J. I., et al. 2010, *ApJ*, **725**, 2349
- Delgado Mena, E., Tsantaki, M., Adibekyan, V. Z., et al. 2017, *A&A*, **606**, A94
- Delisle, J.-B., Ségransan, D., Buchschacher, N., & Alesina, F. 2016, *A&A*, **590**, A134
- Dorn, C., Khan, A., Heng, K., et al. 2015, *A&A*, **577**, A83
- Dotter, A. 2016, *ApJS*, **222**, 8
- Dragomir, D., Teske, J., Günther, M. N., et al. 2019, *ApJ*, **875**, L7
- Dumusque, X., Bonomo, A. S., Haywood, R. D., et al. 2014, *ApJ*, **789**, 154
- Eastman, J. 2017, *Astrophysics Source Code Library* [record ascl:1710.0003]
- Eastman, J., Gaudi, B. S., & Agol, E. 2013, *PASP*, **125**, 83
- Eastman, J. D., Rodriguez, J. E., Agol, E., et al. 2019, *PASP*, submitted [arXiv:1907.09480]
- Findeisen, K., Hillenbrand, L., & Soderblom, D. 2011, *AJ*, **142**, 23
- Ford, E. B. 2006, *ApJ*, **642**, 505
- Fulton, B. J., Weiss, L. M., Sinukoff, E., et al. 2015, *ApJ*, **805**, 175
- Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, *AJ*, **154**, 109
- Fűrész, G. 2008, PhD thesis, University of Szeged, Hungary
- Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, *Space Sci. Rev.*, **123**, 485
- Gelman, A., & Rubin, D. B. 1992, *Stat. Sci.*, **7**, 457
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. 2004, *CRC Texts in Statistical Science* (USA: CRC Press)
- Ginzburg, S., Schlichting, H. E., & Sari, R. 2016, *ApJ*, **825**, 29
- Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, *MNRAS*, **476**, 759
- Girardi, L., Barbieri, M., Groenewegen, M. A. T., et al. 2012, *Astrophys. Space Sci. Proc.*, **26**, 165
- Günther, M. N., Pozuelos, F. J., Dittmann, J. A., et al. 2019, *Nat. Astron.*, **3**, 1099
- Hara, N. C., Boué, G., Laskar, J., & Correia, A. 2017, *MNRAS*, **464**, 1220
- Hara, N., Bouchy, F., Stalport, M., et al. 2020, *A&A*, **636**, L6
- Hara, N. C., Delisle, J.-B., Unger, N., & Dumusque, X. 2022, *A&A*, in press, <https://doi.org/10.1051/0004-6361/202141197>
- Haywood, R. D., Collier Cameron, A., Queloz, D., et al. 2014, *MNRAS*, **443**, 2517
- Hedges, C., Angus, R., Barentsen, G., et al. 2020, *Res. Notes AAS*, **4**, 220
- Hippke, M., & Heller, R. 2019, *A&A*, **623**, A39
- Hippke, M., David, T. J., Mulders, G. D., & Heller, R. 2019, *AJ*, **158**, 143
- Hobson, M. 2019, PhD thesis, Aix-Marseille, France
- Hobson, M., Díaz, R. F., Delfosse, X., et al. 2018, *A&A*, **618**, A103
- Hormuth, F., Brandner, W., Hippler, S., & Henning, T. 2008, *J. Phys. Conf. Ser.*, **131**, 012051
- Howard, A. W., Johnson, J. A., Marcy, G. W., et al. 2010, *ApJ*, **721**, 1467
- Huang, C. X., Vanderburg, A., Pál, A., et al. 2020a, *Res. Notes Am. Astron. Soc.*, **4**, 204
- Huang, C. X., Vanderburg, A., Pál, A., et al. 2020b, *Res. Notes Am. Astron. Soc.*, **4**, 206
- Isaacson, H., & Fischer, D. 2010, *ApJ*, **725**, 875
- Jenkins, J. M. 2002, *ApJ*, **575**, 493
- Jenkins, J. M., Chandrasekaran, H., McCauliff, S. D., et al. 2010, *SPiE Conf. Ser.*, **7740**, 77400D
- Jenkins, J. M., Twicken, J. D., McCauliff, S., et al. 2016, *SPiE*, **9913**, 99133E
- Kley, W., & Nelson, R. 2012, *ARA&A*, **50**, 211
- Lecavelier Des Étangs, A. 2007, *A&A*, **461**, 1185
- Li, J., Tenenbaum, P., Twicken, J. D., et al. 2019, *PASP*, **131**, 024506
- Lillo-Box, J., Barrado, D., & Bouy, H. 2012, *A&A*, **546**, A10
- Lillo-Box, J., Barrado, D., & Bouy, H. 2014, *A&A*, **566**, A103
- Lindgren, L., Hernández, J., Bombrun, A., et al. 2018, *A&A*, **616**, A2
- Lopez, E. D., & Fortney, J. J. 2014, *ApJ*, **792**, 1
- Luque, R., Nowak, G., Pallé, E., et al. 2019, *A&A*, **623**, A114
- Mamajek, E. E., & Hillenbrand, L. A. 2008, *ApJ*, **687**, 1264
- Mandel, K., & Agol, E. 2002, *ApJ*, **580**, L171
- Martin, D. V., El-Badry, K., Hodžić, V. K., et al. 2021, *MNRAS*, **507**, 4132
- Mayo, A. W., Rajpaul, V. M., Buchhave, L. A., et al. 2019, *AJ*, **158**, 165
- Mazevet, S., Licari, A., Chabrier, G., & Potekhin, A. Y. 2019, *A&A*, **621**, A128
- McLaughlin, D. 1924, *ApJ*, **60**
- McNeil, D., & Nelson, R. 2010, *MNRAS*, **401**, 1691
- McQuillan, A., Aigrain, S., & Mazeh, T. 2013, *MNRAS*, **432**, 1203
- Mink, D. J. 2011, *ASP Conf. Ser.*, **442**, 305
- Mortier, A., Bonomo, A., Rajpaul, V., et al. 2018, *MNRAS*, **481**, 1839
- Mousis, O., Deleuil, M., Aguichine, A., et al. 2020, *ApJ*, **896**, L22
- Nielsen, L. D., Gandolfi, D., Armstrong, D., et al. 2020, *MNRAS*, **492**, 5399
- Noyes, R., Hartmann, L., Baliunas, S., Duncan, D., & Vaughan, A. 1984, *ApJ*, **279**, 763
- Osborn, H., Santerne, A., Barros, S., et al. 2017, *A&A*, **604**, A19
- Otegi, J., Bouchy, F., & Helled, R. 2020, *A&A*, **634**, A43
- Owen, J. E., & Wu, Y. 2013, *ApJ*, **775**, 105
- Owen, J. E., & Wu, Y. 2017, *ApJ*, **847**, 29
- Paulson, D. B., Saar, S. H., Cochran, W. D., & Hatzes, A. P. 2002, *AJ*, **124**, 572
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, *J. Mach. Learn. Res.*, **12**, 2825
- Pepe, F., Mayor, M., Galland, F., et al. 2002, *A&A*, **388**, 632
- Perruchot, S., Kohler, D., Bouchy, F., et al. 2008, *SPiE*, **7014**, 70140J
- Ricker, G., Winn, J., Vanderspek, R., et al. 2015, *JATIS*, **1**, 014003
- Rogers, L. A. 2015, *ApJ*, **801**, 41
- Rossiter, R. 1924, *ApJ*, **60**, 15
- Santerne, A., Díaz, R., Moutou, C., et al. 2012, *A&A*, **545**, A76
- Santos, N., Sousa, S., Mortier, A., et al. 2013, *A&A*, **556**, A150
- Schlafly, E. F., & Finkbeiner, D. P. 2011, *ApJ*, **737**, 103
- Schlecker, M., Mordasini, C., Emsenhuber, A., et al. 2021, *A&A*, **656**, A71
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, **500**, 525
- Schlichting, H. E. 2014, *ApJ*, **795**, L15
- Schoenberg, I. J. 1946, *Q. Appl. Math.*, **4**, 112
- Smith, J. C., Stumpe, M. C., Van Cleve, J. E., et al. 2012, *PASP*, **124**, 1000
- Sotin, C., Grasset, O., & Mocquet, A. 2007, *Icarus*, **191**, 337
- Sousa, S., Adibekyan, V., Delgado-Mena, E., et al. 2018, *A&A*, **620**, A58
- Stassun, K. G., & Torres, G. 2016, *AJ*, **152**, 180
- Stassun, K. G., & Torres, G. 2021, *ApJ*, **907**, L33
- Stassun, K. G., Collins, K. A., & Gaudi, B. S. 2017a, *AJ*, **153**, 136
- Stassun, K. G., Corsaro, E., Pepper, J. A., & Gaudi, B. S. 2017b, *AJ*, **155**, 22
- Strehl, K. 1902, *Astron. Nachr.*, **158**, 89
- Stumpe, M. C., Smith, J. C., Van Cleve, J. E., et al. 2012, *PASP*, **124**, 985
- Stumpe, M. C., Smith, J. C., Catanzarite, J. H., et al. 2014, *PASP*, **126**, 100
- Sun, L., Ioannidis, P., Gu, S., et al. 2019, *A&A*, **624**, A15
- Teltng, J. H., Avila, G., Buchhave, L., et al. 2014, *Astron. Nachr.*, **335**, 41
- Torres, G., Andersen, J., & Giménez, A. 2010, *A&ARv*, **18**, 67
- Twicken, J. D., Catanzarite, J. H., Clarke, B. D., et al. 2018, *PASP*, **130**, 064502
- Zechmeister, M., Reiners, A., Amado, P. J., et al. 2018, *A&A*, **609**, A12
- Zhu, W., & Wu, Y. 2018, *AJ*, **156**, 92

- ¹ Department of Physics, Shahid Beheshti University, Tehran, Iran
- ² Laboratoire J.-L. Lagrange, Observatoire de la Côte d'Azur (OCA), Université de Nice-Sophia Antipolis (UNS), CNRS, Campus Valrose, 06108 Nice Cedex 2, France
- ³ Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France
- ⁴ Instituto de Astrofísica de Canarias, 38205 La Laguna, Tenerife, Spain
- ⁵ Departamento de Astrofísica, Universidad de La Laguna, 38206 La Laguna, Tenerife, Spain
- ⁶ Observatoire de Haute-Provence, CNRS, Université d'Aix-Marseille, 04870 Saint-Michel-l'Observatoire, France
- ⁷ Institut d'astrophysique de Paris, UMR 7095 CNRS université Pierre et Marie Curie, 98 bis, boulevard Arago, 75014 Paris, France
- ⁸ Observatoire de Genève, Université de Genève, Chemin Pegasi, 51, 1290 Sauverny, Switzerland
- ⁹ Centro de Astrobiología (CAB, CSIC-INTA), Depto. de Astrofísica, ESAC campus, 28692, Villanueva de la Cañada (Madrid), Spain
- ¹⁰ Center for Astrophysics | Harvard & Smithsonian, 60 Garden St, Cambridge, MA 02138, USA
- ¹¹ Canada France Hawaii Telescope Corporation (CFHT), 65-1238 Mamalahoa Hwy, Kamuela HI 96743 USA
- ¹² Departamento de Matemática y Física Aplicadas, Universidad Católica de la Santísima Concepción, Alonso de Rivera 2850, Concepción, Chile
- ¹³ Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, 4150-762 Porto, Portugal

- ¹⁴ Departamento de Física e Astronomia, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 4169-007 Porto, Portugal.
- ¹⁵ Univ. Grenoble Alpes, CNRS, IPAG, 38000 Grenoble, France
- ¹⁶ NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA
- ¹⁷ University of Maryland, Baltimore County, 1000 Hilltop Cir, Baltimore, MD 21250, USA
- ¹⁸ Space Telescope Science Institute, 3700 San Martin Dr, Baltimore, MD 21218, USA
- ¹⁹ Division of Geological and Planetary Science, California Institute of Technology, Pasadena, CA 91125, USA
- ²⁰ Department of Physics & Astronomy, University of California Irvine, Irvine, CA 92697, USA
- ²¹ Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95060, USA
- ²² Department of Physics & Astronomy, University of Kansas, 1082 Malott, 1251 Wescoe Hall Dr., Lawrence, KS 66045, USA
- ²³ Institute for Astronomy, University of Hawai'i, 2680 Woodlawn Drive, Honolulu, HI 96822, USA
- ²⁴ International Center for Advanced Studies (ICAS) and ICIFI (CONICET), ECyT-UNSAM, Campus Miguelete, 25 de Mayo y Francia, (1650) Buenos Aires, Argentina
- ²⁵ Department of Physics, and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
- ²⁶ Department of Physics and Astronomy, University of New Mexico, 1919 Lomas Blvd NE, Albuquerque, NM 87131, USA
- ²⁷ Department of Astronomy, University of California Berkeley, Berkeley, CA, 94720, USA
- ²⁸ Division of Geological and Planetary Sciences 1200 E California Blvd, Pasadena, CA, 91125, USA
- ²⁹ Department of Earth and Planetary Sciences, University of California, Riverside, CA 92521, USA
- ³⁰ NASA Exoplanet Science Institute/Caltech-IPAC, MC 314-6, 1200 E. California Blvd., Pasadena, CA 91125, USA
- ³¹ Millennium Institute for Astrophysics, Chile
- ³² Instituto de Astrofísica, Facultad de Física, Pontificia Universidad Católica de Chile, Chile
- ³³ Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA
- ³⁴ Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Stanford, CA 94305, USA
- ³⁵ Centre for Astrophysics, University of Southern Queensland, Toowoomba, QLD, Australia
- ³⁶ Department of Physics & Astronomy, University of California Los Angeles, Los Angeles, CA 90095, USA
- ³⁷ Las Campanas Observatory, Carnegie Institution of Washington, Colina el Pino, Casilla 601 La Serena, Chile
- ³⁸ Department of Physics and Astronomy, Johns Hopkins University, 3400 N Charles St, Baltimore, MD 21218, USA
- ³⁹ Gemini Observatory/NSF's NOIRLab, 670 N. A'ohoku Place, Hilo, HI 96720, USA
- ⁴⁰ LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 place Jules Janssen, 92195, Meudon, France
- ⁴¹ Univ. de Toulouse, CNRS, IRAP, 14 avenue Belin, 31400 Toulouse, France
- ⁴² Vanderbilt University, Department of Physics & Astronomy, 6301 Stevenson Center Ln., Nashville, TN 37235, USA
- ⁴³ SETI Institute/NASA Ames Research Center, USA
- ⁴⁴ DTU Space, National Space Institute, Technical University of Denmark, Elektrovej 328, 2800 Kgs. Lyngby, Denmark
- ⁴⁵ Department of Astrophysical Sciences, Princeton University, 4Ivy Lane, Princeton, NJ 08544, USA
- ⁴⁶ NASA Ames Research Center, Moffett Field, CA, 94035, USA
- ⁴⁷ Department of Physics, University of Warwick, Coventry, CV4 7AL, UK
- ⁴⁸ Centre for Exoplanets and Habitability, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, UK
- ⁴⁹ Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
- ⁵⁰ Department of Aeronautics and Astronautics, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Appendix A: RV time series

Table A.1: SOPHIE RVs for HD207897 b

BJD (-2400000 d)	RV (km s ⁻¹)	σ_{RV} (km s ⁻¹)	FWHM (km s ⁻¹)	BIS (km s ⁻¹)	H α	$\sigma_{H\alpha}$	CRX	σ_{CRX}
56117.5531	-6.3203	0.0015	6.6803	-0.0025	0.1294	0.0005	-39.0617	13.9761
56118.5849	-6.3205	0.0015	6.6846	0.0027	0.1305	0.0005	-34.4212	12.8526
56119.5401	-6.3238	0.0016	6.6849	-0.0016	0.1296	0.0006	-12.6147	12.0847
56120.5123	-6.3206	0.0016	6.6712	-0.0006	0.1306	0.0006	-29.3660	11.2578
56137.5869	-6.3240	0.0016	6.7125	0.0011	0.1323	0.0006	-36.7534	14.7557
56138.5331	-6.3249	0.0017	6.6945	0.0065	0.1335	0.0007	-28.7203	14.5124
56139.5731	-6.3300	0.0015	6.6839	-0.0008	0.1296	0.0005	32.8700	12.5074
56140.5330	-6.3310	0.0016	6.6906	0.0043	0.1310	0.0006	-34.3580	10.0305
56150.5049	-6.3162	0.0015	6.6867	-0.0028	0.1343	0.0005	-7.1528	14.0474
56152.5655	-6.3214	0.0018	6.6849	0.0002	0.1334	0.0008	-9.6228	14.1995
56167.5191	-6.3227	0.0015	6.6823	0.0054	0.1302	0.0006	-53.2894	10.1835
56178.3748	-6.3278	0.0015	6.7033	0.0013	0.1323	0.0006	-46.0315	12.8783
56497.5314	-6.3279	0.0018	6.6991	-0.0020	0.1288	0.0008	-27.5805	15.6965
56498.5527	-6.3276	0.0016	6.7072	0.0070	0.1326	0.0006	-6.9804	14.2738
56499.5446	-6.3316	0.0016	6.7074	-0.0023	0.1316	0.0006	-41.0561	12.8978
56500.5510	-6.3333	0.0017	6.6891	-0.0059	0.1321	0.0007	5.3262	13.2646
56517.4678	-6.3314	0.0016	6.7181	0.0051	0.1337	0.0007	15.8414	9.3975
56519.4789	-6.3304	0.0016	6.7067	-0.0048	0.1332	0.0007	-18.4473	11.5302
56520.4837	-6.3319	0.0016	6.7042	-0.0029	0.1345	0.0006	20.1748	12.7335
56521.4875	-6.3261	0.0016	6.6994	-0.0011	0.1338	0.0007	-12.2005	14.9481
56522.4818	-6.3248	0.0016	6.7130	0.0057	0.1331	0.0006	-9.6604	11.5639
56523.4905	-6.3274	0.0015	6.7106	0.0077	0.1319	0.0006	-36.2878	13.6435
56524.4789	-6.3246	0.0024	6.6754	0.0101	0.1384	0.0013	-0.9706	21.2910
56558.3761	-6.3225	0.0015	6.7002	-0.0020	0.1301	0.0006	-12.9535	15.0095
56560.3978	-6.3254	0.0016	6.7038	0.0003	0.1311	0.0006	-30.1229	14.8829
56583.3673	-6.3305	0.0017	6.6867	-0.0021	0.1307	0.0008	-6.3707	16.6035
56592.3536	-6.3289	0.0021	6.6841	-0.0077	0.1321	0.0011	-8.3097	21.0877
56624.3190	-6.3292	0.0019	6.7214	0.0091	0.1346	0.0009	6.39867	17.5343
56625.2867	-6.3321	0.0015	6.7172	-0.0040	0.1310	0.0006	17.7943	10.9701
56626.2718	-6.3297	0.0015	6.7156	0.0024	0.1327	0.0006	-3.3048	12.3435
56628.3163	-6.3305	0.0016	6.7214	0.0007	0.1322	0.0007	-8.6098	14.9817
56629.3041	-6.3354	0.0017	6.7138	-0.0064	0.1322	0.0007	11.1813	13.8973
56630.2650	-6.3339	0.0014	6.7173	0.0018	0.1304	0.0005	-11.7681	12.1465
56631.2725	-6.3342	0.0014	6.7182	-0.0030	0.1297	0.0005	-23.8070	12.1372
56654.2301	-6.3252	0.0019	6.7088	-0.0011	0.1349	0.0009	-15.5221	17.2095
56656.2313	-6.3296	0.0018	6.7171	0.0054	0.1324	0.0008	2.0870	14.7556
56657.2325	-6.3260	0.0015	6.7183	0.0059	0.1325	0.0006	21.1445	11.7855
56664.2359	-6.3336	0.0023	6.7235	0.0026	0.1303	0.0013	-9.4066	21.0183
57214.5707	-6.3257	0.0013	6.7541	-0.0027	0.1291	0.0005	4.0133	13.6434
57217.5644	-6.3202	0.0013	6.7651	0.0025	0.1284	0.0006	5.5657	16.3052
57284.4185	-6.3216	0.0016	6.7057	-0.0038	0.1293	0.0007	-3.4004	12.3004
58862.3082	-6.3280	0.0018	6.7636	0.0076	0.1415	0.0009	-1.5094	21.4660
58875.3024	-6.3284	0.0016	6.8431	-0.0004	0.1440	0.0007	108.6779	21.7024
58877.3789	-6.3325	0.0023	6.7005	0.0187	0.1470	0.0013	69.2134	29.4011
58878.3249	-6.3292	0.0019	6.7894	-0.000	0.1437	0.0009	57.8937	28.6864
58879.3132	-6.3368	0.0022	6.8407	-0.0178	0.1444	0.0011	54.4243	36.3337
58880.2542	-6.3327	0.0012	6.8466	0.0038	0.1414	0.0005	53.1909	18.2402
58881.2536	-6.3336	0.0016	6.7832	-0.0118	0.1455	0.0007	49.2682	15.8376
58882.6853	-6.3330	0.0025	6.7353	-0.0007	0.1445	0.0014	80.9618	43.6620
58883.2559	-6.3304	0.0017	6.7916	-0.0008	0.1442	0.0008	60.3972	18.5206
58885.2949	-6.3384	0.0020	6.7925	-0.0120	0.1415	0.0010	102.8511	29.6527
58886.2969	-6.3374	0.0017	6.8429	-0.0090	0.1438	0.0008	50.0756	21.1131
58888.2703	-6.3319	0.0016	6.8534	-0.0097	0.1424	0.0007	84.0189	25.1804
58890.2920	-6.3314	0.0023	6.7950	-0.0084	0.1472	0.0012	70.5820	30.6721
58892.2593	-6.3200	0.0023	6.7914	0.0029	0.1485	0.0012	61.7681	27.8826
58894.2567	-6.3309	0.0014	6.8221	-0.0060	0.1436	0.0006	85.0683	18.1804
58895.2570	-6.3419	0.0017	6.8537	0.0008	0.1447	0.0008	76.4210	22.0550
58898.2952	-6.3335	0.0021	6.7098	0.0081	0.1446	0.0012	-12.7695	29.9844
58899.2728	-6.3337	0.0019	6.7510	-0.0144	0.1441	0.0009	-4.7509	23.2686
58900.2705	-6.3272	0.0013	6.8566	-0.0026	0.1399	0.0005	17.5256	13.1316
58901.2780	-6.3219	0.0018	6.8054	0.0090	0.1453	0.0009	6.4814	19.9266
58902.2708	-6.3301	0.0012	6.8455	-0.0022	0.1419	0.0005	21.4611	14.2376
58903.2734	-6.3223	0.0016	6.8011	-0.0038	0.1415	0.0007	67.6199	20.3261
59068.5491	-6.3265	0.0011	6.6780	0.0011	—	—	—	—

Table A.2: HIRES RVs for HD207897 b

BJD (-2400000 d)	RV (m s ⁻¹)	σ_{RV} (m s ⁻¹)	S-index	$\sigma_{S-index}$
52832.039671	-6.1499	1.04	—	—
52855.027852	-6.4812	2.42	—	—
53196.053218	-8.9158	1.01	—	—
53602.917000	1.127055	0.819542	0.231000	0.001000
53962.008000	4.990706	0.870220	0.222000	0.001000
56588.935000	3.431498	1.463105	0.225000	0.001000
57237.106000	3.290118	1.190904	0.215400	0.001000
58869.707000	3.595136	1.360255	0.249000	0.001000
59004.080000	-3.829985	1.708751	0.251100	0.001000
59007.977000	0.107424	1.636860	0.247300	0.001000
59008.017000	0.038845	1.893359	0.249900	0.001000
59008.066000	-3.912803	1.738616	0.250000	0.001000
59039.084000	-2.874458	1.880687	0.240000	0.001000
59079.039000	3.561938	1.538733	0.238600	0.001000
59089.002000	0.991416	1.012931	0.233500	0.001000
59090.031000	-1.806837	0.977288	0.236100	0.001000
59091.056000	-3.314227	0.904606	0.233400	0.001000
59092.060000	-3.337441	0.963977	0.237100	0.001000
59093.027000	-0.185500	0.992191	0.239300	0.001000
59094.984000	-4.733336	0.983995	0.233400	0.001000
59097.973000	2.455004	1.046085	0.235400	0.001000
59099.827000	2.774576	1.035973	0.236600	0.001000
59100.985000	4.753109	1.043153	0.237800	0.001000
59102.009000	3.388690	0.992955	0.237100	0.001000
59115.059000	8.057584	0.990576	0.244300	0.001000
59118.985000	0.615674	1.028696	0.240700	0.001000
59119.868000	4.070617	1.021998	0.239300	0.001000
59120.980000	-0.036780	0.993603	0.241600	0.001000
59122.955000	-1.142312	0.922378	0.239500	0.001000
59123.937000	-2.279427	0.982050	0.238200	0.001000
59142.961000	-4.944806	1.161786	0.242200	0.001000
59153.887000	-2.863129	1.141554	0.240700	0.001000
59181.890000	-1.681638	1.441832	0.236500	0.001000
59187.825000	-1.590208	1.406428	0.227200	0.001000
59188.813000	-6.084715	1.833509	0.227800	0.001000
59189.870000	-6.143414	1.421509	0.229800	0.001000

Notes. The first three data were taken with a different CCD detector. For the sake of simplicity, we excluded them in our joint modeling by EXOFASTv2.

Table A.3: APF RVs for HD207897 b

BJD (-2400000 d)	RV (m s ⁻¹)	σ_{RV} (m s ⁻¹)
59002.885	3.030364933	3.218557358
59003.809	5.987553939	2.861260653
59004.916	4.500414023	2.633646965
59005.881	2.367894193	2.662182808
59006.804	14.66981698	4.979400635
59008.864	1.427995591	3.299036264
59009.826	-1.594891397	2.404323578
59010.938	-0.297586458	2.476657391
59011.895	-5.612035257	2.451689482
59012.812	3.174356114	4.957623482
59016.811	11.1418913	3.183858633
59017.938	9.565100481	3.032974005
59018.847	8.944296078	2.887774467
59039.912	-1.683343659	2.742200613
59059.929	-5.287302366	2.546239138
59130.705	8.958821818	3.813406229
59151.657	-1.77725919	2.748860359
59178.67	-3.064406035	2.962447405
59203.816	-1.276284674	3.144914389
59223.847	-3.812726288	3.542819262
59252.786	-4.278596616	4.021532536
59273.802	-3.184576312	3.384922504

Appendix B: Figures

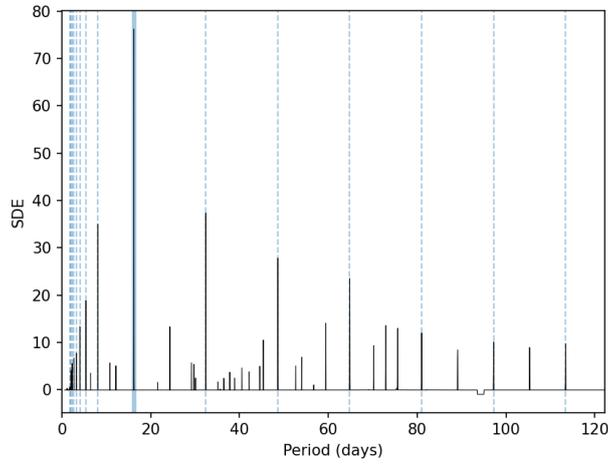


Fig. B.1: TLS periodogram for the TESS light curve of HD207897 b. The blue line shows the highest peak at 16.20 d with SDE=76.2. The dashed blue lines indicate the aliases of this period.

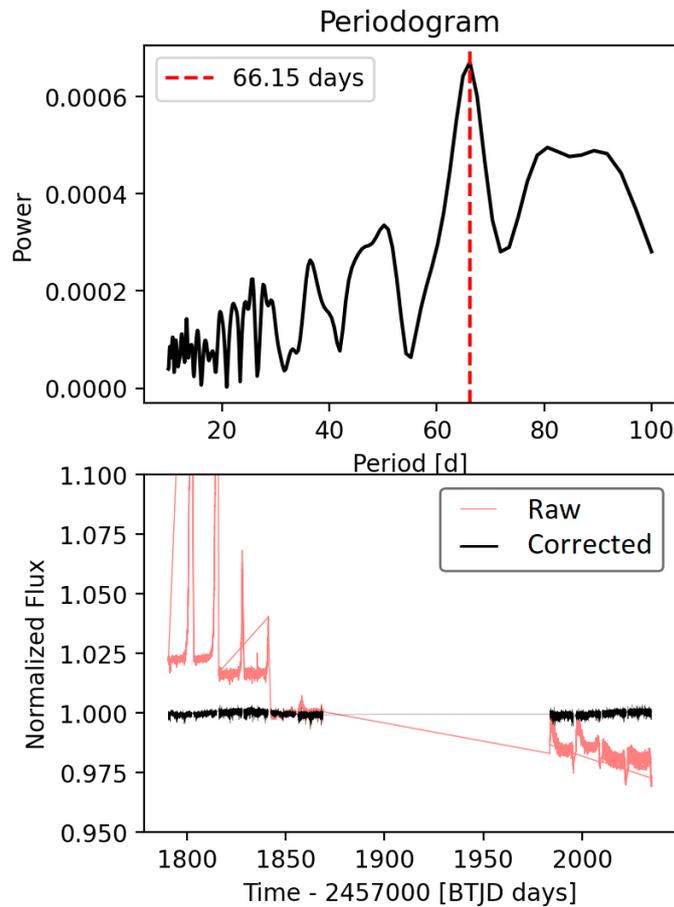


Fig. B.2: Systematics-insensitive periodogram and SAP TESS data for HD207897. *Top*: The SIP periodogram is generated for HD207897. *Bottom*: Raw SAP TESS light curves in red and detrended data against instrument systematics by SIP in black. The SIP periodogram does not exhibit any significant signal for HD207897 b because the highest peak in this periodogram has a power of only 0.0006 at 66.6 d.

3.6.2 Two planets around the HD88986

HD88986 is a multi-planet system with a temperate sub-Neptune and a wide-orbit Jupiter mass planet. Its host star is a G0 type star, of solar metallicity, and it is one of the nearest and brightest exoplanet host stars ($G_{mag}=6.30$, $T_{eff}=5960_{180}^{280}$ K, $d=30.03$ pc). The paper is under preparation and the results will be published in Heidari et al 2023. I summarise the main properties of the planets in this system as the paper has progressed.

We discovered a temperate sub-Neptune and a wide-orbit Jupiter mass planet in this system, using several RV measurements (389 SOPHIE data, 31 ELODIE data, 34 HIRES data), Gaia DR3 data, 21-year photometric observation of automatic photo-electric telescope (APT), 2 sectors of TESS data and 7-day observation of CHEOPS. The sub-Neptune planet detected, HD88986 b, has the longest period among the accurately characterized transiting sub-Neptune with an orbital period of $147.4_{-0.05}^{+0.05}$ d, hence it is a temperate Neptune and an interesting target for the study of the internal structure. The second planet, using Gaia DR3 excess noise, is compatible with an edge-on configuration, and the probability that HD88986 c is a planet with a mass smaller than $13.5 M_{Jup}$ is $\sim 94\%$. The detection of such a system reinforces the relation between the existence of the giant planet in the wide orbit and the smaller inner planet. The HD88986 system, due to the presence of HD88986 b which is a very long-period transiting planet, and also HD88986 c the wide-orbit giant planet, is a prime target for investigating different theories of planetary formation and studying the planetary composition and architecture in multi-planet systems.

The SOPHIE search for northern extrasolar planets-XIX. HD88986: a multi-planet system with a temperate transiting sub-Neptune and a wide-orbit Jupiter mass planet

N. Heidari^{*1,2,3}, I. Boisse³, T. G. Wilson⁴, F. Kiefer⁵ and et al

(Affiliations can be found after the references)

Received XX, 2022; accepted XX, 2022

ABSTRACT

The Transiting Exoplanet Survey Satellite (TESS) revealed numerous single transiting events corresponding to the long-period systems. Such a system can be determined by observing a second transit event and/or measuring a radial velocity orbit. Here, we confirm the planetary nature of one of the TESS single transit events with 189 ± 11 ppm depth and duration of $14.5_{0.9}^{0.8}$ h, in the inner orbit of a wide-orbit Jupiter mass planet. We used intensive radial velocity measurements with time span of about 25 years (389 SOPHIE data, 31 ELODIE data, 34 HIRES data), Gaia DR3 data, 21-year photometric observation of an automatic photoelectric telescope (APT), 2 sectors of TESS data, and a 7-day observation of CHEOPS. We show that the first planet is a sub-Neptune orbiting every $147.4_{-0.05}^{+0.05}$ d around one of the nearest and brightest star HD88986 (G2V type, $G_{mag}=6.30$, distance= 30 pc, $T_{eff}=5960_{180}^{280}$ K). HD88986 b has the longest periods among the accurately characterized transiting sub-Neptune, hence it is temperate ($T_{eff} = 476_{-10}^{+13}$), which makes it an interesting target for the study of the internal structure. The second planet, using Gaia DR3 excess noise, is compatible with an edge-on configuration, and the probability that HD88986 c is a planet with a mass smaller than 13.5 M_{Jup} is $\sim 94\%$. The detection of such a system reinforces the relation between the existence of the giant planet in the wide orbit and the smaller inner planet.

Key words. planets and satellites: detection techniques: photometric, radial velocities stars: individual (HD 88986, and TIC 1042868)

1. Introduction

2. Identification and observations

2.1. High-resolution spectroscopy with SOPHIE

HD88986 has been intensively monitored by the SOPHIE high-precision spectrograph, mounted at the 1.93 m telescope at the Haute-Provence Observatory (OHP, France). The star observations were carried out as part of Recherche de Planètes Extrasolaires (RPE) subprogram 1, also known as SP1, which is a high-precision program to search for Neptunes and Super-Earths orbiting bright stars in the solar neighborhood (Courcol et al. 2015; Heidari et al. 2021). The observation was performed from 7th December 2007 to 24th April 2022, over a time span of 15 years, gathering 390 high-resolution spectra (see Fig. 1). After a few years of our observations, a preliminary analysis of RVs revealed a clear long-term drift along with a periodic signal every 146 days. To confirm the origin of the two signals, we continued our observations. We used SOPHIE high resolution (HR) mode (resolving power of $\lambda/\Delta\lambda \approx 75000$ at 550 nm), with a simultaneous thorium-argon or Fapry-Perot calibration lamp measurements, to track instrumental drift.

In June 2011, hexagonal fibers were installed in the SOPHIE spectrograph. This led to a great improvement in the precision of RV SOPHIE data and also about 50 m s^{-1} shift in the measured RVs of the standard stars (Bouchy et al. 2013). Hence, we separated the data before June 2011 (12 data points, with the name of SOPHIE) from the data after (378 data points, with the name of SOPHIE+). The exposure time was set for both data sets on 600 s to achieve a median signal-to-noise (S/N) ratio of 157.8, with an RVs root means square (RMS) of 13.0 m/s for SOPHIE+ and S/N of 158.3 with an RV RMS for 5.9 m/s for SOPHIE.

2.2. High-resolution spectroscopy with HIRES

The star was observed by HIRES spectrograph from 2nd December 1996 to 19 January 2014, over a time span of 17 years and obtaining 51 high-resolution spectra (see Fig. 1). For more information about the data reduction and detail of observation see Butler et al. (2017). HIRES data experience a small jump ($1.5 \pm 0.1 \text{ m s}^{-1}$) due to the CCD change in August 2004, a long-term drift ($\lesssim 1 \text{ m s}^{-1}$), and a small intra-night drift, which are identified by Tal-Or et al. (2019). We recovered HIRES data after systematic correction by Tal-Or et al. (2019) from the Vizier catalog access tool ¹. The mean uncertainty of this data is 1.2 m/s with a mean RMS of 11.1 m/s. Note that in order to take into account any remaining offsets due to the CCD change in the HIRES data, we fit an offset between the data before (HIRES) and after (HIRES+).

2.3. High-resolution spectroscopy with ELODIE

ELODIE was a high-resolution spectrograph, mounted on the 1.93m telescope at OHP, which was used to discover the first exoplanet in 1995 (Mayor & Queloz 1995). The star was observed by ELODIE from 28 February 1997 to 29 January 2004, gathering 31 high-resolution spectra (see Fig. 1). The RVs are extracted using K0 numerical mask (Baranne et al. 1996), and with an exposure time of between 600 to 900, achieved to the mean uncertainty of 9.0 m s^{-1} and mean RMS of 13.0 m s^{-1} . We note that 3 data points are removed due to their low S/N (< 50).

¹ <https://vizier.u-strasbg.fr/viz-bin/VizieR>

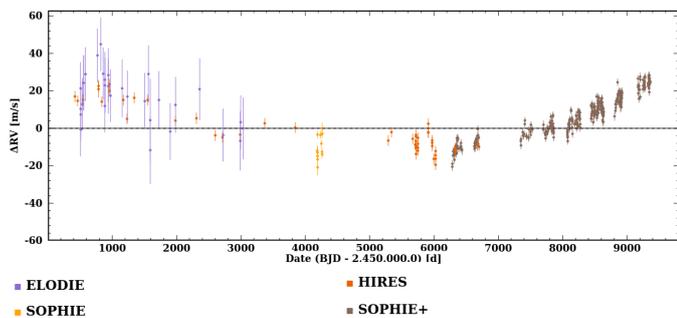


Fig. 1: Radial velocity measurements of HD88986 from ELODIE, HIRES, SOPHIE and SOPHIE+.

2.4. Photometric observations with TESS

HD88986 was observed in TESS sector 21 and 48², in a 2-minute cadence. The photometric data were produced by the Simple Aperture Photometry flux with pre-search data conditioning (PDC-SAP) pipeline (Stumpe et al. 2012; Smith et al. 2012; Stumpe et al. 2014), provided by the Science Processing Operations Center (SPOC; Jenkins et al. 2016). The raw TESS photometric data is plotted in Fig. 2 top panel.

After observation of TESS sector 21 was done, with initial analysis, we identified a single transit candidate with 189 ± 11 ppm depth and duration of $14.5_{0.9}^{0.8}$ h. The single transit events do not provide well-constrained orbital periods. However, to find its estimated orbital period, we followed Winn et al. 2010 and derived an orbital period of $146.91_{47.45}^{56.74}$ d, assuming a circular orbit and using star density from the TESS Input Catalog. This period has a good agreement with the detected period in 146 d from the initial RV analysis.

To confirm the presence and properties of the observed single transit candidate, we conduct custom extractions of the two TESS sectors using the calibrated target pixel files (TPFs) that we retrieved using the default quality bitmask. We extracted target fluxes for a range of custom aperture masks created with radii of two to four pixels in steps of 0.1 pixels centered on the target. It should be noted that as the target does not fall in the exact center of a pixel increasing the aperture mask radius by 0.1 pixels can result in unique non-circular masks. All produced light curves were background-corrected after determining the sky level using custom background masks. We then detrended the data using two methods. Firstly, we conducted principal component analyses on the custom background masks to determine the scattered-light flux contribution to the light curves and then removed these systematics by using the principal components as basis vectors in a linear model. Secondly, we corrected flux modulation due to spacecraft jitter by retrieving the co-trending basis vectors (CBVs) and two-second cadence engineering quaternion measurements for the cameras HD88986 observed in. Following the method used in Delrez et al. (2021), we computed the means and averages of the quaternions over the scientific observational cadences and subsequently used these vectors along with the CBVs to remove any flux trends. The cleaned light curves (see Fig. 2 bottom panel) were then inspected to assess the transits across the aperture mask.

² <https://heasarc.gsfc.nasa.gov/cgi-bin/tess/webtess/wtv.py>

2.5. Photometric observations with CHEOPS

The *CHEOPS* spacecraft is a 30 cm ESA space telescope (Benz et al. 2021) that conducts ultra-high-precision photometry to characterize planets (Bonfanti et al. 2021; Delrez et al. 2021; Lacedelli et al. 2022) and their atmospheres (Lendl et al. 2020; Hooton et al. 2022), but it has also been used to aid in the discovery of new planets (Leleu et al. 2021; Osborn et al. 2022; Serrano et al. 2022; Wilson et al. 2022).

As we indicated in Sect. 2.4, the estimated period for the TESS single transit candidate is $146.91_{47.45}^{56.74}$ d, which is compatible with the candidate detected in 146 d from the initial RV analysis. Hence, it is likely that the observed TESS single transit represents the transit event of the RVs planet candidate in 146 d. In this case, our predicted next transiting event time would fall in the TESS sector 48's gap data. **Therefore, to search for a second transit** of the single transit seen in the TESS data, we obtained one visit of *CHEOPS* observation spanning 167.4 hr between 2022-02-08 and 2022-02-15 with an exposure time of 3.4 s. This allowed us to cover the transit period's uncertainty from RVs by 2.3 σ .

The data were processed with the latest version of the *CHEOPS* Data Reduction Pipeline (DRP v13; Hoyer et al. 2020) that conducts frame calibration, instrumental and environmental correction, and aperture photometry using pre-defined radii ($R = 22.5''$ [RINF], $25.0''$ [DEFAULT], and $30.0''$ [RSUP]) as well as a noise-optimised radius [ROPT]. The DRP produced flux contamination was subtracted from the light curves. We retrieved the data and corresponding instrumental basis vectors, and assessed the quality using the *pyCHEOPS* Python package (Maxted et al. 2022), and found that the DEFAULT aperture minimized the root mean square (RMS) noise. Therefore, we used these data for further analysis.

In previous studies it has been noted that environmental effects (i.e. spacecraft temperature and illumination) and the presence of nearby contaminants can induce flux modulation in light curves (Morris et al. 2021; Maxted et al. 2022; Wilson et al. 2022). In order to correct for these effects and search for the smallest transit signals in our transit search analysis, we conduct a principal component analysis on the auto-correlation function of the *CHEOPS* frames using the methodology detailed in Wilson et al. (2022). The process has been shown to monitor PSF shape changes and so any effects that alter the *CHEOPS* PSF, such as environmental and contamination effects, are measured by this tool and can be removed by using the produced principal components as the basis vectors in a linear model detrending. Further examples of applications of this tool can be seen in Fridlund et al. (submitted), Hawthorn et al. (submitted), Hoyer et al. (submitted), and Ehrenreich et al. (in prep.).

2.6. Gaia Astrometry

The European Space Agency's Gaia satellite provides a precise catalog of stars together with astronomical measurements. In Gaia EDR3 ~ 1 million new sources are detected including one new source (star A: 741184091113735424) at 1.4" west, of HD88986 (see Fig. 3). This star has $G=12.3$ mag and no values are reported for R_p magnitude and renormalized unit weight error (RUWE) which is a measurement of goodness of the astrometrical solution of the star. The poor behavior of this star could be due to blending in HD88986. We will take into account this light contribution to photometric data in our analysis in Sect. 3.3. In addition, we used Gaia Astrometry data to verify the origin of long-term drift seen in RVs (see Fig. 1) in Sect. 3.3.4.

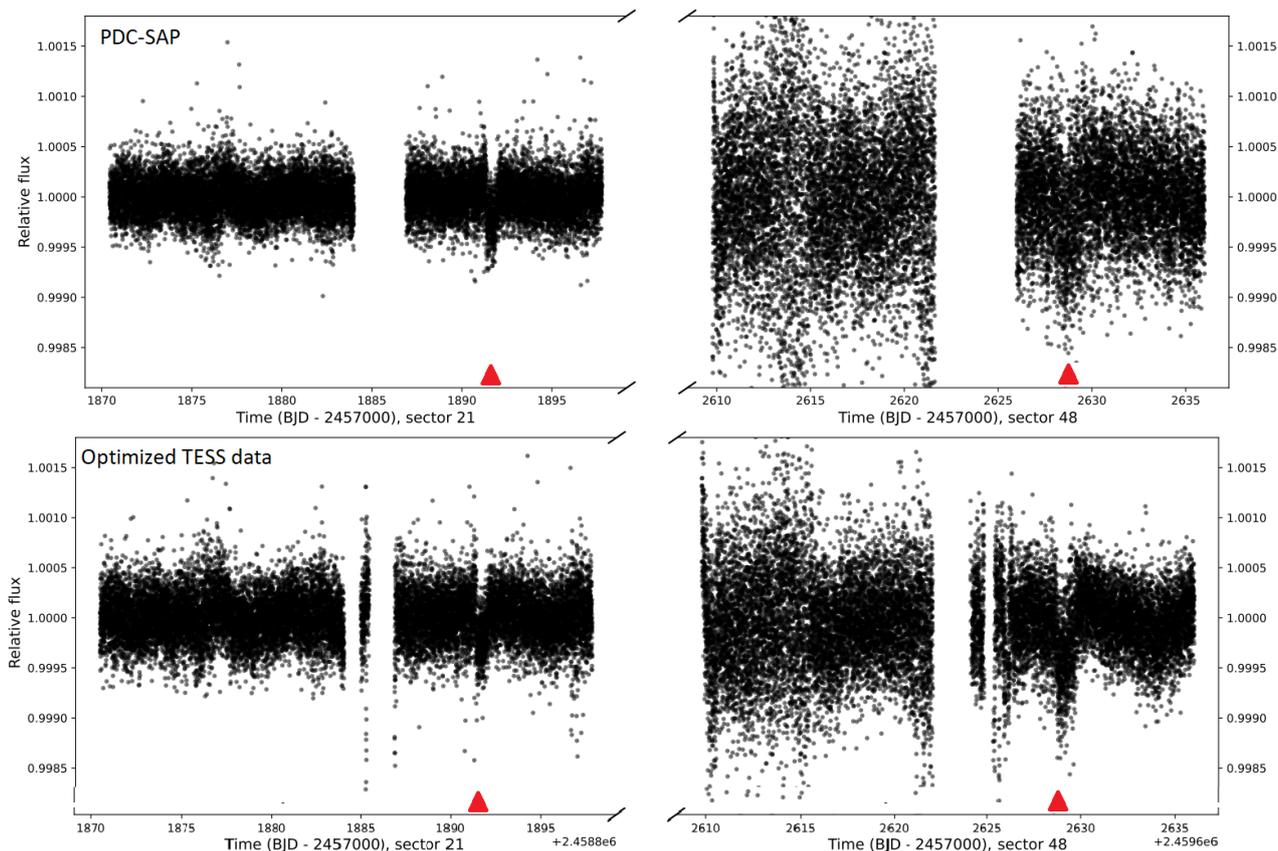


Fig. 2: TESS photometry light curve for HD88986, sector 21 on the left and sector 48 on the right. *Top panel*: The black points show the TESS PDC-SAP raw light curve. *Bottom panel*: The optimized light-curve used in this work (see Sect. 2.4). Red triangles show the positions of transit events.

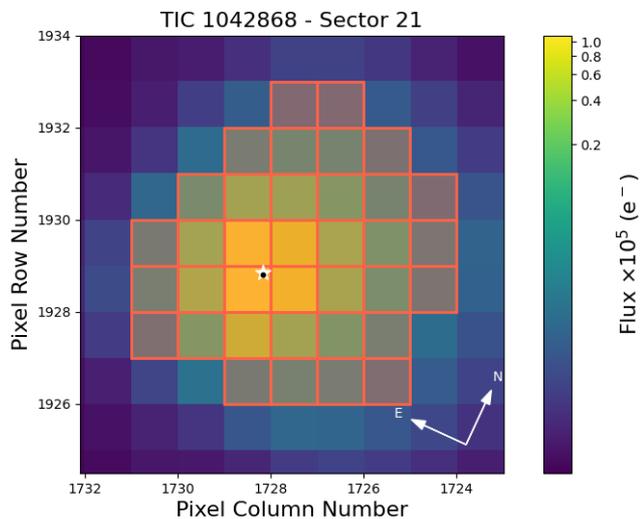


Fig. 3: Target pixel file of TESS for HD88986. While the star is shown by a white star mark, the star A (see Sect. 2.6) is illustrated by a black dot. Also, the red square presents the pipeline mask.

2.7. APT photometric observations

To help our search for confirming the origin of long-term drift observed in RVs (see Fig. 1), we also used 1335 good photometric observations of HD 88986 covering 21 observing sea-

sons from 1995-1996 to 2019-2020, except the four observing seasons 2015-16 through 2018-19. The observations were acquired with the T8 0.80 m automatic photoelectric telescope (APT) at Fairborn Observatory in southern Arizona. The T8 APT is equipped with a two-channel photometer that uses two EMI 9124QB bi-alkali photomultiplier tubes to measure the stellar brightness simultaneously in the Strömgren *b* and *y* passbands.

The observations are made differentially with respect to three nearby comparison stars. We measure the difference in brightness between our program star HD 88986 (star d) and the comparison stars (stars a, b, c) and create differential magnitudes in the following six combinations: d-a, d-b, d-c, c-a, c-b, and b-a. Intercomparison of these six light curves shows that the comparison star a (HD 89557) is the only one that appears to be constant to the limit of our precision, so we present our results as differential magnitudes in the sense star d minus star a, which we designate as d-a.

To improve the photometric precision of the individual nightly observations, we combine the differential *b* and *y* magnitudes into a single $(b + y)/2$ “passband”, designated by the .by filename extension on the APT data files. So our differential observations of HD 88986 are contained in the data file d-a.by. The precision of a single .by observation with T8, as measured from pairs of constant comparison stars, typically ranges between 0.001 mag and 0.0015 mag on good nights. The T8 APT is described in Henry (1999), where further details of the telescope, precision photometer, and observing and data reduction procedures can be found.

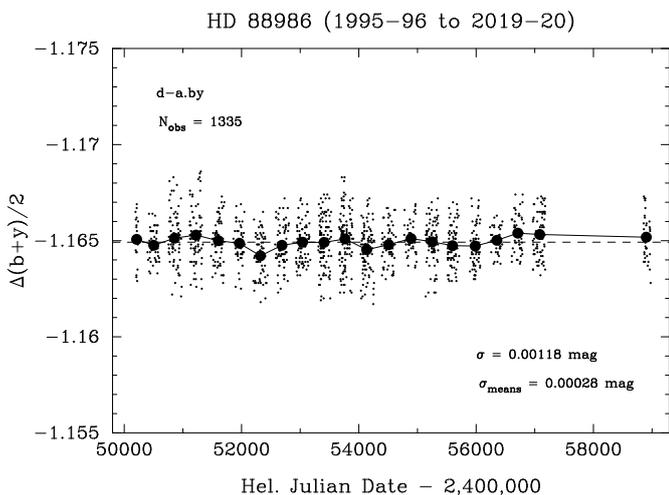


Fig. 4: The nightly Strömrgren $(b+y)/2$ band photometry (d-a.by) of HD 88986 from 21 observing seasons from 1995-96 to 2020-21 (small filled circles) scatter about their mean (dashed line) with a standard deviation of 0.00118 mag. Seasonal means from the 21 seasons (large filled circles) scatter about their mean with a standard deviation of 0.00028 mag.

Figure 4 plots the 1335 nightly d-a.by observations from the 21 observing seasons as small filled circles. The mean of all the nightly observations, -1.16492 mag, is plotted as the dashed line in the figure. The standard deviation of the nightly observations from their mean is 0.00118 mag, consistent with the precision of the measurements. The 21 seasonal means of the d-a.by observations are plotted as large filled circles. The standard deviations of the individual seasonal means are roughly the size of the plot symbols. The standard deviation of the 21 seasonal means from the mean of the seasonal means is 0.00028 mag, indicating that there is no long-term variability in HD 88986 to the limit of our photometric precision.

Table 1 summarizes the d-a.by observations of HD 88986. The standard deviations of the nightly observations for each observing season are given in column 4 and range from 0.00068 to 0.00146 mag, indicating little or no short-term variability within each observing season. The seasonal means plotted in Figure 1 are listed in column 5. Frequency analysis of each individual observing season using the method of Vaníček (1971) confirms the lack of any periodic variability (column 6). Henry et al. (2022) show extensive examples of this method of period analysis.

Figure 5 presents a frequency analysis of the full 21-year d-a.by data set. The top panel replots all nightly d-a.by observations. The middle panel plots the frequency spectrum of those observations and reveals no evidence for any significant periodicity between 10 and 300 days. The two small arrows mark the frequencies corresponding to the radial velocity period of 146.30 days presented in this paper (left arrow) and the $\log R'_{HK}$ period of 25 days (right arrow). The bottom panel presents a phase curve of the d-a.by photometry on the 146.30-day RV period. A least-squares sine fit gives a peak-to-peak amplitude of only 0.00024 ± 0.00009 mag. A separate sine fit on the 25.0-day period of the $\log R'_{HK}$ data (not shown) gives a peak-to-peak amplitude of 0.00013 ± 0.00009 mag. There is clearly no significant photometric variability in the vicinity of either period.

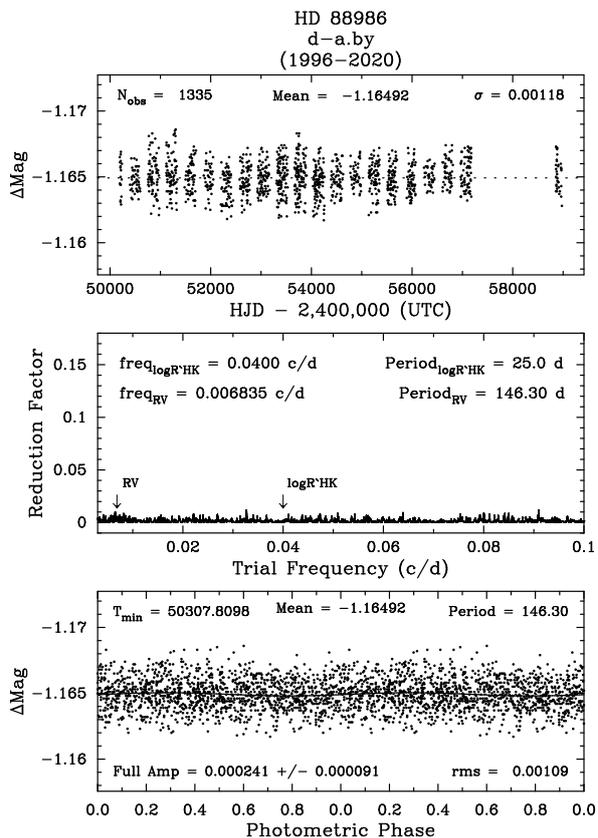


Fig. 5: Period analysis of all 21 observing seasons of the d-a.by data set for HD 88986 shows no evidence of any significant periodicity. See the text for an explanation of the three panels.

3. Analysis and results

3.1. Stellar parameters

HD88986 is a G2V type star with a V band magnitude of 6.5. To obtain the stellar atmospheric parameters, we co-added the spectra after correcting the RV variation of the star, barycentric Earth radial velocity, and background pollution due to the calibration lamp. It results in a high S/N per pixel spectrum of 3174.6 at 550 nm. Then, we calculated the T_{eff} and $[Fe/H]$ using the procedure described in Santos et al. (2013) and Sousa et al. (2018). The resulting T_{eff} and $[Fe/H]$, together with other stellar parameters of HD88986 are presented in Table 1.

As an independent determination of the basic stellar parameters, we performed an analysis of the broadband spectral energy distribution (SED) of the star together with the *Gaia* EDR3 parallax (with no systematic offset applied; see, e.g., Stassun & Torres 2021), in order to determine an empirical measurement of the stellar radius, following the procedures described in Stassun & Torres (2016); Stassun et al. (2017); Stassun & Torres (2018). We pulled the $B_T V_T$ magnitudes from *Tycho-2*, the JHK_S magnitudes from *2MASS*, the W1–W4 magnitudes from *WISE*, the $uvby$ Strömrgren magnitudes from Paunzen (2015), and the $GG_{BP}G_{RP}$ magnitudes from *Gaia*. We also used the UV measurement at 274 nm from the TD1 UV satellite. Together, the available photometry spans the full stellar SED over the wavelength range $0.2\text{--}22\ \mu\text{m}$ (see Figure 6).

We performed a fit using Kurucz stellar atmosphere models, with the free parameters being the effective temperature (T_{eff}), surface gravity ($\log g$), and metallicity ($[Fe/H]$), for which we adopted the spectroscopically determined values. The remaining

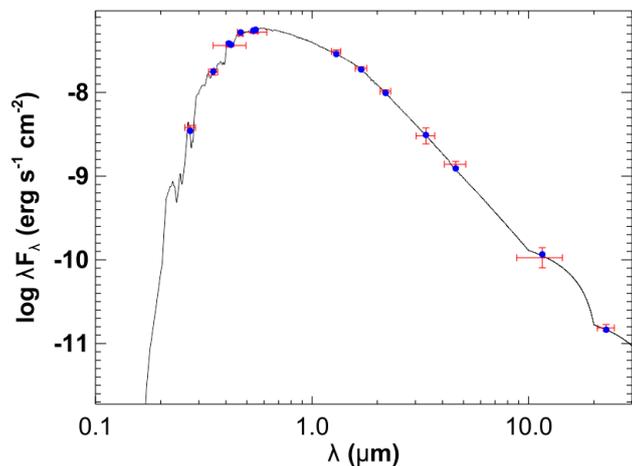


Fig. 6: Spectral energy distribution of HD88986. Red symbols represent the observed photometric measurements, whereas the horizontal bars represent the effective width of the passband. Blue symbols are the model fluxes from the best-fit Kurucz atmosphere model (black).

free parameter is the extinction A_V , which we fixed at zero due to the proximity of the system to Earth. The resulting fit (Figure 6) has a reduced χ^2 of 1.4. Integrating the model SED gives the bolometric flux at Earth, $F_{\text{bol}} = 7.28 \pm 0.17 \times 10^{-8} \text{ erg s}^{-1} \text{ cm}^{-2}$. Taking the F_{bol} and T_{eff} together with the *Gaia* parallax, gives the stellar radius, $R_{\star} = 1.543 \pm 0.010 R_{\odot}$. In addition, we can estimate the stellar mass from the empirical relations of Torres et al. (2010), giving $M_{\star} = 1.19 \pm 0.07 M_{\odot}$, which is consistent with the value of $1.25 \pm 0.05 M_{\odot}$ determined empirically via R_{\star} and $\log g$.

3.2. Radial velocity data analysis

3.2.1. RV data reduction on SOPHIE

The radial velocities (RV) were derived by SOPHIE data reduction (DRS, Bouchy et al. 2009) pipeline, including spectrum extraction, telluric line removal, CCD charge transfer inefficiency (CTI) correction, computation of cross-correlation function (CCF) of spectra with a G2 binary mask, and barycentric Earth RV correction. Then, the pipeline fits Gaussians on CCFs and extracts RVs (Baranne et al. 1996; Pepe et al. 2002). Prior to the cross-correlating calculation to extract RVs, we corrected the spectra from the atmospheric dispersion effect. Atmospheric dispersion can introduce a slope on the continuum and thus change the mean RV value of target (Pepe & Lovis 2008). To correct this, we used several templates which are a high-resolution spectrum of a list of the standard stars with different spectral types at low air mass (Modigliani et al. 2019; Wehbe et al. 2020). Then, we applied the template correction method that was first developed and tested on HARPS data. This method scales the target spectrum by multiplying it by the flux ratio of spectra and the corresponding template with the same spectral type. Applying this method led to 3 cm/s and 7.83 m/s improvements in the mean RMS of RVs and FWHM activity indicators, respectively.

Once the RVs are extracted, we corrected the nightly instrumental drift, measured by simultaneously calibration lamp observation. For this type of observation, SOPHIE fiber A observes the target and its fiber B monitor the calibration lamp. In this case, no sky observation is possible. However, detecting and

Table 1: Stellar properties of HD88986

Other identifiers		
	TIC 1042868	
	HD 88986	
	HIP 50316	
	Gaia DR3 741184091114529792	
	2MASS J10162809+2840571	
Parameter	HD88986	References
Astrometric properties		
Parallax (mas)	30.06070 ± 0.04920	<i>GaiaDR2*</i>
	$29.94818^* \pm 0.04120$	<i>GaiaEDR3</i>
Astrometric excess noise (mas)	0.135	<i>GaiaEDR3</i>
Significance of	20.6	<i>GaiaEDR3</i>
Distance	30.0307 ± 0.0412	Gaia DR2
α (hms)	10 : 16 : 28	Gaia DR2
δ (dms)	28 : 40 : 56.94	Gaia DR2
Photometric properties		
B-V	0.635 ± 0.006	HIP
V(mag)	6.46 ± 0.010	HIP
Gaia(mag)	6.2983 ± 0.0004	Gaia DR2
<i>Gaia</i> _{BP} (mag)	6.6534 ± 0.0037	Gaia DR2
<i>Gaia</i> _{RP} (mag)	5.8299 ± 0.0059	Gaia DR2
TESS(mag)	5.8706 ± 0.0061	TESS
J(mag)	5.247 ± 0.024	2MASS
H(mag)	4.946 ± 0.023	2MASS
K_s (mag)	4.884 ± 0.020	2MASS
W_1 (mag)	4.895 ± 0.239	WISE
W_2 (mag)	4.762 ± 0.085	WISE
W_3 (mag)	4.933 ± 0.014	WISE
W_4 (mag)	4.873 ± 0.029	WISE
Spectroscopic properties		
Spectral type	G0V	HIP
ξ_r (kms ⁻¹)	1.11 ± 0.02	Sec 3.1
$\log(R'_{HK})$	-5.07 ± 0.1	Sec 3.2.2
$v \sin i$ (kms ⁻¹)	3.3	SOPHIE DRS
[Fe/H] dex	0.06 ± 0.02	Sec 3.1
Bulk properties		
Mass (M_{sun})	1.19 ± 0.02	Sec 3.1
Radius (R_{sun})	1.673 ± 0.069	Sec 3.1
T_{eff} (K)	5861 ± 17	Sec 3.1
P_{rot} (days)	25^{+8}_{-6}	Sec 3.2.2

Notes. * We applied the offset correction for Gaia DR2 and Gaia EDR3 as prescribed in Lindegren et al. (2018) and Lindegren (2020), respectively.

flagging moon-contaminated spectra is essential for the accurate analysis of RV data. To find these affected spectra, we considered two criteria:

1. Moonlight contributes significantly to the target spectrum. It could be assumed when: 1) either the Moon phase is more than 68% on time of observation and the Sky-level is more than the mean of all observations, or 2) the Separation of the target and moon is less than 30 degrees. Note that the sky level is a criterion for estimating sky background light. It is calculated by collecting the CCD pixel of the main target those above a threshold and considering the rest as sky level. The sky level is calculated by the SOPHIE DRS and it is available in the header of each fits file spectrum. If this condition is fulfilled, the second criterion will be checked.

2. Barycentric Earth radial velocity (BERV) on the day of the observation, and in the direction of the target is close to target radial velocity $|RV_{\text{target}} - \text{BERV}| < 2 * \text{FWHM}$.

We note that mean of sky-level used here, is calculated after removing 3 σ outliers. Following this criteria, 29 spectra were identified to be contaminated by moonlight. Thus, we discarded them from our data. We note including/excluding these data does not change our final results. Moreover, 9 and 2 measurements

are also removed due to the low S/N ($S/N \lesssim 50$) and detected as outliers, respectively.

SOPHIE experience a long-term variation (Courcol et al. 2015), which tracks by observing every night so-called ‘constant stars’. We built a master time series from these constant stars on RVs, FWHM, and Bisector span and subtract it from the HD888986 data. The detail of our update on building master constant time series is described in thesis chapter 2.

3.2.2. Stellar rotation and activity

We used $\log(R'_{HK})$, $H\alpha$ index, S-index, full-width-at-half-maximum (FWHM), and CCF bisector to study host star activity and its rotational period. We obtained the bisector spans (BIS) and FWHM from SOPHIE DRS, which are calculated following Queloz et al. (2001). An S-index value is provided for the HIRES spectra. To extract the $\log(R'_{HK})$ and $H\alpha$ index, we followed Noyes et al. (1984) and Boisse et al. (2011), respectively. The key step before driving the $\log(R'_{HK})$ and $H\alpha$ index is subtracting the background light due to the diffuse of light from the Th-Ar or FP calibration lamp to the star spectra. To correct this, we directly measure this contamination with calibration lamp illusion for each lamp of Thorium-Argon and Fabry-Perot on fiber B, and no light on fiber A (Lovis et al. 2011). The same method was already applied for only the FP calibration lamp by Hobson (2019).

To estimate the rotational period of the star, we summed 96 HD888986 spectra which fulfilled two criteria. First, the spectra with $SN > 50$ in the first (bluest) order of spectra where CaII H&K lines are placed. Second, the spectra with less contamination due to the background light. This led to the value of $\log(R'_{HK}) = -5.07 \pm 0.10$. This value is in good agreement with the value of $\log(R'_{HK}) = -5.22$ and $\log(R'_{HK}) = -5.07$ reported by Radick et al. (2018) and Hall et al. (2007), respectively. **Finally**, I estimated rotational period of 25^{+8}_{-6} d following Noyes et al. (1984).

To constrain the stellar rotational period, we investigated the periodogram of RVs and activity indicators, using the website of the Data and Analysis Center for Exoplanets (DACE, Delisle et al. 2016)³. We excluded **212 data points** from $\log(R'_{HK})$ time series **because of** the $SN < 48$ in order 1 and highly contaminated by the calibration lamp. From the $H\alpha$ index time series, we discarded **210 data points** which suffer from both contaminations of the telluric line and also dependency on $SN (< 150$ in order 36). Additionally, **158 FWHM** data for those with $SN < 150$ is removed.

The periodogram of RVs, the RV residuals of Keplerian models, and the activity indicators are presented in Fig 7. We subsequently fit a keplerian model on the highest peak at each RV periodogram, which revealed three signals at the long-period of >2000 below the FAP level of 0.01 %, and two further signals at 146.4 d, and 29.5 d both below the FAP level of 1 % (see Fig. 7 first, second, and third panels). We note that the signal at 105 d is an alias of the 146.4 d signal (see Fig. 7 second panel). After removing the signal at 29.5 d, the RV residuals show no more periodic signals. The s-index and FWHM activity indicators show periodic signals inside of the estimated rotational period of the star, which is 29.6 with a FAP level of below 1 % for the S-index and 2 periodic signals of 33.2 and 35.6 days for FWHM. Note that the periodogram of the s-index also shows a weak activity signal at 141.1 d with FAP below 10 %. However, none of these signals correspond to the planet candidate in 146 d. Given the

two periodic signals of 32 days in the FWHM and 29.6 d in the S-index, and also the value of the estimated star rotation period, it is likely that the RV signal at 29.5 d is due to the stellar rotational period.

Additionally, we searched the SAP and PDC-SAP TESS light curves to find the signature of the rotational period of the star. No convincing signal was found. Note that the APT data also did not show any photometric variability related to the stellar rotational period (see Sect. 2.7).

3.2.3. RVs analysis

The RVs analysis has been carried out with *juliet* (Espinoza et al. 2019) which uses: *radvel* (Fulton et al. 2018) to model RVs, and *george* (Ambikasaran et al. 2015) and *celerite* (Foreman-Mackey et al. 2017) to model possible activity effect on the data through gaussian process methods (GPs). We modeled combined data by the following model:

$$M(t) = K(t) + \epsilon_i(t) + \mu_i, \quad (1)$$

where $K(t)$ is the Keplerian model and the $\epsilon_i(t) \sim \mathcal{N}(0, \sigma(t)^2 + \sigma_w^2)$, is a white-gaussian noise for instrument i which $\sigma(t)$ is the uncertainty of each RV points at the time of t and σ_w is a jitter term. Also, μ_i is a systematic RV offset of instrument i . Moreover, to explore the possible effect of stellar activity on the planet parameters, we used a GP model with a quasi-periodic Kernel as follows:

$$\kappa_{QP}(\tau) = \sigma_{GP}^2 \exp(-\alpha_{GP}\tau^2 - \Gamma \sin^2(\pi\tau/P_{rot})), \quad (2)$$

where σ_{GP} is the amplitude of the GP component with the unit of m/s for RVs. Γ is the amplitude of the GP for the sine-squared component of the kernel and it is unitless. α is the inverse length-scale of the GP exponential component in the unit of d^{-2} and P_{rot} is the period of the quasi-periodic (QP) GP kernel and given in d and τ is the time lag.

We tested different models on the data with a circular orbit. First, we applied the two-Keplerian model (2KP) on long-period and 146 d planet candidates. Then, we take into account the stellar noises through 1) the two-Keplerian model on long-period and 146 d planet candidates with a QP-GP model (2KP-QP), 2) the three-Keplerian model on long-period, 146 and 29.2 d signals (3KP). Our priors and the results of different models on RVs are presented in Table 2. All the parameters derived by the different models are well-consistent. However, since the model of 2KP+GP shows higher log evidence ($\ln Z$), is strongly favored over the others. Therefore, we consider the results of this model for the rest of this work. The Keplerian model on the two planet candidates is plotted in Fig. 8.

We will examine the different GP regressors (e.g. FWHM activity indicators) for our GP kernel.

3.3. Photometry data analysis

3.3.1. CHEOPS

To assess the existence of a transiting body with the *CHEOPS* observations, we conduct a statistically rigorous analysis using a newly developed tool (Hara et al. (in prep.)). In brief, we use the PSF-based PCA components (Wilson et al. 2022) produced above in combination with the instrumental basis vectors to construct a linear noise model that is fit simultaneously with either

³ Available at <https://dace.unige.ch>

Parameters	Prior	Combined data 2kp	Combined data 2KP+GP	Combined data 3kp
Posterior parameters for HD88986 b:				
P_b d	$\mathcal{U}(135, 155)$	146.58 $^{0.50}_{0.54}$	146.60 $^{0.50}_{0.51}$	146.42 $^{0.57}_{0.56}$
T_{0b} d	$\mathcal{U}(58870, 58870 + 150)$	58903.2 $^{4.5}_{4.6}$	58903.56 $^{4.13}_{4.75}$	58902.1 $^{4.4}_{4.9}$
K_b (m s $^{-1}$)	$\mathcal{U}(0, 10)$	1.53 \pm 0.26	1.54 $^{0.25}_{0.26}$	1.39 \pm 0.25
Posterior parameters for HD88986 c:				
P_c d	$\mathcal{U}(3000, 45000)$	12125.00 $^{338.48}_{338.49}$	12119.48 $^{341.28}_{285.30}$	12098.04 $^{323.88}_{281.67}$
T_{0c} d	$\mathcal{U}(14635, 59635)$	53488.6 $^{95.92}_{112.82}$	2453496.34 $^{96.26}_{110.93}$	53524.79 $^{90.60}_{108.44}$
K_c d	$\mathcal{U}(5, 100)$	39.02 $^{1.79}_{1.64}$	39.27 $^{1.8}_{1.7}$	39.67 $^{1.72}_{1.61}$
Posterior parameters for activity signal:				
$P_{activity}$ d	$\mathcal{U}(25, 35)$	—	—	29.20 $^{0.02}_{0.02}$
$K_{activity}$ d	$\mathcal{U}(0, 10)$	—	—	1.28 $^{0.24}_{0.23}$
Telescope Parameters:				
$\sigma_{SOPHIE+}$	$\mathcal{U}(1e - 3, 100.)$	2.87 $^{0.14}_{0.13}$	2.83 \pm 0.13	2.74 $^{0.13}_{0.12}$
σ_{SOPHIE}	$\mathcal{U}(1e - 3, 100.)$	7.11 $^{1.98}_{1.42}$	7.11 $^{2.08}_{1.44}$	6.61 $^{1.99}_{1.38}$
σ_{ELODIE}	$\mathcal{U}(1e - 3, 100.)$	3.99 $^{3.77}_{3.96}$	4.01 $^{3.73}_{3.98}$	2.32 $^{4.77}_{2.30}$
σ_{HIRES_b}	$\mathcal{U}(1e - 3, 100.)$	4.30 $^{1.10}_{0.79}$	4.34 $^{1.04}_{0.82}$	4.48 $^{1.14}_{0.82}$
σ_{HIRES_a}	$\mathcal{U}(1e - 3, 100.)$	6.8 $^{1.08}_{0.87}$	6.90 $^{1.05}_{0.90}$	6.98 $^{1.11}_{0.87}$
$\mu_{SOPHIE+}$	$\log\mathcal{U}(28995, 29196),$	29115.82 $^{1.94}_{1.81}$	29116.05 $^{1.96}_{1.84}$	29116.50 $^{1.91}_{1.81}$
μ_{SOPHIE}	$\mathcal{U}(28974, 29174)$	29089.27 $^{3.05}_{2.89}$	29089.27 $^{3.11}_{2.86}$	29089.09 $^{2.84}_{2.69}$
μ_{ELODIE}	$\mathcal{U}(28903, 29103),$	28971.13 $^{2.14}_{2.31}$	28970.81 $^{2.34}_{2.22}$	28970.62 $^{2.10}_{2.14}$
μ_{HIRES_b}	$\mathcal{U}(-100, 100),$	-14.93 $^{1.75}_{1.77}$	-15.21 \pm 1.73	4.48 $^{1.14}_{0.82}$
μ_{HIRES_a}	$\mathcal{U}(-100, 100)$	32.30 $^{2.01}_{1.98}$	32.41 $^{2.18}_{1.99}$	6.98 $^{1.11}_{0.87}$
$\ln Z$	—	-1265.01	-1259.60	-1263.15

Table 2: Median values and 68% confidence interval for the planetary system of HD88986 with different RV models. 2KP: two-Keplerian model on long-period and 146 d signals. 3KP: three-Keplerian model on long-period, 146 d and 29.2 d signals. 2KP+GP: two-Keplerian model on long-period and 146 d signals, with a QP-GP model (see Sect. 3.2.3). Gaussian priors labels of \mathcal{N} , \mathcal{U} , and $\log\mathcal{U}$ represent normal, uniform, and Jeffrey distribution, respectively.

a 0 or 1 planet transit model that allows us to compute the True and False Inclusion Probabilities (TIP and FIP; Hara et al. 2022) for the presence of transit in the data. These are calculated using the Bayes Evidences and posterior distributions of the 0 and 1 planet fits. For this study, we conduct this analysis twice; one with a constrained period prior to the transit model from the RV data and the other with a free prior. For both cases, for all transit T_0 values within the *CHEOPS* dataset we find $FIP \sim 1$, which statistically means that there is no transit in the lightcurve.

Despite the fact that the *CHEOPS* data did not show any transit features, through this observation we covered the TESS gap data in sector 48 which greatly help in concluding whether the RVs planet in 146.4 d is transiting or not.

3.3.2. TESS

We used the extracted photometry by the optimized aperture as described in Sect. 2.4. Two transit-like features clearly appear in the light curve, one at $T_0 = 58891.629$ and the other at $T_0 = 59628.77$. i.e. 5×147.4 days further (see Fig. 2).

To investigate the relation between these two possible transits, we fitted a transit model jointly on both transits through Juliet. To do this, we considered a wide uniform prior on R_p/R_* , a/R_* , and b . Uniform priors were also applied on quadratic limb-darkening parameters q_1 and q_2 as introduced in Kipping (2013). To take into account the correlated noise, we used a GP model with an approximate Matern kernel introduced in Foreman-Mackey et al. (2017). The reason for this choice is that there is no evidence of existing quasi-periodic oscillations in the

TESS light curves. Then, we compared the consistency of the model on the two transits. As a result, the two transit features are compatible with each other. **I will add the results after further exploring the two transit-like features.**

3.3.3. Transit attribution

Here, we closely investigate the relation between RVs planet at 146.4 d with the two detected transit-like features in the TESS data at $T_0 = 58891.629$ and $T_0 = 59628.77$, from now referred to as transit-1 and transit-2, respectively.

We used the posterior distribution of the expected transit time from the RVs model (Sect. 3.2.3), and also the posterior distribution of the predicted period from the transit model (Sect. 3.3.2). As it is shown in Fig. 9, the T_0 of transit-1 at 58891.629 is well inside of the posterior distribution of the expected transit time from the RVs. Additionally, the period of RV planet candidate in 146.4 days is located inside of the predicted period of transit model. This figure confirms that both conjunction times inferred from the RV model and T_0 inferred from transit fit are consistent with each other. Given these conjunctions, and also the estimated transit period following the theoretical formula of Winn et al. 2010, hereafter, we attribute the RV planet candidate in 146.4 d, HD88986 b, to the transit-1.

By accepting the attribution of transit-1 to HD88986 b, the 5th transit event should be in the TESS data in sector 48, and its covered gap by the *CHEOPS* data. Our analysis in Sect. 3.3 showed no transit events in sector 48 and the *CHEOPS* data, except the transit-2. This transit event is in 147.4 d periods of

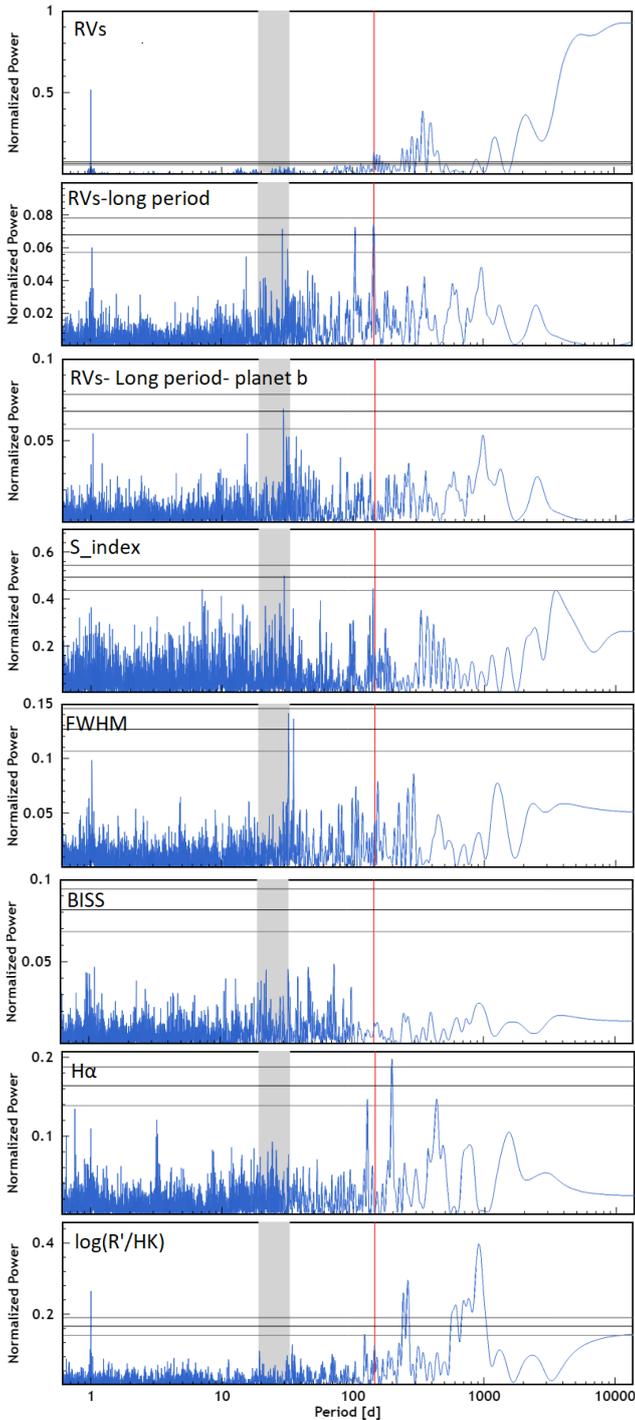


Fig. 7: Periodogram of RVs and activity indicators of HD88986. From *top to bottom*: RVs, residual of RVs after fitting a Keplerian fit on the long-period, residuals of RVs after 2-Keplerian fits on the long-period candidate and 146.1 d, S-index, FWHM, bisector, $H\alpha$ and $\log(R'_{HK})$ activity indicator. The vertical red line illustrates the planet candidates on 146.1 days which have no corresponding peak in activity indicators. The vertical gray strip marks the estimated rotational period of the star. Also, the horizontal lines show the FAP level of 10%, 1%, and 0.1 %, respectively.

transit-1 which is consistent with the period of HD88986 b in 2σ . Therefore, it is highly likely that the transit-2 is related to the HD88986 b and the true period of the planet is 147.4 d.

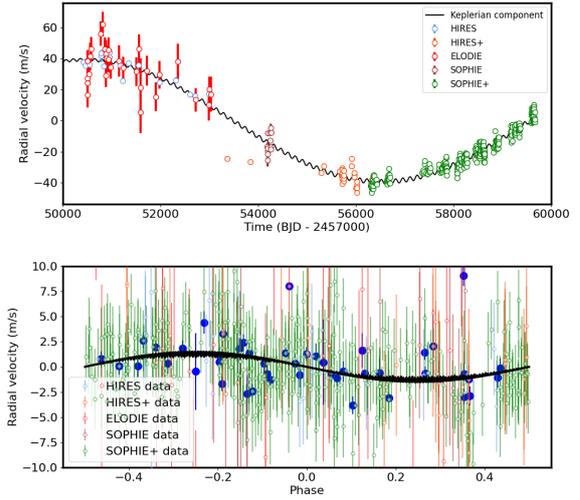


Fig. 8: *Top*: radial velocity time series of HD88986, over-plotted by Keplerian model. *Bottom*: the combined RVs phase-folded to the planet period 146.6 d. The black line represents the best fit Keplerian model on the data. Also, the blue points are the binned RV data.

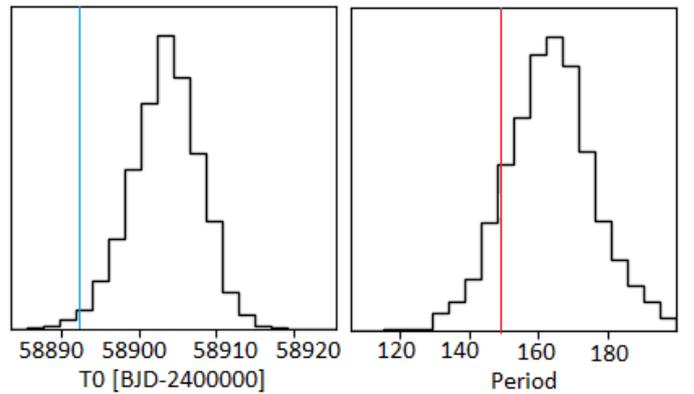


Fig. 9: (*Left*;) the expected T_0 from the RV fit for the planet candidate on 146.4 d. The blue color is the T_0 of the first transit feature detected in the TESS data. (*Right*;) the predicted period for the first transit from the transit modeling. The red color is the derived period from the RVs fit. This figure confirms that both conjunction times inferred from the RV model and T_0 inferred from transit fit are consistent with each other.

3.3.4. Constraining a long-term companion

We used the Gaia data simulator from the *gaston* code first developed for the Gaia DR1 (Kiefer et al. 2019, 2021) to test whether astrometric excess noises from the Gaia DR3 could lead to mass constraints on the companion of HD 88986. The astrometric excess noise is a measurement of supplementary motion, beyond proper motion and parallax is the astrometric data of a source. The astrometric excess noise is obtained from the RMS of residuals after fitting out the ra-dec position, proper motion, and parallax to the simulated astrometric Gaia measurements:

$$\sigma_{\text{AL}}^2 + \varepsilon_{\text{attitude}}^2 + \varepsilon_{\text{DR3}}^2 = \frac{\sum_j R_j^2}{N - 5} \quad (3)$$

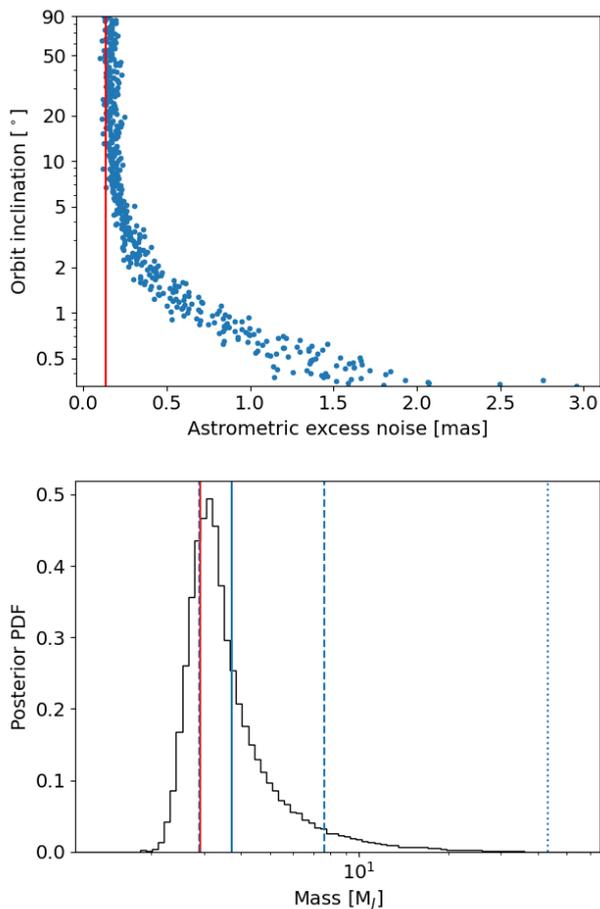


Fig. 10: Top: the relation between the simulated astrometric excess noise and the orbital inclination. The red line indicates the measured astrometric excess noise $\varepsilon_{\text{DR3}}=0.135$ mas. Bottom: the companion mass posterior distribution. The dotted line shows the 3- σ upper-limit, the dashed lines show the 1- σ confidence interval, the solid blue line is the median mass, and the solid red line shows the radial-velocity $m \sin i$.

where R_j are the N along-scan (AL) angle residuals of the astrometric fit; σ_{AL} is the typical AL angle measurement noise; $\varepsilon_{\text{attitude}}$ is the spacecraft attitude excess noise (e.g. Lindegren et al. 2021), and ε_{DR3} is the astrometric excess noise. The AL angle measurement noise has a conservative value of $\sigma_{\text{AL}}=0.1$ mas for targets with a G-magnitude within 6-11 (Fig. A.1 from Lindegren et al. 2021). The typical attitude noise in the DR3 is $\varepsilon_{\text{attitude}}=0.076$ mas (Lindegren et al. 2021).

HD88986 has a magnitude of G -mag ~ 6.3 and a color $Gb-Gr$ of ~ 0.8 . In the Gaia-DR3 catalog, the median astrometric excess noise at that magnitude and color for sources fitted with 5-parameters, as HD88986, is 0.16 mas. The astrometric excess noise of HD88986 is $\varepsilon_{\text{DR3}}=0.135$ mas. It is thus smaller than the median astrometric excess noise for the same type of sources. The Gaia DR3 astrometry of this target is thus compatible with a single star without a companion, but it will allow deriving an upper-limit constraint on the mass of the RV-detected companion.

We will follow the method from Kiefer et al. (2019, 2021), using the code `gaston`, adapted to the (E)DR3 (Lagrange et al. submitted). The general principle of the method is the same as with the DR1. Fixing P , $m \sin i$, e , ω , & T_0 within their priors

Parameter	1 σ	3 σ limits
I_c (degree)	52^{+26}_{-30}	>3.6
M_c (M_{jup})	4^{+4}_{-1}	<43
aphot (mass)	$0.9^{+1.0}_{-0.2}$	<10.7

Table 3: Resulting constraints on the orbital inclination, companion mass, and photocenter semi-major axis of HD 88986 from Gaia DR3.

derived from radial velocities only, we run several simulations of Gaia measurements of the target along the orbital motion of the system due to HD 88986 b and derive simulated values of astrometric excess noise that we compare with the actual ε_{DR3} .

We sample orbital inclination uniformly between 0 and 90° by an MCMC routine based on the `emcee` code (Foreman-Mackey et al. 2013). The orbital inclination changes the amplitude of the astrometric motion due to a different mass of the companion determined from $M=m \sin i/\sin i$ and thus changes the value of the astrometric excess noise. We calculate a log-likelihood that directly compares the simulated astrometric excess noise to its measured value in the DR3 within the MCMC routine.

$$\ln \mathcal{L} = \frac{1}{2} \left(\frac{(\varepsilon_{\text{simu}} - \varepsilon_{\text{DR3}})^2}{\sigma_{\varepsilon}^2} + \ln \sigma_{\varepsilon}^2 \right) \quad (4)$$

where σ_{ε} is a hyper-parameter of the MCMC. This allows a direct constraint on the possible orbital inclination and true mass of the RV companion.

Noises, epochs, scan angles, and the number of measurements used in the simulations are updated with respect to the new data reduction of DR3 (Gaia Collaboration et al. 2022). An epoch is a date when the star is transiting the Gaia field of view; several measurements, typically 9, are performed during a single transit. Those epochs can be found for any target in the Gaia Observation Forecast Tool (or GOST⁴). The median astrometric excess noise (0.16 mas) at the target's G -mag (6.3) and $Gb-Gr$ (0.8) within the 5-parameters solution dataset is used as a proxy for typical systematic noise, including instrumental and global modelization noise, that remains for all sources (Lindgren et al. 2021). In our simulations, it is modeled as a Gaussian noise changing every epoch of observation. The spacecraft attitude noise is also added to the model as a systematic Gaussian dispersion that changes every observation epoch with a standard deviation of 0.076 mas. A Gaussian measurement noise of $\sigma_{\text{AL}}=0.1$ mas is added to each of the N_{AL} astrometric measurements performed at a given epoch.

Table 3 summarises the results of astrometric excess noise fitting for this star. Fig. 11 shows the relation between astrometric excess noise and inclination in the simulations, and plots the posterior distribution of companion mass. The posterior distribution on mass gives an upper limit on the mass of the companion below 43 M_{jup} at 3- σ . The Gaia EDR3 data are nevertheless compatible with an edge-on configuration and the probability that HD88986 b is a planet with a mass smaller than 13.5 M_{jup} is $\sim 94\%$.

⁴ <https://gaia.esac.esa.int/gost/index.jsp>

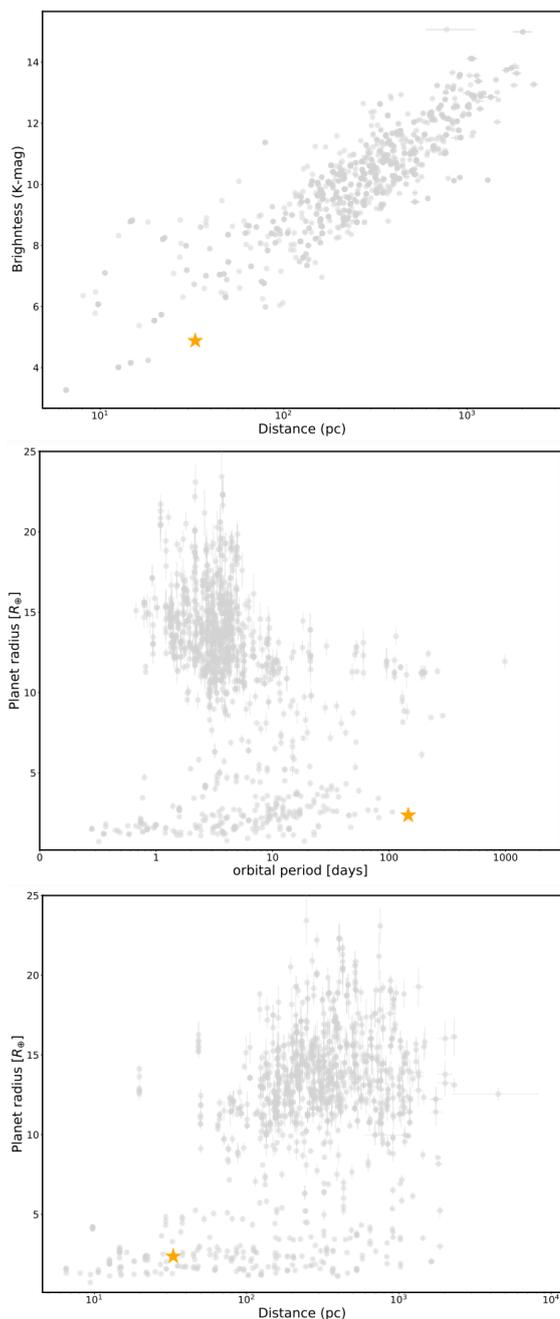


Fig. 11: HD88986 b and its host star in the context of known exoplanets from NASA archival data (February 2022) with accurate mass and radius (Otegi et al. 2020). *Top*: brightness versus distance of exoplanet host stars to Earth. *Middle*: planets radius versus their orbital periods. *Bottom*: planets radius versus the distance of their host stars to the Earth. These three views highlight how HD88986 b occupies an exciting parameter space for future studies.

4. summary and to-do list

We discovered a temperate sub-Neptune and a Jupiter mass planet orbiting HD88986, using several observations. HD88986 is a G0 type star, of solar metallicity, and it is one of the nearest and brightest exoplanet host stars ($G_{mag}=6.30$, $T_{eff}=5960^{280}_{180}$ K, $d=30.03$ pc). The sub-Neptune planet detected, HD88986 b is a transiting planet on the period of 147.4 d, which places it as

a temperate Neptune (assuming no albedo and perfect redistribution). In Fig. 11, we highlight the position of HD88986 b and its host star, in the context of known exoplanets from NASA archival data (February, 2022) with accurate mass and radius (Otegi et al. 2020). These three views highlight how HD88986 b occupies an exciting parameter space for future studies. This planet due to the brightness of its host star is a good candidate for atmospheric characterization via transmission spectroscopy.

We also using Gaia DR3 excess noise, determined the edge-on configuration and mass of the second planet which is smaller than 13.5 M_{Jup} with a probability of $\sim 94\%$. The occurrences rate of the systems with long-period giant planets (cold Jupiter) that harbor low-mass planets are high (Zhu & Wu 2018; Schlecker et al. 2021). The detected system is interesting for studying the relation between cold Jupiter and the inner orbit low-mass planet with properties of the host star

To explore the full parameters of the system, we will soon model jointly RVs and photometric data. We will also infer the interior structure of both planets and will investigate the system parameters with theories of planet formation and evolution.

References

- Ambikasaran, S., Foreman-Mackey, D., Greengard, L., Hogg, D. W., & O’Neil, M. 2015, IEEE transactions on pattern analysis and machine intelligence, 38, 252
- Baranne, A., Queloz, D., Mayor, M., et al. 1996, Astronomy and Astrophysics Supplement Series, 119, 373
- Benz, W., Broeg, C., Fortier, A., et al. 2021, Experimental Astronomy, 51, 109
- Boisse, I., Bouchy, F., Hébrard, G., et al. 2011, Astronomy & Astrophysics, 528, A4
- Bonfanti, A., Delrez, L., Hooton, M. J., et al. 2021, A&A, 646, A157
- Bouchy, F., Díaz, R., Hébrard, G., et al. 2013, Astronomy & Astrophysics, 549, A49
- Bouchy, F., Hébrard, G., Udry, S., et al. 2009, Astronomy & Astrophysics, 505, 853
- Butler, R. P., Vogt, S. S., Laughlin, G., et al. 2017, The Astronomical Journal, 153, 208
- Courcol, B., Bouchy, F., Pepe, F., et al. 2015, Astronomy & Astrophysics, 581, A38
- Delisle, J.-B., Ségransan, D., Buchschacher, N., & Alesina, F. 2016, Astronomy & Astrophysics, 590, A134
- Delrez, L., Ehrenreich, D., Alibert, Y., et al. 2021, Nature Astronomy, 5, 775
- Espinoza, N., Kossakowski, D., & Brahm, R. 2019, Monthly Notices of the Royal Astronomical Society, 490, 2262
- Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, The Astronomical Journal, 154, 220
- Fulton, B. J., Petigura, E. A., Blunt, S., & Sinukoff, E. 2018, Publications of the Astronomical Society of the Pacific, 130, 044504
- Hall, J. C., Lockwood, G., & Skiff, B. A. 2007, The Astronomical Journal, 133, 862
- Hara, N. C., Unger, N., Delisle, J.-B., Díaz, R. F., & Ségransan, D. 2022, A&A, 663, A14
- Heidari, N., Boisse, I., Orell-Mique, J., et al. 2021, arXiv preprint arXiv:2110.08597
- Henry, G. W. 1999, Publications of the Astronomical Society of the Pacific, 111, 845
- Henry, G. W., Fekel, F. C., & Williamson, M. H. 2022, The Astronomical Journal, 163, 180
- Hobson, M. J. 2019, PhD thesis, Aix-Marseille
- Hooton, M. J., Hoyer, S., Kitzmann, D., et al. 2022, A&A, 658, A75
- Hoyer, S., Guterman, P., Demangeon, O., et al. 2020, A&A, 635, A24
- Jenkins, J. M., Twicken, J. D., McCauliff, S., et al. 2016, in Software and Cyberinfrastructure for Astronomy IV, Vol. 9913, International Society for Optics and Photonics, 99133E
- Kipping, D. M. 2013, Monthly Notices of the Royal Astronomical Society, 435, 2152
- Lacedelli, G., Wilson, T. G., Malavolta, L., et al. 2022, MNRAS, 511, 4551
- Leleu, A., Alibert, Y., Hara, N. C., et al. 2021, A&A, 649, A26
- Lendl, M., Csizmadia, S., Deline, A., et al. 2020, A&A, 643, A94
- Lindgren, L. 2020, Astronomy & Astrophysics, 633, A1
- Lindgren, L., Hernández, J., Bombrun, A., et al. 2018, Astronomy & Astrophysics, 616, A2
- Lovis, C., Dumusque, X., Santos, N., et al. 2011, arXiv preprint arXiv:1107.5325

- Maxted, P. F. L., Ehrenreich, D., Wilson, T. G., et al. 2022, *MNRAS*, 514, 77
- Mayor, M. & Queloz, D. 1995, *Nature*, 378, 355
- Modigliani, A., Sownowska, D., & Lovis, C. 2019, *ESPRESSO Pipeline User Manual*
- Morris, B. M., Delrez, L., Brandeker, A., et al. 2021, *A&A*, 653, A173
- Noyes, R., Hartmann, L., Baliunas, S., Duncan, D., & Vaughan, A. 1984, *The Astrophysical Journal*, 279, 763
- Osborn, H. P., Bonfanti, A., Gandolfi, D., et al. 2022, arXiv e-prints, arXiv:2203.03194
- Otegi, J., Bouchy, F., & Helled, R. 2020, *Astronomy & Astrophysics*, 634, A43
- Paunzen, E. 2015, *Astronomy & Astrophysics*, 580, A23
- Pepe, F. & Lovis, C. 2008, *Physica Scripta*, 2008, 014007
- Pepe, F., Mayor, M., Galland, F., et al. 2002, *Astronomy & Astrophysics*, 388, 632
- Queloz, D., Henry, G. W., Sivan, J.-P., et al. 2001, *Astronomy & Astrophysics*, 379, 279
- Radick, R. R., Lockwood, G. W., Henry, G. W., Hall, J. C., & Pevtsov, A. A. 2018, *The Astrophysical Journal*, 855, 75
- Santos, N., Sousa, S., Mortier, A., et al. 2013, *Astronomy & Astrophysics*, 556, A150
- Schlecker, M., Mordasini, C., Emsenhuber, A., et al. 2021, *Astronomy & Astrophysics*, 656, A71
- Serrano, L. M., Gandolfi, D., Hoyer, S., et al. 2022, arXiv e-prints, arXiv:2207.13920
- Smith, J. C., Stumpe, M. C., Van Cleve, J. E., et al. 2012, *Publications of the Astronomical Society of the Pacific*, 124, 1000
- Sousa, S., Adibekyan, V., Delgado-Mena, E., et al. 2018, *Astronomy & Astrophysics*, 620, A58
- Stassun, K. G., Collins, K. A., & Gaudi, B. S. 2017, *The Astronomical Journal*, 153, 136
- Stassun, K. G. & Torres, G. 2016, *The Astronomical Journal*, 152, 180
- Stassun, K. G. & Torres, G. 2018, *The Astrophysical Journal*, 862, 61
- Stassun, K. G. & Torres, G. 2021, *The Astrophysical Journal Letters*, 907, L33
- Stumpe, M. C., Smith, J. C., Catanzarite, J. H., et al. 2014, *Publications of the Astronomical Society of the Pacific*, 126, 100
- Stumpe, M. C., Smith, J. C., Van Cleve, J. E., et al. 2012, *PASP*, 124, 985
- Tal-Or, L., Trifonov, T., Zucker, S., Mazeh, T., & Zechmeister, M. 2019, *Monthly Notices of the Royal Astronomical Society: Letters*, 484, L8
- Torres, G., Andersen, J., & Giménez, A. 2010, *The Astronomy and Astrophysics Review*, 18, 67
- Vaniček, P. 1971, *Astrophysics and Space Science*, 12, 10
- Wehbe, B., Cabral, A., Martins, J., et al. 2020, *Monthly Notices of the Royal Astronomical Society*, 491, 3515
- Wilson, T. G., Goffo, E., Alibert, Y., et al. 2022, *MNRAS*, 511, 1043
- Zhu, W. & Wu, Y. 2018, *The Astronomical Journal*, 156, 92

¹ Department of Physics, Shahid Beheshti University, Tehran, Iran. e-mail: neda.heidari@lam.fr

² Laboratoire J.-L. Lagrange, Observatoire de la Côte d'Azur (OCA), Université de Nice-Sophia Antipolis (UNS), CNRS, Campus Valrose, 06108 Nice Cedex 2, France.

³ Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France.

⁴ SUPA, School of Physics and Astronomy, University of St. Andrews, North Haugh, Fife KY16 9SS, UK.

⁵ Sorbonne Université, CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98 bis bd Arago, 75014 Paris, France.

3.6.3 A transiting planet candidate

Photometry: HD6660 (TOI-2009) is a nearby (20.52 pc), bright ($G_{mag} = 8.03$) K4V star. From the TESS input catalog the star has $T_{eff} = 4755 \pm 99$ K, $R_* = 0.70 \pm 0.04 R_{Sun}$ and $M_* = 0.76 \pm 0.08 M_{Sun}$. This star was observed by TESS in sector 17 from October 7th to November 2nd, 2019. It is scheduled to be re-observed in sector 57 from September 30 to October 29, 2022. After the observation of sector 17, the community and also the MIT Quick Look pipeline (QLP- Huang et al., 2020b; Huang et al., 2020a) detected the signature of a single transit event in the TESS data. Then, an alert was issued and the target was named TOI-2009.

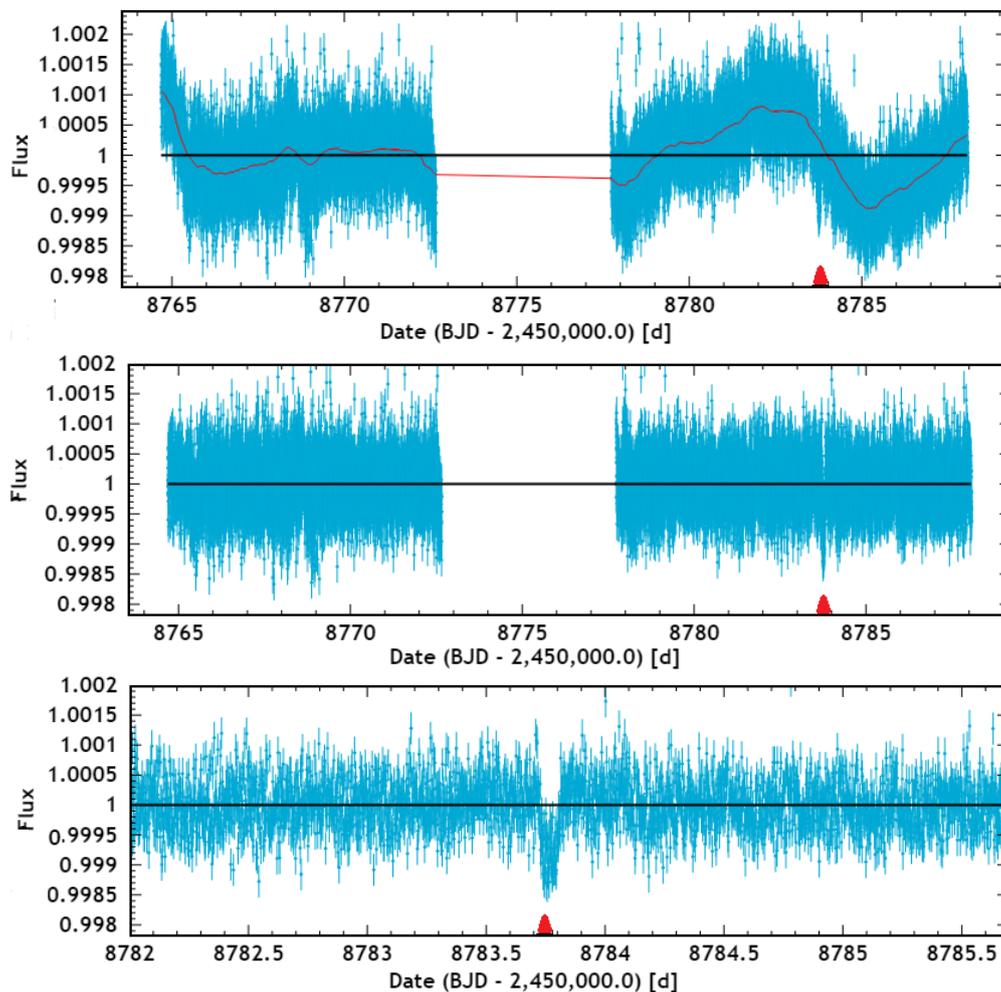


Figure 3.13 – TESS data of HD6660. *Top:* PDC-SAP flux over-plotted with the best-detrended model in red. *Middle:* the final detrended light curve. *Bottom:* Zoom in on the single transit event. In the overall plot, the transit event is marked by a red triangle.

Here, I used the 2-minutes cadence TESS light curve extracted by the PDC-SAP pipeline via the DACE website. I removed 5σ outliers and then detrended the light curve with 8-hour boxes (see Fig. 3.13 top and middle). A zoomed view of the single transit event is shown in Fig. 3.13 bottom. According to the ExoFOP website², the transit occurs at $T_0 = 58783.772719$ (BJD) with a depth= 630.72 ± 66.55 and a duration of 2.57 ± 0.32 hours. I estimated the planet's orbital period from the transit duration following Winn, 2010a and derived an orbital period of $6.2_{3.0}^{4.7}$ d. I note that this period is calculated under the assumption of circular orbit and using the star density from the TESS Input Catalog. The lower band of the orbital period can be excluded due to the fact that there is only one transit during the ~ 24 d TESS observation, which is divided into ~ 8 d and ~ 10 d. However, the exact planet period can only be determined by either observing a second transit event, which might be possible in the TESS sector 57 observation, or by measuring an RV orbit.

RVs: This star was monitored by the SOPHIE spectrograph between 2014 to 2022, collecting 46 high-resolution spectra. Two data points were excluded, one because it was identified as an outlier and the other because it did not meet the required $S/N_{550} > 50$. For data processing, I followed the steps mentioned in Sect. 3.3. The final RV RMS is 6.8 m s^{-1} with a mean error bars of 1.5 m s^{-1} .

The RVs are shown in Fig. 3.14 top, and present a linear drift of $1.22 \pm 0.36 \text{ m s}^{-1} \text{ yr}^{-1}$. The linear drift is also seen in the $H\alpha$ activity indicator (Fig. 3.14, bottom), but no correlation ($R= 0.1$) is found between RVs and $H\alpha$. After removing the drift from the RVs and $H\alpha$, I searched for periodic signals by performing the periodogram for the RVs and activity indicators (see Fig. 3.14, panels 3-5). No long-period signal is seen in the periodogram of activity indicators. Therefore, the origin of the RV drift is currently uncertain. In the RVs periodogram, I found two bulk signals with the highest peaks at 15.8 d and 29.1 d with a FAP of below 10 %. These signals have no corresponding peak in the activity indicators. However, more observations of the target are required to determine the cause of the drift, and statistically, the RV signals become stronger, allowing the period of the single transit planet to be investigated. Hence, I proposed that a higher priority be assigned to this star to be observed by SOPHIE. We are currently observing this star.

2. <https://exofop.ipac.caltech.edu/tess/target.php?id=243187830>

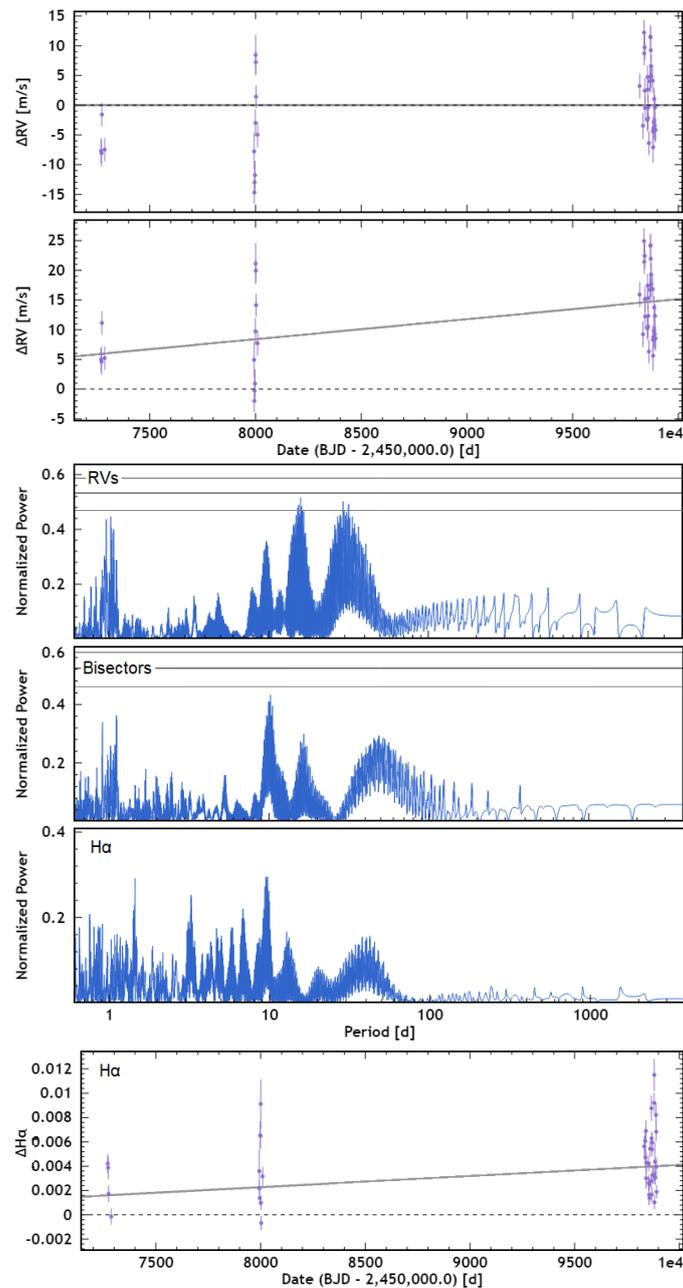


Figure 3.14 – The RV time series, RVs and activity indicators periodogram, and H α time series of HD6660.

3.7 Other contributions to SOPHIE RV programs

During the development of this thesis, I acquired several skills including reducing data and improving the RV's precision, and also detecting and characterizing low-mass

planets. These skills allowed me to contribute to other SOPHIE exoplanet programs, which are described in the following sections.

3.7.1 RV search around binary stars

One of the SOPHIE exoplanet programs is an RV survey to find circumbinary planets, named BEBOP (Binaries Escorted By Orbiting Planets, (Martin et al., 2019)). The goal of BEBOP is to understand more about the occurrence rate, multiplicity, formation, and evolution of circumbinary planets, as well as the relationship between binary and planet properties. Since 2016, BEBOP surveys a series of binary stars, with a target sample of 94 single-lined and 8 double-lined binaries. They are composed of a sun-like primary with a very low-mass secondary star. The binary orbits range from 5 to 50 days and the planets are expected in the 40 to 300 days.

I developed and performed an automated pipeline to extract all RVs of the BEBOP star catalog using optimized DRS (see chapter 2), as well as identify and correct several instrumental effects from data. This pipeline has the following steps for the single-lined binaries:

- Extracting the RV data with a wide CCF window (typically=100 Km/s) and performing the CTI correction. This wide CCF is essential for extracting the RVs of binaries due to their high movements. If the observation is in simultaneous calibration mode, the pipeline flags the moon contaminated dates; otherwise, for simultaneous sky observation mode it applies the moon correction method if the spectrum is detected as contaminated (see the Sect. 2.4 for the full description of optimized conditions for detecting moon contaminated spectra).
- If target observation is performed with simulation calibration lamp observation, the pipeline subtracts the measured drift. For those target observations with simultaneous sky monitoring, I developed an automatic python program to estimate the spectrograph drift at the time of observation. To do so, I use two consecutive calibration exposures before and after target observation, with regard to the different calibration lamps (Th-Ar or FP). Th-Ar and FP calibration lamps both measure the spectrograph drift but there is a conceptual difference. When Th-Ar exposure is taken, a new wavelength solution generates and the drift will be calculated with respect to the last Th-Ar exposure. The FP, on the other hand, measures the drift with respect to a zero-point which is typically set from wavelength solution in the afternoon calibration.
- Performing the optimized master constant correction on RV (see Sect. 2.6,) by considering the observation mode (HR or HE).

- Saving the final reduced RVs in the ASCII files.

For the double-lined binaries, except for extracting the RVs, all the other steps will be done, and measurements from each step will be stored in ASCII files. For extracting the RVs, one should carefully consider the relative movement of the two CCFs from each other (Fernandez et al., 2017), which we measure outside of the pipeline and one by one. Additionally, for these stars, the pipeline will also provide a CTI correction estimated following Santerne et al., 2012.

Our work on the BEBOP program marks the first RV detection of a circumbinary planet (Triaud et al., 2022), with gathering 144 RV measurements. This planet was already known with 3 transits within the KEPLER light curve (Doyle et al., 2011). However, the RV method was unable to detect it previously. The planet has a mass of $0.313 \pm 0.039 M_{Jup}$ and radius of $0.753 \pm 0.002 R_{Jup}$. It orbits with a period of 228.3 ± 1.8 days around both stars and the host stars also orbit around each other with a period of 41.07 ± 54 d. Note that this work has been the subject of several Press Releases that one of which is provided in Appendix 5 (in french).

In this work which I carried out its full data processing, we show that it is possible to achieve photon-noise precision and accuracy on single-lined binaries, with a final precision of 1.5 m/s. To demonstrate the great improvement of the signal by the processing methods which I have developed and applied, I plotted the Kepler16 time series and associated periodogram before and after the reductions in Fig. 3.15. As clearly appears the signal after reduction reaches the FAP of 0.01 %.

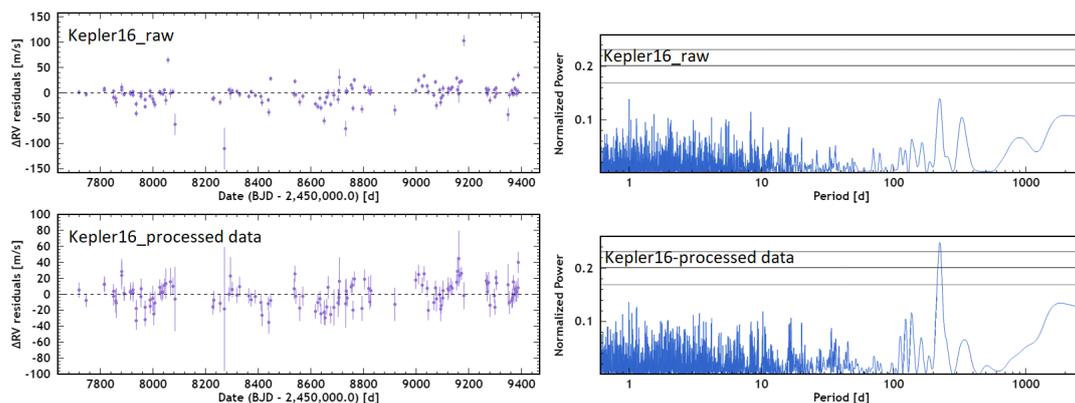


Figure 3.15 – The RV data of Kepler16 and its corresponding periodogram before (*first row*) and after processing (*second row*) are displayed. The great improvements in the signal clearly appear by comparing the periodogram of raw and processed data.

These figures, clearly show the efficiency of the data reduction methods and

pipeline which I have been developing during this thesis. In every semester, using this automated pipeline, I carry out the data processing of the BEBOP stars. These great improvements in the data, lead to the detection of several new candidates and consequently well characterize them.

3.7.2 High precision RV search around M dwarfs

M dwarfs are suitable for searching temperate low-mass planets for two primary reasons. First, these stars have low mass (from 0.07 to 0.6 M_{sun} (Reid et al., 2013)) compared to other stars. According to equation 1.5, because the induced RV amplitude is inversely proportional to the mass of the star, the smaller mass of M dwarfs will help in the detection of smaller planets from their RV measurements. Second, due to the low luminosity of M dwarfs, the habitable zone of these stars is close to the host star, and planets therein have short periods. Since according to equation (1.5), shorter orbital period planets can produce stronger signals, this will facilitate the detection of habitable zone planets around M dwarfs. Therefore, M dwarfs are good targets for detecting low-mass planets in habitable zone (Gillon et al., 2017; Dittmann et al., 2017).

SOPHIE has a dedicated program to search for low-mass planets around M dwarfs, named SP3 (PI: X. Delfosse). The target list of this sub-program contains 180 M dwarfs, at a distance of less than 12 pc, with a magnitude of $V < 14.0$. During the development of my thesis, I contributed to this study. I have developed an automated pipeline for processing data from this program, which contains the same steps as the BEBOP program (see Sect. 3.7.1). I also studied two targets of this program in detail.

Gl686: The star has an M1 spectral type with a magnitude of $V = 9.5$. We observed the target between August 2011 and September 2018 and collected 178 high-resolution spectra. The target observations were performed with simultaneous sky observation to monitor the sky's background brightness.

The cross-correlation function is widely used to derive RVs. However, for M and late-type K dwarfs might not be a sufficient method. The lower effective temperatures of these stars enhance the number of different molecular lines and produce complicated spectra. In this case, the binary mask which is used for the cross-correlation method might miss some Doppler information. For improving the accuracy of RVs, we extracted data through the template matching algorithm. This template extracts RVs through a χ^2 -minimization approach of the difference between RVs determined by the CCF method as a first guess and a reconstructed stellar and telluric template from the observed spectra (Astudillo-Defru et al., 2017c; Astudillo-Defru et al., 2015).

Table 3.3 – Resulting parameters of fitting Keplerian model.

parameters	Units	This work (combined data)	Lalitha et al., 2019
P	day	15.53 ± 0.00	15.53 ± 0.00
K	m/s	2.72 ± 0.10	$3.02^{+0.18}_{-0.20}$
e		0.18 ± 0.03	$0.08^{+0.08}_{-0.09}$
a	Au	0.12	$0.92^{+0.002}_{-0.002}$
$M \sin i$	M_{Earth}	5.91 ± 0.25	$6.64^{+0.53}_{-0.54}$

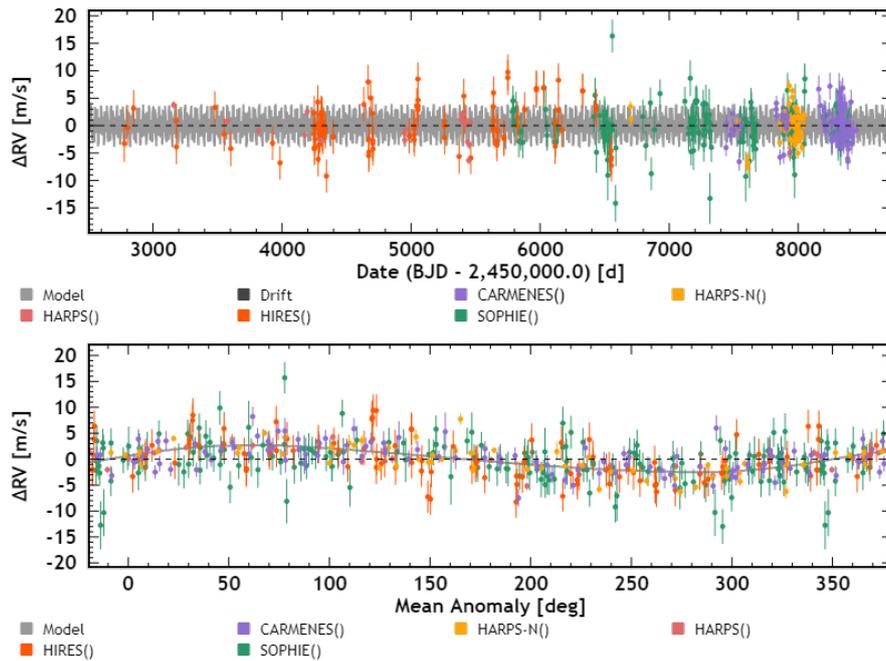


Figure 3.16 – *Top*: the combined RV time series of Gl686 from different instruments. *Bottom*: Keplerian fit on the planet signal at 15.53 d.

Additionally, this star has also been monitored by HIRES, HARPS, and HARPS-N for over 20 years, and the discovery of a super-Earth was reported by Affer et al., 2019 and Butler et al., 2017. Later, Lalitha et al., 2019 also combined all the data with CARMENES and refined the orbital parameters of Gl686 b.

Here, I combined the all dataset with SOPHIE and refine the parameter. The results are presented in the Table. 3.3. As it is shown in this table, my results are in 1σ agreements with Lalitha et al., 2019. I found a slightly smaller and more precise RV semi-amplitude of 2.72 ± 0.10 m/s, resulting in a lower minimum mass of the planet $5.91 \pm 0.25 M_E$. The mass uncertainty has been reduced by more than 1σ (from 0.53 to 0.25), which is very important in the internal structure modeling of planets.

Within the residuals, I found several signals below FAP 0.01 %: 1) A bulk of signals on 40.9 days which are consistent with the rotational period of star from activity indicators. It is also noted by Lalitha et al., 2019. 2) Several signals at longer-period (1026.8 d, 2188.4 d). The highest one is on 1026.8 d which is also compatible with the reported signal by Lalitha et al., 2019 at 1161_{-81}^{+53} d. This signal could be due to an imperfect fitted offset between different instruments; existing of a second long-period planet; or an activity cycle. Further observations are needed to conclude this signal which is why we continue to observe this star.

StarA: This star is a single-lined binary system that is located about 6.14 pc away from our Sun. The starA and its B stellar companion are separated by 103 AU and both are flare stars. While starA has a K7 spectral type with a magnitude of $V=7.0$, starB has a spectral type of M0V with a magnitude of $V=7.1$.

I used 133 RV measurements for this star between October 2011 to March 2019. The target observations were done with simultaneous calibration lamp observation (FP or Th-Ar), to track the spectrograph drift. The same as star Gl686, we used template matching to extract the RVs.

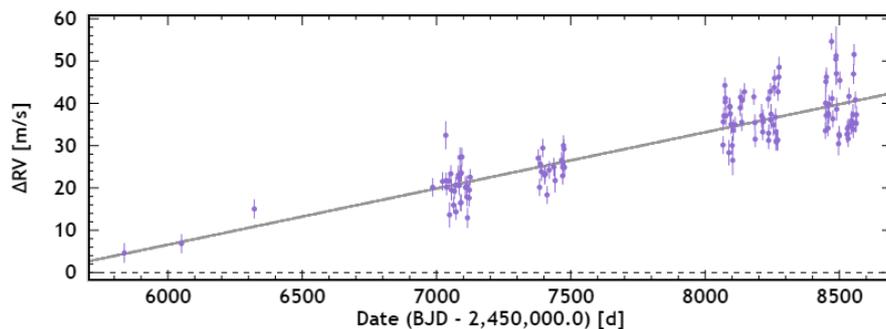


Figure 3.17 – RV time series of starA overplotted by the best fitted linear drift. the observed drift is likely due to the Keplerian orbit of the StarB, which is not completed yet.

I removed a clear linear trend from RVs with a slope of $4.84 \pm 0.26 \text{ m s}^{-1} \text{ y r}^{-1}$ (Fig. 3.17). This trend is likely due to the Keplerian orbit of the starB, which is not completed yet.

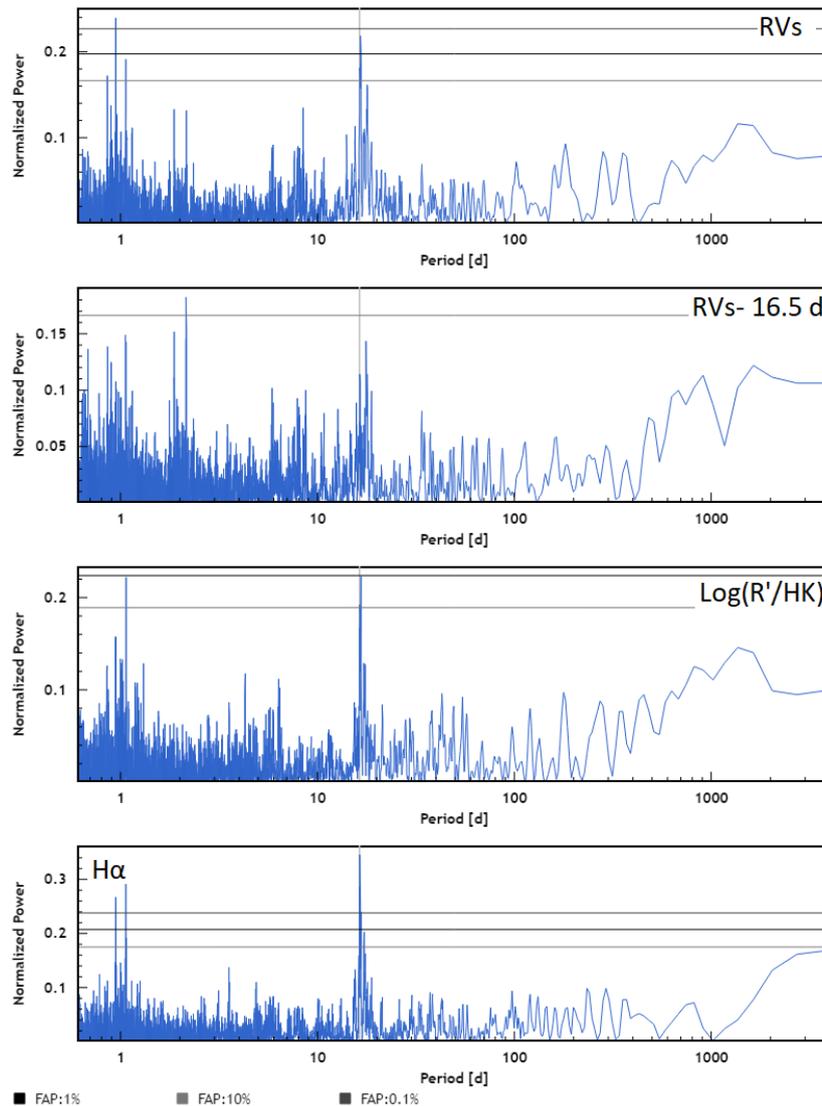


Figure 3.18 – Periodogram of RVs and activity indicators of starA. The highest peak at the RVs periodogram at 16.5 days has corresponding peaks in the activity indicators.

Fig. 3.18, illustrate the RVs periodogram of starA (*first panel*), its residual (*second panel*), $H\alpha$, and $\log(R'_{HK})$ activity index (*third and forth panels*, respectively). The highest peak in the periodogram of RVs is at 16.5 d with FAP below 0.1 % which has corresponding signals at $H\alpha$ and $\log(R'_{HK})$ activity indicators. Both activity indicators of $H\alpha$ and $\log(R'_{HK})$ have a correlation of 0.3 and 0.5 with RVs, respectively. Therefore, it is likely that this signal at 16.5 d is due to the activity of the star. After removing the signal at 16.5 d (Fig. 3.18 middle), the periodogram show a signal at 2.14 d with FAP below 10 % which has no corresponding peak at activity indicators.

I performed another test which was subdividing the RVs into two consecutive subsets, containing 47 points in 2015-2016 and 83 points in 2017-2018 (Fig. 3.19), and removing 3 points at the beginning. If the signal at 2.14 d is due to the activity of the host star, it is likely because of the changing activity of the star during 3 years (a total time span of observation), it will disappear in one of the subsets. As a result, both data sets showed a signal at 2.14 d. If we assume the signal is a planet, the Keplerian fit shows a minimum mass of $4.03 M_E$ with an eccentric orbit. However, since the signal is not significant yet, more observation of the target is essential to confirm the origin of this signal. I also searched the TESS photometric observation of the target which did not lead to the finding of the planet sign.

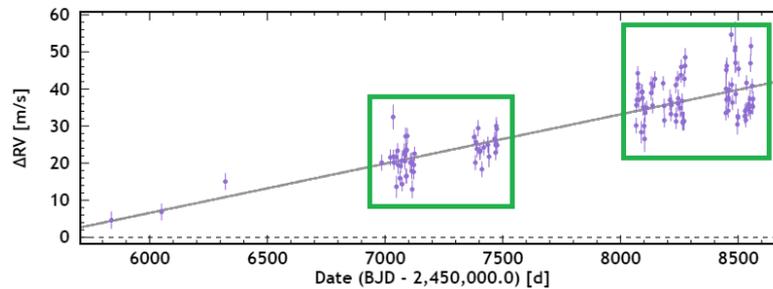


Figure 3.19 – RV time series starA, subdivided by two consecutive subsets.

3.7.3 High precision RV search for super-Earth in systems harboring giant planets

The occurrence rate of the systems with the long-period giant planets (mass $> 0.3 M_J$) which harbor small planets in the inner orbit is high (Zhu et al., 2018; Schlecker et al., 2021). Our solar system is one example of such a system with Jupiter in the long-period and other small planets in shorter periods. Since 2017, the SOPHIE consortium has been following a program, known as SP2+ (PI: G. Hebrard), to detect inner orbit small planets in systems that are already known to harbor long-period giant planets. The goal of this program is to examine the different theories of planet formation, study the architecture of such a multi-planet system, and also the role of the presence of gas giants on the composition of inner orbit planets. Such studies will greatly help to understand the formation's path of our solar system. The observation time, dedicated to this program, is $\sim 6-7$ nights per semester.

The SP2+ stars initially were observed simultaneously with the sky monitoring (indicated by $N_{obs(SKY)}$ in Table 3.4). After the detection of a giant planet in a system, and if the system presents a high RMS in the RV residuals after the Keplerian fit, we continue to observe the star to search for inner orbit low-mass planets. To do this, we follow the observational strategy for detecting low-mass planets with SOPHIE,

explained in Sect 3.1.3. This also included the star observation simulations with FP calibration lamp observation (indicated by $N_{obs(FP)}$ in Table 3.4). This change in the observation strategy is essential for achieving the required precision for detecting low-mass planets.

Table 3.4 – Summary SP2+ candidates.

star	$N_{obs(SKY)}$	$N_{obs(FP)}$	$K_J(\text{Publication})$ (m/s)	$K_J(\text{this work})$ (m/s)	Other candidates (d)	$P_{H\alpha}$ (d)	P_{FWHM} (d)
star1	67	31	98.6±0.5	94.5±0.5	2271.9*	—	>1000*
Star2	71	101	56.53±0.22	58.57±0.61	—	—	—
star3	56	14	—	—	—	2.3 [×]	275
star4	19	27	408±7	405.9±0.7	—	—	—
star5	44	30	726.4±7.1	710.6±1.3	—	101 [†]	>100*
star6	69	48	56.4±0.9	55.3±0.7	—	—	>200*
star7	88	24	173.9±1.3	172.5±0.5	—	11*	395*
star8	35	24	374.2±2.4	382.2±0.9	—	—	—
star9	41	47	39.0±0.9	33.7±0.7	—	3.9 [×] ,31 [×] ,39 [×]	31 [×]
star10	19	27	38.2±1.6	39.3±0.6	—	—	50 [†] ,170*
star11	19	30	578±20	512.9±0.6	>2000*	—	—
star12	83	26	91.1±2.1	91.8±1.5	—	—	101 [†]
			56.6±3.3	53.3±1.9	—	—	—
star13	61	34	108.1±1.2	108.0±0.7	—	17 [×]	—
star14	27	31	—	—	—	—	—

Note: *,[†],[×] denote a periodic signal below FAP 0.1%, 1%, 10%, respectively. Also, K_J presents the semi-amplitude of the Jupiter-like planet in the long orbit.

For this program, I developed an automated pipeline to extract data and process them. To do this, I followed the same steps of the BEBOP pipeline. Then, I analyzed the data of 14 stars of this program and present the results in Table 3.4.

I found my results in good agreement with the publications of the giant planet, but with more precise parameters. Indeed, with more observation points, I achieved more robust results with a smaller uncertainty. I also detected several possible small planet signals but none of them were statistically significant. Such signals might become stronger with further observations. Therefore, we continue to observe those systems. I also note that, for star11, I re-confirmed the existence of a signal with the long-orbital period of >2000, which in the publication, the authors had attributed to a stellar companion with the upper mass of $0.35 M_{Sun}$. Also, for star1, the long-period signal in RVs has a corresponding peak in activity. Hence, it is likely that the signal is an activity cycle of the star.

Chapter 4

Conclusions and future perspectives

Contents

4.1	Overview of low-mass planet populations as of 2022	144
4.1.1	Low-mass planets radius gap	145
4.1.2	Radius-period diagram	146
4.1.3	Mass-radius diagram & internal composition model	147
4.2	Key questions in exoplanet science as 2022	148
4.3	What to expect from the next decades?	150
4.3.1	Future space telescopes	150
4.3.2	Future ground-based telescopes	154
4.4	My future role in answering exoplanet science key questions	155

In this chapter, I draw a picture of low-mass planet knowledge as of 2022. Then, I review some important questions in the exoplanet field and show a big image of the field in the future and my role in it.

4.1 Overview of low-mass planet populations as of 2022

During 3 years of developing this thesis from 2019 to 2022, the number of detected low-mass planets with precise mass and radius rapidly increased.

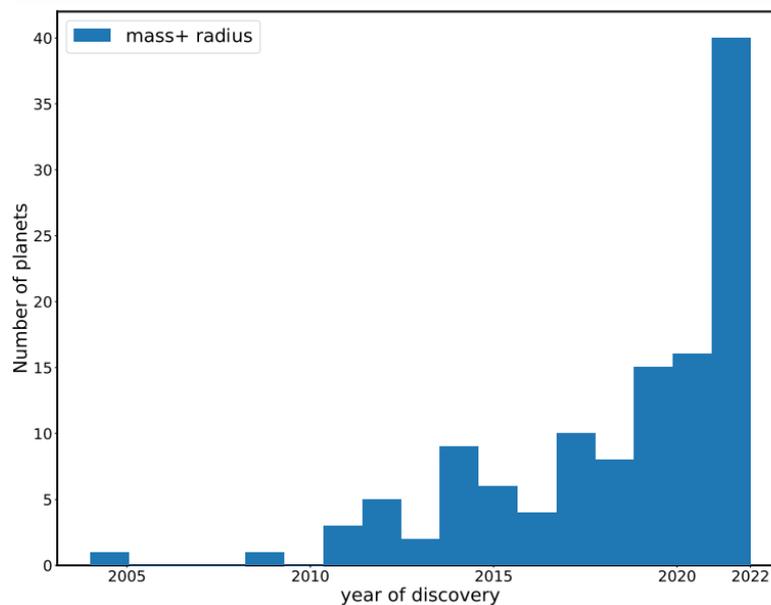


Figure 4.1 – Histogram of the known low-mass exoplanets with 50 % precision on mass and radii versus the year of discovery.

As of July 2022, the total number of detected planets are 5054¹, 562 of which have both radius and mass determined with a fraction of 163 being low-mass planets ($M_p < 30 M_E$). To study the low-mass planet populations in this section and its following, I employed a cut of on mass and radius with relative uncertainties smaller than 50 % ($\sigma_M/M < 50\%$, $\sigma_R/R < 50\%$), remaining 120 precisely measured low-mass planets. Figure 4.1, illustrates the histogram of the known low-mass exoplanets with precise mass and radii versus discovered years. The plot clearly shows the number of low-mass planets has dramatically increased in recent years. This is mostly due to the TESS

1. NASA exoplanet catalog: <https://exoplanetarchive.ipac.caltech.edu/>

space mission which targets bright stars that allow us to follow up by RV method to confirm their nature and determine their fundamental parameters (i.e. mass, radius, and density).

4.1.1 Low-mass planets radius gap

Fig. 4.2 illustrates the histogram of the radius of low-mass planets with mentioned precision. The histogram shows that the low-mass planets tend to have smaller radius ($< 4 R_E$), with a shape of bimodality which has two peaks at super-Earth ($< 1.5 R_E$) and sub-Neptune ($\sim 2-4 R_E$) and a gap in $\sim 1.5-2 R_E$ with a few planets in there (e.g. TOI-402 (Dumusque et al., 2019) and GJ 9827 (Niraula et al., 2017)). The presence of this gap is one of the most important open questions in planetary formation which for the first time was confirmed by Fulton et al., 2017 in the California-Kepler Survey.

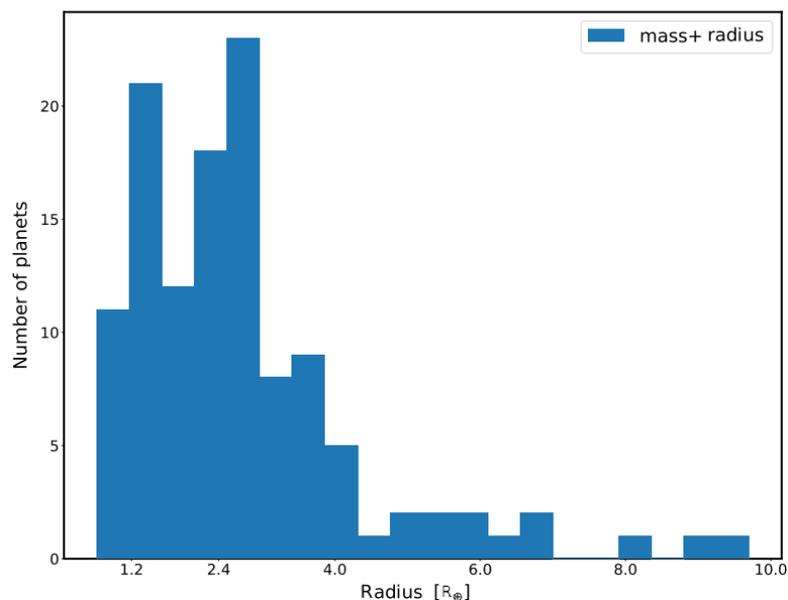


Figure 4.2 – Histogram of radius of low mass known exoplanet those with precise mass and radii.

There are two main theories to explain this bimodal distribution: photoevaporation (Owen et al., 2013; Owen et al., 2017; Lopez et al., 2014) and core-powered mass loss (Ginzburg et al., 2016; Ginzburg et al., 2018). Both theories have the same principle: heating of the upper atmosphere causes a hydrodynamic outflow and consequently the planet's mass loss. However, the energy source of this heat for the two theories is different. In photoevaporation theory, extreme ultraviolet (XUV) photons of the star, heat the upper atmosphere and cause planetary mass loss. While, in the core-powered

mass-loss theory, infrared radiation from the cooling of the planetary interior and bolometric irradiation of the star provide the energy sources.

Up to 2022, despite the additional detection of low-mass planets, it is still unclear which model (even maybe the combination of both) is the right framework. Petigura et al., 2022 examined both theories by investigating the relation between host stars in M_* , metallicity, and age from Gaia DR2, and radius and mass of small planets. They concluded that both theories sufficiently match with observations.

Additional observations and improvements of the host star properties (it can be with Gaia DR3 which recently came out or Gaia DR4 in the coming years), and precisely detecting and characterizing more low-mass planets, will allow us to examine the two theories on the larger sample. Particularly, refinement of Gaia data on T_{eff} and $[Fe/H]$ will be useful in calculating of mass and age of additional early stars. Further planet detections and well-characterizing them in fundamental parameters such as mass and radius, help in constraining the relation between the planet core mass distribution and host stars' properties which have remained flexible for both models (Petigura et al., 2022).

4.1.2 Radius-period diagram

Studying the radius-period diagram is very useful to distinguish features and test different theories of formation and evolution. Fig. 4.3 shows the radius-period diagram of low-mass planets in logarithmic scale. Their distribution in this diagram illustrates the lack of planets between ~ 2 - $10 R_E$ in the shorter orbital periods, known as a hot Neptune desert (Mazeh et al., 2016). In other words, planets in shorter orbital periods tend to be either rocky with radii $< 2 R_E$ or hot Jupiters with radius $> 10 R_E$. This gap does not seem to be due to the observational bias because many smaller planets (radii $< 2 R_E$), which are more difficult to detect, have already been discovered close to their host stars. The lack of planets with intermediate size in such a close orbit of their host star has been interpreted as an inability of low-mass planets to keep the H/He envelope in the face of high irradiation of the host star (Lopez, 2017; Owen et al., 2017).

We are slowly detecting the planet inside of hot Neptune desert such as LTT 9779 b (Jenkins et al., 2020), TOI-849 b (Armstrong et al., 2020), TOI-824 b (Burt et al., 2020), and NGTS-4 b (West et al., 2019). Such planets tend to be more dense compare to other planets with the same size which is expected due to the their atmospheric loss. Detection and well characterization of these potential remnant planetary cores will greatly help us to test the different theories of photo-evaporation, its dependency on the spectral type of star, and the defining boundaries for this desert.

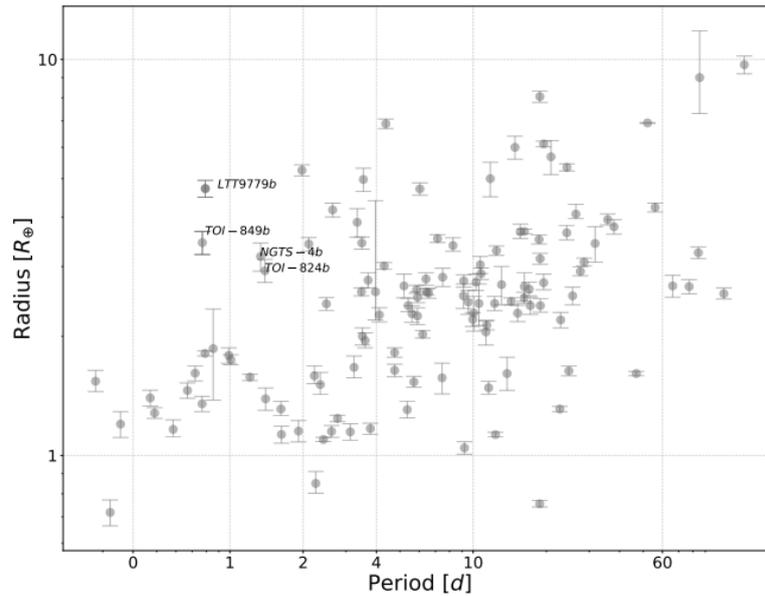


Figure 4.3 – Radius-period diagram of known planets with precise mass and radii. The unpopulated space in the plot with a shorter orbital period and planet radius of between ~ 2 - $10 R_E$ illustrates the hot Neptune desert.

4.1.3 Mass-radius diagram & internal composition model

Fig. 4.4 illustrates the mass-radius diagram of the precisely known low-mass planets which is over-plotted by the internal composition model following Zeng et al., 2016. The three dashed lines are related to the three different core mass compositions of Fe, silicate, and water. The planets above these lines have some percentage of mass fraction of H_2 envelope. Such a diagram has the key information for understanding and constraining the history of the formation and evolution of planetary systems (e.g. nature of the accreted material, orbital migration, atmospheric escape). As an example, two planets of Kepler-36 b ($M = 4.45_{0.27}^{0.33} M_E$, $R = 1.486 \pm 0.035 R_E$) and c ($M = 8.08_{0.46}^{0.60} M_E$, $R = 3.679 \pm 0.054 R_E$) orbit at 13.8 d and 16.2 d (Carter et al., 2012), respectively, around their host star. These two planets stand in a close distance from each others at semi-major axis of 0.12 AU and 0.13 AU, respectively, but with extremely different density of $7.46_{0.59}^{0.74} \text{ gr/cm}^{-3}$ and $0.89_{0.05}^{0.07} \text{ gr/cm}^{-3}$. While planet b has a composition of an iron-rich core, planet c must be volatile-rich. The different compositions of these two planets which are at a very close distance from each others show that they have a different history of formation and evolution (e.g. Quillen et al., 2013; Lopez et al., 2013).

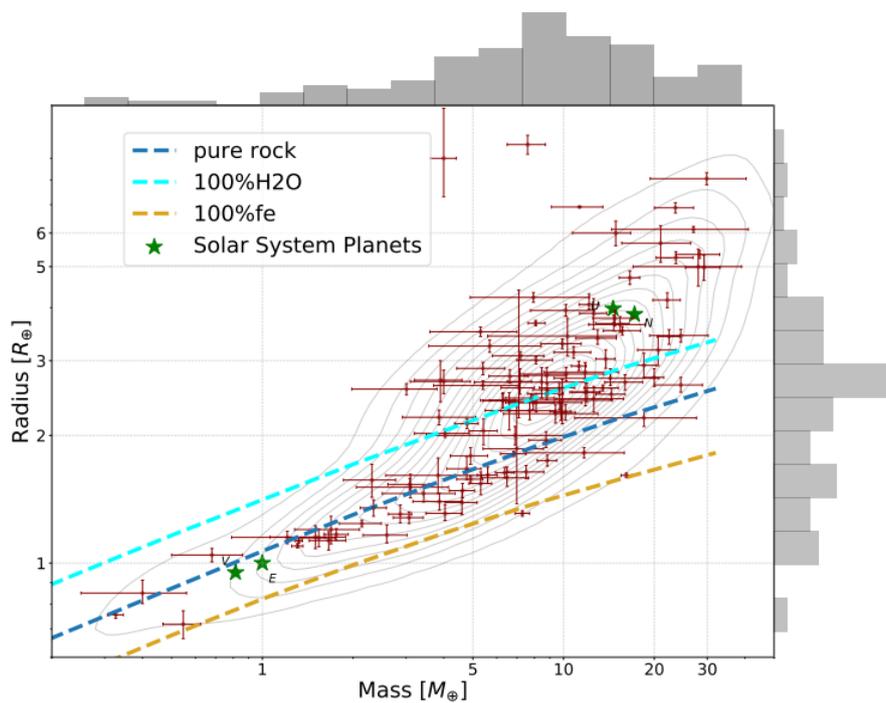


Figure 4.4 – Mass-radius diagram of known low-mass planets with precise mass and radii, over-plotted by composition model following Zeng et al., 2016. For the sake of better comparison, I have also included the solar system planets. The contour represents the number-density of planets which is weighted by their measurement uncertainties. The plot is highly inspired by Otegi et al., 2020.

Also, in Fig. 4.4, the histogram of mass illustrates the lack of enough detection with $M < 1 M_E$. This is because of the difficulty of their detection due to their small Keplerian amplitude which for now is mostly limited by instrumental accuracy and intrinsic activity of star itself. Populating mass-radius diagram in this region helps us in understanding the formation and evolution of our own planet, Earth, at a different stage of evolution. We can also probe the history of atmospheric evolution between Venus, Earth, and Mars.

4.2 Key questions in exoplanet science as 2022

In the past three decades, the exoplanet field has never stopped surprising us, thanks to the enormous progress in terms of both facilities and analyzing the data. This trend is not expected to fade any time soon. There are a lot of things to be learned and discovered. Here, I highlight a few of the key open questions in the field:

- How unique is our Solar System?
- What is the history of Earth's formation and evolution? How long did Earth's magma ocean period and other evolution steps last?
- How common are planets located within the habitable zone? To date, only a few planets have been discovered within the habitable zone, which are mostly around M dwarfs and might suffer from the activity of their host stars, making them unsuitable for harboring life. A focus on discovering them is essential.
- How can we search for life? What is the composition of the atmosphere of habitable zone planets? Do they contain signs of life as we know by today (i.e., biosignature gases)?
- What is the origin of super-Earths and sub-Neptunes (see Fig. 4.5)? Why they are absent in our solar system and very common elsewhere in the galaxy? Why some planets as small as super Earth are puffy (gaseous)?

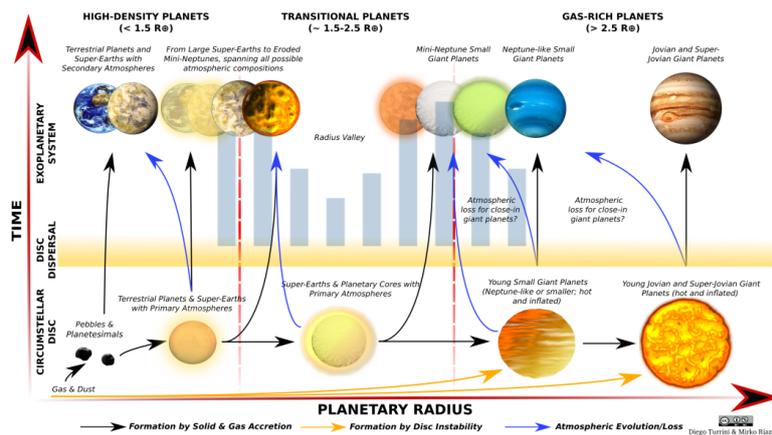


Figure 4.5 – Schematic figure of different theories from our current understanding of the formation and evolution of three planet categories (super-Earths, sub-Neptunes, and hot Jupiters) which are absent in our solar system. Credited by Tinetti et al., 2020.

- What is the origin of the radius gap within $\sim 1.5\text{--}2.0 R_E$. Which theory of photo-evaporation or core-powered mass loss?
- How can we effectively face the challenge of interference with stellar activities?
- What is the relation between star and planetary composition or number of planets in a system? What is the history of planetary evolution and formation when the planet and host star have a different compositions?

- What is the evolution of planetary systems and their atmosphere over time? As an example in our solar system, Venus, Earth, and Mars are inside habitable zone but with different atmospheres.
- What are the dependency of Neptune desert boundaries to the spectral type of star and its fundamental star parameters such as mass and age (e.g. McDonald et al., 2019)?
- What is the origin of hot Jupiters very close to their host stars?
- How can we efficiently detect longer period planets, exo-moons, and exo-rings? Increasing the diversity of discovered planets and well characterizing them is essential to test different theories of planet formation and evolution and also model their internal structure. As an example, among the 5000+ discovered planets, many characterized planets in terms of both mass and radius have an observational bias toward shorter orbital periods (<40 days). In other words, we have mostly discovered worlds that were formed through evaporation processes rather than sculpting (i.e. other planetary formation processes), and with a lower probability of forming exo-moons (Sucerquia et al., 2020).

4.3 What to expect from the next decades?

For the time being, accurate answers to these questions are not possible due to the lack of a sufficiently large and diverse sample of exoplanets with precise parameters of both planets (e.g. mass, radius) and their host stars (e.g. age). Given the strong community investment and future ground and space telescopes, with the launching or installation within the next decades, we will confidently move forward to answer the mentioned questions, and detect Earth twins, while at the same time detecting several unusual planets.

4.3.1 Future space telescopes

In the coming decades, we expect a big revolution in space telescopes with different missions (or proposals) such as JWST (the first JWST data, released at the time of writing of this chapter), PLATO, ARIEL, LUVOIR, HabEX, and LIFE.

JWST: James Webb Space Telescope (JWST, Gardner et al., 2006), or simply WEBB, is a collaborative mission between NASA, ESA, and the Canadian Space Agency (CSA),

whose launched on January 2022. It has a broad range of goals from studying the history of the universe to exoplanet and solar system formation. Its goal in the exoplanet area is to study the chemical and physical properties of the planetary system. It has dedicated more than 20 % of its observational time during the first cycle to the study of exoplanets. JWST is an infrared telescope and through this, it is suitable for studying the exoplanet's atmosphere via transmission and emission spectroscopy. As WEBB also carries a coronagraph, it is able to directly image exoplanets. WEBB will greatly help in well characterizing the planets. It will help to answer several exoplanet challenges including the challenge of what are the chemical compositions of exoplanet atmospheres and their surface? How thick is their atmosphere? Do the rocky (or terrestrial) planets have an atmosphere or not (see Sect. 1.3.4)? What is the relation between the planetary atmosphere and their formation path? and so on.

WEBB during its commissioning studied the atmosphere of WASP-96b (Hellier et al., 2014) through transmission spectroscopy. This planet is a gas giant with a mass of about half of Jupiter and a radius of $1.2 R_J$, orbiting every 3.4 d around its host star. The extremely high temperature of the planet (1285.0 K), makes it 'puff' out (i.e. less massive than Jupiter but with a larger radius). The target was already well-known to have water in its atmosphere (Yip et al., 2020). The WEBB clearly re-confirms the existence of the water feature. As this data was recently published, many research currently is ongoing to study the WEBB data. WEBB will also highlight the study of the atmosphere of two super-Earths with rocky composition, 55 Cancri e (Demory et al., 2011) and LHS 3844 b (Vanderspek et al., 2019) during its initial operations. Both these planets are close to their host stars and indeed very hot, with a temperature of 2697 K (at maximum hemisphere-averaged temperature by Demory et al., 2016) and the orbital period of 18 hours for 55 Cancri e, and 805 K with the orbital period of 11 hours for the LHS 3844 b. The reason for these choices are their highest S/N of thermal emission among the other small rocky planets and also several questions about them (e.g. Hammond et al., 2017 and Kreidberg et al., 2019 with the references in there). Such a study will provide us with a new perspective of the small planets in general. It also helps us in understanding the early formation history of our own planets, Venus and Mars, and the fact that they are very different in atmosphere.

PLATO: The European PLANetary Transits and Oscillations of stars (PLATO 2, Rauer et al., 2014) ESA mission, whose launch is scheduled for 2026, has the following scientific goals:

- Detecting and precisely characterizing many exoplanet systems to study their internal structure, architecture, and dependency on their formation and evolution to the host star. The expected precision in radius will be up to 3%, and in the planet mass and the stellar age up to 10%.

- Detecting and characterizing planets as small as Earth and up to an orbital period of 1 year by transit method inside the habitable zone.
- Study the internal structure of stars and their evolution by age through studying stellar oscillations.
- Providing good targets for future study of the planetary atmospheres.

To achieve these goals, several lessons are to be learned from TESS and Kepler. Despite Kepler which targets the faint stars, PLATO will observe bright stars with a magnitude of $V < 11$ at a very high photometric precision of about 50 ppm (Marchiori et al., 2019). This allows one to have RV follow-up observations to determine the planet mass and mean density. Also, despite the TESS which due to its observation strategy (~ 27 days) mostly detects planets with shorter orbital periods, PLATO is sensitive to both longer and shorter orbital periods. This is because of PLATO's observation strategy which will be long with a duration from a few months up to 3 years. The final observation strategy will be decided two years before the launch in 2026.

ARIEL: The Atmospheric Remote-Sensing Infrared Exoplanet Large (ARIEL², Pascale et al., 2018) survey mission, is a mission to study and characterize at least 1000 known planets' chemical composition and thermal structures through transmission spectroscopy method. ARIEL will address three main questions:

- What are the physical processes for the formation and evolution of the planetary atmosphere?
- What is the composition of planets?
- What is the formation and evolution of planetary systems?

ARIEL is planned to be launched in 2029. Note that ARIEL is an exoplanet dedicated mission and compared to the James Webb Space Telescope, provides more observing time for the community.

ARIEL and PLATO are space missions for 2020s, while there are several proposals under review for the 2030s-2040s. Here are some examples:

LIFE: Large Interferometer For Exoplanets (LIFE³, Quanz et al., 2021) ESA mission, is a space mission to study and characterize terrestrial exoplanet atmosphere through

2. <https://arielmission.space/>

3. <https://life-space-mission.com/>

the direct imaging method. Its targets are exoplanets with $0.5R_E < R_p < 1.5R_E$ and receiving 0.35-1.7 Earth insolation flux. Note that there are two major differences between ARIEL and JWST missions with LIFE mission. 1) LIFE will use direct imaging to study exoplanet atmospheres while ARIEL and JWST use/will use the transit method. Probing planetary atmospheres using direct imaging, allows us to access the deeper atmospheric regions. The deeper atmosphere is a region that transfers energy through the upper atmosphere to outer space. This region by transmitting the energy highly affects the composition of the upper layer. The study of the deeper atmosphere is very important because it can tell us about the composition of the upper atmosphere, the existence of surfaces on the planets, and approximate surface conditions (Yu et al., 2021). 2) Objective targets for ARIEL and JWST are hot/warm close-in transiting planets while LIFE will focus on temperate Earth-like planets. Its science goals are:

- Constrain the diversity of terrestrial exoplanets and their atmospheres.
- Assess the habitability of rocky planets in the habitable zone.
- Searching for bio-marks in the planetary atmospheres.

HabEx: Habitable Exoplanet (HabEx⁴, Gaudi et al., 2020) observatory, is a NASA's future concept mission with a similar science goals to the LIFE mission. HabEx will study exoplanet atmospheres through the direct imaging. Note that HabEx is not an exoplanet dedicated mission. It also will perform several solar system, Galactic, and extragalactic observations through a wide wavelength range from the ultraviolet (UV) to the near-infrared (near-IR).

LUVOIR: Large UV/Optical/Infrared Surveyor telescope (LUVOIR⁵) is also a future NASA concept mission. It covers a large wavelength range between the far-ultraviolet to the near-infrared. It has a variety of science goals in different fields. In the exoplanet domains, it will study exoplanets through both transit spectroscopy and direct imaging. The LUVOIR science goals, in addition to the science goals of the LIFE mission, are (see Team et al., 2019 for a full description of science goals):

- Study atmospheric composition for a wide range of exoplanets with different sizes, orbits, and host stars (LIFE mission will focus only on habitable planets).
- Detecting new terrestrial exoplanets in the habitable zones of A-type stars and investigating their atmosphere and their habitability through direct imaging.

4. <https://www.luvoirtelescope.org/>

5. <https://www.luvoirtelescope.org/>

- Measuring the compositions of vaporized silicate and metal atmospheres from extremely hot molten rocky planets.
- Detection of exo-moons through spectroastrometry (Agol et al., 2015).

4.3.2 Future ground-based telescopes

On the ground also, many current and future spectrographs with many innovative and new generations of instruments will be contributed to answering big exoplanet questions (see Sect. 4.2) in the next decades. I highlighted the current spectrograph such as SPIROU, NIRPS and ESPRESSO in chapter 1, and here I will mention a few new spectrographs which will be mounted during the next decade.

METIS, HARMONI, and ANDES: Extremely Large Telescope (Rodler, 2018) will be a new state-of-the-art ESA telescope with a 39 m primary mirror and several different instruments such as the new generation of high-resolution spectrographs METIS, HARMONI and in the longer term ANDES. One of the main scientific goals of ELT is to study exoplanets by directly imaging them, including habitable zone planets. It will also be able to investigate transiting planets' atmospheres with high precision via transmission spectroscopy. The first light of this telescope is expected to be in 2027.

HARPS3: The HARPS3 (Thompson et al., 2016) will be a fiber-fed, high-resolution echelle spectrograph. It will be mounted on the 2.5m Isaac Newton Telescope in La Palma in the Canary Islands. HARPS3 is very similar to HARPS and HARPS-N. It covers wavelengths between 380-690 nm with a high-resolution power of $R = 115,000$. It is also able to do spectro-polarimetry. The scientific goal of this new spectrograph, also known as Terra Hunting Experiment⁶, is to observe 40 bright G and K spectral type stars, for 6-8 consecutive months, more than 10 years to detect Earth-mass planets with orbital periods close to one year. The estimated time of 10 years is required for gathering enough measurements to detect and reach the desired precision of small keplerian amplitudes of Earth-mass planets by overcoming instrumental and stellar noises.

EXPRES: EXTreme PREcision Spectrometer (EXPRES⁷, Jurgenson et al., 2016) will be a new high resolution spectrograph mounted on the 4.3-m Lowell Observatory

6. <https://www.terrahunting.org/>

7. <http://exoplanets.astro.yale.edu/instrumentation/expres.php>

Discovery Channel Telescope (DCT). The spectrograph will cover a wavelength range of 380 to 680 nm with a resolution power of $R = 150,000$. The scientific goal of this spectrograph is that through its high-resolution power and inherent precision in the wavelength solution of the order of 1 cm/s, detect Earth-mass planets and face stellar activity challenges.

4.4 My future role in answering exoplanet science key questions

As we approach the end of this thesis, one can realize that only the surface of the exoplanet knowledge has been scratched. As we saw in sect. 4.2, a lot of questions remain to be answered. These questions might be answered by new current and future facilities, and developed methods in the next decades.

Given the skills that I acquired during my thesis, including analyzing the RVs and photometric data, I am able to contribute to the attempt of finding answers to the big questions in the field, in the future. To do so, I propose to significantly increase the number of long-period planets and accurately characterize them in mass and radius, leading to access to their mean density. Such a focus on detecting of long-period planets can have several dramatic impacts as follows:

- Planets with longer orbital periods stand in less explored areas of parameter space. Hence, detecting the long-period planets will lead to breaking observational bias and significantly help in constraining planetary formation and migration models. Additionally, the planets with shorter orbital periods are mostly shaped by evaporation, while planets with longer periods are mostly shaped by sculpting. This will greatly help in constraining planetary formation models.
- More chance for discovering interesting planets such as planets with exomoons.
- Increasing the chance of discovering habitable zone planets.
- Expanding the number of detected planets and widen their diversities which allows us to apply statistics and measure planetary occurrence rates.

For detecting long-period planets, considering the currently available data, the strategy that I will take will be focusing on detecting the TESS single transit events. The reason is that the TESS observation strategy for most of the sky is short (~ 27 days).

Therefore, the longer period planets will appear by a single transit event. I will develop a pipeline to search for a single transit in the TESS data. The developed pipeline and methods will have further applications on PLATO photometric data in the future. After the availability of the PLATO data, I will work on discovering the long-period planets in this data (for more information about my future project see [Appendix 5](#)).

The perspective of understanding planetary formation and evolution and detecting life seems very promising. Various current and future ground and space-based facilities move forward our way to the detection of the Earth twins. The method developed in this thesis will contribute to the precise detection and characterization of such planets.

Overview of the results

This thesis was focused on study of the diversity of low-mass planets composition, by expanding the sample of precisely characterized low-mass planets and making the link between the observations and the formation paths of these bodies. To accomplish this goal, I used both photometric and RV data, which combining them for transiting planets, allows one to measure mass and radius, leading to inferring density and bulk composition.

First, I faced the challenge of instrumental limitations as one of the biggest obstacles in low-mass planet detection. I investigated the SOPHIE data reduction system and added several new features to it. This attempt led to overcoming various instrumental limitations, great improvement of 22 cm/s and 18.6 m/s in RVs and FWHM RMS, respectively, and pushing further the limit of SOPHIE for detecting low-mass planets (see chapter 2).

Second, I faced the challenge of low-mass planets detection and precisely characterizing them. I carried out the full photometry and RV analysis of the SOPHIE SP1 targets, a dedicated program to probe low-mass planets detection. As a result, we discovered and fully characterized two low-mass planets: HD207897 b and HD88986 b, in addition to several other candidates. HD207897 b is one of the relatively rare dense ($5.1 \pm 0.7 \text{ g cm}^{-3}$) rocky sub-Neptunes. The existence of this dense planet at only 0.12 AU from its host star is unusual in the currently observed sub-Neptune population. The most likely scenario is that this planet has migrated to its current position. The other detected planet, HD88986 b, has the longest period among the accurately characterized transiting sub-Neptune with an orbital period of $147.4^{+0.05}_{-0.05}$ d. Hence it is a temperate planet and an interesting target for the study of internal planet composition (see chapter 3).

To conclude, through the development of this thesis, I have been able to explore the discovery of low-mass planets via photometry and RV methods, from raw spectra to planet characterization. The techniques developed and the results gathered in this thesis will greatly contribute to the accurate detection and well characterization of future planets detection by SOPHIE. Discovered low-mass planets will contribute to a deeper understanding of the low-mass planet population. Finally, all of these studies allow us to improve our understanding of how planets form and evolve as well as highlight a promising prospect that will be investigated in the upcoming years.

Bibliography

- Affer, L. et al. (2019). “HADES RV program with HARPS-N at the TNG-IX. A super-Earth around the M dwarf Gl 686”. In: *Astronomy & Astrophysics* 622, A193 (cit. on p. 138).
- Agol, E. et al. (2015). “The center of light: Spectroastrometric detection of exomoons”. In: *The Astrophysical Journal* 812.1, p. 5 (cit. on p. 154).
- Aigrain, S. et al. (2012). “A simple method to estimate radial velocity variations due to stellar activity using photometry”. In: *Monthly notices of the royal astronomical society* 419.4, pp. 3147–3158 (cit. on p. 24).
- Armstrong, D. J. et al. (2020). “A remnant planetary core in the hot-Neptune desert”. In: *Nature* 583.7814, pp. 39–42 (cit. on p. 146).
- Artigau, É. et al. (2014). “SPIRou: the near-infrared spectropolarimeter/ high-precision velocimeter for the Canada-France-Hawaii telescope”. In: *Ground-based and airborne instrumentation for astronomy v. Vol. 9147*. International Society for Optics and Photonics, p. 914715 (cit. on p. 28).
- Astudillo-Defru, N. et al. (2015). “The harps search for southern extra-solar planets-xxxvi. Planetary systems and stellar activity of the m dwarfs gj 3293, gj 3341, and gj 3543”. In: *Astronomy & Astrophysics* 575, A119 (cit. on p. 137).
- Astudillo-Defru, N. et al. (2017a). “The HARPS search for southern extra-solar planets-XLII. A system of Earth-mass planets around the nearby M dwarf YZ Ceti”. In: *Astronomy & Astrophysics* 605, p. L11 (cit. on p. 25).
- Astudillo-Defru, N. et al. (2017b). “The HARPS search for southern extra-solar planets-XLII. A system of Earth-mass planets around the nearby M dwarf YZ Ceti”. In: *Astronomy & Astrophysics* 605, p. L11 (cit. on p. 54).
- Astudillo-Defru, N. et al. (2017c). “Magnetic activity in the HARPS M dwarf sample-The rotation-activity relationship for very low-mass stars through R’HK”. In: *Astronomy & Astrophysics* 600, A13 (cit. on p. 137).
- Baglin, A. et al. (2006). “CoRoT: a high precision photometer for stellar evolution and exoplanet finding”. In: *36th COSPAR Scientific assembly*. Vol. 36, p. 3749 (cit. on pp. 13, 35).
- Baluev, R. V. (2008). “Assessing the statistical significance of periodogram peaks”. In: *Monthly Notices of the Royal Astronomical Society* 385.3, pp. 1279–1285 (cit. on pp. 18, 19).
- (2009). “Accounting for velocity jitter in planet search surveys”. In: *Monthly Notices of the Royal Astronomical Society* 393.3, pp. 969–978 (cit. on p. 19).

- Baluev, R. V. (2013). “Detecting non-sinusoidal periodicities in observational data: the von Mises periodogram for variable stars and exoplanetary transits”. In: *Monthly Notices of the Royal Astronomical Society* 431.2, pp. 1167–1179 (cit. on p. 19).
- (2015). “Keplerian periodogram for Doppler exoplanet detection: optimized computation and analytic significance thresholds”. In: *Monthly Notices of the Royal Astronomical Society* 446.2, pp. 1478–1492 (cit. on p. 19).
- Barros, S. C. et al. (2022). “Detection of the tidal deformation of WASP-103b at 3σ with CHEOPS”. In: *Astronomy & Astrophysics* 657, A52 (cit. on p. 39).
- Benz, W. et al. (2021). “The CHEOPS mission”. In: *Experimental Astronomy* 51.1, pp. 109–151 (cit. on pp. 38, 39).
- Bluhm, P. et al. (2021). “An ultra-short-period transiting super-Earth orbiting the M3 dwarf TOI-1685”. In: *Astronomy & Astrophysics* 650, A78 (cit. on p. 26).
- Boisse, I. et al. (2009). “Stellar activity of planetary host star HD 189 733”. In: *Astronomy & Astrophysics* 495.3, pp. 959–966 (cit. on pp. 25, 50).
- Boisse, I. et al. (2012a). “SOAP-A tool for the fast computation of photometry and radial velocity induced by stellar spots”. In: *Astronomy & Astrophysics* 545, A109 (cit. on p. 26).
- Boisse, I. et al. (2012b). “The SOPHIE search for northern extrasolar planets-V. Follow-up of ELODIE candidates: Jupiter-analogs around Sun-like stars”. In: *Astronomy & Astrophysics* 545, A55 (cit. on p. 54).
- Bonfils, X. et al. (2007). “The HARPS search for southern extra-solar planets-X. A $m \sin i = 11$ planet around the nearby spotted M dwarf GJ 674”. In: *Astronomy & Astrophysics* 474.1, pp. 293–299 (cit. on p. 25).
- Borucki, W. J. et al. (2010). “Kepler planet-detection mission: introduction and first results”. In: *Science* 327.5968, pp. 977–980 (cit. on pp. 14, 35).
- Borucki, W. J. et al. (2011). “Characteristics of planetary candidates observed by Kepler. II. Analysis of the first four months of data”. In: *The Astrophysical Journal* 736.1, p. 19 (cit. on p. 40).
- Bouchy, F. et al. (2001). “Fundamental photon noise limit to radial velocity measurements”. In: *Astronomy & Astrophysics* 374.2, pp. 733–739 (cit. on p. 22).
- Bouchy, F. et al. (2009a). “The SOPHIE search for northern extrasolar planets-I. A companion around HD 16760 with mass close to the planet/brown-dwarf transition”. In: *Astronomy & Astrophysics* 505.2, pp. 853–858 (cit. on pp. 27, 88).
- Bouchy, F. et al. (2013). “SOPHIE+: First results of an octagonal-section fiber for high-precision radial velocity measurements”. In: *Astronomy & Astrophysics* 549, A49 (cit. on pp. 27, 47, 50).
- Bouchy, F. et al. (Jan. 2009b). “Charge Transfer Inefficiency effect for high-precision radial velocity measurements”. In: *EAS Publications Series*. Ed. by P. Kern. Vol. 37. EAS Publications Series, pp. 247–253 (cit. on p. 49).
- Bouchy, F. et al. (2017). “Near-InfraRed Planet Searcher to Join HARPS on the ESO 3.6-metre Telescope”. In: *The Messenger* 169, pp. 21–27 (cit. on p. 28).
- Burt, J. A. et al. (2020). “TOI-824 b: a new planet on the lower edge of the hot Neptune desert”. In: *The Astronomical Journal* 160.4, p. 153 (cit. on p. 146).

- Butler, R. P. et al. (2017). “The LCES HIRES/Keck precision radial velocity exoplanet survey”. In: *The Astronomical Journal* 153.5, p. 208 (cit. on p. 138).
- Cameron, A. C. (2012). “Astrophysical false positives”. In: *Nature* 492.7427, pp. 48–49 (cit. on p. 34).
- Carter, J. A. et al. (2012). “Kepler-36: A pair of planets with neighboring orbits and dissimilar densities”. In: *Science* 337.6094, pp. 556–559 (cit. on pp. 42, 147).
- Charbonneau, D. et al. (1999). “Detection of planetary transits across a sun-like star”. In: *The Astrophysical Journal* 529.1, p. L45 (cit. on p. 13).
- Charbonneau, D. et al. (2002). “Detection of an extrasolar planet atmosphere”. In: *The Astrophysical Journal* 568.1, p. 377 (cit. on p. 13).
- Chatterjee, S. et al. (2008). “Dynamical outcomes of planet-planet scattering”. In: *The Astrophysical Journal* 686.1, p. 580 (cit. on p. 17).
- Cheetham, A. et al. (2019). “Spectral and orbital characterisation of the directly imaged giant planet HIP 65426 b”. In: *Astronomy & Astrophysics* 622, A80 (cit. on p. 40).
- Christensen-Dalsgaard, J. (2004). “Physics of solar-like oscillations”. In: *Solar Physics* 220.2, pp. 137–168 (cit. on p. 23).
- Cloutier, R. et al. (2019). “Confirmation of the radial velocity super-Earth K2-18c with HARPS and CARMENES”. In: *Astronomy & Astrophysics* 621, A49 (cit. on p. 64).
- Cosentino, R. et al. (2012). “Harps-N: the new planet hunter at TNG”. In: *Ground-based and Airborne Instrumentation for Astronomy IV*. Vol. 8446. International Society for Optics and Photonics, p. 84461V (cit. on p. 27).
- Courcol, B. et al. (2015). “The SOPHIE search for northern extrasolar planets-VII. A warm Neptune orbiting HD 164595”. In: *Astronomy & Astrophysics* 581, A38 (cit. on pp. 69, 70).
- Courcol-Rozès, B. (2016). “Détection et propriétés des exo-Neptunes et Super-Terres: synergies entre SOPHIE et CHEOPS”. PhD thesis. Aix-Marseille (cit. on pp. 72, 88).
- Cram, L. et al. (1979). “Model chromospheres of flare stars. I-Balmer-line profiles”. In: *The Astrophysical Journal* 234, pp. 579–587 (cit. on p. 25).
- Crossfield, I. J. et al. (2020). “Phase curves of hot neptune LTT 9779b suggest a high-metallicity atmosphere”. In: *The Astrophysical Journal Letters* 903.1, p. L7 (cit. on p. 40).
- Cumming, A. et al. (1999). “The Lick planet search: detectability and mass thresholds”. In: *The Astrophysical Journal* 526.2, p. 890 (cit. on p. 18).
- Da Silva, J. G. et al. (2011). “Long-term magnetic activity of a sample of M-dwarf stars from the HARPS program-I. Comparison of activity indices”. In: *Astronomy & Astrophysics* 534, A30 (cit. on p. 25).
- Da Silva, J. G. et al. (2012). “Long-term magnetic activity of a sample of M-dwarf stars from the HARPS program-II. Activity and radial velocity”. In: *Astronomy & Astrophysics* 541, A9 (cit. on pp. 22, 25).
- Danby, J. M. (1988). “Fundamentals of celestial mechanics, Willmann-Bell”. In: *Inc., Richmond, VA* 2 (cit. on p. 16).
- Das, T. et al. (2018). “Development of a stabilized Fabry-Perot based wavelength calibrator for precision Doppler spectroscopy”. In: *Ground-based and Airborne Instrumentation for Astronomy VII*. Vol. 10702. SPIE, pp. 1899–1909 (cit. on p. 49).

- Del Moro, D. et al. (2004). “Dynamics and structure of supergranulation”. In: *Solar Physics* 221.1, pp. 23–32 (cit. on p. 23).
- Del Moro, D. (2004). “Solar granulation properties derived from three different time series”. In: *Astronomy & Astrophysics* 428.3, pp. 1007–1015 (cit. on p. 23).
- Deleuil, M. et al. (2021). “III. 1 Transit features detected by the CoRoT/Exoplanet Science Team”. In: *The CoRoT Legacy Book*. EDP Sciences, pp. 117–122 (cit. on p. 35).
- Delisle, J.-B. et al. (2016). “Analytical determination of orbital elements using Fourier analysis-I. The radial velocity case”. In: *Astronomy & Astrophysics* 590, A134 (cit. on p. 92).
- Delisle, J.-B. et al. (2020). “Efficient modeling of correlated noise-I. Statistical significance of periodogram peaks”. In: *Astronomy & Astrophysics* 635, A83 (cit. on p. 19).
- Demory, B.-O. et al. (2011). “Detection of a transit of the super-Earth 55 Cancri e with warm Spitzer”. In: *Astronomy & Astrophysics* 533, A114 (cit. on p. 151).
- Demory, B.-O. et al. (Jan. 2016). “Variability in the super-Earth 55 Cnc e”. In: *Monthly Notices of the Royal Astronomical Society* 455.2, pp. 2018–2027 (cit. on p. 151).
- Des Etangs, A. L. (2007). “A diagram to determine the evaporation status of extrasolar planets”. In: *Astronomy & Astrophysics* 461.3, pp. 1185–1193 (cit. on p. 102).
- Desidera, S. et al. (2004). “No planet around HD 219542 B”. In: *Astronomy & Astrophysics* 420.3, pp. L27–L30 (cit. on p. 21).
- Dittmann, J. A. et al. (2017). “A temperate rocky super-Earth transiting a nearby cool star”. In: *Nature* 544.7650, pp. 333–336 (cit. on p. 137).
- Diaz, R. et al. (2012). “The SOPHIE search for northern extrasolar planets-IV. Massive companions in the planet-brown dwarf boundary”. In: *Astronomy & Astrophysics* 538, A113 (cit. on pp. 50, 59).
- Doyle, L. R. et al. (2011). “Kepler-16: a transiting circumbinary planet”. In: *Science* 333.6049, pp. 1602–1606 (cit. on p. 136).
- Dravins, D. et al. (1981). “Solar granulation-Influence of convection on spectral line asymmetries and wavelength shifts”. In: *Astronomy and Astrophysics* 96, pp. 345–364 (cit. on p. 23).
- Dravins, D. (1982). “Photospheric spectrum line asymmetries and wavelength shifts”. In: *Annual Review of Astronomy and Astrophysics* 20.1, pp. 61–89 (cit. on p. 23).
- Dumusque, X. et al. (2011a). “Monteiro MJPGF (2011a) Planetary detection limits taking into account stellar noise. I. Observational strategies to reduce stellar oscillation and granulation effects”. In: *Astron Astrophys* 525, A140 (cit. on pp. 23, 25).
- Dumusque, X. et al. (2014). “SOAP 2.0: a tool to estimate the photometric and radial velocity variations induced by stellar spots and plages”. In: *The Astrophysical Journal* 796.2, p. 132 (cit. on pp. 24, 26).
- Dumusque, X. et al. (2011b). “Planetary detection limits taking into account stellar noise-I. Observational strategies to reduce stellar oscillation and granulation effects”. In: *Astronomy & Astrophysics* 525, A140 (cit. on p. 23).
- Dumusque, X. et al. (2012). “An Earth-mass planet orbiting α Centauri B”. In: *Nature* 491.7423, pp. 207–211 (cit. on p. 22).

- Dumusque, X. et al. (2015a). “Characterization of a spurious one-year signal in HARPS data”. In: *The Astrophysical Journal* 808.2, p. 171 (cit. on pp. 70, 84, 85, 93).
- Dumusque, X. et al. (2015b). “HARPS-N observes the Sun as a star”. In: *The Astrophysical Journal Letters* 814.2, p. L21 (cit. on p. 27).
- Dumusque, X. et al. (2019). “Hot, rocky and warm, puffy super-Earths orbiting TOI-402 (HD 15337)”. In: *Astronomy & Astrophysics* 627, A43 (cit. on p. 145).
- Eggenberger, A. et al. (Jan. 2010). “Detection and Characterization of Extrasolar Planets through Doppler Spectroscopy”. In: *EAS Publications Series*. Ed. by T. Montmerle et al. Vol. 41. EAS Publications Series, pp. 27–75 (cit. on p. 20).
- Faria, J. P. d. S. (2018). “Exoplanet detection in metal-poor stars”. In: *Ph. D. Thesis* (cit. on p. 23).
- Fernandez, M. et al. (2017). “IN-SYNC VI. Identification and radial velocity extraction for 100+ double-lined spectroscopic binaries in the APOGEE/IN-SYNC fields”. In: *Publications of the Astronomical Society of the Pacific* 129.978, p. 084201 (cit. on p. 136).
- Ferraz-Mello, S. (1981). “Estimation of periods from unequally spaced observations”. In: *The Astronomical Journal* 86, p. 619 (cit. on p. 18).
- Fischer, D. A. et al. (2015). “Exoplanet detection techniques”. In: *arXiv preprint arXiv:1505.06869* (cit. on p. 24).
- Fischer, D. A. et al. (2016). “State of the field: extreme precision radial velocities”. In: *Publications of the Astronomical Society of the Pacific* 128.964, p. 066001 (cit. on pp. 26, 27).
- Foreman-Mackey, D. et al. (2021). “exoplanet: Gradient-based probabilistic inference for exoplanet data & other astronomical time series”. In: *arXiv preprint arXiv:2105.01994* (cit. on p. 101).
- Fortier, A. et al. (2014). “CHEOPS: a space telescope for ultra-high precision photometry of exoplanet transits”. In: *Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave*. Vol. 9143. SPIE, pp. 750–761 (cit. on p. 38).
- Fortney, J. J. et al. (2013). “A framework for characterizing the atmospheres of low-mass low-density transiting planets”. In: *The Astrophysical Journal* 775.1, p. 80 (cit. on p. 41).
- Foukal, P. et al. (1986). “The influence of faculae on total solar irradiance and luminosity”. In: *The Astrophysical Journal* 302, pp. 826–835 (cit. on p. 24).
- Fulton, B. J. et al. (2017). “The California-Kepler survey. III. A gap in the radius distribution of small planets”. In: *The Astronomical Journal* 154.3, p. 109 (cit. on p. 145).
- Gan, T. et al. (2021). “HD 183579b: a warm sub-Neptune transiting a solar twin detected by TESS”. In: *Monthly Notices of the Royal Astronomical Society* 507.2, pp. 2220–2240 (cit. on p. 26).
- Gardner, J. P. et al. (2006). “The James Webb Space Telescope”. In: *Space Science Reviews* 123.4, pp. 485–606 (cit. on p. 150).
- Gaudi, B. S. (2012). “Microlensing surveys for exoplanets”. In: *Annual Review of Astronomy and Astrophysics* 50.411, p. 2012 (cit. on p. 15).
- Gaudi, B. S. et al. (2020). “The Habitable Exoplanet Observatory (HabEx) mission concept study final report”. In: *arXiv preprint arXiv:2001.06683* (cit. on p. 153).

- Gillon, M. et al. (2017). “Seven temperate terrestrial planets around the nearby ultra-cool dwarf star TRAPPIST-1”. In: *Nature* 542.7642, pp. 456–460 (cit. on p. 137).
- Ginzburg, S. et al. (2016). “Super-Earth atmospheres: self-consistent gas accretion and retention”. In: *The Astrophysical Journal* 825.1, p. 29 (cit. on p. 145).
- Ginzburg, S. et al. (2018). “Core-powered mass-loss and the radius distribution of small exoplanets”. In: *Monthly Notices of the Royal Astronomical Society* 476.1, pp. 759–765 (cit. on p. 145).
- Halbwachs, J. et al. (2005). “Statistical properties of exoplanets-IV. The period–eccentricity relations of exoplanets and of binary stars”. In: *Astronomy & Astrophysics* 431.3, pp. 1129–1137 (cit. on p. 17).
- Hammond, M. et al. (2017). “Linking the climate and thermal phase curve of 55 Cancri e”. In: *The Astrophysical Journal* 849.2, p. 152 (cit. on p. 151).
- Hanslmeier, A. et al. (1991). “The variation of the solar granulation structure in active and non-active regions”. In: *Astronomy and Astrophysics* 251, pp. 307–311 (cit. on p. 24).
- Hara, N. et al. (2020). “The SOPHIE search for northern extrasolar planets-XVI. HD 158259: A compact planetary system in a near-3: 2 mean motion resonance chain”. In: *Astronomy & Astrophysics* 636, p. L6 (cit. on p. 90).
- Hébrard, G. et al. (2008). “Misaligned spin-orbit in the XO-3 planetary system?” In: *Astronomy & Astrophysics* 488.2, pp. 763–770 (cit. on p. 58).
- Heidari, N. et al. (Feb. 2022). “HD 207897 b: A dense sub-Neptune transiting a nearby and bright K-type star”. In: *Astronomy & Astrophysics* 658, A176, A176 (cit. on p. 102).
- Hellier, C. et al. (2014). “Transiting hot Jupiters from WASP-South, Euler and TRAPPIST: WASP-95b to WASP-101b”. In: *Monthly Notices of the Royal Astronomical Society* 440.3, pp. 1982–1992 (cit. on p. 151).
- Hobson, M. J. (2019). “Exoplanet detection around M dwarfs with near infrared and visible spectroscopy”. PhD thesis. Aix-Marseille (cit. on p. 67).
- Horne, J. H. et al. (1986). “A prescription for period analysis of unevenly sampled time series”. In: *The Astrophysical Journal* 302, pp. 757–763 (cit. on p. 18).
- Howell, S. B. et al. (2014). “The K2 mission: characterization and early results”. In: *Publications of the Astronomical Society of the Pacific* 126.938, p. 398 (cit. on p. 35).
- Huang, C. X. et al. (2017). “Dynamically hot super-earths from outer giant planet scattering”. In: *The Astronomical Journal* 153.5, p. 210 (cit. on p. 42).
- Huang, C. X. et al. (2020a). “Photometry of 10 Million Stars from the First Two Years of TESS Full Frame Images: Part I”. In: *Research Notes of the AAS* 4.11, p. 204 (cit. on p. 132).
- (2020b). “Photometry of 10 Million Stars from the First Two Years of TESS Full Frame Images: Part II”. In: *Research Notes of the AAS* 4.11, p. 206 (cit. on p. 132).
- Huélamo, N. et al. (2008). “TW Hydrae: evidence of stellar spots instead of a Hot Jupiter”. In: *Astronomy & Astrophysics* 489.2, pp. L9–L13 (cit. on pp. 22, 25).
- Inamdar, N. K. et al. (2015). “The formation of super-Earths and mini-Neptunes with giant impacts”. In: *Monthly Notices of the Royal Astronomical Society* 448.2, pp. 1751–1760 (cit. on p. 42).

- Isaacson, H. et al. (2010). “Chromospheric activity and jitter measurements for 2630 stars on the California planet search”. In: *The Astrophysical Journal* 725.1, p. 875 (cit. on p. 71).
- Izidoro, A. et al. (2017). “Breaking the chains: hot super-Earth systems from migration and disruption of compact resonant chains”. In: *Monthly Notices of the Royal Astronomical Society* 470.2, pp. 1750–1770 (cit. on p. 42).
- Jenkins, J. S. et al. (2020). “An ultrahot Neptune in the Neptune desert”. In: *Nature Astronomy* 4.12, pp. 1148–1157 (cit. on p. 146).
- Jenkins, J. M. et al. (2016). “The TESS science processing operations center”. In: *Software and Cyberinfrastructure for Astronomy IV*. Vol. 9913. International Society for Optics and Photonics, 99133E (cit. on p. 37).
- Johnson, L. J. et al. (2021). “Forward modelling of Kepler-band variability due to faculae and spots”. In: *Monthly Notices of the Royal Astronomical Society* 504.4, pp. 4751–4767 (cit. on p. 24).
- Jontof-Hutter, D. (2019). “The compositional diversity of low-mass exoplanets”. In: *Annual Review of Earth and Planetary Sciences* 47, pp. 141–171 (cit. on pp. 41, 42).
- Jurgenson, C. et al. (2016). “EXPRES: a next generation RV spectrograph in the search for earth-like worlds”. In: *Ground-based and Airborne Instrumentation for Astronomy VI*. Vol. 9908. SPIE, pp. 2051–2070 (cit. on p. 154).
- Jurić, M. et al. (2008). “Dynamical origin of extrasolar planet eccentricity distribution”. In: *The Astrophysical Journal* 686.1, p. 603 (cit. on p. 17).
- Kaib, N. A. et al. (2013). “Planetary system disruption by Galactic perturbations to wide binary stars”. In: *Nature* 493.7432, pp. 381–384 (cit. on p. 17).
- Kane, S. R. et al. (2014). “Orbital dynamics of multi-planet systems with eccentricity diversity”. In: *The Astrophysical Journal* 784.2, p. 104 (cit. on p. 17).
- Kempton, E. M.-R. et al. (2018). “A framework for prioritizing the TESS planetary candidates most amenable to atmospheric characterization”. In: *Publications of the Astronomical Society of the Pacific* 130.993, p. 114401 (cit. on p. 40).
- Koen, C. (1990). “Significance testing of periodogram ordinates”. In: *The Astrophysical Journal* 348, pp. 700–702 (cit. on p. 18).
- König, P.-C. et al. (2022). “A warm super-Neptune around the G-dwarf star TOI-1710 revealed with TESS, SOPHIE and HARPS-N”. In: *arXiv preprint arXiv:2204.08984* (cit. on p. 54).
- Kopal, Z. (1955). “The classification of close binary systems”. In: *Annales d’Astrophysique*. Vol. 18, p. 379 (cit. on p. 29).
- Kreidberg, L. et al. (2019). “Absence of a thick atmosphere on the terrestrial exoplanet LHS 3844b”. In: *Nature* 573.7772, pp. 87–90 (cit. on p. 151).
- Lafarga, M. et al. (2020). “The CARMENES search for exoplanets around M dwarfs—Radial velocities and activity indicators from cross-correlation functions with weighted binary masks”. In: *Astronomy & Astrophysics* 636, A36 (cit. on pp. 21, 82).
- Lagrange, A.-M. (2014). “Direct imaging of exoplanets”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 372.2014, p. 20130090 (cit. on p. 15).

- Lalitha, S. et al. (2019). “The CARMENES search for exoplanets around M dwarfs-Detection of a mini-Neptune around LSPM J2116+ 0234 and refinement of orbital parameters of a super-Earth around GJ 686 (BD+ 18 3421)”. In: *Astronomy & Astrophysics* 627, A116 (cit. on pp. 138, 139).
- Lee, E. J. et al. (2016a). “Breeding super-Earths and birthing super-puffs in transitional disks”. In: *The Astrophysical Journal* 817.2, p. 90 (cit. on p. 42).
- (Feb. 2016b). “Breeding Super-Earths and Birthing Super-puffs in Transitional Disks”. In: *ApJ* 817.2, 90, p. 90 (cit. on p. 42).
- Léger, A. et al. (2009). “Transiting exoplanets from the CoRoT space mission-VIII. CoRoT-7b: The first super-Earth with measured radius”. In: *Astronomy & Astrophysics* 506.1, pp. 287–302 (cit. on pp. 14, 35).
- Lendl, M. et al. (2020). “The hot dayside and asymmetric transit of WASP-189 b seen by CHEOPS”. In: *Astronomy & Astrophysics* 643, A94 (cit. on p. 38).
- Lomb, N. R. (1976). “Least-squares frequency analysis of unequally spaced data”. In: *Astrophysics and space science* 39.2, pp. 447–462 (cit. on p. 17).
- Lopez, E. D. (2017). “Born dry in the photoevaporation desert: Kepler’s ultra-short-period planets formed water-poor”. In: *Monthly Notices of the Royal Astronomical Society* 472.1, pp. 245–253 (cit. on p. 146).
- Lopez, E. D. et al. (2013). “The role of core mass in controlling evaporation: The Kepler radius distribution and the Kepler-36 density dichotomy”. In: *The Astrophysical Journal* 776.1, p. 2 (cit. on p. 147).
- (2014). “Understanding the mass-radius relation for sub-Neptunes: radius as a proxy for composition”. In: *The Astrophysical Journal* 792.1, p. 1 (cit. on pp. 42, 145).
- Lovis, C. et al. (2011). “The HARPS search for southern extra-solar planets-XXVIII. Up to seven planets orbiting HD 10180: probing the architecture of low-mass planetary systems”. In: *Astronomy & Astrophysics* 528, A112 (cit. on pp. 56, 67).
- Macintosh, B. et al. (2015). “Discovery and spectroscopy of the young jovian planet 51 Eri b with the Gemini Planet Imager”. In: *Science* 350.6256, pp. 64–67 (cit. on p. 24).
- Malmberg, D. et al. (2007). “Close encounters in young stellar clusters: implications for planetary systems in the solar neighbourhood”. In: *Monthly Notices of the Royal Astronomical Society* 378.3, pp. 1207–1216 (cit. on p. 17).
- Marchiori, V. et al. (2019). “In-flight photometry extraction of PLATO targets-Optimal apertures for detecting extrasolar planets”. In: *Astronomy & Astrophysics* 627, A71 (cit. on p. 152).
- Martin, D. V. et al. (2019). “The BEBOP radial-velocity survey for circumbinary planets-I. Eight years of CORALIE observations of 47 single-line eclipsing binaries and abundance constraints on the masses of circumbinary planets”. In: *Astronomy & Astrophysics* 624, A68 (cit. on p. 135).
- Masuda, K. (Mar. 2014). “Very Low Density Planets around Kepler-51 Revealed with Transit Timing Variations and an Anomaly Similar to a Planet-Planet Eclipse Event”. In: *The Astrophysical Journal* 783.1, 53, p. 53 (cit. on p. 42).
- Maunder, E. W. (1904). “Note on the distribution of sun-spots in heliographic latitude, 1874-1902”. In: *Monthly Notices of the Royal Astronomical Society* 64, pp. 747–761 (cit. on p. 24).

- Mayor, M. et al. (1995). “A Jupiter-mass companion to a solar-type star”. In: *Nature* 378.6555, pp. 355–359 (cit. on p. 13).
- Mazeh, T. et al. (2016). “Dearth of short-period Neptunian exoplanets: A desert in period-mass and period-radius planes”. In: *Astronomy & Astrophysics* 589, A75 (cit. on p. 146).
- McDonald, G. D. et al. (2019). “The sub-Neptune desert and its dependence on stellar type: controlled by lifetime X-ray irradiation”. In: *The Astrophysical Journal* 876.1, p. 22 (cit. on p. 150).
- Meunier, N. et al. (2010a). “Using the Sun to estimate Earth-like planets detection capabilities-II. Impact of plagues”. In: *Astronomy & Astrophysics* 512, A39 (cit. on p. 24).
- Meunier, N. et al. (2013). “Using the Sun to study the impact of stellar activity on exoplanet detectability”. In: *Astronomische Nachrichten* 334.1-2, pp. 141–144 (cit. on p. 22).
- Meunier, N. et al. (2010b). “Reconstructing the solar integrated radial velocity using MDI/SOHO”. In: *Astronomy & Astrophysics* 519, A66 (cit. on p. 24).
- Modigliani, A. et al. (2019). *ESPRESSO Pipeline User Manual* (cit. on pp. 53, 85).
- Morbidelli, A. et al. (2016). “Challenges in planet formation”. In: *Journal of Geophysical Research: Planets* 121.10, pp. 1962–1980 (cit. on p. 40).
- Moriarty, J. et al. (2016). “The Kepler dichotomy in planetary disks: Linking Kepler observables to simulations of late-stage planet formation”. In: *The Astrophysical Journal* 832.1, p. 34 (cit. on p. 42).
- Morris, R. L. et al. (2017). “Kepler Data Processing Handbook: Photometric Analysis”. In: *ksci*, p. 6 (cit. on p. 37).
- Niraula, P. et al. (2017). “Three super-Earths transiting the nearby star GJ 9827”. In: *The Astronomical Journal* 154.6, p. 266 (cit. on p. 145).
- Noyes, R. W. et al. (1984a). “The relation between stellar rotation rate and activity cycle periods”. In: *The Astrophysical Journal* 287, pp. 769–773 (cit. on p. 25).
- Noyes, R. et al. (1984b). “Rotation, convection, and magnetic activity in lower main-sequence stars”. In: *The Astrophysical Journal* 279, pp. 763–777 (cit. on pp. 98, 101).
- Oelkers, R. J. et al. (2018). “Precision Light Curves from TESS Full-frame Images: A Different Imaging Approach”. In: *The Astronomical Journal* 156.3, p. 132 (cit. on p. 36).
- Oshagh, M. et al. (2013). “SOAP-T: a tool to study the light curve and radial velocity of a system with a transiting planet and a rotating spotted star”. In: *Astronomy & Astrophysics* 549, A35 (cit. on p. 26).
- Otegi, J. et al. (2020). “Revisited mass-radius relations for exoplanets below $120 M_E$ ”. In: *Astronomy & Astrophysics* 634, A43 (cit. on p. 148).
- Owen, J. E. et al. (2013). “Kepler planets: a tale of evaporation”. In: *The Astrophysical Journal* 775.2, p. 105 (cit. on p. 145).
- (2017). “The evaporation valley in the Kepler planets”. In: *The Astrophysical Journal* 847.1, p. 29 (cit. on pp. 145, 146).

- Pascale, E. et al. (2018). “The ARIEL space mission”. In: *2018 5th IEEE International Workshop on Metrology for Aerospace (MetroAeroSpace)*. IEEE, pp. 31–34 (cit. on p. 152).
- Pasquini, L. et al. (1991). “H-alpha absolute chromospheric fluxes in G and K dwarfs and subgiants”. In: *Astronomy and Astrophysics* 251, pp. 199–209 (cit. on p. 25).
- Pepe, F. et al. (2002). “The CORALIE survey for southern extra-solar planets VII—Two short-period Saturnian companions to HD 108147 and HD 168746”. In: *Astronomy & Astrophysics* 388.2, pp. 632–638 (cit. on pp. 19, 82).
- Pepe, F. et al. (2021). “ESPRESSO at VLT—On-sky performance and first results”. In: *Astronomy & Astrophysics* 645, A96 (cit. on p. 27).
- Pepe, F. et al. (2008). “From HARPS to CODEX: exploring the limits of Doppler measurements”. In: *Physica Scripta* 2008.T130, p. 014007 (cit. on p. 52).
- Perruchot, S. et al. (2008). “The SOPHIE spectrograph: design and technical key-points for high throughput and high stability”. In: 7014, 70140J (cit. on pp. 47, 48, 50).
- Perruchot, S. et al. (2011). “Higher-precision radial velocity measurements with the SOPHIE spectrograph using octagonal-section fibers”. In: 8151, p. 815115 (cit. on p. 27).
- Petigura, E. A. et al. (2022). “The California-Kepler Survey. X. The Radius Gap as a Function of Stellar Mass, Metallicity, and Age”. In: *The Astronomical Journal* 163.4, p. 179 (cit. on p. 146).
- Piro, A. L. et al. (2020). “Exploring Whether Super-puffs can be Explained as Ringed Exoplanets”. In: *The Astronomical Journal* 159.4, p. 131 (cit. on p. 42).
- Pollacco, D. et al. (2008). In: *Monthly Notices of the Royal Astronomical Society* 385.3, pp. 1576–1584 (cit. on p. 58).
- Pont, F. et al. (2005). “A planet-sized transiting star around OGLE-TR-122—Accurate mass and radius near the Hydrogen-burning limit”. In: *Astronomy & Astrophysics* 433.2, pp. L21–L24 (cit. on p. 33).
- Pu, B. et al. (2015). “Spacing of Kepler planets: Sculpting by dynamical instability”. In: *The Astrophysical Journal* 807.1, p. 44 (cit. on p. 42).
- Quanz, S. P. et al. (2021). “Large Interferometer For Exoplanets (LIFE): I. Improved exoplanet detection yield estimates for a large mid-infrared space-interferometer mission”. In: *arXiv preprint arXiv:2101.07500* (cit. on p. 152).
- Queloz, D. et al. (2001). “No planet for HD 166435”. In: *Astronomy & Astrophysics* 379.1, pp. 279–287 (cit. on p. 23).
- Queloz, D. et al. (2009). “The CoRoT-7 planetary system: two orbiting super-Earths”. In: *Astronomy & Astrophysics* 506.1, pp. 303–319 (cit. on pp. 14, 35).
- Quillen, A. C. et al. (2013). “Origin scenarios for the Kepler 36 planetary system”. In: *Monthly Notices of the Royal Astronomical Society* 435.3, pp. 2256–2267 (cit. on p. 147).
- Quirrenbach, A. et al. (2014). “CARMENES instrument overview”. In: *Ground-based and airborne instrumentation for astronomy V*. Vol. 9147. SPIE, pp. 531–542 (cit. on p. 27).

- Quirrenbach, A. et al. (2018). “CARMENES: high-resolution spectra and precise radial velocities in the red and infrared”. In: *Ground-based and Airborne Instrumentation for Astronomy VII*. Vol. 10702. SPIE, pp. 246–263 (cit. on p. 27).
- Rajpaul, V. et al. (2015a). “A Gaussian process framework for modelling stellar activity signals in radial velocity data”. In: *Monthly Notices of the Royal Astronomical Society* 452.3, pp. 2269–2291 (cit. on p. 25).
- Rajpaul, V. et al. (2015b). “Ghost in the time series: no planet for Alpha Cen B”. In: *Monthly Notices of the Royal Astronomical Society: Letters* 456.1, pp. L6–L10 (cit. on p. 22).
- Rauer, H. et al. (2014). “The PLATO 2.0 mission”. In: *Experimental Astronomy* 38.1, pp. 249–330 (cit. on p. 151).
- Reid, N. I. et al. (2013). *New light on dark stars: red dwarfs, low-mass stars, brown dwarfs*. Springer Science & Business Media (cit. on p. 137).
- Reinhold, T. et al. (2020). “The Sun is less active than other solar-like stars”. In: *Science* 368.6490, pp. 518–521 (cit. on p. 24).
- Ricker, G. R. et al. (2014). “Transiting exoplanet survey satellite”. In: *Journal of Astronomical Telescopes, Instruments, and Systems* 1.1, p. 014003 (cit. on p. 36).
- Ricker, G. et al. (2015). “JATIS, 1, 014003, doi: 10.1117/1”. In: *JATIS* 1, p. 014003 (cit. on p. 36).
- Rodler, F. (2018). “Exoplanet Research in the Era of the Extremely Large Telescope (ELT)”. In: *Handbook of Exoplanets*, p. 194 (cit. on p. 154).
- Rosenblatt, F. (1971). “A two-color photometric method for detection of extra-solar planetary systems”. In: *Icarus* 14.1, pp. 71–93 (cit. on p. 28).
- Roy, A. et al. (2020). “Solar Contamination in Extreme-precision Radial-velocity Measurements: Deleterious Effects and Prospects for Mitigation”. In: *The Astronomical Journal* 159.4, p. 161 (cit. on p. 58).
- Saar, S. H. et al. (1997). “Activity-related radial velocity variation in cool stars”. In: *The Astrophysical Journal* 485.1, p. 319 (cit. on p. 22).
- Saar, S. H. et al. (1999). “Time evolution of the magnetic activity cycle period. II. Results for an expanded stellar sample”. In: *The Astrophysical Journal* 524.1, p. 295 (cit. on p. 24).
- Santerne, A. et al. (2011). “SOPHIE velocimetry of Kepler transit candidates-IV. KOI-196b: a non-inflated hot Jupiter with a high albedo”. In: *Astronomy & Astrophysics* 536, A70 (cit. on p. 58).
- Santerne, A. et al. (2012). “SOPHIE velocimetry of Kepler transit candidates-VII. A false-positive rate of 35% for Kepler close-in giant candidates”. In: *Astronomy & Astrophysics* 545, A76 (cit. on pp. 81, 136).
- Santos, N. et al. (2013). “SWEET-Cat: A catalogue of parameters for Stars With ExoplanETs-I. New atmospheric parameters and masses for 48 stars with planets”. In: *Astronomy & Astrophysics* 556, A150 (cit. on p. 17).
- Santos, N. et al. (2014). “The HARPS search for southern extra-solar planets-XXXV. The interesting case of HD 41248: stellar activity, no planets?” In: *Astronomy & Astrophysics* 566, A35 (cit. on p. 56).

- Scargle, J. D. (1982). “Studies in astronomical time series analysis. II-Statistical aspects of spectral analysis of unevenly spaced data”. In: *The Astrophysical Journal* 263, pp. 835–853 (cit. on p. 18).
- Schlecker, M. et al. (2021). “The New Generation Planetary Population Synthesis (NGPPS)-III. Warm super-Earths and cold Jupiters: a weak occurrence correlation, but with a strong architecture-composition link”. In: *Astronomy & Astrophysics* 656, A71 (cit. on p. 141).
- Schrijver, C. (2002). “Solar spots as prototypes for stellar spots”. In: *Astronomische Nachrichten* 323.3-4, pp. 157–164 (cit. on p. 24).
- Schrijver, C. et al. (2000). “Book Review: Solar and stellar magnetic activity/Cambridge U Press, 2000”. In: *Irish astronomical journal* 27, p. 234 (cit. on p. 23).
- Seager, S. et al. (2003). “A unique solution of planet and star parameters from an extrasolar planet transit light curve”. In: *The Astrophysical Journal* 585.2, p. 1038 (cit. on pp. 29, 34).
- Smith, J. C. et al. (2012). “Kepler presearch data conditioning II-A bayesian approach to systematic error correction”. In: *Publications of the Astronomical Society of the Pacific* 124.919, p. 1000 (cit. on p. 37).
- Sousa, S. et al. (2018). “SWEET-Cat updated-New homogenous spectroscopic parameters”. In: *Astronomy & Astrophysics* 620, A58 (cit. on p. 17).
- Southworth, J. et al. (2017). “Detection of the Atmosphere of the 1.6 M_E Exoplanet GJ 1132 b”. In: *The Astronomical Journal* 153.4, p. 191 (cit. on p. 42).
- Spalding, C. et al. (2016). “SPIN–ORBIT MISALIGNMENT AS A DRIVER OF THE KEPLER DICHOTOMY”. In: *The Astrophysical Journal* 830.1, p. 5 (cit. on p. 42).
- Spronck, J. F. P. et al. (Oct. 2015). “Fiber Scrambling for High-Resolution Spectrographs. II. A Double Fiber Scrambler for Keck Observatory”. In: *PASP* 127.956, p. 1027 (cit. on p. 27).
- Stevens, D. J. et al. (2013). “A posteriori transit probabilities”. In: *Publications of the Astronomical Society of the Pacific* 125.930, p. 933 (cit. on p. 41).
- Sucerquia, M. et al. (2020). “Can close-in giant exoplanets preserve detectable moons?” In: *Monthly Notices of the Royal Astronomical Society* 492.3, pp. 3499–3508 (cit. on p. 150).
- Süveges, M. (2014). “Extreme-value modelling for the significance assessment of periodogram peaks”. In: *Monthly Notices of the Royal Astronomical Society* 440.3, pp. 2099–2114 (cit. on p. 18).
- Tarbell, T. et al. (1987). “On the relation between magnetic field structures and granulation”. In: *The Astrophysical Journal* 317, pp. 892–899 (cit. on p. 24).
- Tarbell, T. et al. (1989). *Statistical properties of solar granulation derived from the SOUP instrument on Spacelab 2*. Tech. rep. PHILLIPS LAB HANSCOM AFB MA (cit. on p. 23).
- Team, L. et al. (2019). “The LUVOIR mission concept study final report”. In: *arXiv preprint arXiv:1912.06219* (cit. on p. 153).
- Thompson, S. J. et al. (2016). “HARPS3 for a roboticized Isaac Newton Telescope”. In: *Ground-based and Airborne Instrumentation for Astronomy VI*. Vol. 9908. SPIE, pp. 1949–1961 (cit. on p. 154).

- Tinetti, G. et al. (2020). “Ariel: Atmospheric Remote-sensing Infrared Exoplanet Large-survey-enabling planetary science across light-years. Definition study report”. In: (cit. on p. 149).
- Triaud, A. H. et al. (2022). “BEBOP III. Observations and an independent mass measurement of Kepler-16 (AB) b—the first circumbinary planet detected with radial velocities”. In: *Monthly Notices of the Royal Astronomical Society* 511.3, pp. 3561–3570 (cit. on pp. 59, 136, 177).
- Twicken, J. D. et al. (2010). “Photometric analysis in the Kepler Science Operations Center pipeline”. In: *Software and Cyberinfrastructure for Astronomy*. Vol. 7740. International Society for Optics and Photonics, p. 774023 (cit. on p. 37).
- VanderPlas, J. T. (2018). “Understanding the lomb–scargle periodogram”. In: *The Astrophysical Journal Supplement Series* 236.1, p. 16 (cit. on p. 19).
- VanderPlas, J. T. et al. (2015). “Periodograms for multiband astronomical time series”. In: *The Astrophysical Journal* 812.1, p. 18 (cit. on p. 18).
- Vanderspek, R. et al. (2019). “TESS Discovery of an Ultra-short-period Planet around the Nearby M Dwarf LHS 3844”. In: *The Astrophysical Journal Letters* 871.2, p. L24 (cit. on p. 151).
- Vogt, S. S. (1994). *HIRES User’s Manual*. 67. Lick Observatory (cit. on p. 27).
- Wang, L. et al. (Mar. 2019). “Dusty Outflows in Planetary Atmospheres: Understanding “Super-puffs” and Transmission Spectra of Sub-Neptunes”. In: *ApJ* 873.1, L1, p. L1 (cit. on p. 42).
- Wehbe, B. et al. (2020). “The impact of atmospheric dispersion in the performance of high-resolution spectrographs”. In: *Monthly Notices of the Royal Astronomical Society* 491.3, pp. 3515–3522 (cit. on p. 53).
- West, R. G. et al. (2019). “NGTS-4b: A sub-Neptune transiting in the desert”. In: *Monthly Notices of the Royal Astronomical Society* 486.4, pp. 5094–5103 (cit. on p. 146).
- Wilson, R. E. et al. (1971). “Realization of accurate close-binary light curves: application to MR Cygni”. In: *The Astrophysical Journal* 166, p. 605 (cit. on p. 29).
- Winn, J. N. (2010a). *Exoplanet transits and occultations*. Vol. 55. University of Arizona Press Tucson (cit. on p. 133).
- (2010b). “Transits and occultations”. In: *arXiv preprint arXiv:1001.2010* (cit. on p. 31).
- Wolszczan, A. et al. (1992). “A planetary system around the millisecond pulsar PSR1257+12”. In: *Nature* 355.6356, pp. 145–147 (cit. on p. 13).
- Yip, K. H. et al. (2020). “On the Compatibility of Ground-based and Space-based Data: WASP-96 b, an Example”. In: *The Astronomical Journal* 161.1, p. 4 (cit. on p. 151).
- Yu, X. et al. (2021). “How to identify exoplanet surfaces using atmospheric trace species in hydrogen-dominated atmospheres”. In: *The Astrophysical Journal* 914.1, p. 38 (cit. on p. 153).
- Zechmeister, M. et al. (2009). “The generalised Lomb-Scargle periodogram—a new formalism for the floating-mean and Keplerian periodograms”. In: *Astronomy & Astrophysics* 496.2, pp. 577–584 (cit. on p. 18).
- Zeng, L. et al. (2016). “Mass–radius relation for rocky planets based on PREM”. In: *The Astrophysical Journal* 819.2, p. 127 (cit. on pp. 147, 148).

- Zeng, L. et al. (2019). “Growth model interpretation of planet size distribution”. In: *Proceedings of the National Academy of Sciences* 116.20, pp. 9723–9728 (cit. on p. [42](#)).
- Zhu, W. et al. (2018). “The Super Earth–Cold Jupiter Relations”. In: *The Astronomical Journal* 156.3, p. 92 (cit. on p. [141](#)).

Appendix A

The skills that I acquired during this thesis, allowed me to plan for my future project with the title of "detection and characterization of TESS single transit events: a strategy to find habitable zone planets". This project is accepted by the French National Center for Space Studies (CNES) and will be funded for two years (2022-2024). I have attached the full description of the project in the subsequent.

Detection and characterization of *TESS* single transit events: a strategy to find habitable zone planets

Introduction:

Among $\sim 5000^1$ discovered planets (and many more candidates), transiting planets (~ 3780) have a considerable impact in improving our understanding of the formation and evolution of planetary systems. Such planets, **when orbiting a bright host star** which allows radial velocity (RV) follow-up, can be accurately characterized in terms of fundamental parameters such as mass and radius, allowing modeling of their internal structure and composition. Moreover, these golden objects give us a great opportunity to gather information about the composition and temperature of their atmospheres, as well as the presence of molecular species, including gases that could be of the biotic origin or surface features.

Among all the planets detected, **the number of habitable-zone planets transiting bright stars is still small**. The majority of transiting planets discovered by the *Kepler* space telescope orbit faint stars, therefore there is a limitation in characterizing their mass and atmosphere (e.g. Earth-size planet Kepler-1649, Vanderburg et al. 2020). The main objective of the Transiting Exoplanet Survey Satellite mission (*TESS*; Ricker et al. 2015) is to search for planets that transit bright stars, typically 30-100 times brighter than the stars *Kepler* surveyed. This offers us a unique opportunity for discovering and characterizing the mass and atmosphere of the planet candidates by follow-up observations. **However, due the *TESS* observation strategy, longer period planets - including habitable zone planets, mostly show up as single transit events. Some of these potential events can easily be missed by the *TESS* pipeline.**

The proposal objective:

I propose a project to detect and characterize *TESS* candidates with an emphasize on single transit events by developing and performing a fully automated pipeline. Since *TESS* has an observing baseline of only ~ 27 days for a large part of the sky, $\sim 90\%$ of its discovered planets have an orbital period shorter than 20 days². *TESS* also reveals numerous (estimated at over **1000**, Villanueva et al. 2019) single transit events; in the case of a planetary origin, they would represent long-period ($P \geq 45$ days) planets. These systems would be part of a **lesser explored part of the parameter space** (see Fig. 1), thus their investigation is potentially very rewarding. They represent a unique opportunity to **investigate a number of questions** related to planet formation, evolution, and different migration scenarios that have not yet been answered. According to the description of the habitable zone by Kopparapu et al. (2013), habitable-zone planets around M4 or earlier spectral-type stars have a longer period than ~ 11 days, most of them thus appearing as a single transit in the *TESS* data. Accordingly, searching for single transit events allow us to **probe planets in the habitable zones**. Moreover, single transits could be used by many upcoming surveys (e.g. PLATO, ARIEL, CHEOPS) for follow-up long-period transiting planets around bright stars. Finally, such a study in addition of help in detection and characterization of unique planets, through **expanding of planetary**

¹<http://exoplanet.eu>

²<https://exoplanetarchive.ipac.caltech.edu/>

population allow us to study planetary occurrence rate and statistic. To date, only a **few** *TESS* single transit events have been characterized as planets (e.g. S. Gill et al 2020a,b, M. Lendl et al 2020). Therefore, it is worth the additional effort required to detect and characterize long period planets which are more challenging because of not providing well-constrained orbital period, but scientifically particularly valuable.

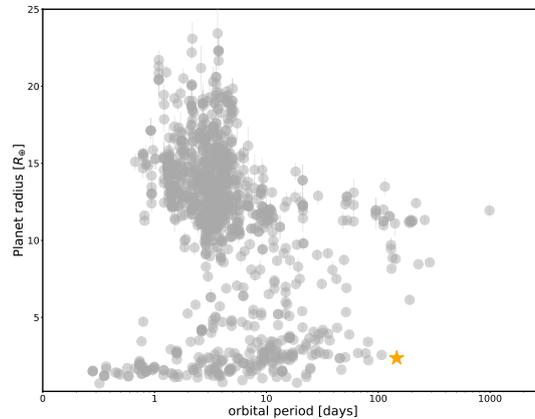


Figure 1 Orbital period versus planet radius of known exoplanets from NASA archival data (February, 2022) with accurate mass and radius (Otegi et al. 2020). HD88986 b (Heidari et al, in prep) is illustrated with a star mark. This figure highlight how HD88986 b, as an example of *TESS* single transit event, has a long period and occupies an exciting parameter space.

Method/work plan:

In this project, I will focus on bright star ($T \leq 9$) in the *TESS* catalog and develop a new automated pipeline to systematically analyze the short cadence of the *TESS* light curves, with aim of finding the transiting planet with a focus on single transit events and validating them. The single transit events are mostly either missed by the *TESS* pipeline, as for most transit detections the pipeline requires at least 2 transit events to achieve a certain S/N, or missed because of stellar variability or instrumental artifacts. Clearly, the single transit events need to be handled differently to prevent losing potential discoveries. This pipeline will execute the following steps: I) Get the *TESS* light curve and remove stellar and instrumental effects. II) Search for transiting events. III) Search for single transit events. Except for the “standard” exoplanet-detection pipelines like those run by Science Processing Operations Center (SPOC) for *Kepler* & *TESS* which are not public, there is only one public code (to my knowledge) called MonoTools³, for searching single transit event. I will do this step through convolution of several transit model on the light curve at various positions and use the goodness of fit to determine whether a single transit is present. IV) If the detected signal has an certain S/N, the pipeline will find a best-fit transit model and vote the candidate through exploring different scenarios of false positive (e.g. odd-even transit depth test, a different image centroid test). V) The single transit events do not provide well-constrained orbital periods. However, in the case of a single transit alert, I will program the pipeline to estimate the planet period from the knowledge of the host star in the *TESS* catalog and transit parameters (Winn et al. 2010) or using the Bayesian transit fitting method NamastePymc3 (Osborn et al. 2016; Gill et al. 2020). VI) Search for complementary photometric and spectroscopic public/private observations (through G. Hébrard, who is the supervisor of this work, I have access to observing open time dedicated

³<https://github.com/hposborn/MonoTools>

on transit confirmation on SOPHIE and SPIRou spectrograph) in order to 1. Discriminate between false positives and real transiting exoplanets. 2. Measure/increase the precision on the planetary radii, mass measurements, and explore the architecture of the planetary systems. If necessary, complementary observations can be performed through submitting observing proposals for the CHEOPS ESA mission which I am the **PI on a DDT on CHEOPS allocated of 7 continues days of HD88986 observations** (Heidari et al, in prep). I can also submit follow-up proposals on other high precision RV spectrograph such as ESPRESSO and HARPS.

After developing and testing the pipeline, I will make the code public on GitHub. Such code not only will be useful for the *TESS* space mission but also will be applicable to future space telescopes such as PLATO. I expect developing and testing the pipeline will take around 4 months time, then I will run the pipeline on the *TESS* data and search for candidates (5 months). The remainder of the time (~ 15 months) I will work on validating candidates and write observation proposals and publications.

Environment

This study will be based at the Institut d’astrophysique de Paris (IAP). The IAP has a large program of detecting and characterizing transiting planets. Thus, it provides an ideal environment for me in the interpretation of data under the supervision of G. Hébrard, who has a long time experience on detection and characterization of transiting planets especially those with longer periods (e.g. Hébrard et al 2019, 2010). He is responsible of transiting planet follow-up with SOPHIE and SPIRou high precision spectrographs which allows me to benefit using the radial velocity data. In IAP, I can also benefit of collaboration with A. Lecavelier who is one of the two French members of the CHEOPS board team (the other is M. Deleuil which I already have a strong collaboration with her) for follow-up observation by CHEOPS. The strong existing scientific interaction between IAP and several institutions in France (e.g., LAM, IPAG) and around the world (e.g. *TESS* science team) will provide me with a large collaboration.

Expected impact

I propose a project to detect and characterize the transiting planet with a focus on single transit events. The proposed project in addition to extending the number of planets and accurately characterizing them will allow us to advance our understanding through the detection of long-period planets which are less explored. Habitable zone planets which are still a few can be found by more chance on this lesser explored part of the parameter space. The project can also provide targets for current and future ground and space telescopes (e.g., VLT, JWST, PLATO, ARIEL). Moreover, the developed code will become public and be useful for users for future space missions.

Appendix B

Our work on the BEBOP program marks the first radial-velocity detection of a circumbinary planet (Triaud et al., [2022](#)). This work which I carried out all its data processing (see Sect. [3.7.1](#)), has been the subject of several Press Releases. In the following, I have provided one of them (in french).

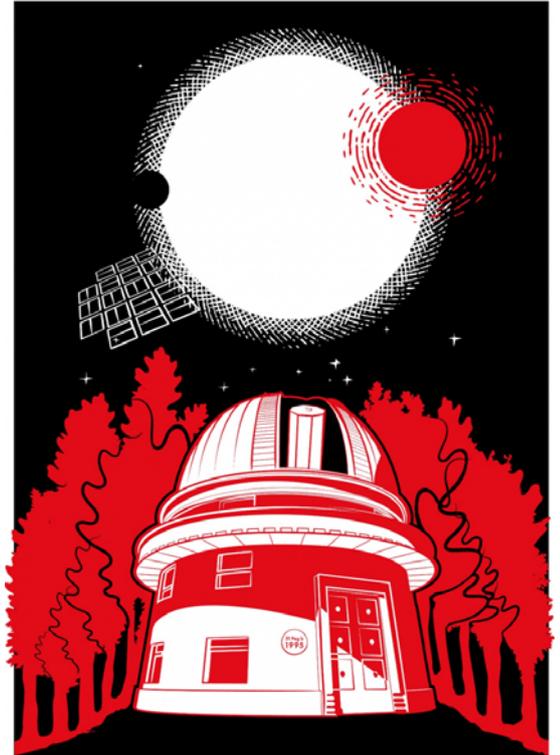
Communiqué de presse du 25/02/2022 – OSU Institut Pythéas

University of Birmingham News Release

Une planète semblable à "Tatooine" observée par un télescope de l'Observatoire de Haute Provence (OHP) de l'OSU Institut Pythéas

Une exoplanète rare qui tourne autour de deux étoiles à la fois a été détectée à l'aide d'un télescope de l'OHP grâce à une collaboration étroite entre l'université de Birmingham et le Laboratoire d'Astrophysique de Marseille.

La planète, appelée Kepler-16b, n'avait jusqu'à présent été observée qu'à l'aide du télescope spatial Kepler. Elle est en orbite autour de deux étoiles. Les deux étoiles sont également en orbite l'une autour de l'autre, ce que l'on appelle un système stellaire binaire. Kepler-16b est située à quelque 245 années-lumière de la Terre et, comme la planète Tatooine de Luke Skywalker, dans l'univers de la Guerre des étoiles, elle présenterait deux couchers de soleil si l'on pouvait se tenir à sa surface.



© Amanda Smith – University Of Birmingham

C'est avec le fameux télescope de 193 cm de l'OHP équipé de son spectrographe SOPHIE que cette nouvelle observation a été effectuée. L'équipe a pu détecter la planète en utilisant la méthode des vitesses radiales, dans laquelle les astronomes observent les effets d'une planète sur le mouvement de l'étoile autour de laquelle elle est en orbite.

La détection de Kepler-16b grâce à la méthode des vitesses radiales met en évidence qu'il est aussi possible de détecter des planètes circumbinaires à l'aide de méthodes traditionnelles d'observation au sol moins coûteuses que l'utilisation de télescopes spatiaux. Notons aussi que cette méthode permet de détecter plus facilement la présence d'autres planètes dans un système, et qu'elle permet de mesurer la propriété la plus fondamentale d'une planète ; à savoir sa masse.

Après avoir démontré l'efficacité de cette stratégie à de Kepler-16b, l'équipe poursuit la recherche de planètes circumbinaires encore inconnues jusqu'à présent. L'objectif est de contribuer à répondre aux questions sur la formation des planètes. Habituellement, on pense que la formation des planètes a lieu dans un disque protoplanétaire - un disque de poussière et de gaz qui entoure une jeune étoile. Cependant, ce processus pourrait ne pas être possible dans un système circumbinaire.

En se basant sur ce modèle standard, il est difficile de comprendre comment des planètes circumbinaires peuvent exister. La présence de deux étoiles interfère en effet avec le disque protoplanétaire, ce qui empêche la poussière de s'agglomérer en planètes, un processus appelé accrétion.

Le Professeur Amaury Triaud, de l'Université de Birmingham, premier auteur de l'article précise : *"on peut aussi supposer que la planète s'est formée loin des deux étoiles, là où leur influence est plus faible, puis qu'elle s'est déplacée vers l'intérieur dans un processus appelé migration induite par le disque. Nous pourrions également arriver à la conclusion que nous devons revoir notre compréhension du processus d'accrétion planétaire."*

Neda Heidari, étudiante en thèse au LAM, qui a analysé les mesures de vitesses radiales explique *"Les planètes circumbinaires fournissent l'un des indices les plus clairs que la migration engendrée par les disques est un processus crédible, et qu'il se produit régulièrement."*

Le Dr Isabelle Boisse du LAM, scientifique responsable de l'instrument SOPHIE qui a été utilisé pour collecter les données de cette étude explique : *"Cette découverte met clairement en évidence que les télescopes au sol restent tout à fait pertinents pour mener des recherches sur les exoplanètes et qu'ils peuvent permettre de développer de nouvelles stratégies d'observation tout à fait passionnantes. Dans la continuité de cette étude nous allons maintenant analyser les données prises sur de nombreux autres systèmes d'étoiles binaires, et rechercher de nouvelles planètes circumbinaires."*

Le Dr Alexandre Santerne, lui aussi du LAM et responsable de l'obtention des données, explique : *"Kepler-16b a été découvert pour la première fois il y a 10 ans par le satellite Kepler de la NASA en utilisant la méthode des transits. Ce système a été la découverte la plus inattendue faite par Kepler. Nous avons choisi d'utiliser le télescope de 193 cm de l'OHP afin de démontrer que la méthode des vitesses radiales pouvait également permettre d'étudier des systèmes tels que Kepler-16."*

Notes :

- Référence de l'article scientifique : Triaud et al. (2022). BEBOPIII : Observations and an independent mass measurement of Kepler-16 (AB) b - the first circumbinary planet detected with radial velocities'. BMC Biology.
- L'équipe est composée de scientifique provenant de : University of Birmingham (Royaume Uni), Aix-Marseille Université (France), Centre National de la Recherche Scientifique (France), Shahid Beheshti University (Iran), Université de Nice-Sophia Antipolis (France), Ohio State University (USA), Universidad de Coimbra (Portugal), Observatoire de Paris (France), University of Warwick (Royaume Uni), Université de Grenoble Alpes (France), University of St Andrews (Royaume Uni), Institut d'Astrophysique de Paris (France), Universidad de Porto (Portugal), Observatoire de Genève (Suisse), Keele University, Laboratório Nacional de Astrofísica (Brazil), Queen Mary University London (Royaume Uni), Université Paris Est Créteil (France), Université Paris-Saclay (France) et University of Toronto (Canada).

Contact scientifique

Isabelle Boisse

Laboratoire d'astrophysique de Marseille (LAM / CNRS / Aix-Marseille Université)
isabelle.boisse@lam.fr

Alexandre Santerne

Laboratoire d'astrophysique de Marseille (LAM / CNRS / Aix-Marseille Université)
alexandre.santerne@lam.fr