EXPLORING THE ORIGINS AND DIVERSITY OF PLANETARY SYSTEMS: INSIGHTS FROM YOUNG GIANT PLANETS

A thesis submitted by

Alexis Heitzmann, M. Sc.

For the award of

Doctor of Philosophy

2022

Abstract

STUDYING PLANETARY SYSTEMS in their infancy is essential to understand of the origin of gas giant exoplanets found to closely orbit their host stars. Tied with the study of young stars is the challenge of stellar activity, arising from their fast rotation and strong magnetic fields. This cocktail yields a wealth of phenomena able to mask signatures from orbiting planets, greatly hampering our capacity to detect them.

In this Ph. D. project, we first assessed our chances to detect non-transiting gas giants in close orbit around young and very active stars, using legacy data from a non-stabilised spectrograph. We performed an injection and recovery study using Doppler Imaging and Gaussian Processes as stellar activity mitigation strategies. We were able to recover planets down to 4 times below the stellar noise and conclude that both a robust statistical framework and high quality datasets are crucial. We then precisely measured, despite the stellar activity, the spin–orbit alignment for the youngest Jupiter-sized planet known to transit, the 17 Myr old HIP 67522 b, and showed that its origin pathway is unlikely to be the commonly invoked high eccentricity migration. Finally, we conducted a 2-year radial velocity follow up yielding the discovery of TOI-4562 b, identified in Transiting Exoplanet Survey Satellite data. TOI-4562 b is a highly eccentric ($e = 0.77 \pm 0.03$), 225 day period, 300 ± 190 Myr old Jupiter analogue. We strongly suspect the presence of a companion that could be at the origin of TOI-4562 b eccentricity. We hope to reveal the 3D architecture of this system with on-going observing campaigns.

This work paves the way for further characterisation of young short-orbit gas giants, key to constrain models explaining their formation and evolution.

Certification of Thesis

I, Alexis Heitzmann, declare that the Ph. D. thesis entitled "Exploring the origins and diversity of planetary systems: Insights from young giant planets" is not more than 100 000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references, and footnotes.

This thesis is the work of Alexis Heitzmann except where otherwise acknowledged, with the majority of the contribution to the papers presented as a thesis by publication undertaken by the student. The work is original and has not previously been submitted for any other award, except where acknowledged.

Signed: October 27, 2022

Endorsed by:

Associate Professor Stephen C. Marsden
Principal Supervisor

Dr. George Zhou
Associate Supervisor

Dr. Duncan Wright
Associate Supervisor

Dr. Pascal Petit
Associate Supervisor (external)

Student and supervisors' signatures of endorsement are held at USQ.

Statement of Contribution

This section details contributions by the various authors for each of the peer-reviewed or in preparation papers presented in this thesis by publication.

Chapter 3: Heitzmann, A., Marsden, S. C., Petit, P., et al. (2021a). Planets around young active solar-type stars: assessing detection capabilities from a non-stabilized spectrograph. MNRAS, 505(4), 4989–5011

Author	Percent Contribution	Tasks Performed
A. Heitzmann	80	Designed the injection/recovery planet experiment, simulated all the datasets, developed and tested various self-implemented method for the data analysis, including the Gaussian Process code, adapted the ZDIpy code to suit the analyses of the simulated datasets, run the analysis on all datasets and interpreted the results, wrote all paper drafts.
S.C Marsden, P. Petit, M.W. Mengel, D. Wright, M. Clerte, I. Millburn, C.P. Folsom, B.C. Addison, R.A. Wittenmyer & I.A Waite.	20	Provided original legacy datasets, guidance on the Doppler Imaging and Zeeman Dopper Imaging techniques and code, suggested edit to the manuscripts.

Chapter 4: Heitzmann, A., Zhou, G., Quinn, S. N., et al. (2021b). The Obliquity of HIP 67522 b: A 17 Myr Old Transiting Hot, Jupiter-sized Planet. ApJ, 922(1), L1

Author	Percent Contribution	Tasks Performed
A. Heitzmann	80	Came up with the project idea, designed and wrote the code for data analysis from scratch, performed data analysis and interpretation, wrote all paper drafts.
G. Zhou, S.N. Quinn, S.C. Marsden, D. Wright, P. Petit, A.M. Vanderburg, L.G. Bouma, A.W. Mann & A.C. Rizzuto.	20	Observation scheduling, data collection, guidance for data analysis, provided comments and edits to the draft.

Chapter 5: Heitzmann, A., Zhou, G., Quinn, S. N., et al. (submitted to the Astrophysical Journal, August 2022). TOI-4562 b: A highly eccentric cool Jupiter analog orbiting a young field star.

Author	Percent Contribution	Tasks Performed
A. Heitzmann	80	Combined data from all sources and performed the data analysis and interpretation, wrote all paper drafts.
G. Zhou, S.N. Quinn, C.X. Huang, J. Dong, L.G. Bouma, R.I Dawson, S.C. Marsden, D. Wright, P. Petit, K.A. Collins, K. Barkaoui, R.A. Wittenmyer, E. Gillen, M. Battley, D.L Pollacco, A. Triaud, R. Brahm, M. Hobson, C. Hellier, D.W. Latham, J.J. Lissauer, E.R. Newton, R. Rampalli, D.R Ciardi & A.M. Vanderburg.	20	Identification of the candidate, observation scheduling, data collection (photometry, spectroscopy) , provided comments and edits to the draft.

The publications presented below are ones to which I contributed as a Ph. D. candidate, but not as first author. A short description of my contribution is given alongside each reference.

1. Addison, B. C., Wright, D. J., Nicholson, B. A., et al. (2021). TOI-257b (HD 19916b): a warm sub-saturn orbiting an evolved F-type star. MNRAS, 502(3), 3704–3722:

   I took part in *TOI-257b (HD 19916b): a warm sub-saturn orbiting an evolved F-type star*, through my involvement with the MINERVA Australis telescope array (detailed in appendix B). I conducted manual observations of targets shortly after MINERVA Australis' commissioning (2019).

2. Addison, B. C., Horner, J., Wittenmyer, R. A., et al. (2021). The youngest planet to have a spin-orbit alignment measurement AU mic b. *The Astronomical Journal*, 162(4), 137

   For *The youngest planet to have a spin-orbit alignment measurement AU Mic b*, I performed the simulations shown in Figure 6 and described in section 3.1 of the paper. This simulation assesses the influence on the Rossiter-McLaughlin effect measured from a single large starspot on the stellar surface. I used my own model and code, originally developed for chapter 4 and shown in appendix C.

3. Dong, J., Huang, C. X., Zhou, G., et al. (2021). TOI-3362b: A Proto Hot Jupiter Undergoing High-eccentricity Tidal Migration. ApJ, 920(1), L16

   For *TOI-3362b: A proto hot Jupiter undergoing high-eccentricity tidal migration*, I carried out observations with MINERVA Australis. At the time of the observations of this target, the entire telescope array was capable to observe in a fully automated manner requiring no human intervention. However, some 'difficult' targets, having bright stars nearby rendering the telescope autonomous guiding complicated required manual observations to which I contributed.

4. Zhou, G., Wirth, C. P., Huang, C. X., et al. (2022). A Mini-Neptune from TESS and CHEOPS Around the 120 Myr Old AB Dor Member HIP 94235. AJ, 163(6), 289

   My contribution to *A mini-neptune from TESS and CHEOPS around the 120 Myr old AB Dor member HIP 94235* is also through manual observations of this target using MINERVA Australis.

5. Psaridi, A., Bouchy, F., Lendl, M., et al. (2022). Three new brown dwarfs and a massive hot Jupiter revealed by TESS around early-type stars. *arXiv e-prints*, (pp. arXiv:2205.10854)

   For *Three new brown dwarfs and a massive hot Jupiter revealed by TESS around early-type stars*, I handled the planning of the observations of TOI-1107 with MINERVA Australis. I designed the automated scheduler from the observatory which enabled efficient observations campaigns for targets such as the one presented in this paper.

ACKNOWLEDGMENTS

Contents

List of Figures

(Excluding publications included in chapters 3 to 5

Units & Abbreviations

Chapter 1: Introduction

People from all communities, whether aboriginal Australians, Mayans or present-day Europeans have looked up to the stars and wondered about our place in the cosmos. A recurring question has concerned humanity's loneliness on this life-harbouring, terrestrial planet that we call Earth. The quest of foreign extra-solar worlds is arguably one of the most meaningful in human history, and we have come a long way since the first people started scrutinising the night sky.

The advent of the scientific method and the availability of ever improving technology allowed us to move past the old western centred view that placed the Earth at the centre of the Solar system and the Universe. Following these realisations, it has been theorised that planets must exist beyond our own Solar system, orbiting other stars in the Milky Way. This was confirmed in 1992 when astronomers discovered two planets orbiting a pulsar (a fast rotation radio emitting neutron star), PSR B1257+12 c & d (Wolszczan & Frail, 1992). In 1995, the discovery of 51 Pegasi b (Mayor & Queloz, 1995), the first exoplanet to orbit a Sun-like Star (of comparable mass, radius and surface temperature as the Sun) shook the community and opened up a new field in astronomy: exoplanet science.

In March 2022, the community passed the staggering milestone of 5000

discovered exoplanets. Although current detection capabilities and barriers such as stellar intrinsic variability (see section 2.4.3) still keep true Earth analogues out of our reach, we are finding an unexpected variety of exo-worlds. Unforeseen features of planetary systems architectures have emerged from the large sample of known planets. Such features include: the existence of giant planets orbiting their host stars in a few days to a few weeks (see section 2.2); the paucity of planets with radii between 1.5 and 2 Earth masses ($M_\oplus$) called the radius valley (Owen & Wu, 2013; Fulton et al., 2017; Lee & Connors, 2021); the commonness of the previously unknown few $M_\oplus$ terrestrial super-Earths and the seemingly high number of giant planets found in the ~10 to ~100 days Period Valley (Wittenmyer et al., 2010; Santerne et al., 2016, see section 2.3.2). New models are emerging to describe the processes shaping the observed diversity of planetary systems and their peculiar architectures. Astronomical observations are mandatory to validate these models and build a coherent and comprehensive picture of the formation and evolution of planetary systems.

In this Ph. D. thesis, I specifically focus on larger than Neptune gas giants that share a common crucial characteristic, their young age. I aim to bring my contribution to the understanding of planetary systems through exoplanet characterisation and assessment of our limitations when attempting to discover such young candidates. As I will develop upon, characterising planets orbiting very young stars is a promising although challenging way to find the missing pieces of the puzzle of planetary systems formation and evolution.

IN THIS INTRODUCTORY CHAPTER, we will first outlay the research objectives and the research questions motivating this Ph. D. thesis (section 1.0.1).

Then we will give background on the origin of exoplanets (section 2.1) and the different pathways explaining the formation and evolution of giant planets in short-orbit (section 2.2). Section 2.3 will discuss our current understanding of gas giant origin through the lens of two distinct properties of planetary systems: obliquities (section 2.3.1), and occurrence and eccentricity of warm Jupiters (section 2.3.2), both relevant to the work presented in chapters 4 and 5. Section 2.4 will explain why focusing on young planets is paramount in this context although tied to strong difficulties: the rarity of young planets (section 2.4.1), the determination of stellar ages (section 2.4.2) and the intrinsic activity of young stars (section 2.4.3). We will conclude this introductory chapter in section 2.4.4 by discussing the strategies aiming to mitigate the effect of stellar activity, relevant to chapter 3.

1.0.1 Thesis objectives and research questions

The main objective of this Ph. D. is to contribute to exoplanetary science by investigating how giant planets orbiting young stars can help our understanding of the mechanisms behind the observed diversity of planetary systems. This revolves around one of the major questions in exoplanetary science:

How do planets form and evolve?

More specifically, I attempted to answer the following two questions:

1. **Can we reliably detect gas giants in short orbit around young and very active stars using radial velocities[1]?**

   Stellar variability can mimic planetary signatures. This renders exoplanet detection, especially around young and intrinsically active stars, very

---

[1] We refer the reader to appendix A for a detailed description of the derivation of stellar radial velocities and how this is used for exoplanet detections.

challenging. Many strategies have emerged to hope to mitigate this stellar noise. However, different approaches can yield different results and need to be benchmarked. It is also important to have a robust statistical framework to avoid false positive detections. This is what we explored in the first publication of my candidacy, described in chapter 3. For this work, we hid simulated datasets of planets in real observations of HD 141943 to extensively test two mitigation strategies, Doppler Imaging and Gaussian Processes in order to assess detection capabilities for large planets closely orbiting young and highly active stars.

2. **Can the characterisation of young planets provide meaningful insights into the processes shaping the observed variety in the gas giants exoplanet population?**

All the proposed mechanisms to explain the current exoplanet demographics, including the origin of short-orbit gas giants (or hot Jupiters), are only testable through observations. Catching young planets as they experience these mechanisms (e.g. migration, photo-evaporation, scattering, etc.) appears as a promising avenue to constrain models and the associated processes timescales. Because one single planet or system is unlikely to unveil the full picture, efforts are required to sample the young end of the planet characteristics distribution, despite the challenges associated with young stars (see section 2.4.3). We contributed to this effort by (i) performing the youngest obliquity measurement to date for the hot Jupiter-sized planet HIP 67522 b and (ii) discovering TOI-4562 b, a 'warm' Jupiter on a very eccentric, 225 days orbit around its young host star. These contributions are described in chapters 4 and 5, respectively.

CHAPTER 2: BACKGROUND AND STATE OF THE FIELD

The population of more than 5000 exoplanets discovered to date is surprisingly diverse. In Figure 2.1, we show all known exoplanets with orbital periods $P_{orb} < 10^4$ days with measured radii (mostly resulting from the transit method) or minimum masses (mostly resulting from the radial velocity, or RV, method). Planets with known bulk densities, i.e., with measured mass and radius, are shown with dark edges. We can differentiate terrestrial planets without a significantly large atmosphere (e.g. Mercury, Venus, Earth or Mars), shown as brown dots[1] on the figure from the gas & icy giants with an extended H/He atmosphere (e.g. Jupiter, Saturn, Uranus or Neptune). The latter are defined as planets with masses > 0.3 M_J (or alternatively, radii > 6 $R_{\oplus}$) and will be referred to as gas giants. They can be divided into three sub-categories, as a function of their orbital distance from the central star (function of P_{orb}). The first kind are hot Jupiters (and by extension hot Saturn, Neptunes, etc. depending on their size), orbiting their star in $\lesssim$ 10 days and depicted as orange circles[1] in Figure 2.1. We then find the warm Jupiters (yellow squares[1]), with P_{orb} between $\sim$ 10 and $\sim$ 250 days. Finally, we call cold gas giants planets with $P_{orb} > 250$ days (shown as blue diamonds[1]). In

---

[1] This colour code will be consistent throughout all figures of this manuscript.
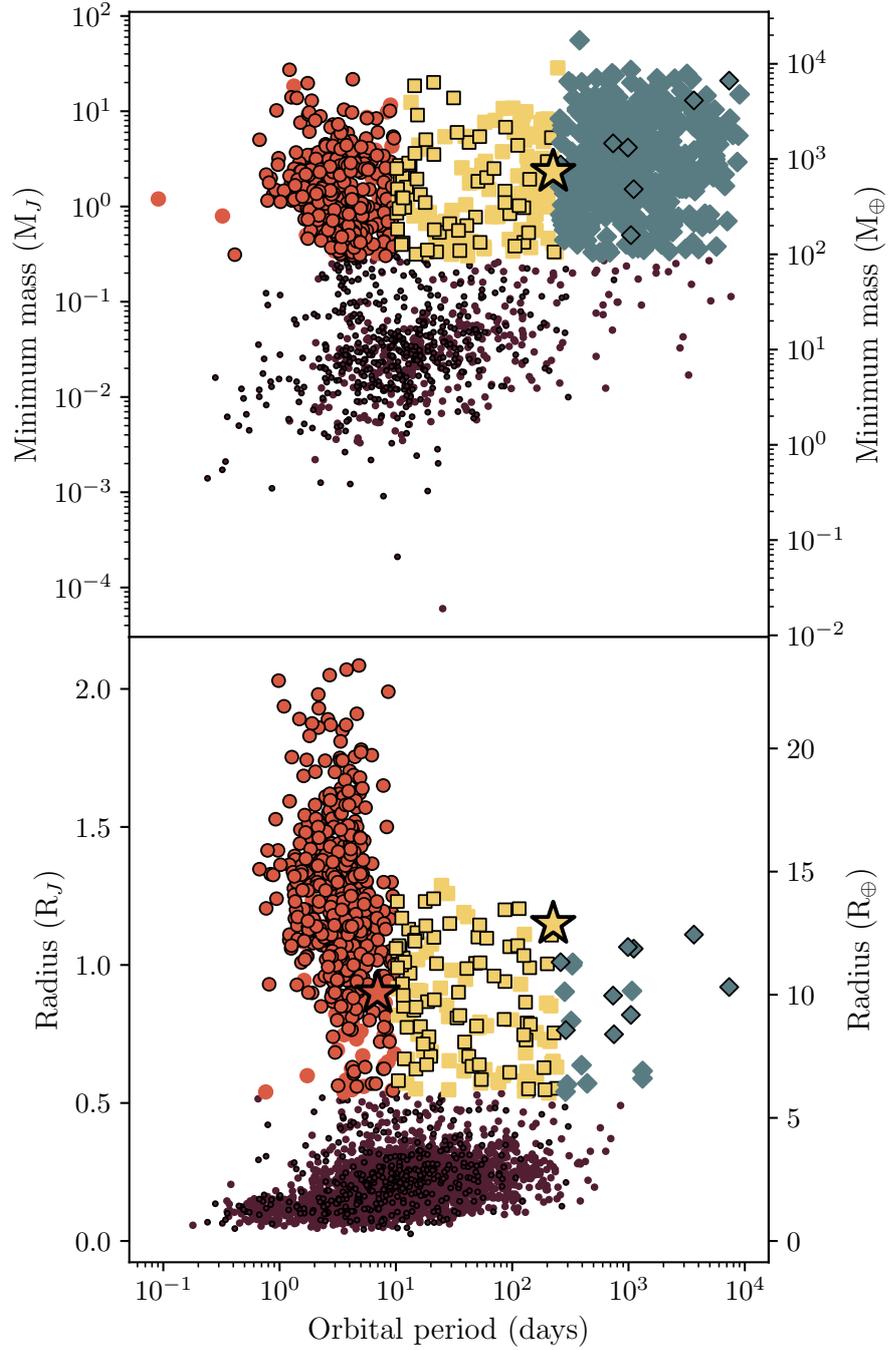
Figure 2.1: Distribution of all discovered exoplanets (data from the NASA exoplanet archive, June 2022) with $P_{orb} < 10^4$ days (4821 exoplanets). We show P_{orb} versus planet mass (top) and radius (bottom). Brown dots are 'small' planets, defined as having either mass < 0.3 M_J (top) or radius < 6 $R_\oplus$ (bottom). Among larger/heavier planets, we find hot gas giants shown as orange circles ($P_{orb} < 10$ days), warm gas giants as yellow circles (10 days $< P_{orb} <$ 250 days) and cold gas giants as blue diamonds ($P_{orb} > 250$ days). Markers with contours are planets with both mass and radius measurements. The two star labels show HIP 67522 b (orange) and TOI-4562 b (yellow), subject of the work presented in chapters 4 and 5. We note that because the mass of HIP 67522 b is unknown, the planet does not appear in the upper plot.

this section, we will describe the current paradigm on planet formation, and then focus on the still debated origin of the short-orbit gas giants (i.e., hot and warm Jupiters).

2.1   The origin of planets

Stars originate from the collapse of interstellar gas clouds. These clouds are mainly composed of molecular hydrogen (H_2) plus traces of other heavier elements and molecules, relics from supernovae, the catastrophic death of previous generations of stars. Following a perturbation, a cloud will become unstable and experience gravitational collapse. Clumps will aggregate and become locally denser as more material is accreted onto them. When the pressure and temperature at the centre of these spherical bodies becomes sufficient, hydrogen fusion is triggered. This produces an outward energy flow, in the form of photons (i.e., light particles) exerting radiation pressure on the in-falling material and the newly born star becomes stable, in thermodynamic equilibrium. The remaining material surrounding the star possesses some angular momentum inherited from the initial collapse and will end up in the most stable configuration, a disc. This protoplanetary disc extends to several tens of Astronomical Units (AU). Gas and dust present in such discs will start to aggregate into larger and larger objects, until a planetary core is built up. This process is believed to be completed in a few millions of years. From this point on, protoplanets are embedded in the disc and a distinction will be made between future terrestrial and gas giant planets.

2.2   Gas giants and the puzzle of their formation

Broadly speaking, gas giants are composed of a massive core (5 to 20 $M_\oplus$) made of metals, silicates and ices, surrounded by a massive envelope largely composed of H and He.

Historically, the only observing ground for understanding their formation has been our own Solar system. One noticeable feature in the Solar system is the separation between rocky planets, orbiting within two AUs of the Sun, and the gas giants, all with orbital distances larger than 5 AUs. At distances greater the than so-called 'ice line' ($\gtrsim 3.1$ AUs for the Solar system, Martin & Livio 2012), water and other volatile species are found in solid form, broadly referred as 'ices'. In the current accepted model of giant planet formation called core accretion (Pollack et al. 1996, see Chabrier et al. 2014 for a review), these ices enhance the protoplanetary disc's surface gravity allowing quick growth of protoplanetary cores up to 20 $M_\oplus$. During the following few millions of years, these protoplanet cores will steadily grow and slowly accrete gas from their surroundings. Once the resulting gaseous envelope reaches a mass similar to that of the core, runaway accretion is triggered, where the rate of gas accretion onto the planet will exponentially grow. This process terminates when no more gas is available in the vicinity of the young gas giant, either due to a gap carved in the disc or depletion of the disc's gas by the central star. Finally, the planet will start to cool down and contract to reach its final state.

Following this core accretion paradigm, gas giants originate from beyond the ice line. In 1995, the discovery of the first exoplanet orbiting a Sun-like star, 51 Pegasi b (Mayor & Queloz, 1995), dramatically challenged this picture. This giant planet, half the mass of Jupiter and twice its size, orbits its

host star in a mere 4.23 days and lies 0.052 AU from its host star, far inside the ice line. As it turns out, more than 10% of discovered exoplanets are these hot Jupiters (visible on Figure 2.1). Their extremely short orbits (down to 16 hours for TOI-2109 b, the most extreme case, Wong et al. 2021) mean that these worlds receive extreme irradiation from their host stars, yielding surface temperatures up to a few thousand Kelvin. Routinely discovered, observations are highly biased towards their detections in (i) RVs, where their close-orbit and high masses yield a large radial velocity signature, and (ii) transits, due to their size and close-orbit resulting in deep, highly probable transits. Correcting for such biases, it is estimated that $\lesssim 1$ % of stars are hot Jupiter hosts (Wright et al., 2012; Petigura et al., 2018; Zhou et al., 2019). This seemingly small population of exoplanets, along with their slightly cooler counterparts, the warm Jupiters, found on both circular and eccentric orbits (see section 2.3.2) are difficult to reconcile with the simplistic 'gas giants are born outside the ice-line' view.

Different mechanisms are invoked to explain the existence of gas giants closely orbiting their parent star. In the coming sections 2.2.1 to 2.2.3, we will review the three main scenarios attempting to explain the existence of short orbit gas giants. We encourage the reader to consult Dawson & Johnson (2018) for an extensive review on the origin of hot Jupiters.

2.2.1 In-situ formation

In-situ formation proposes the seemingly simple idea that hot Jupiters are formed in the vicinity of their host star, where we later observe them, as illustrated on the bottom left of Figure 2.2. However, two constraints apply to the core accretion model for it to thrive in the inner region of the protoplan-
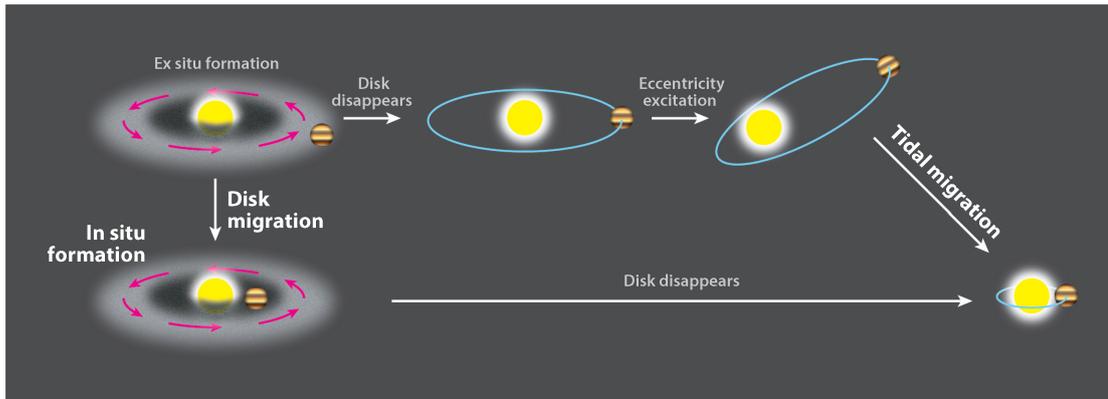
Figure 2.2: Illustration from Dawson & Johnson (2018) showing the three main mechanisms invoked to explain the presence of gas giants in short orbit: In-situ formation, gas disc migration and high eccentricity migration.

etary disc. First, the gas present in the disc will be depleted in a few million years by the host star, destroying any chances for a future gas giant to grow an atmosphere if its core is not fully grown by then. Secondly, enough material is required to reach the core mass required to trigger runaway gas accretion. Simulations struggle to generate such massive cores close to the central star, where the absence of condensed compounds yield a much lower disc surface density (by up to a factor of 4 compared to beyond the ice line, Lecar et al. 2006). This could be reconciled with mechanisms such as pebble accretion enabling an efficient transport of material to the inner regions of the disc and a quicker core growth (for a review, see Johansen & Lambrechts, 2017). In–situ formation is still debated (Batygin et al., 2016) as we do observe super-Earths with relatively short periods (< 100 days) and masses comparable to a gas giant core.

The competing explanations for the existence of gas giants in short orbits involve a stage of migration, where the gas giants do originate from beyond the

ice line but later experience a reduction of their semi-major axis (i.e., orbital distance). Two main mechanisms are invoked to transform primordially cold Jupiters into short orbit hot/warm Jupiters: gas disc migration (section 2.2.2) and high eccentricity migration (section 2.2.3).

2.2.2 GAS DISC MIGRATION

This second scenario (reviewed in Baruteau et al., 2014), called gas disc migration, was first proposed as an explanation for the short-orbit of 51 Pegasi b (Lin et al., 1996). In this picture, the gas from the disc exerts drag on the newly-born planet, reducing its angular momentum and therefore its orbital distance, until it reaches its final orbital configuration close to the central star. The strongest constraints imposed on gas disc migration to be the origin of hot Jupiters is that the entire process has to take place before the gas dissipates, i.e., within the first few million years of the system's lifetime. For this scenario, shown on the top left of Figure 2.2, tight constraints on the migration timescales are difficult to obtain as the migration rate is highly dependent on both planet mass and local disc properties (e.g. density and temperature profiles, turbulent viscosity and magnetic fields). Complementary mechanisms are also required to halt the hot Jupiter in its approach and avoid it to fall onto the central star. It could be stopped by the star's magnetic field driven cavity (Trilling et al., 1998; Valsecchi et al., 2015) or via a transfer of angular momentum from the star to the migrating planet as a result of tidal interactions (Rice et al., 2008; Chang et al., 2010).

2.2.3 High eccentricity migration

This last scenario occurs in two distinct phases, illustrated in the top part of Figure 2.2. First, the newly formed cold gas giant experiences dynamical torques, transforming its primordially circular orbit into a highly elliptical one. This can occur sporadically in planet–planet scattering (Weidenschilling & Marzari, 1996; Rasio & Ford, 1996; Ford & Rasio, 2006; Chatterjee et al., 2008), via close encounters with other planets (in closely packed systems, Jurić & Tremaine 2008) or stellar fly-bys (Shara et al., 2016; Rodet et al., 2021). It can also be the result of secular interactions, or slower angular momentum exchanges with a other planet yielding high eccentricity, either periodically through e.g. von Zeipel–Lidov–Kozai cycles (von Zeipel 1910; Lidov 1962; Kozai 1962, see Naoz 2016 for a review) or chaotically in secular chaos (Wu & Lithwick, 2011; Hamers et al., 2017).

In the second phase, the highly elliptical orbit is circularised through tides raised on the planet by the star periodically at periapse (i.e., closest approach), where tidal forces are the strongest. Once the planet is free from the influence of the perturber responsible of the high eccentricity, the evolution of the orbit is traceable. The planet follows a path of constant angular momentum, where the evolution of its semi-major axis (a) and eccentricity (e) evolve such that $a(t)[(1 - e(t)^2)] = \text{constant} = a_{final}$. This final semi-major axis a_{final} determines the circularization timescale τ_{circ}, such that $\tau_{circ} \propto a_{final}^8$ (Eggleton et al., 1998).

Timescales for the eccentricity excitation can be extremely fast (few 10^3 years) for a favourable configuration in planet–planet scattering and up to some 10^6 years in some cases via Zeipel–Lidov–Kozai cycles. Adding in the circularisation timescales, which itself can vary from 10^5 to few 10^9 years,

identifying high eccentricity migration as a mechanism for producing short orbit gas giants solely based on timescales is non trivial. However, a decisive property of this scenario that differentiates it from both in-situ formation and gas disc migration is that it has to occur after disc dissipation as the gas from the disc acts as a damper of eccentricity and prevents orbits from becoming dramatically elliptical. Let us review the available clues brought from observations to help differentiate between formation pathways.

2.3 INSIGHTS FROM EXOPLANET CHARACTERISTICS

Each of the described formation channels differ by the timescale at which they operate and by the different, sometimes overlapping, distribution of gas giants located inside the ice line that they produce. To confront these predictions, astronomy relies on observations. Evidence to discriminate between formation/evolution channels can be broadly divided in three categories: properties of hot Jupiters (occurrence, eccentricity, obliquity, inflated radius, semi-major axis or atmosphere) and other types of planets (occurrence and eccentricity of longer periods warm/cool Jupiters, occurrence of smaller rocky planets, occurrence of short periods Neptune like planets, a.k.a hot Neptunes), stellar properties of hot Jupiter hosts (age, metallicity, stellar type) and architecture of the systems hot Jupiters belong to (presence/lack of nearby/distant companions).

In their review paper, Dawson & Johnson (2018) identified, among all these properties, two for which there are 'additional or complementary observations needed'. These are hot Jupiter's obliquities (section 2.3.1) and host star ages (section 2.4). The latter represents the backbone of this work, and as we will see, ages and in particular young ages are key to understand the pro-
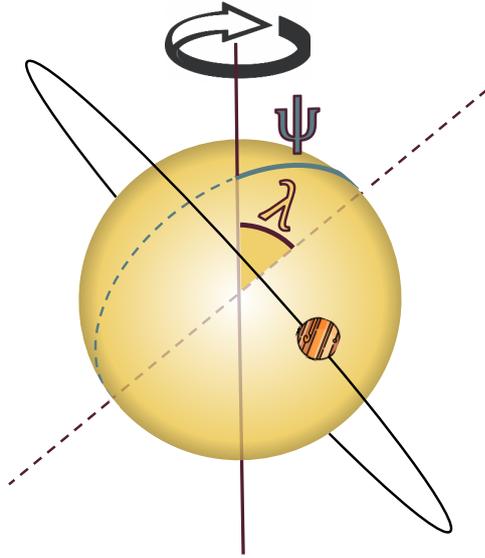
Figure 2.3: Illustration of a rotating star with an orbiting planet. The spin–orbit angle (or true obliquity Ψ), angle between the star spin axis and the planet orbital axis, is shown in blue. We also show the observable, sky-projected obliquity (λ), in yellow. Here, the line of sight is perpendicular to the plane of the figure.

cesses shaping the observed gas giant population. We will now focus on the two characteristics relevant to chapters 4 and 5, i.e., exoplanets obliquity and warm Jupiters occurrence rate and eccentricity.

2.3.1   Obliquity

Obliquity, or spin-axis angle, is the angle between the star's spin axis and the planet's orbital angular momentum vector, i.e., the rotation axis of the orbit. In Figure 2.3, we illustrate the obliquity ψ and the sky-projected obliquity λ. The relationship between ψ and λ is a function of the stellar inclination relative to the line of sight $i_\star$ and the rotation axis of the orbit relative to the line of sight i_{orb}. It is given by the following relation:

$$\cos \psi = \cos i_\star \cos i_{orb} + \sin i_\star \sin i_{orb} \sin \lambda \qquad (2.1)$$

Measuring obliquity (extensively described in chapter 4) involves spectroscopic (linked to RVs) observations during the planet's transit. A detectable transit means $i_{orb} \sim 90°$, therefore the previous equation can be reduced to:

$$\cos \psi \approx \sin i_\star \sin \lambda \qquad (2.2)$$

$i_\star$ can be estimated from the projected rotational velocity $v \sin i_\star$, the stellar rotation period P_{rot} and the stellar radius $R_\star$, as $i_\star = \sin^{-1}\left(\frac{v \sin i_\star}{2\pi R_\star / P_\star}\right)$. However, imprecise measurements of $v \sin i_\star$[2] (especially for slowly rotating stars like mature Solar-type analogues) or P_{rot} as well as asymmetries in the error bars of $R_\star$ can yield imprecise $i_\star$ estimations and even wrong $i_\star$ posterior distributions when improper assumptions are made. The Bayesian framework from Masuda & Winn (2020) addresses these issues and yields robust estimates of $i_\star$, especially when dealing with young and fast rotating stars. For the rest of the discussion, we will only refer to the projected obliquity, λ, and define aligned planets as having $\lambda \lesssim 30°$ and misaligned planets with $\lambda \gtrsim 30°$. Contrary to the minor obliquities exhibited by the Solar system planets, the exoplanet population shows a much more spread distribution of obliquities including dramatic misalignments.

The simple picture is the following: we do not expect in-situ formation or gas disc formation to produce misaligned orbits, as no strong torques are exerted on the planets. On the other hand, high eccentricity migration is predicted to yield a whole range of obliquities. For instance, in Zeipel–Lidov–Kozai cycles, obliquity and eccentricity are traded between the migrating planet and the pertuber (Chatterjee et al., 2008; Marzari & Nelson, 2009). In

---

[2] $v \sin i_\star$ translates in a Doppler-broadening of spectral features. It is therefore measured from the width of observed spectral absorption lines. When small, i.e., a few $\mathrm{km\,s^{-1}}$, $v \sin i_\star$ becomes hard to precisely measure as other phenomena such as macro-turbulence and instrumental effects yield broadening of the spectral lines comparable to that resulting from $v \sin i_\star$.
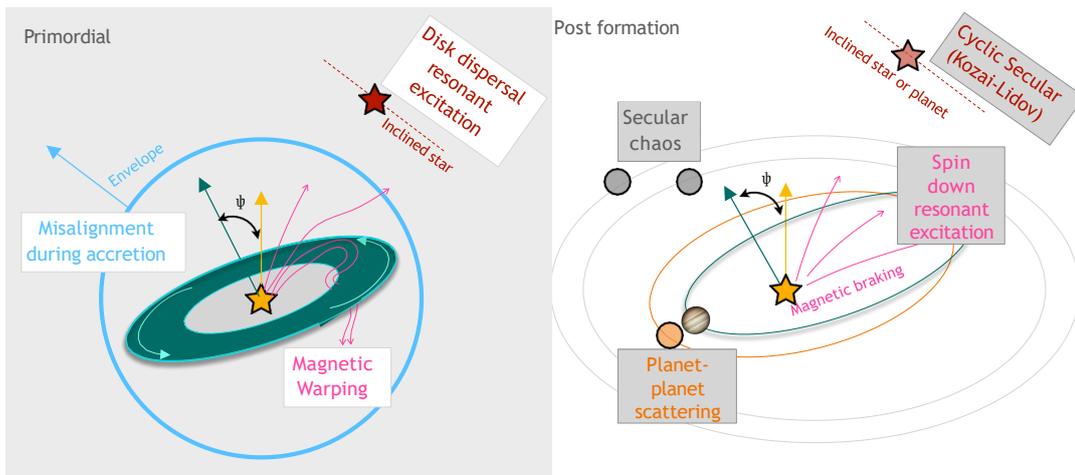
Figure 2.4: Illustration from Albrecht et al. (2022) depicting the processes that can lead to spin–orbit misalignment. **Left** Primordial, or pre-planet formation mechanisms. **Right** Post-formation processes.

this context, we can naively expect planets originating from high-eccentricity migration to have misaligned orbits and planets formed in-situ or that experienced gas disc migration to have a small spin–orbit angle. In reality, obliquity is an imperfect tracer of a planet's origin for two reasons:

- Primordially misaligned discs can muddy the waters (Batygin, 2012). Different mechanisms such as internal gravity waves (Rogers et al., 2012), the presence of a stellar companion, or magnetic star–disk torques (Spalding & Batygin, 2015), can result in a primordial misalignment of the disk, in which forming planets are embedded. It remains unclear how commonly protoplanetary discs are (mis)aligned. The left of Figure 2.4 shows the possible primordial processes yielding misalignments.

- The predicted higher obliquity of planets that experienced high-eccentricity migration can be erased by tides. Overtime, the planet exerts tidal

forces on the star, which will eventually realign the star, removing any trace of earlier misalignment. Such post-formation mechanisms are illustrated on the right of Figure 2.4.

One piece of evidence for tidal realignment is that the hot Jupiters orbiting hotter stars are found to be more systematically misaligned compared to systems orbiting cooler stars (Schlaufman, 2010; Winn et al., 2010). The common explanation argues that cooler stars have larger convective envelopes, more sensitive to tidal forces, and will therefore enable more efficient tidal realignment of planets (Albrecht et al., 2012). This dichotomy might not be observed among eccentric hot Jupiters (Rice et al., 2022), although four eccentric hot Jupiters orbiting cooler stars, very recently discovered by Mancini et al. (2022) have been found to be aligned, potentially challenging the previous result given the small sample size. This last example outlines the fact that the current picture remains unclear. As detailed in Table 3 from Albrecht et al. (2022), different scenarios proposed to explain the observed obliquity distribution would result in various observable trends with respect to planetary system characteristics (age, mass, orbital separation, multiplicity...). A larger sample of planetary obliquities spanning a wide variety of such characteristics is therefore key in order to use obliquity as a tracer of the origin of gas-giants.

It is crucial to better sample the evolution of planet obliquities to get a stronger grasp on timescales. As it clearly appears in Figure 2.5, young planet obliquities are missing and focusing on these is a way to lift the ambiguity arising from tidal realignment, by catching systems early after the disc dissipation when their primordial spin–orbit angle are still observable. Obliquity determination of planets younger than 200 Myr (and with age uncertainties
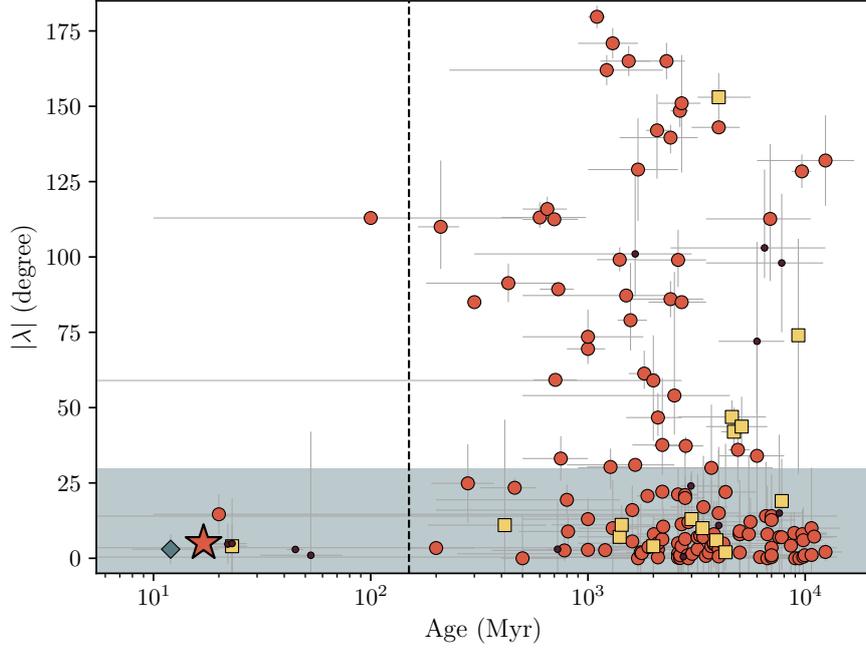
Figure 2.5: Distribution of planets with measured sky-projected obliquity and age estimate. Hot gas giants are shown as orange circles, warm giants as yellow squares, cold gas giants as blue diamonds and planets smaller than 6 $R_\oplus$ as brown dots. The grey region comprises aligned systems (i.e., $\lambda \lesssim 30°$). The vertical dashed line separates young (i.e., < 150 Myr) from more mature planets. Young planets include, from youngest to oldest: β Pic b, HIP 67522 b, WASP-25 b, AU Mic b, V 1298 Tau b & c, DS Tuc Ab, TOI-942 b and WASP-33 b. We note that only one cold gas giant has a spin–orbit alignment, β Pic b. This is because cold gas giants are extremely unlikely to transit and therefore to yield a λ measurement. β Pic b's obliquity was obtained from direct imaging. The red star shows HIP 67522 b, focus of chapter 4. The dataset originates from Table 4 and 5 in Albrecht et al. (2022).

less than their estimated age) only started in 2019 (see table A2 in Albrecht et al. 2022). Measuring obliquities requires a good signal to noise ratio in relatively short (i.e., few minutes to few tens of minutes) consecutive observations, making brighter stars the best suited targets. The TESS[3] mission has discovered more than 200 planets and holds close to 6000 candidates orbiting the brightest and nearest stars. This has resulted in a small but promising sample of extremely young planets (i.e., $\lesssim 150$ Myr old, shown on the left of the vertical dashed line in Figure 2.5) with measured obliquities; AU Mic b (Palle et al., 2020; Addison et al., 2021; Martioli et al., 2020; Hirano et al., 2020), DS Tuc Ab (Newton et al., 2019; Zhou et al., 2020; Montet et al., 2020), TOI-942 b (Wirth et al., 2021; Zhou et al., 2021) and V 1298 Tau b (Johnson et al., 2021; Gaidos et al., 2021) & c (David et al., 2019; Feinstein et al., 2021). To this list, we can add WASP-25 b (Brown et al., 2012) and WASP-33 b (Johnson et al., 2015). Although we should also mention β Pic b, the youngest planet to have an obliquity measurement obtained from direct imaging (Kraus et al., 2020), it is somewhat in a different category with its $\sim 11\,M_J$ mass and 9 AU semi-major axis. The work described in chapter 4 adds to this list the obliquity measurement of the youngest planet known to transit, the hot Jupiter-sized HIP 67522 b (Rizzuto et al., 2020; Heitzmann et al., 2021b). This measurement is described in section 2.4.3.

2.3.2 OCCURRENCE AND ECCENTRICITY OF WARM JUPITERS

Warm Jupiters are loosely defined as Jupiter-like planets with orbital periods in the ~ 10 to ~ 250 days range. Explanations for their presence inside the

---

[3] The Transiting Exoplanet Survey Satellite is a photometric space telescope from NASA (https://tess.mit.edu/). Its mission is to survey the entire sky looking for transiting planets around the brightest ($V_{mag} \lesssim 13$) stars.
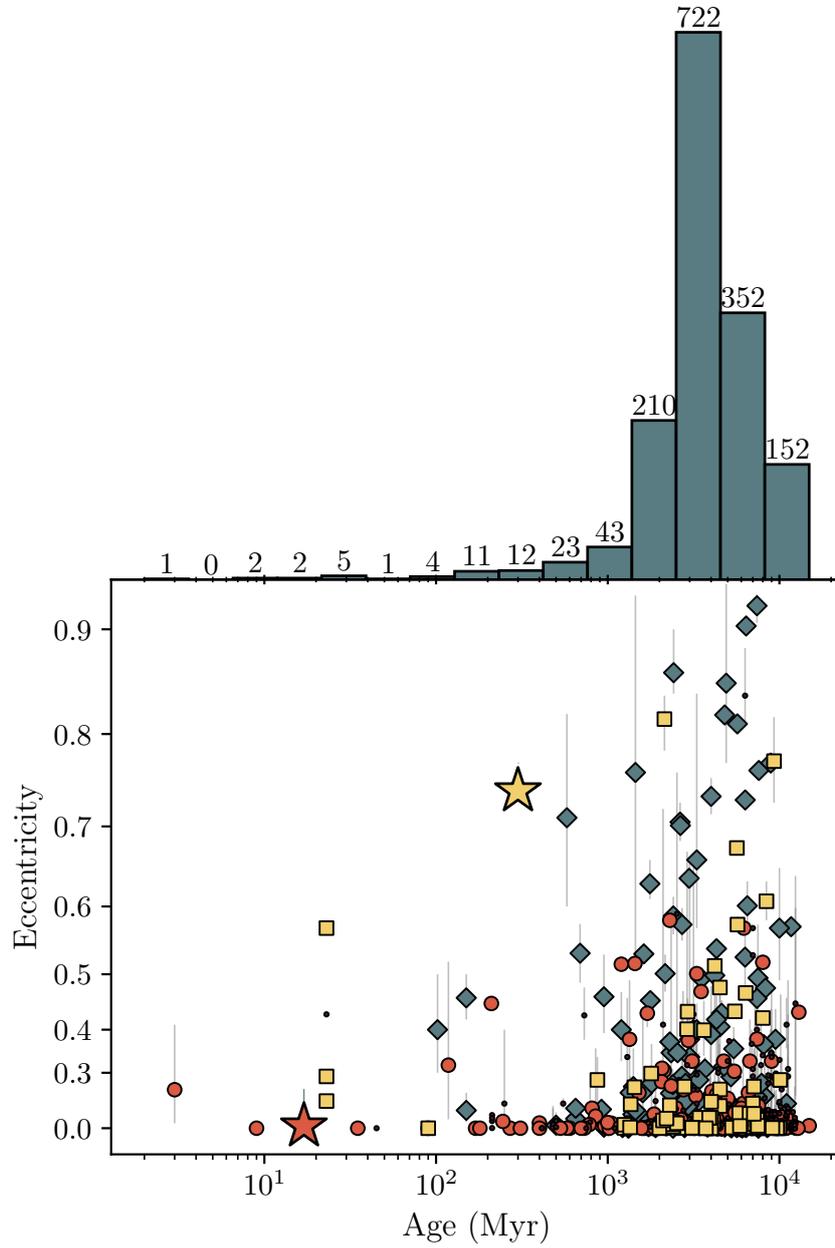
Figure 2.6: **Top** 1-D age distribution of the planets plotted in the bottom plot. **Bottom** Age versus eccentricity for all planets with $P_{orb} < 10^4$ days and $\sigma_{age} \lesssim$ age. Hot gas giants are shown as orange circles, warm giants as yellow squares, cold gas giants as blue diamonds and planets smaller than $6\ R_\oplus$ as brown dots. The orange and yellow stars respectively indicate HIP 67522 b and TOI-4562 b, coloured according to their orbital periods.

ice line invokes the same mechanisms previously described for their hotter cousins, the hot Jupiters. The warm Jupiter population exhibits two notable features. The first is the Period Valley (e.g. Jones et al., 2003; Wittenmyer et al., 2010), showing a paucity of these planets compared to hot Jupiters and wide orbit gas giants. But despite the presence of this Period Valley, the number of warm Jupiters discovered is significantly larger than expected. Secondly, one third of warm Jupiters are found on moderately ($e > 0.3$) eccentric orbits, and a small fraction shows large to extreme eccentricity ($e > 0.6$), as shown on Figure 2.6. Reconciling both these features with a single pathway for their origin is extremely challenging.

If high-eccentricity migration is the generator for hot Jupiters, then eccentric warm Jupiters are future hot Jupiters in transition, awaiting tidal circularization. Simulations using a wide range of initial configurations (Petrovich, 2015; Anderson et al., 2016; Hamers et al., 2017) predict a number of warm Jupiters much smaller than what is observed, even when invoking eccentricity oscillations (Petrovich & Tremaine, 2016). Conversely, highly eccentric planets possibly on a path to become hot Jupiters, i.e., having the correct combination of e and a, (see the purple region in Figure 2.7), are missing (Socrates et al., 2012) with only a few found to date: HD 80606 b (Naef et al., 2001), HD 17156 b (Fischer et al., 2007) or TOI-3362 b (Dong et al., 2021). We note, that a couple of other planets could be added to this list if the previously mentioned eccentricity oscillations generated by an external pertuber are at play. The low-eccentricity warm Jupiter population is also incompatible with the high eccentricity scenario. Most of these planets lie too far from their host star to experience circularisation in a timescale shorter than their lifetime (planets located outside the purple region in Figure 2.7).
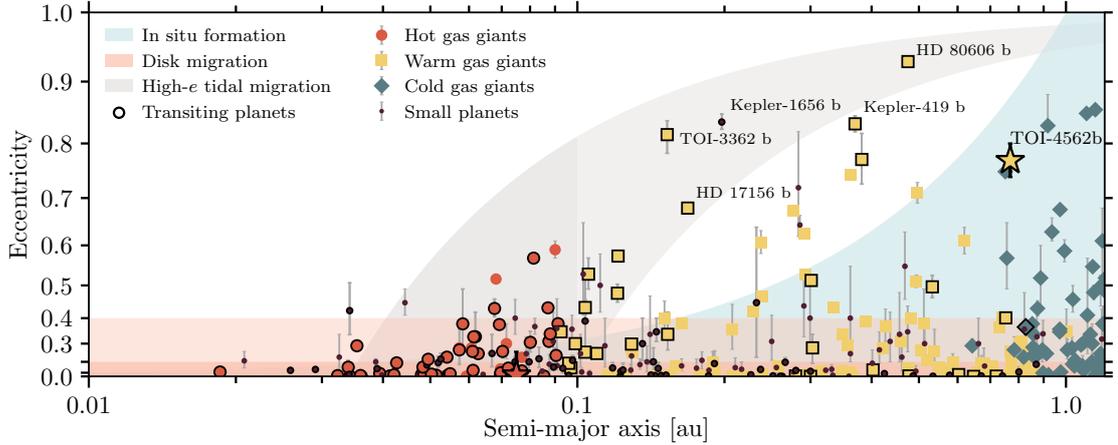
Figure 2.7: Eccentricity versus semi-major axis for all confirmed planets (obtained from the NASA exoplanet archive 13 Feb. 2022) with $M_{\mathrm{p}} < 13\,M_{\mathrm{J}}$. The y-axis is scaled to e^2 to emphasize non-circular planets. Shaded areas highlight different formation scenarios. Planets in the grey region are on the path of high-eccentricity migration, with a final semi major axis between 0.034 and 0.1 au. The upper and lower bounds of this region are set by the Roche limit and the circularization timescale respectively. Disc migration, expected to only marginally excite orbital eccentricity is shown as the red shaded region (with its most extreme cases up to $e \sim 0.4$ shown in lighter red). Finally, in-situ formation, with eccentricity excited by, e.g., planet–planet scattering is shown in blue. Transiting planets are represented with thicker borders. Hot gas giants are shown as orange circles, warm giants as yellow squares, cold gas giants as blue diamonds and planets smaller than $6\,R_{\oplus}$ as brown dots. Only planets with uncertainties on e smaller than 50% of the measured e are shown. We note that a similar figure is also available in the work shown in chapter 5 and was first published in Dong et al. (2021).

Finally, many warm Jupiters are found to have either lower mass smaller companions (Huang et al., 2016) or giant external companions (Antonini et al., 2016), mostly inconsistent with the violent dynamical interactions stemming from high-eccentricity migration.

In-situ formation suffers the same issues as for hot Jupiters, i.e., material is lacking at these orbital distances and accretion is not efficient enough. If solved, in-situ formation followed by planet scattering induced by an external companion (blue region in Figure 2.7) could explain a large fraction of warm Jupiters (Anderson et al., 2020).

Gas disc migration could be at the origin of the circular warm Jupiters population (red region in Figure 2.7), in a scenario where the planets are stopped in their migration process (Coleman & Nelson, 2016). Still, both in-situ and gas disc migration are unable to generate eccentricities larger than 0.4 (in the most optimistic cases, Debras et al. 2021) without obvious external companions or complex multi planet configurations.

Once again, the underlying picture is unclear, but the community seems to agree that no mechanisms alone can reproduce the observed population of warm Jupiters. A combination of high-eccentricity migration, gas disc migration and in-situ formation followed by scattering might be at play. Similarly as for obliquities, young planets could prove very informative to add constraints on these evolution pathways. Again, this population is lacking as clearly shown in Figure 2.6. Another avenue is to intensively characterise planets that hosts outer companions. Knowledge of the mass and eccentricity of a warm Jupiter's companion, as well as the relative inclination on the two (or more) planets can be of tremendous help in dynamical simulations aiming to discriminate between formation scenarios. This has been done for four systems so far:

- In Masuda (2017), the authors argue that both Kepler-693 b and Kepler-448 b could be consistent with high eccentricity migration, but propose a variation where a planet would be born inside the ice-line and then experience high eccentricity migration.

- Dawson et al. (2014) mapped the 3D architecture of the Kepler-419 system and the configuration of the two planets does not appear to favour a high-eccentricity scenario to explain the existence of its warm Jupiter Kepler-419 b .

- In May 2022, Angelo et al. (2022) identified a 1919 day orbit outer companion to Kepler-1656 b, an eccentric warm sub-Jovian. Their simulations describe the most plausible explanation for its origin as in-situ formation followed by gentle eccentricity excitation. Until now, due to its position in the a–e parameter space (see Figure 2.7), Kepler-1656 b was believe to be the result of a more violent, high eccentricity migration.

In chapter 5, we present a newly discovered highly eccentric young planet, TOI-4562 b (identified by the yellow star in Figures 2.6 and 2.7), the second longest period planet from TESS, at the edge of the warm Jupiter population, that also show an ambiguous formation history. As we will see, there is convincing indication of the presence of an external companion, of which further characterisation would provide important clues similar to the four systems discussed above. As we pointed out in this paragraph and in section 2.3.1, young planets can deliver crucial insights into the formation and evolution of gas giants. In the coming section, we will explain why so few young exoworlds have been discovered to date.

2.4 Young planets: A promising but challenging avenue

An efficient way to discriminate between origin channels is to catch these processes at every step of the way to see what mechanisms are truly at play. Young stars (and consequently young planets) provide two decisive advantages. First, the final configuration of these systems are reached after a few hundreds of millions of years, so young targets allow us to catch the transformations that these planet undergo as they occur. Secondly, further evolution

of the systems, happening over longer timescale, can erase traces of key features of a particular formation/evolution channel (e.g. tidal realignment and circularization, see section 2.2.3). In Figure 2.6, we show the age–eccentricity distribution for planets with a correct age estimate (i.e., $\sigma_{\text{age}} < $ age). We also show the age 1-D distribution, where the lack of young known planets clearly stands out.

Because the overwhelming majority of planets discovered to date orbit mature stars, searching for young planets to sample the primordial population is of utmost importance. In this section, we discuss the strong barriers preventing their reliable systematic characterisation.

2.4.1 THE RARITY OF YOUNG DETECTABLE PLANETS

Only 17 planetary systems younger than 250 Myr old (containing 26 planets) have been found via the transit method. The Kepler and K2 surveys yielded only 5 of these systems, because of their lack of coverage of young stellar associations and open clusters, which as we will see is the most reliable way to obtain precise ages estimates. By surveying the entire sky since July 2018, TESS drastically changed the state of the field, discovering 11 of such young systems and is expected to continue to deliver young planets in the coming years.

The first difficulty is that detectable, young planets are rare. Based on timescales for the formation and evolution processes, we loosely define young stars as being at most a few hundreds of Myr old. This is two orders of magnitude smaller than a Solar type star's typical lifetime. Given an evenly distributed population of stellar ages in the Galaxy, young stars are therefore intrinsically rare (given a Sun-like star lifetime of ~10 Gyr, < 500 Myr old stars represent 5% of the population). We have seen that hot/warm Jupiters are present

around only up to a few percent of stars. Assuming a transit probability of 10% ($p_{transit} \approx \frac{R_\odot}{a}$, assuming a $R_p \ll R_\star$ and a circular orbit) for a planet orbiting a Sun-like star at 0.05 AU leaves us with an rough estimate of one star in $\sim 10^4$ hosting a young, detectable, short-orbit gas giant.

2.4.2 THE DIFFICULTY OF DETERMINING STELLAR AGES

Large protostellar molecular clouds can generate gravitationally bound groups of stars. Over time, these tight stellar clusters will get disrupted as a result of the intrinsic velocities of some members able to escape and interactions with external sources (i.e., interstellar clouds or other associations). Associations and open clusters are remnants of these co-moving groups of stars that haven't yet totally been disrupted and therefore are young ($\lesssim 1$ Gyr and many younger than a few hundreds of Myr). Our knowledge of stellar evolution combined with population studies in these moving groups allows us to know their age with good precision. Although only a few planets have been found to belong to known moving groups, this is expected to change thanks to the Gaia mission (Gaia Collaboration et al., 2016). By delivering parameters (masses, radii, effective temperatures, distances, astrometric solutions, proper motions, etc.) and spectra for hundreds of millions of stars (Gaia Collaboration et al., 2021), Gaia has the potential to reveal the full extent of some known moving groups as well as identifying new ones (Ujjwal et al., 2020). Very recently, Gaia data allowed the discovery a new young association, the 40 Myr old Cep-Her complex, located in the Kepler field. Four previously discovered Kepler planets were found to belong to that cluster (Bouma et al., 2022a,b), adding to the pool of very young known planets.

Unfortunately, not all young stars belong to open clusters and associations.

Field stars are not known to be co-moving with a particular group of stars and their age is a priori unknown. Although some features such as fast rotation period, strong photometric (and spectroscopic, see section 2.4.3) variations or excess UV/X-ray flux are telltale signs of youth, precise age estimate for these wandering stars is notoriously challenging. A common method is to fit the stellar parameters to isochrones[4]. This method is best suited for main sequence stars and its precision deteriorates for cooler stars that evolve more slowly. Other age indicators exist, most of them calibrated using young open clusters and associations. One is the evolution of P_{rot}, termed gyrochronology, using the fact that the remnant angular momentum carried by young stars from the collapse of the primordial cloud will diminish over time, in part due to magnetic coupling between the star and the surrounding disc. Relations between P_{rot}, colour and age (also calibrated using moving groups) can be used to roughly estimate the ages of young stars (e.g. Barnes, 2007; Mamajek & Hillenbrand, 2008). Lithium is another common age indicator. For stars of mass < 1.5 $M_\odot$, efficient transport of elements in the outer convective layer will carry lithium deep into the star interior where it will experience proton capture, transform into beryllium and later decaying into helium. Lithium in low-mass stars ends up being depleted and lithium abundance can be used as a proxy for age. For Solar analogues, this correlation between Lithium depletion and age is found to be the strongest up to ages of ~2 Gyr (Carlos et al., 2019). Although much more imprecise than cluster member ages, combining these indicators can give a good idea of the age of a single star. In chapter 5, we used all the above described indicators to estimate the age of TOI-4562 b.

---

[4]Isochrones, from *iso* same and *chronos* age, are functions traced by stars of similar ages across all stellar types on a T_{eff} vs luminosity diagram, or Hertzsprung–Russell diagram.

In our discussion, we have mainly focused on transiting planets, and we could imagine that radial velocity surveys, not suffering from the low transit probability constraint (although systems close to face on orbit are more easily detectable) would be more fruitful. Unfortunately, the radial velocity method has yielded only a handful of short-orbit very young planets (V 830 Tau b; Donati et al. 2016, 2017, CI Tau b; Flagg et al. 2019, TAP 26 b; Yu et al. 2017, HD 70573; Setiawan et al. 2007 b and β Pic c; Lagrange et al. 2019), of which two have been contested (Damasso et al., 2020; Donati et al., 2020). This is because radial velocity surveys are less systematic than transit surveys but more importantly because of the major obstacle imposed by the stellar activity exhibited by young stars. We will elaborate on this aspect in the coming section 2.4.3. Evaluating how this stellar activity is hampering our chances to detect young gas giants in short-orbit constitute the first part of my thesis, focus of chapter 3.

2.4.3 The challenge of stellar activity

Young stars exhibit very different behaviours than their more mature counterparts. The initial contraction they experience in the first millions of years of their existence results in a boosted rotational rate. Typically, a young Sun-like star of a few tens of millions of years will exhibit a P_{rot} on the order of days (versus 25 days for the Sun's equatorial P_{rot}). Interior plasma motions such as convection or meridional circulation combined with rotation generate strong surface magnetic fields through dynamo processes. In turn, these magnetic fields yield a wealth of phenomena that manifests themselves in various forms, each on different scales, duration and influence on observations. Below we give an overview of the various manifestations of stellar activity and
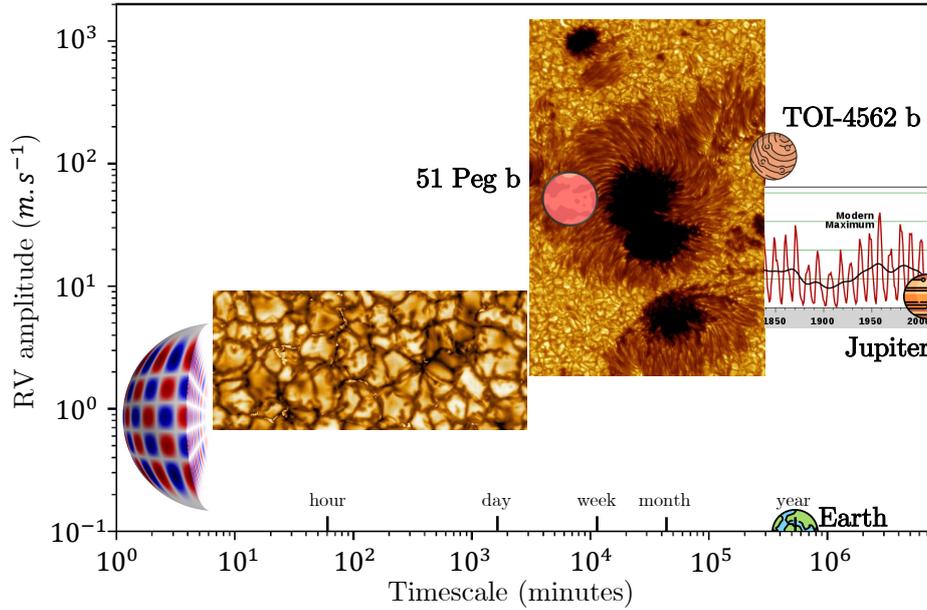
Figure 2.8: Timescale versus amplitude of the RV signature for the different stellar activity phenomena described in section 2.4.3. These are from left to right: Stellar oscillations, granulation, spots, plages and faculae and magnetic cycles. The size of each image delimits the zone of influence of the associated stellar variability manifestation. For reference, we added the planetary RV signatures of (from left to right): 51 Pegasi b, TOI-4562 b, Earth and Jupiter. Granulation - author: Daniel K. Inouye Solar Telescope (NSO/AURA/NSF), obtained from nsf.edu on June, 5th 2022 / Sunspots - author: Royal Swedish Academy of Sciences, obtained from https://ttt.astro.su.se/isf/NatureNov2002/ on June, 5th 2022 / Magnetic cycles - author: 'DeWikiMan', obtained from Wikipedia on June, 5th 2022. Icons credits: https://www.flaticon.com, made by Freepik, Victoruler and Icongeek26.

their influence on RVs (see Meunier 2021 for a review). These phenomena can be grouped by their different timescales and the amplitude of their signature in RV relative to typical short-orbit gas giants, i.e., with P_{orb} from a few days to a few months and RV signatures between a few 10^1 to a few 10^3 m s^{-1}, as illustrated in Figure 2.8.

STELLAR OSCILLATIONS   The convective envelope of stars generate acoustic oscillations modes (called p modes) that can make the stellar surface oscil-

late and generate a RV signature on the scale of minutes (Michel et al., 2008; O'Toole et al., 2008). Much shorter than exoplanet orbital periods, this effect can be mitigated by averaging data over a few tens of minutes (Dumusque et al., 2011b). The amplitude of this phenomenon in RV goes from tens of $cm\,s^{-1}$ to a few $m\,s^{-1}$, three orders of magnitude smaller than a typical hot Jupiter signature around a Sun-like star and up to four orders of magnitude smaller than the RV amplitude due to spots and plages (see below).

GRANULATION   Stellar convection is organised in vertical cells. Plasma at the centre of these cells rise towards the surface (i.e., in direction the observer), inducing a blueshift of the emitted light. Conversely, plasma on the edge of these cells is sinking down and escaping light will be redshifted. The upward motion has a stronger contribution than the sinking flow of matter, resulting in a net blueshift called convective blueshift. Integrated over the stellar disc, the effect of granulation on RVs, long known (Dravins, 1987), is on the order of a few $m\,s^{-1}$ and operates on timescales ranging from minutes to a couple of days. In the case of young active stars, this effect is negligible, dwarfed by active regions that we will describe next. We note that granulation and stellar oscillations are a current major obstacle for the search of Earth analogues (Cegla, 2019; Palumbo et al., 2022), as these phenomenon are not confined to young stars and can be an order of magnitude stronger than the RV signature induced by Earth-like planets.

SPOTS, PLAGES AND FACULAE   Easily identifiable at the surface of the Sun, large dark and bright surface features (i.e. dark spots and bright plages or faculae) are active regions where the strong local magnetic field acts to inhibit convection (we refer the reader to Berdyugina 2005 for an extensive discus-

sion on stellar spots in cool stars). Cumulatively, they can cover up to a few percents of the stellar surface. Due to the intrinsic rotation of the star, light emitted from the hemisphere rotating towards the observer will be blushifted (as a function of longitude), and light issued from the receding hemisphere redshifted. Active regions, by locally enhancing or decreasing brightness, will break the symmetry with respect to the rotational axis of the stellar disc, yielding a net blue or red shift that will modify the shape of observed spectral lines and therefore induce non-negligible RV variations able to mimic planetary signatures. This is illustrated in Figure 2.9. These regions also act to inhibit the convective blueshift. Put together, the effect of active regions is of a few m s^{-1} for quiet, mature stars (Dumusque et al., 2011a), e.g. 1 to 2 m s^{-1} for the Sun at its activity maximum. In the case of extremely active and fast rotating young stars, their effect in RVs is commonly in the few hundreds of m s^{-1} range and up to the km s^{-1} level for the most extreme cases (e.g. in Donati et al., 2016, for the young V 830 Tau), deeply burying any RV signature induced by an orbiting planet. For reference, a 1 M_J planet on a 5 to 10 day orbit around a 1 $M_\odot$ star will yield a RV signature of ∼100 m s^{-1}. Dark and bright features evolve on the scale of days to months, and are rotationally modulated as they cross the stellar disc as the star rotates, therefore recurring at a period matching P_{rot} and its harmonics ($P_{rot}/2$, $P_{rot}/3$, etc.). Brightness feature's frequency of modulation can slightly vary if they are located at high latitudes while the star is exhibiting differential rotation. The combination of their evolutionary timescale and periodicity is difficult to deal with and acts on the same timescales as orbiting planet candidates. At first order for young stars, active regions are the main contribution to stellar intrinsic variability.

MAGNETIC CYCLES   The longest known process associated with stellar magnetic fields are magnetic cycles. For the Sun, this 11 year cycle is associated with variation in the number and location of spots, as well as the frequency and intensity of flares, coronal loops and coronal mass ejections. Magnetic cycles have been observed to occur in many other stars (e.g. κ^1 Ceti; Saar & Baliunas 1992, ε Eridani; Jeffers et al. 2014b, τ Boötis; Mengel et al. 2016; Jeffers et al. 2018). These magnetic cycles occur on the scale of years and their associated changes in stellar activity result in RV variations at the level of tens of m s^{-1} (Gomes da Silva et al., 2012). Given their long timescale and relatively small RV effect in the context of young active stars, magnetic cycles are not of major concern here.

Stellar activity is a major obstacle for exoplanet detection, and presents currently one of the major challenge in the field. Intensive research is conducted on the question of its mitigation which is the subject of the next section and chapter 3.

2.4.4 MITIGATING STELLAR ACTIVITY

Different strategies are employed to attempt bypassing stellar activity and recover potentially hidden exoplanets. Various methods are being developed, and although they seem to be efficient in some cases, Zhao et al. (2022) points out that they do not all yield consistent results.

As briefly touched upon, stellar variability impacts spectroscopic observations by altering the shape of spectral lines. As illustrated in Figure 2.9, brightness inhomogeneities at the surface of the star (i.e., bright and dark spots) will create bumps and dips in the spectral lines relative to their longitude, breaking the symmetry of the lines. As a result, the centre of the mean line profile
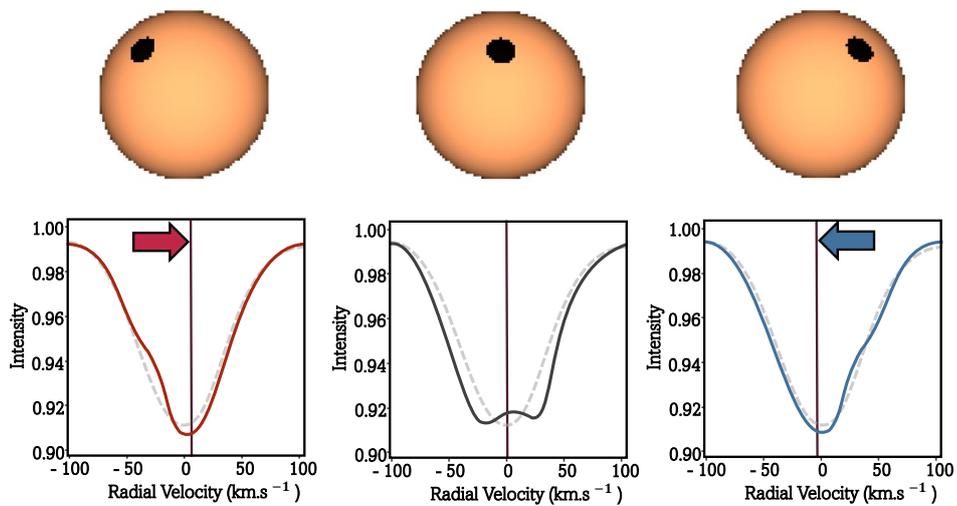
Figure 2.9: Schematics of the influence of a large dark spot on the surface of a rotation star on the recovery of the stellar RV. The bottom plots show mean line profiles (in red, black and blue), i.e., weighted sum of a set of line in the stellar spectra (usually performed via a cross correlation function or least square deconvolution). Deformations of the line profiles are induced by a lack of light coming from the hemisphere with the spot. The red and blue arrow represent the offset due to the deformation when measuring the center of mass of the line profile (vertical brown lines), tracing the RV value. In dashed grey are line profiles resulting from an unspotted stellar surface.

(CCF or LSD profile, see Appendix A for details), fitted with a symmetrical profile to recover the RV, will be offset.

It is worth noting that the majority of the research conducted on stellar activity filtering is not in the context of very active young stars, but rather concerns small terrestrial planets (e.g., Earth-like) orbiting quieter, mature stars like the Sun. In these cases, both the stellar variability and the RV planetary signatures are 3 to 4 orders of magnitudes smaller than for young stars.

To finish this introductory chapter, let us review the principal diagnostics and treatments applied to stellar intrinsic variability. We will highlight the most relevant to this work.

OBSERVATIONS   Observations can be averaged out over timescales longer than the phenomenon we want to mitigate, e.g. acquiring spectra over a few dozens of minutes to mitigate stellar oscillations. Another strategy is to observe planets in quadrature, when the RV signal is the strongest (only doable when the period is know, e.g. for planets known to also transit).

ACTIVITY INDICATORS   Stellar active regions are associated with locally stronger magnetic fields. This enhanced magnetic activity increases the temperature in the chromosphere, exciting atoms and triggering emissions. These emissions are identifiable in spectra in the core of lines such as the calcium H & K lines found at 3969 and 3934 Å or the He I line at 6563 Å (Boisse et al., 2009). Although activity indicators do correlate with RVs, they are not reliable as a sole mean to remove the stellar activity contribution from the data. We note that other diagnostics inferred from the shape of spectral lines do exist, such as the bisector or the full width at half maximum, that can be used to diagnose stellar variability.

AT THE SPECTRUM LEVEL   We saw that activity induced distortions in RVs express themselves at the spectra level. We can attempt to account for such distortions directly at the line level (Dumusque, 2018; Simola et al., 2019) or isolate groups of lines that are the least influenced by activity to be used for the CCF/LSD computation (Meunier et al., 2017; Bellotti et al., 2022). This is a rather recent approach, as the signal contained in single lines is insufficient to derive precise RVs and can only be useful when using new generation spectrographs. Current research on individual lines also aim to understand the influence of stellar activity on lines formations parameters, e.g. line depth (Cretignier et al., 2020) or line formation temperature (Al Moulla et al., 2022).

AT THE RV TIME SERIES LEVEL   For all the methods mentioned in this paragraph, the expected RV contribution from the stellar activity is modelled and then removed (subtracted) from the observed reduced RVs, leaving behind residuals potentially bearing exoplanet signature(s). Because of the periodic modulation of active regions due to the stellar rotation, the first promising attempt to model the RV contribution of stellar intrinsic variability used sinusoids close to P_{rot} and its harmonics (Boisse et al., 2011). One downside is that this approach does not model the evolution of surface features. A second technique is Doppler Imaging (Vogt & Penrod, 1983; Brown et al., 1991; Donati & Brown, 1997; Donati et al., 2014), thereafter DI, where the star is modelled with active regions on its surface. A mean line profile is derived from the model and fitted to the observed line profiles. The RV time-series obtained from the modelled line profiles is then subtracted from observations and the residuals can be searched for planets. The last method is the now

routinely used Gaussian Processes (GP), seeking to model the stellar activity contribution in observed RV time series as a collection of *correlated* Gaussian distributed variables. DI and GP are the two stellar activity mitigation strategies used in the work presented in chapter 3 and are described in further details in sections 2.4.4.1 and 2.4.4.2.

CHROMATICITY   Simultaneous observations in different wavelengths can also help mitigate stellar activity. The contrast between surface spots and the photosphere is wavelength dependent and less important towards longer wavelengths (Desort et al., 2007; Reiners et al., 2010). Leveraging this effect, the chromatic index (Zechmeister et al., 2018) measures the RV variations as a function of wavelength to identify RV signatures originating from stellar activity. Simultaneous visible and infrared observations can also be used to more efficiently model the stellar activity. Cale et al. (2021) employed an extended GP model with 2 components to account for simultaneous visible and infrared data, each having a different RV amplitude. The authors point out that varying amplitudes is a starting point, as they also witness variations in the feature content of the RV signatures in the visible vs infrared (due to the chromatic dependence of effects like the convective blueshift or limb darkening). More sophisticated models have yet to be explored, which could also include activity indicators within already existing framework such as multi-variate Gaussian Processes (e.g., Barragán et al. 2022).

To conclude this introductory chapter, we will present Doppler Imaging and Gaussian Processes, the two strategies employed in chapter 3 to mitigate stellar activity.

As described in section 2.4.3, stellar activity induces brightness features (spots, plages, faculae) on the stellar surface and excluding the Sun, it is not possible to resolve surface stellar features, although some evolved stars such as Betelgeuse (Young et al., 2000), or ζ Andromedae (Roettenbacher et al., 2016) have been resolved thanks to interferometry. Nonetheless, these brightness features generate a RV signature that can, under certain conditions, be recovered from observed spectra.

As previously explained, any darker/brighter area on the surface will diminish/enhance the amount of blueshifted/redshifted incoming light and therefore appear in the spectrum. Thus, in every spectral line, and consequently on mean line profiles (with the simplifying assumption that all lines are affected by stellar activity in the same way) is
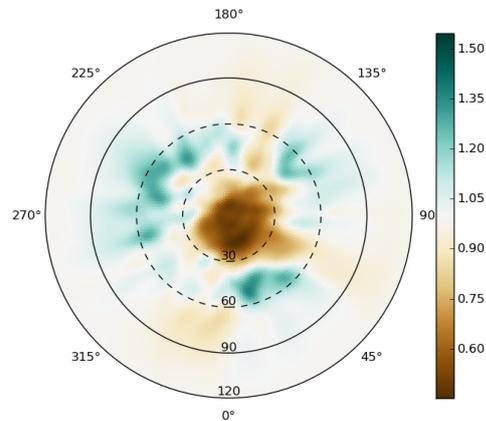
Figure 2.10: Brightness map of LO Peg obtained via Doppler Imaging (from Heitzmann et al. (2021a)). Brighter and darker regions are shown in blue and brown, respectively.

encrypted the information of the surface features location and intensity. More specifically, bumps and dips in these distorted line profiles are the direct translation of stellar spots (cool & dark patches) and plages/faculae (hot & bright patches) respectively.

The idea behind Doppler Imaging is thus to infer the location of these patches from the spectra in order to generate a brightness map of the stel-

lar surface (see fig. 2.10). Although one stellar surface brightness distribution results in a single possible profile (forward problem), various reconstructed images can be the result of a single line profile (inverse problem). To lift some part of the degeneracy for this ill-posed problem, we measure profiles at different phases of the stellar rotation. The longitude of the feature is derived by identifying the moment when the distortion (bump/dip) crosses the line centre and its latitude is inferred from the time spent by this distortion on the line profile. Still, many configurations of surface features can yield the same line profile time-series and different strategies, mentioned below and further described in Strassmeier (2009) are employed to 'choose' the most appropriate solution.

The equation to translate a surface map into a profile is:

$$R_{calc}(\lambda, \phi) = \frac{\iint I_l \left[M, \theta, \lambda + \Delta\lambda_D \left(M, \phi \right) \right] . \cos\theta . \mathrm{d}M}{\iint I_c(M, \theta) . \cos\theta . \mathrm{d}M} \qquad (2.3)$$

With R_{calc} the residual intensity of the line depth at the stellar rotation phase ϕ and wavelength λ. I_l is the line intensity (function of: the position on the stellar disk M, the angle between the normal to the surface, the line of sight θ and the wavelength of light λ corrected from $\Delta\lambda_D(M, \phi)$, the Doppler shift induced by rotation) and I_c the intensity of the continuum, the part of the spectrum where no lines are present.

The DI process then goes as follows. A brightness (or temperature, see Strassmeier 2009) map is first generated as a grid of pixel and each surface element is given a value representing their brightness (or temperature). Synthetic line profiles resulting from that distribution are then computed using equation 2.3. For each considered phase of the stellar rotation, the synthetic lines profiles are then fitted to the observed line profiles and the error of the

fit is evaluated.

From there, we iteratively generate images leading to the smallest fitting error (χ^2 minimisation) and a criteria on the most reasonable distribution of features is used to assess convergence. This can be minimum information content of the image (maximum entrop; Vogt et al. , 1987), the smoothest map possible (Piskunov, 1991) or the choice of having only two possible values for each surface elements represtening spot and surface temperature respectively (Collier Cameron et al., 1990). After this whole process, we have a snapshot of the star, showing the brightness features at its surface (shown in fig. 2.10) and a fit to the line profiles recovered from the most likely brightness image. Doppler Imaging allows to infer various stellar parameters, such as the inclination of the rotational axis with respect to the line of sight (i), the rotational velocity ($v. \sin i$), the star's rotation period (P_{rot}) and differential rotation ($d\Omega$).

Under the assumption that brightness features are the main components of stellar activity induced RV at the time and space scales of interest, fits of the line profiles provide a strong approach to filter out the stellar intrinsic varability. This first method was introduced in Donati et al. (2014), and we apply it in chapter 3. Assuming that Doppler Imaging only negligibly absorb the planetary induced shift of each profile, retrieving the RV displacement of each fit grants access to the temporal evolution of the RV signal solely due to activity. After subtracting these modelled RVs from the raw RVs, we are left with residuals RV that can be analysed to search for exoplanets.

2.4.4.2   GAUSSIAN PROCESSES (GP)

Gaussian processes (GP) in the context of stellar activity filtering first appeared in Haywood et al. (2014) and Rajpaul et al. (2015). The idea is to

consider RV variations as a combination of one (or more) planetary component(s) easily modelled by a classical Newtonian mechanics model and a stellar activity component modelled by a GP.

A GP is a Gaussian distributed collection of random variables with the property that the joint distribution of any finite subset of n variables is also a Gaussian (n-dimensional) distribution. The GP is a correlated (or red) noise model where each time sample is considered to be a Gaussian distributed random variable with their value being the RV. The degree of correlation between each time sample constitute the core of this model and is contained in a covariance matrix, itself described by a correlation function. The GP requires two main elements: The correlation function (or kernel) characterised by some hyperparameters and describing the correlation/covariance between the random variables and the measurements that will constrain the model (i.e., fix some of the variables). In the case of stellar activity, different kernel function can be used such as a damped harmonic oscillator, a square-exponential function, Matern kernels, etc. Here we chose the quasi-periodic function:

$$k(t, t') = \theta_1^2 \exp\left[-\frac{(t - t')^2}{\theta_2^2} - \frac{\sin\left(\frac{\pi(t-t')}{\theta_3}\right)}{\theta_4^2}\right] \tag{2.4}$$

where the hyperparameters can be *interpreted* as: θ_1 the GP semi-amplitude of the jitter, θ_2 the periodic component accounting for the recurrence of surface features as the star rotates, θ_3 the decay timescale of features (spot lifetime) and θ_4 a parameter encoding the high-frequency content of the model.

This approach is Bayesian, where the correlation function is the prior knowledge (left plot of fig. 2.11), and the prediction is the posterior distribution (right plot of fig. 2.11) which is constrained by the data (black dots
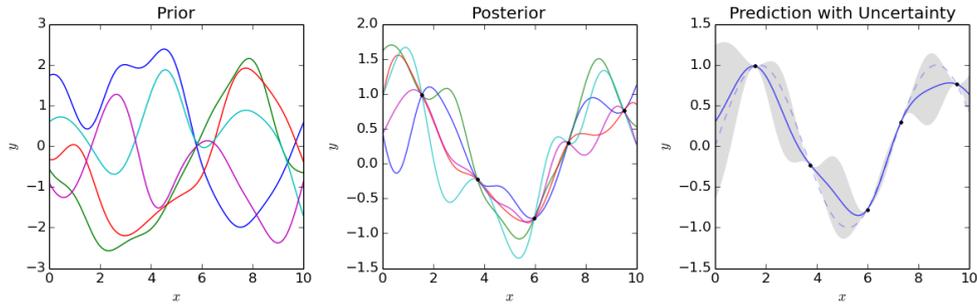
Figure 2.11: Illustration of a GP model. On the left are samples (colored lines) of the prior distribution (only guided by the covariance function). The center plot shows the posterior distribution, that is after having fixed some variables (measured data points). The right plot displays the final model with confidence intervals (grey areas).Credit: author: 'Cdipaolo96', image obtained from Wikipedia.

on both the centre and right plot of fig. 2.11). The model is not definite but rather indirectly shaped by the covariance function and trained with the data.

Then, a parameter space search is performed via an optimisation algorithm. We used PYMULTINEST (Buchner et al., 2014) in Heitzmann et al. (2021a), a python implementation of MULTINEST (Feroz et al., 2009) and EMCEE (Foreman-Mackey et al., 2013) in Heitzmann et al. (2021b) and Heitzmann et al. (2022). The role of the optimisation algorithm is to maximise a likelihood (or cost) function that depends on all parameters of the model. Both the planet parameters and the hyperparameters of the GP (describing the stellar activity) are searched for simultaneously. The posterior distribution of all parameters is the final product, from which the value and standard deviation of each parameters are inferred. This method is very robust and does not require straightforward assumption as any classical modelling would. That makes it particularly suited for stellar activity filtering as it is extremely challenging to come up with a physical model thoroughly describing

the RV signature resulting from stellar activity. For an extensive introduction to Gaussian processes in the context of stellar activity mitigation, we refer the reader to the thesis of Dr. R. D. Haywood[5].

The final aspect of the GP approach is model selection. That is to compare the likelihood between a model containing stellar activity only (i.e., only the GP) and other models containing activity + n planets. As different models do not have the same number of parameters, with each planets on a Keplerian orbit adding 5 parameters, models are not directly comparable. That is because a model with more parameter is more flexible and susceptible to a better fit by design and the addition of extra parameters need to be accounted for. To compare models, it is therefore necessary to evaluate the Bayes factor, comparing the normalised (or marginalised) likelihood of each model. Marginalised likelihood, also called evidence, require an integration over the entire parameter space, very expensive computationally. Approximations do exist, and are compared, in the context of exoplanets, in Nelson et al. (2020). For the work described in chapter 3, we decided to use a nested sampling approach (Feroz & Hobson, 2014) . Extensive details on Bayesian model selection and our strategy are provided in Appendix A of Heitzmann et al. (2021a), presented in chapter 3.

IN THIS SECOND CHAPTER, we have given background on the origin of exoplanets, specifically of gas giants found orbiting their host stars inside the so-called 'ice-line', the hot and warm Jupiters. We covered the mechanisms invoked to explain these exoplanet populations and discussed the clues that can arise from a well-sampled distribution of young planet characteristics. Fi-

---

[5] https://research-repository.st-andrews.ac.uk/handle/10023/7798?mode=full&submit_simple=Show+full+item+record.

nally, we described the challenge associated with the discovery and characterisation of young planets. In chapters 3 to 5, we will showcase the core of the work completed during this Ph. D.

Chapter 3: Planets around young active Solar-type stars: Assessing detection capabilities from a non-stabilised spectrograph.

This first piece of work was inspired by the idea that legacy spectroscopic datasets on young active stars, showing large projected rotational velocity ($v \sin i$) exist as a result of studies aiming to map the distribution of stellar surface features and large scale magnetic field. In this paper, we assess the limitations when attempting to detect giant planets in close-orbits using these legacy datasets. We performed injection/recovery of 37 simulated planetary radial velocity signals (with planets of various masses $\lesssim 7\,M_{\mathrm{J}}$ and orbital periods $\lesssim 7$ days) in the velocities of the young and active Sun-like star HD 141943. To recover the planets' signatures, we applied two stellar activity mitigating strategies. The first was DI, where we modelled the stellar surface to remove the stellar component in the RV and search the residuals using a periodogram. The second was a GP that I implemented, combined with a nested sampling approach, allowing a robust Bayesian framework.

IN THIS PAPER we found that using our legacy dataset consisting of 24 datapoints spread over 10 days we could recover planets of semi-amplitude close to $\frac{1}{4}$ of the stellar activity levels, with $P_{\mathrm{orb}} \lesssim \frac{1}{2}$ of the dataset time span and avoiding cases where P_{orb} is close to P_{rot} or its harmonics. The DI technique shows less robustness than the GP and we advocate for the latter if both can not be performed simultaneously. A study was performed in Jeffers et al. (2014a) using the same star. Both our techniques show an improvement in the recovery of planets compared to this previous work. Finally, we conclude that more datapoints, well sampling the stellar rotation period, and of greater RV precision can drastically improve the detection capabilities and note that such datasets are largely within the reach of state-of the art spectrographs.

Planets around young active solar-type stars: assessing detection capabilities from a non-stabilized spectrograph

A. Heitzmann [ORCID],[1]★ S. C. Marsden [ORCID],[1] P. Petit,[2] M. W. Mengel [ORCID],[1] D. Wright,[1] M. Clerte,[1] I. Millburn,[1]
C. P. Folsom [ORCID],[3,4] B. C. Addison [ORCID],[1] R. A. Wittenmyer[1] and I. A. Waite[1]

[1]*Centre for Astrophysics, University of Southern Queensland, West Street, Toowoomba, QLD 4350, Australia*
[2]*Institut de Recherche en Astrophysique et Planétologie, Université de Toulouse, CNRS, CNES, 14 Avenue Edouard Belin, F-31400 Toulouse, France*
[3]*Department of Physics and Space Science, Royal Military College of Canada, PO Box 17000 Station Forces, Kingston, ON K7K 0C6, Canada*
[4]*Tartu Observatory, University of Tartu, Observatooriumi 1, Tõravere, 61602 Tartumaa, Estonia*

ABSTRACT

Short-orbit gas giant planet formation/evolution mechanisms are still not well understood. One promising pathway to discriminate between mechanisms is to constrain the occurrence rate of these peculiar exoplanets at the earliest stage of the system's life. However, a major limitation when studying newly born stars is stellar activity. This cocktail of phenomena triggered by fast rotation, strong magnetic fields, and complex internal dynamics, especially present in very young stars, compromises our ability to detect exoplanets. In this paper, we investigated the limitations of such detections in the context of already acquired data solely using radial velocity data acquired with a non-stabilized spectrograph. We employed two strategies: Doppler Imaging and Gaussian Processes and could confidently detect hot Jupiters with a semi-amplitude of $100 \, \mathrm{m \, s^{-1}}$ buried in the stellar activity. We also showed the advantages of the Gaussian Process approach in this case. This study serves as a proof of concept to identify potential candidates for follow-up observations or even discover such planets in legacy data sets available to the community.

Key words: techniques: radial velocities – planets and satellites: detection – planets and satellites: formation – stars: activity – stars: individual: HD 141943 – stars: pre-main-sequence.

1 INTRODUCTION

In the quest to understand the processes governing planetary system development, the peculiar case of short-orbit gas giants (i.e. Jovian and sub-Jovian exoplanets, orbiting their star with periods of less than a few weeks that we will refer as hot Jupiters or HJs hereafter) is a real challenge as classical theories describing their formation and evolution do not predict their presence in the vicinity of their parent star. Although representing a significant fraction of all exoplanets discovered (between 10 and 15 per cent[1]), their true occurrence rate is estimated to be around 1 per cent for mature solar-type stars (Wright et al. 2012). Even though this discrepancy can be explained through observing biases, their scarcity raises the question of the formation channel generating this population.

In the most accepted explanation, future HJs form in the colder region of the protoplanetary disc (more than a few au) and later experience orbital decay to eventually reach a close-in orbit. Two migration mechanisms are proposed: gas disc migration (see Baruteau et al. 2014, for a review), where the planet migrates inwards as the result of angular momentum exchange between the gas giant and the disc, and high-eccentricity tidal migration. In this last scenario, the planet is sent to a highly eccentric orbit following a strong perturbation (planet–planet scattering, e.g. Chatterjee et al. 2008,

or secular interactions, see Beaugé & Nesvorný 2012; Petrovich 2015; Petrovich & Tremaine 2016; Hamers et al. 2017 for the different proposed mechanisms). Now being close enough to the star at periastron, tidal forces exerted by the star act to circularize the planet's orbit.

Confronting migration theories is *in situ* formation, where the HJ forms in the vicinity of the host star and remains in close orbit. This explanation has been historically rejected as it sets restrictive constraints on the inner stellar disc, i.e. there must be enough available material to form the core of these gas giants and that core forming process needs to be completed before the star depletes all the gas from the area for the future HJ to successfully accrete its gaseous envelope. Due to these constraints, it is unlikely to occur according to the *Solar nebula theory*, assuming a disc composition similar to the one that gave birth to our Solar system (Perryman 2011). Now realizing that our Solar system may be far from being the norm in the great diversity of planetary systems, *in situ* formation has come back under the spotlight (Batygin, Bodenheimer & Laughlin 2015; Boley, Contreras & Gladman 2015). Recent studies, such as Bailey & Batygin (2018) or Dawson & Johnson (2018), suggest that HJs could have a different origin in different systems and/or that a combination of the proposed mechanisms could be at play.

In their review paper, Dawson & Johnson (2018) propose to test the different theories by searching for correlations between properties of HJs and their parent stars. Among the 15 studied properties, two are flagged as requiring further observations: HJ obliquities and host star ages. This paper focuses on the latter.

★ E-mail: alexis.heitzmann@usq.edu.au
[1]https://exoplanetarchive.ipac.caltech.edu/

46

Studying young stars is a privileged approach as it would help us to discriminate between early-stage mechanisms such as *in situ* formation or gas disc migration versus more prolonged and late-stage mechanisms, like high-eccentricity migration. However, Dawson & Johnson (2018) warn that high-eccentricity migration-driven HJs could arrive in close-in orbit fairly early in the system's formation, showing that the dependence of these mechanisms with stellar age is not yet completely clear. Therefore, young stars ($<$20–50 Myr) and, even more so, younger ($<$10 Myr) low-mass ($<$3 $M_\odot$) T Tauri pre-main-sequence (PMS) stars as defined in Appenzeller & Mundt (1989) are probably the best candidates.

Unfortunately, with the exception of direct imaging surveys, the youngest stars are typically systematically avoided when searching for exoplanets, as they exhibit particularly strong intrinsic variability, or stellar activity. For such stars, this activity-induced correlated noise results primarily from surface brightness features, linked to complex internal processes and a strong magnetic field. Surface features yield spurious radial velocity (RV) signatures that generally completely mask exoplanet signatures, hence preventing their discovery. Additionally, Nava et al. (2020) showed that activity can generate unexpected spurious peaks in a periodogram analysis, which could lead to false positives if no adequate treatment of the activity is applied.

Filtering or mitigating this stellar activity becomes crucial if one hopes to find traces of exoplanets orbiting young active stars. It is also important to note that effective activity mitigating strategies are key in the search of Earth-sized planets around less active stars. In those cases, both the activity level and the planetary signature are up to two orders of magnitude smaller, but present a similar situation in relative terms. However, it is still slightly different as additional phenomena are also at play (i.e. granulation and pulsations). The exoplanet community is actively trying to develop and assess these strategies (see Cabot et al. 2021).

Available data on planets with periods less than 15 d orbiting very young stars ($<$50 Myr) are very scarce. Six planet-hosting stars have been found from transits (David et al. 2016, 2019; Newton et al. 2019; Bouma et al. 2020; Plavchan et al. 2020; Rizzuto et al. 2020) and three from RV searches: CI Tau b (Johns-Krull et al. 2016; Flagg et al. 2019), V830 Tau b (Donati et al. 2015, 2016, 2017), and TAP 26 b (Yu et al. 2017b). Recently, however, the existence of both V830 Tau b and CI Tau b has been challenged by Damasso et al. (2020) and Donati et al. (2020). V830 Tau b and TAP 26 b were found by the MaTYSSE (Magnetic Topologies of Young Stars and the Survival of massive close-in Exoplanets) observation programme in a sample of 33 weak-line T Tauri stars (Yu 2017a). If real, these two planets would indicate a fraction of HJs as high as 6 per cent for newly born stars. In this context, it is crucial to carry on the search for close-in gas giants around young stars to better estimate their occurrence rate at that stage.

In this paper, we investigated the case of searches for short-period gas giants orbiting very young and active stars solely using RV data. More specifically, we injected various RV signatures mimicking single circular planet systems behind real data of the young active G dwarf HD 141943 (not known to host a massive planetary companion) and assessed our detection limits using two distinct strategies: Doppler Imaging (DI) activity filtering (Section 3.1) and Gaussian Process (GP) Regression (Section 3.2).

Although already used in the past (DI + GP in Donati et al. 2016, 2017; Yu et al. 2017b, 2019; Klein et al. 2020 and GP in most exoplanet searches for the past few years), testing the respective performance of these two methods in legacy data sets has not been performed. We note that the underlying data were not

optimized to search for exoplanets and were obtained using a non-stabilized spectrograph (e.g. with $\approx$50–100 m s^{-1} uncertainty on radial velocities). The limitations we describe should therefore be significantly improved with RV stabilized data sets. However, they provide a strong baseline for what is achievable and are useful to investigate other data sets of this nature already available [i.e. in the Bcool (Marsden et al. 2014) or TOUPIES (Folsom et al. 2016, 2018) surveys]. We compared our results to the planet 'hide and seek' study done on the same star with no specific treatment for stellar activity (Jeffers et al. 2014).

This paper is organized as follows: Details on the techniques used to reduce the data, more specifically to get from raw spectra to radial velocities, are given in Section 2. We then cover the methods addressing stellar variability in Section 3. Section 4 of this paper focuses on our reanalysis of HD 141943's raw data set. Section 5 explains how we set up our simulated data sets, and results from the analysis are laid out in Section 6. Finally, we give our conclusions and future prospects in Sections 7 and 8.

2 DATA ANALYSIS

2.1 From spectra to line profiles

Both methods we utilized to disentangle stellar activity from planetary signals take RV time series as input. The extraction of RV values from raw stellar spectra was performed by finding the centroid (described in Section 2.2) of a 'mean line profile' obtained using least-squares deconvolution (LSD; Donati et al. 1997; Kochukhov, Makaganiuk & Piskunov 2010). LSD convolves an observed stellar spectrum with a spectral line mask. Given an appropriate mask, the result is an enhanced peak signal-to-noise ratio (S/N) 'mean line profile' exhibiting stellar activity-induced line features. We chose the stellar mask best matching our star in the list of masks designed in the scope of the Bcool survey (Marsden et al. 2014) using VALD (Kupka 2000) for a star with an effective temperature of $T_{\rm eff} = 6000$ K, a surface gravity of log $g = 4.5$ cm s^{-2}, and [Fe/H] $= +0.2$. Only spectral lines deeper than 20 per cent of the maximum line depth were kept for the LSD computation, yielding a total of 4097 lines. The outcome was an S/N increase from $\approx$50–100 for the observed spectra (depending on the spectrum and spectral order considered) to $\approx$1000 for the LSD mean line profiles.

2.2 From line profiles to radial velocities

Classically, each RV is taken to be the mean of a Gaussian profile fitted to the obtained line profile. However, for active stars, the distortion and here the 'flat bottom' (see the centre plot on Fig. 1) of the line show that a Gaussian fit is not suitable. We considered two alternatives.

First, we chose a generalized normal distribution (GND; Nadarajah 2005), as shown in green on the central plot of Fig. 1 and described by the following p.d.f:

$$\mathrm{GND}(x) = \frac{\beta}{2\sigma\Gamma\left(\frac{1}{\beta}\right)} \exp\left(-\left|\frac{x-\mu}{\sigma}\right|^\beta\right), \qquad (1)$$

where Γ denotes the gamma function, μ the position parameter (mean), σ the scale parameter, and β the shape parameter. $\beta <$ 2 results in wings more extended than a normal distribution and a sharper distribution peak. When $\beta = 2$, the GND becomes a Gaussian distribution (where σ is the standard deviation). For $\beta > 2$, the distribution yields wings less extended than a normal distribution

3 STELLAR ACTIVITY FILTERING

Stellar activity distorts line profiles, causing a shift in the line's centroid and therefore in the measured RV. Modelling the activity thus aims to correct for these distortion-induced shifts.

3.1 Method #1: filtering activity using DI

Section 3.1.1 describes DI, representing the core of our filtering process, following Donati et al. (2014). Sections 3.1.2 and 3.1.3 describe magnetic imaging (Zeeman DI, ZDI) and differential rotation, complementary to the DI technique. The actual filtering process is described in Section 3.1.4.

3.1.1 Doppler imaging

DI is a tomographic technique that, for rapidly rotating stars ($v \sin i \gtrsim 10$ km s^{-1}), uses spectroscopic observations to infer the brightness features at their surface (Brown et al. 1991; Donati & Brown 1997). Practically, a time series of observed pseudo-line profiles obtained through LSD is iteratively adjusted using a tomographic algorithm. Irregularities in the profiles are interpreted as surface bright/dark spots that enhance/block Doppler-shifted light due to stellar rotation. Then, iteratively, synthetic profiles, derived from the DI surface maps, are fitted to the observed ones. To reach a unique solution to the ill-posed problem of DI inversion (as a single line profile can be generated from different surface map solutions), a maximum entropy selection of the solution is adopted (i.e. minimizing the information content of the brightness map), while ensuring that the χ^2 is kept below a defined threshold. This is done following the routine of Skilling & Bryan (1984) and using the entropy as defined in Hobson & Lasenby (1998). Further details can be found in appendix B of Folsom et al. (2016). The model output is constituted of a synthetic set of LSD profiles, and of the brightness surface map producing this spectral information.

Synthetic line profiles are obtained by integrating the Doppler-shifted flux (due to the rotation of the star) emerging from each point of the visible hemisphere. This flux is scaled according to the local surface cell projected area, brightness, and limb darkening. The local line profiles are calculated using a Voigt profile, a convolution of a Gaussian, and a Lorentzian profile.

Output products of DI include a set of synthetic profiles and a surface brightness map (or a magnetic map for ZDI; see the next section). The use of DI also enables us to constrain the stellar fundamental parameters by selecting the parameter values that optimize the brightness model (i.e. inclination of the stellar rotational axis with respect to the line of sight i, line-of-sight projected equatorial rotation velocity $v \sin i$, stellar equatorial rotation period P_{eq}, stellar mean radial velocity $\overline{RV}$, and differential rotation $d\Omega$) and line profile parameters (i.e. line depth and Gaussian and Lorentzian equivalent widths). The DI analysis of HD 141943 is described in Section 4.2 and Fig. 3.

3.1.2 Zeeman Doppler imaging

Although ZDI is not part of the filtering process, it is similar to the stellar mapping process and is therefore described here.

Similarly to DI, ZDI (e.g. Semel 1989) is a technique that uses polarimetric information (i.e. Stokes V LSD profiles) to reconstruct the magnetic field structure at the surface of the star. Here, we used a spherical harmonic expansion to describe the large-scale components of the magnetic field (i.e. poloidal and toroidal; Donati et al. 2006).

Figure 1. Diagram of the data analysis procedure, from raw spectra to periodic signature identification. Each block is a step of the process. Bold text and the associated numbers in parenthesis, respectively, indicate the method used to progress from one block to the next and the section of this paper detailing the corresponding process. Each point on the bottom plot results from the analysis of a single spectrum using the entire procedure described here.

and tends to a uniform distribution as $\beta \to \infty$. This grants more flexibility to the distribution, resulting in a better fit to broadened profiles. Error bars on the GND parameters are given by the fitting method. The centroid μ and the associated error bars for each LSD profile constituted our RV time series.

Secondly, we derived RV values using the first-order moment (generalized centroid, FOM hereafter) of each LSD profile, computed as

$$RV = \frac{\int (I_c - I(\nu))\, \nu\, d\nu}{\int (I_c - I(\nu))\, d\nu}, \tag{2}$$

with $I(\nu)$ the intensity of the profile at radial velocity ν and I_c the continuum level. Here, error bars were propagated using the LSD-derived uncertainties. We note that results given by FOM are sensitive to integration limits (i.e. the limits on the line profiles used to compute it). This is further described in Section 5.3.

48

Figure 2. Reduced χ^2 surface of the differential rotation (rad d^{-1}) (*y*-axis) versus equatorial rotation period (*x*-axis), equivalent to the rotation frequency $\Omega_{eq} = 2\pi/P_{eq}$. Contours show confidence levels at 1σ, 3σ, and 5σ. Colour bar shows the reduced χ^2 values.

The Zeeman effect allows one to infer the strength and direction of the surface magnetic field, provided one has high enough S/N line profiles, rendered possible by the LSD technique. Like DI, solving for a magnetic field configuration is an ill-posed problem and ZDI also relies on maximum entropy image reconstruction. The ZDI analysis of HD 141943 is described in Section 4.2 and Fig. 4.

3.1.3 Surface differential rotation

The information used to generate a snapshot of the stellar surface through DI and ZDI often spans multiple stellar rotation cycles. Thus, the effect of differential rotation needs to be accounted for. The code we used models that differential rotation as a simplified solar-like differential rotation law:

$$\Omega(\theta) = \Omega_{eq} - d\Omega \sin^2 \theta, \tag{3}$$

with $\Omega(\theta)$ the rotation rate at latitude θ, Ω_{eq} ($= \frac{2\pi}{P_{eq}}$) the rotation rate at the equator, and dΩ the difference in rotation rate between the equator and the poles (i.e. the differential rotation). Following Petit, Donati & Collier Cameron (2002) and Donati, Collier Cameron & Petit (2003), we explored the dΩ and Ω_{eq} parameter space, by running DI inversions for various values of the two parameters, looking for the doublet that optimizes the DI model, i.e. the dΩ and Ω_{eq} values that minimize the χ^2 of our model at fixed entropy level. The resulting χ^2 surface is used to derive our uncertainty on these two parameters.

We performed our DI, ZDI, and differential rotation analyses using the PYTHON ZDIPY code (see appendix from Folsom et al. 2018, for a more detailed description of the code). The code has been adapted to run on Fawkes, the High Performance Computing (HPC) facility at the University of Southern Queensland. The HPC allowed us to quickly explore our parameter space. Practically, we varied the stellar parameters (up to 3 at a time) to find the best solution. By best solution we mean the set of parameters that fit our line profiles down to the target χ^2 (<1 due to the LSD process; see Cang et al. 2020, for a similar case and follow-up explanations) and also maximize the entropy value. The differential rotation analysis of HD 141943 is described in Section 4.2 and Fig. 2.

3.1.4 Filtering the activity

Following Donati et al. (2014), we removed the stellar activity contribution by subtracting the RV time series derived from our modelling of the activity alone (i.e. the synthetic profile centroids obtained from DI) from the values obtained from the raw/observed LSD profiles (i.e. the observed/raw LSD profile centroids). We assumed that for very active stars, stellar variability is in a first approximation entirely due to features present on the stellar surface. We then searched for periodicity in the resulting filtered RVs, utilizing a Lomb–Scargle (LS) periodogram (Lomb 1976; Scargle 1982). We point out that residuals exhibit some red noise leftovers, while LS periodograms are designed for uncorrelated/white noise (VanderPlas 2018). We keep this approach here for maximal consistency with previous papers of Donati et al. (2014, 2016) and Yu et al. (2017b, 2019).

The nature of stellar variability (i.e. correlated/red noise), combined with the imperfect filtering (see Fig. 5) of the activity using DI, results in residuals exhibiting some red noise leftovers. As LS periodograms are designed for uncorrelated/white noise (VanderPlas 2018), this approach is limited and should not be used alone to claim a planet detection. To assess significance of a detection, we use the false alarm probability (FAP).[2] To compute the FAP levels, we used the Baluev approximation (see Baluev 2008). We also tried a bootstrap approach, which yielded very similar results.

3.2 Method #2: modelling the activity using a GP regression

Our second approach uses a GP regression to model the stellar activity-induced RV and its temporal evolution as first suggested in Haywood et al. (2014) and Rajpaul et al. (2015). The GP regression treats stellar activity as Gaussian red (correlated) noise. This Bayesian approach is driven by the data points, considered to be random correlated Gaussian variables and the covariance matrix **C**, specifying the correlation between each pair of data points. Following Haywood et al. (2014), we computed each entry $\mathbf{C_{ij}}$ of this covariance matrix using the following physically driven quasi-periodic kernel made of a sinusoidal component to account for the rotation of the star combined with an exponential component for the surface feature appearance/decay:

$$\mathbf{C_{ij}} = \theta_1^2 \cdot \exp\left[-\frac{(t_i - t_j)^2}{\theta_2^2} - \frac{\sin^2\left(\frac{\pi(t_i - t_j)}{\theta_3}\right)}{\theta_4^2}\right] + \left(\sigma_i^2 + \sigma_s^2\right)\delta_{ij}, \tag{4}$$

where the four hyper-parameters can be interpreted as follows:

   (i) θ_1 (km s^{-1}): Semi-amplitude of the activity RV signature.
   (ii) θ_2 (d): Decay parameter, or typical surface feature lifetime.
   (iii) θ_3 (d): Recurrence time-scale, expected to be very close to P_{eq}.
   (iv) θ_4 [0:1]: Smoothing parameter or amount of high frequency in the signal. Smaller and larger values of θ_4, respectively, indicate variations on longer and shorter time-scales. From experience (Jeffers & Keller 2009; Haywood et al. 2018), light curves and RV time series exhibit values of around 0.3 to 0.4 for this parameter. We chose a uniform prior that guarantees to largely encompass these values.

---

[2]The FAP limit indicates the likelihood that a peak caused by random fluctuations in the data would reach a given height/power (see dashed lines in Figs 5, B2, and C1). However, it does not indicate the probability of a data set to have a periodic component given the data.

49

σ_i is the uncertainty of data point i and σ_s is an extra white, uncorrelated noise parameter accounting for variations due to other sources and not explicitly captured by the model. σ_i and σ_s were added in quadrature and only applied to the diagonal of our matrix (i.e. variance of the data points).

Our global model is the sum of the GP model accounting for the stellar activity (RV_{GP}), a sinusoid for the circular planetary signature (RV_{pla}), and a constant offset (RV_0):

$$RV_{tot} = RV_0 + RV_{GP}(t, \theta_1, \theta_2, \theta_3, \theta_4, \sigma_i, \sigma_s)$$
$$+ RV_{pla}(t, K, P_{orb}, \Phi). \qquad (5)$$

We ended up with a parameter space to explore containing 5 (θ_1, θ_2, θ_3, θ_4, and σ_s) + 3 × n parameters + RV_0, for n planets (i.e. 9 parameters for single planet model). Then, two aspects need to be considered in order to confidently claim the presence of a periodic signal in the data. The first part is parameter estimation, where we explore the parameter space yielding posterior distributions from which the most likely set of parameters, as well as their mean and uncertainty values, can be recovered. The second aspect is model selection, where we assess how much more likely a model containing one planet is, compared to one with only stellar activity. Commonly, parameter space exploration is performed using Monte Carlo approaches. Despite the efficiency of some algorithms (e.g. EMCEE from Foreman-Mackey et al. 2013), the bottleneck of planet searches is usually model selection.

Model selection is performed by comparing the *marginal* likelihood (or evidence, $\mathcal{Z}$) of different models (i.e. activity only, activity with one planet, two planets, etc.). A detailed description of the evidence is given in Appendix A. Accurate estimation of this evidence is computationally expensive as it implies multidimensional integration over potentially large parameter spaces. Recently, Nelson et al. (2020) compared different methods for computing the evidence applied by different research groups. Although this was preliminary and would require follow-up studies to completely generalize the results, some approaches proved to be more consistent than others. Following their results, we developed our GP code using PYMULTI-NEST (Buchner et al. 2014), a PYTHON implementation of MULTINEST (Feroz, Hobson & Bridges 2009). This Importance Nested Sampling algorithm estimates the evidence and provides, as a by-product, the posterior probabilities and can therefore also be used for parameter estimation.

For the rest of this paper, when comparing models, we will refer to the Bayes factor (BF) and/or the associated probability (p) in favour of a single planet model (model $\mathcal{M}_1$) over an activity-only model (model $\mathcal{M}_0$):

$$BF = \frac{\mathcal{Z}_1}{\mathcal{Z}_0}, \qquad (6)$$

with $\mathcal{Z}_0$ and $\mathcal{Z}_1$ the marginal likelihood for $\mathcal{M}_0$ and $\mathcal{M}_1$, respectively. We used the metric of Jeffreys (1961) (see Table A1) to assess significance from the BF.

3.2.1 Likelihood and priors

Two ingredients are needed to recover the posterior probabilities: likelihoods and prior probabilities.

In our case, the natural logarithm of the likelihood [i.e. probability of the data given the model and its parameters, p($\mathbf{y}|\theta, \mathcal{M}_i$) or $\mathcal{L}$] is given by

$$2 \ln \mathcal{L} = -n \ln(2\pi) - \ln(|\mathbf{C}|) - \mathbf{y}^T (\mathbf{C})^{-1} \mathbf{y}, \qquad (7)$$

Table 1. Prior distributions of parameters used for the GP regression. The right column gives the prior for each parameter of the model. J(*min*, *max*) stands for Jeffrey's priors, MJ(*max*, *knee*) for Modified Jeffery's priors, $\mathcal{N}$(*mean*, *std*) for Gaussian priors, and $\mathcal{U}$[*min*, *max*] for Uniform priors. $\overline{\sigma}_{RV}$ is the mean of the RV uncertainties. RV_{max} is the largest absolute value in the data set and RV_{std} is the standard deviation of all RV values.

Parameters	Priors
Stellar activity	
θ_1 (km s^{-1})	MJ($1.5\times RV_{max}$, $\overline{\sigma}_{RV}$)
θ_2 (d)	J(1, 100)
θ_3 (d)	$\mathcal{N}$(2.2, 0.05)
θ_4 [0:1]	$\mathcal{U}$[0:1]
Planet	
K (km s^{-1})	MJ($2\times RV_{max}$, $\overline{\sigma}_{RV}$)
P_{orb} (d)	J(0.1, 15)
Φ [0:1]	$\mathcal{U}$[0:1]
Telescope and noise	
RV_0 (km s^{-1})	$\mathcal{U}$[$-RV_{max}$:RV_{max}]
σ_s (km s^{-1})	MJ(RV_{std}, $\overline{\sigma}_{RV}$)

with $\mathbf{y}$ the vector (of length n) containing the residuals after having removed both RV_{pla} and RV_0 from the original RVs and $\mathbf{C}$ the covariance matrix computed using our GP kernel from equation (4).

Our priors, physically motivated following Gregory (2007), are listed in Table 1. Because the evidence is dependent on prior probabilities, we emphasize the importance of favouring uninformative priors, such as uniform, Jeffrey's (uniform prior in logarithmic space; see Gregory 2007) or Modified Jeffrey's (Jeffrey's prior, approaching a uniform distribution for values $\ll$ to the *knee* parameter of the modified Jeffrey's prior to handle priors that have 0 as a lower boundary; also see Gregory 2007), or at least priors independent of the studied data when previous and statistically valid knowledge is available in the literature. Using informative priors, without justification, would act to artificially boost the evidence. This is especially true for parameters that are not shared by the compared models (the planetary parameters in our case). The only informative prior we used here is θ_3 as P_{eq} has been constrained from DI.

We ran PYMULTINEST with an efficiency of 0.3 and 2000 live points (see Nelson et al. 2020). For each run, the parameter search drew between ≈50 000 samples from the posterior for the model with no activity and ≈150 000 for the single-planet model. Details of the results for all data sets are in Table 4.

4 ANALYSIS OF HD 141943

Before attempting to recover injected planets behind HD141943's activity, we analysed the raw observations (data set #5 containing no planet) to recover stellar parameters and ensure the star does not host any planet *that we can detect*.

4.1 Spectropolarimetric data set

Spectroscopic Stokes *I* (intensity) and *V* (polarized) observations of HD 141943 used in this study were acquired using the SEMPOL instrument, visitor polarimeter operating together with the University College London Echelle Spectrograph (Donati et al. 2003) and mounted on the 3.9 m Anglo-Australian Telescope (AAT) in Siding

Table 2. Fundamental parameters of HD 141943.

Parameter	HD 141943
Spectral type	G2V
Distance (pc)	60.028 ± 0.083^d
Age (Myr)	$17–32^b$
$M_\star$ (M$_\odot$)	1.3^a
Photospheric temperature $T_{\rm eff}$ (K)	5850 ± 100^a
Spot temperature (K)	$\approx 3950^a$
$R_{\rm eq}$ (R$_\odot$)	$1.5^{+0.06}_{-0.05}$ ^{c}
i (°)	70 ± 10^a
$v \sin i$ (km s^{-1})	35.6 ± 0.7 ^{e}
Equatorial rotation period $P_{\rm eq}$ (d)	2.198 ± 0.002^e
dΩ (rad d^{-1})	$0.1331^{+0.0095}_{-0.0094}$ ^{e}

[a]M11A.
[b]Hillenbrand et al. (2008).
[c]*Gaia* DR2: Gaia Collaboration (2016, 2018).
[d]*Gaia* EDR3: Gaia Collaboration (2016, 2021).
[e]This study.

Spring, Australia. Available data comprise 92 spectra spread over 11 d between 2007 March 30 and April 9, covering 4.68 stellar revolutions, offering a well-sampled rotational phase coverage as required for DI and ZDI (further details on the data can be found in Marsden et al. 2011a) and a suitable time-scale to search for HJs. The 92 spectra were taken in chunks of four 30 min consecutive exposures, each in different polarization states to perform ZDI. As each 2 h observing run represents a very short time frame compared to $P_{\rm eq}$ (stellar equatorial rotation period) and any simulated HJs' orbital period, this data set can be treated as containing 23 epochs rather than 92. A previous DI and ZDI analysis of this data set is available in Marsden et al. (2011a, b) (M11A/B hereafter). The reduction of raw spectra was done using the ESPRIT pipeline (Echelle SPectra Reduction: an Interactive Tool; Donati et al. 1997).

4.2 Stellar parameters and surface mapping

HD 141943 is a young ($\approx$17–32 Myr, M11A and Hillenbrand et al. 2008), nearby [60 ± 0.08 pc, estimated using Bailer-Jones et al. (2020) with the *Gaia* EDR3 data (Gaia Collaboration 2016, 2021)] active G PMS star. This Sun-like star has a mass of 1.3 M$_\odot$ and a radius of $1.5^{+0.06}_{-0.05}$ R$_\odot$ (*Gaia* DR2; Gaia Collaboration 2018). Soummer et al. (2014) also identified a surrounding near-edge-on debris disc, consistent with a planetesimal belt populated by two dust components at respective grain temperatures of 60 and 202 K. The extended list of stellar parameters can be found in Table 2.

We inferred stellar parameters by analysing the raw HD 141943 data set, containing no injected planet. These are marked with the superscript 'd' in Table 2: $v \sin i = 35.6 \pm 0.7$ km s^{-1}, $i = 43 \pm 10°$, $P_{\rm eq} = 2.198 \pm 0.002$ d, and d$\Omega = 0.1331^{+0.0095}_{-0.0094}$ rad d^{-1}.

These parameters are close although not exactly matching the previous analysis from M11A/B (see the first two lines from Table 3). This discrepancy could be explained by the fact that the DI/ZDI code used between M11A/B is slightly different than ours. Mainly, ZDIPY lets us map bright and dark surface features (spots) in contrast with only dark spots in M11A/B. The inclination is the parameter with the largest difference (43° versus 70°), and also the hardest to constrain. To further investigate, we derived the best solution when fitting both (i) only for dark spots and (ii) using both dark and bright spots but forcing i to match M11A/B's value (i.e. 70°). Obtained Doppler maps and best parameters for the three cases (dark + bright, only dark and

Table 3. Set of parameters resulting from four different analysis: (i) bright and dark feature mapping, (ii) from M11A, (iii) bright and dark feature mapping with an inclination angle constrained to 70° matching M11A's value, and (iv) only dark features.

Best value	$v \sin i$ (km s^{-1})	$P_{\rm eq}$ (d)	dΩ (rad d^{-1})	i (°)
This work	35.6	2.198	0.13	43
M11A/B	35	2.182	0.36	70
Fixed i	35.6	2.197	0.12	70 (fixed)
Dark only	35.4	2.214	0.02	42

dark + bright with imposed 70° inclination) are given in Fig. 3 and Table 3, respectively.

These three cases yielded similar results, however, noting the negligible differential rotation when fitting only the dark features. Forcing i to 70° did not change the overall solution, and we found good agreement between the dark + bright non-forced and forced analyses. The contrast difference on the Doppler maps as seen on the bottom-left map of Fig. 3 is due to the effect of projection imposed by i. Spot locations are consistent across all maps and with M11A. The difficulty to constrain the inclination angle prevents a reliable deduction of the stellar radius $R_{\rm eq}$ and we therefore used the *Gaia* DR2's value given in Table 2. Our main objective for this paper was to filter out as efficiently as possible any rotationally modulated signal in RVs. Since setting i to 43° optimizes this task, we adopted this value for the inclination in the rest of this study.

Fig. 4 shows the radial (top), azimuthal (middle), and meridional (bottom) magnetic field distribution, derived with ZDI using Stokes V LSD profiles. We find a magnetic field with 52 and 48 per cent distribution for the poloidal and toroidal components, respectively, well agreeing with the 47 and 52 per cent from M11A. The mean strength is $\approx$52 G, much lower than M11A's value of 91 G. This can be explained again by the difference in inclination angle. Indeed, reapplying ZDI with a forced $i = 70°$ yields a field strength of 85 G, better agreeing with M11A.

4.3 Planet search

Before injecting planets in the HD 141943 data set, we ensured it did not exhibit any sign of hosting a planet.

The top panel of Fig. 5 shows the periodogram of the raw RVs, where we identified $P_{\rm eq}$ and its harmonics, the strongest signature being present at $P_{\rm eq}/2$. Secondly, third and fourth panels are periodograms of the filtered RVs, respectively, from dark and bright, dark and bright with imposed $i = 70°$ features, and only dark analysis. All show similar features but one peak (around 2.7 d) did show different heights across analysis, and was above the 0.001 FAP threshold for the dark spot-only analysis. However, it did not reach overwhelming significance. This data set did not allow us to assess the impact of the varying DI solutions (dark, dark + bright, and dark + bright with forced inclination) on the planet retrieval as it has no injected planet. To test that, we performed a second analysis using these three configurations for data set 22 (see Section 5 for details on simulated data sets), containing a simulated planet in the 'uncertain' range of detection. We found that the different DI solutions did not change our conclusions regarding the planet search (detailed analysis available in Appendix C).

The GP confirmed that we were not able to detect any significant planet in the raw data set. We found a BF in favour of the single planet model over the activity-only model of only 0.3 ($p \approx 0.23$).

51
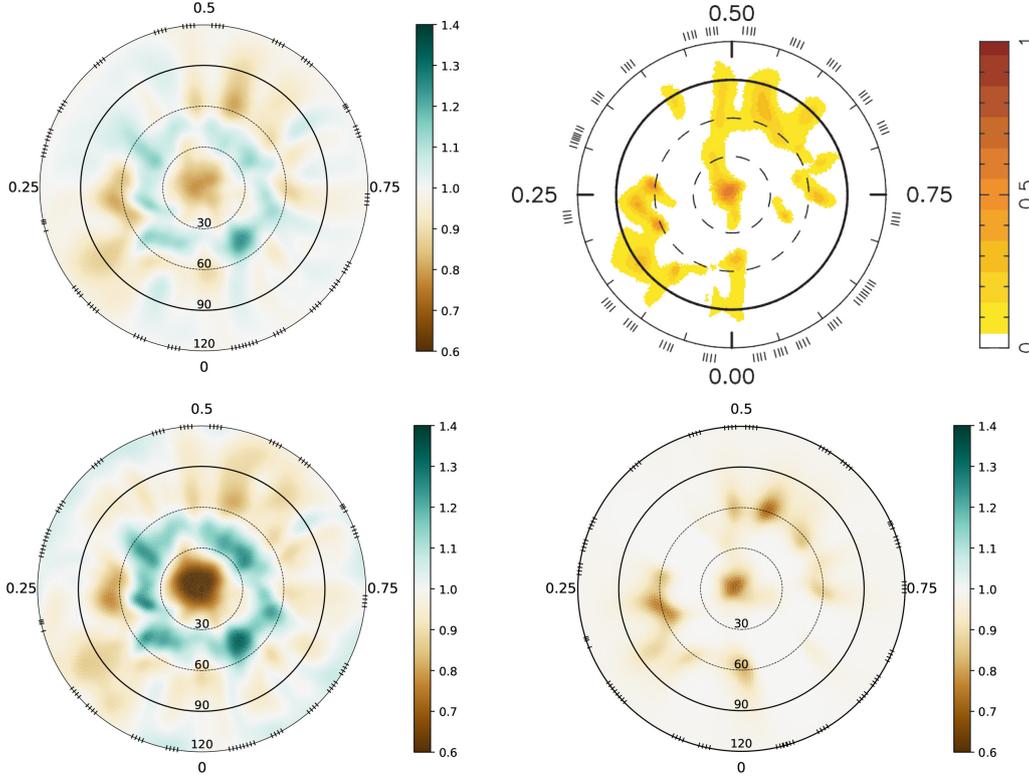
Figure 3. Comparison of the surface brightness maps for HD 141943's original data set (#5). Each map is a maximum entropy reconstructed image of the brightness features at the surface of the star. The blue and brown patches indicate regions that are warmer or colder than the photosphere, respectively. The maps are polar projected, with the centre being the visible rotation pole and the full ring labelled 90 the equator. Values on the outermost ring give the rotational phase and the ticks indicate the phase of each observation. Top left: Map obtained using the parameters from the first line in Table 3 and mapping for bright and dark features. Top right: Map extracted from M11A (second line of Table 3) and mapping only dark features. In their approach, the colour scale expresses the spot filling factor. Bottom left: Map obtained using the parameters from the third line of Table 3, with a forced inclination parameter of 70° and mapping for bright and dark features. Bottom right: Map obtained using the parameters from the fourth line of Table 3 and mapping only dark features.

5 SIMULATED EXOPLANET DATA SETS

We created 37 data sets, each containing a single planet on a circular orbit around HD 141943, following a procedure described in Sections 5.1 and 5.2. Each planet was incorporated into the raw spectra studied in the previous section.

5.1 Injected planets

Injected planets were chosen to be massive short-period exoplanets, with masses ranging from 0.38 up to $5.9M_J$ and periods shorter than 6 d. We set the orbits to be circular, as it is believed to be the case for most HJs, especially for orbits shorter than 3 d (Dawson & Johnson 2018). We should none the less bear in mind that eccentricity can be a crucial aspect in favour of high-eccentricity migration and should not be overlooked especially when attempting to detect the slightly cooler warm Jupiters ($P_{orb} > 10$ d). The RV shift induced by each planet was defined as

$$\mathrm{RV_{pla}}(t) = K \sin\left[2\pi\left(\frac{t}{P_{orb}} - \Phi + 0.5\right)\right], \tag{8}$$

with K the semi-amplitude of the signal, P_{orb} the planet's orbital period, and Φ the phase. $\Phi \in [0: 1]$ and was defined such that when $\Phi = 0$, the planet crosses the plane containing the line of sight. We set $\Phi = 0$ to match the mid-point of the observations ($\mathrm{BJD_{\star_{mid}}} = 2454195.153776$). For the rest of this paper, we will refer to semi-amplitude values (K) for the planets rather than mass. The equivalence between K and mass is described in Section 6.2.5.

5.2 Complete data sets

To build our data sets, we generated an RV time series using equation (8) at times matching our observing epochs and then shifted each spectrum accordingly in wavelength space. In order to explore our planetary parameter space (made of K, P_{orb}, and Φ) without being overwhelmed with the number of data sets to analyse ($n_K \times n_{P_{orb}} \times n_\Phi$), we used the following strategy:

First, we created seven data sets (#1 to #8, excluding #5, the original one) at a fixed period (3.653 d) with K ranging from 50 to 500 m s^{-1} and random Φ. This initial analysis provided an estimate of the limiting semi-amplitude range for detectability.

52

Radial Field

Azimuthal Field
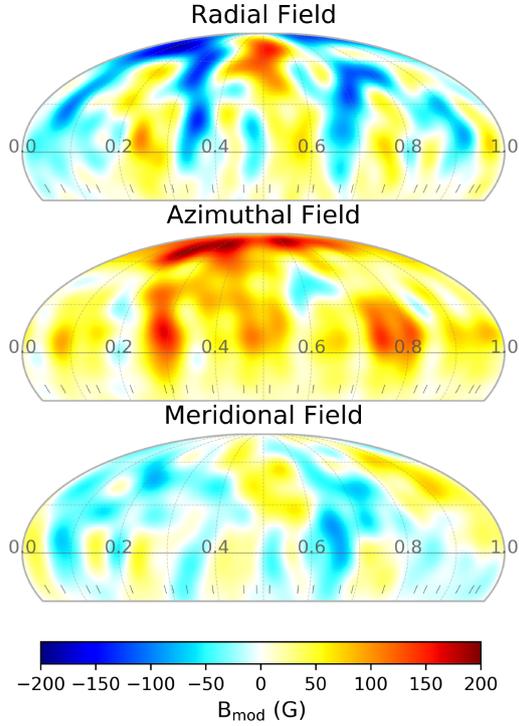
Meridional Field

Figure 4. Maximum entropy image reconstruction of the radial (top), azimuthal (middle), and meridional (bottom) components of the magnetic field for HD 141943. Positive/negative field modulus values (in Gauss) are displayed in yellow/red and blue, respectively. The horizontal line shows the equator with the number describing the phase. Ticks translate each measurement's epoch.

Then, we generated additional data sets 4–5 at a time, filling areas in the parameter space that seemed relevant, i.e. around the noise limit, around P_{eq} and harmonics and to cover empty areas of the parameter space. After generating each batch of data sets, these were randomly assigned a mock value before analysis to avoid biases.

We ended up with 37 data sets (38 when including the original data set) spanning the following ranges: 42–532 m s^{-1} in K, 0.288–5.69 d for P_{orb}, and 0.02–0.99 in Φ. See Table 4 for specific details.

5.3 RV extraction

We tested the RV extraction from the line profiles using both a first-order moment approach and a GND fit. For the original data set (#5 with no injected planet) and given our limits on the integration for the FOM approach the average difference between FOM- and GND-extracted RVs (from the observed profiles) is 8 m s^{-1} with a maximum difference of 144 m s^{-1} for the most extreme point. The uncertainty on the RVs also differed as GND yielded uncertainties twice as large as those from FOM (148.5 ± 2.0 m s^{-1} versus 70.4 ± 1.8 m s^{-1}).

As previously mentioned, FOM-derived RVs and uncertainties are dependent on the number of points taken into account for its computation (i.e. chosen integration limits for the line profile) and where to cut in the wings of the line profile can be somewhat arbitrary.

On both sides of each profile (and for all of them), we cut at 43 km s^{-1} relative to the line centre. We tested different limits and chose the one that gave the least average difference with the GND approach, which is an analytical function and therefore not prone to this effect.

6 RESULTS

All data sets are extensively described in Table 4 and will be referred to using their number (#1, #2, ..., #38).

6.1 Stellar parameters

Stellar parameters inferred from the data sets containing injected planets are consistent with our refined parameters derived from the raw data set (#5; see Section 4.2). The mean values of the retrieved stellar parameters across all data sets, along with the largest deviation from the mean (given by the ±), are: $v \sin i = 35.6^{+0.27}_{-0.45}$ km s^{-1}, $i = 43^{+5}_{-3}$ °, $P_{eq} = 2.1995^{+0.007}_{-0.006}$ d, and $d\Omega = 0.119 \pm 0.08$ rad d^{-1}. In all cases, spot distributions are similar, with slight differences in terms of contrast. This can be explained by the fact that the fit to the line profiles was sometimes performed to a slightly different χ^2 level. Typically, the presence of a planet with a semi-amplitude significantly larger than the activity level (e.g. for data set #6) slightly impacts the performance of the DI. However, even such a large planet signature did not hamper the capacity of the DI to identify spot locations and recover the planet.

6.2 Planet detection: methods performance

6.2.1 Method 1: DI activity filtering

Results are shown in Fig. 6. Each marker represents a data set with its corresponding number in order to easily refer to Table 4 containing more details. Marker positions indicate the injected planets' K and P_{orb} (as we did not identify a systematic impact of Φ on retrievals, it was omitted for clarity). Each data set is identified by a specific marker/colour combination: A green circle when a periodogram peak was identified above an FAP threshold of 0.001 (0.1 per cent) and with a deviation from the true period of <10 per cent, an orange square when two peaks above an FAP of 0.001 and of similar height were found or when the right peak was identified but with a deviation from the true period of >10 per cent, preventing a safe conclusion, a grey cross when no signature above FAP = 0.001 could be identified, and finally a red cross when a peak was present above the FAP threshold but was not matching the injected planetary period, i.e. a false positive.

This approach yielded 16 positive detections, 6 inconclusive findings, 11 non-detections, and 4 false positives. These four data sets confirm that using FAP as a measure of significance is not the safest approach, as discussed in Section 3.1.4. Rigorous estimation of the significance was performed with the GP analysis.

All six simulated planets with semi-amplitudes larger than 150 m s^{-1} were well retrieved and half (8/18) between 100 and 150 m s^{-1}. This fraction increased to 60 per cent (8/13) when removing all data sets close to P_{eq} and its harmonics. Only 1/13 planets below 100 m s^{-1} were found (noting that #25 was a very weak detection with an FAP of 3.5×10^{-4}).

We note that all analyses that identified the right peak but with deviation of more than 5 per cent (up to 10.6 per cent for #21) from the true period (#19, #21, #25, and #35) had fewer than 2.5 orbital periods within our observation time span. This inaccuracy was due to the width of the peaks in the periodogram arising as the period

53
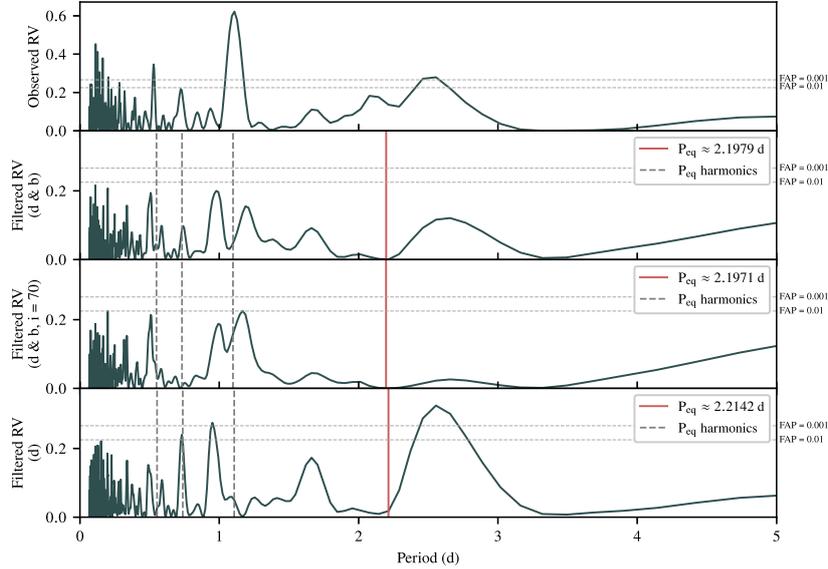
Figure 5. LS periodograms for the original data set (#5, containing no planet). First panel: Observed (raw) RVs. All other panels are for filtered RVs (i.e. raw RVs – synthetic RVs, where synthetic RVs are derived from the DI fitting). Second panel: Filtered RVs using the dark and bright (d and b) features for DI. Third panel: Filtered RVs using the dark and bright features for DI and with the 70° constraint on i (d and b, $i = 70$). Fourth panel: Filtered RVs using only dark features (d) for DI. We note that the peak around 2.6 d is likely caused by the rotation period at intermediate latitudes (offering a maximal visibility given the inclination angle of the star). This value is larger than the equatorial period depicted by the red, vertical lines, as expected in the case of a differentially rotating surface (see Section 3.1.3).

represented a significant fraction of the time span. It is safest, given our number of samples, to cover at least 2 to 3 orbital periods to achieve sufficient precision on P_{orb}.

The six 'to be confirmed' (orange squares) data sets #2, #15, #32, #34, and #36 all exhibited two competing peaks above the FAP threshold and at a similar height (FAP of 1.8×10^{-4} and 7.2×10^{-4} for #2, 1.6×10^{-4} and 1.4×10^{-4} for #15, 1.1×10^{-11} and 1.1×10^{-11} for #32, 1.25×10^{-6} and 2.6×10^{-4} for #34, and 1.15×10^{-10} and 2.43×10^{-9} for #36), preventing us from being able to choose the correct period. For #2, #15, and #32, the peaks are just above our detection threshold and it is therefore not surprising to find competing features. For #32 and #36, however, the competing peaks were both very significant. We are not sure as to how to interpret these, which vanished as we filter the signature associated with one of the two peaks. As shown in Fig. B2, these spurious peaks did not seem to correspond to any harmonics of neither the planet nor the star. See periodograms. Although complex interactions between the uneven data sampling and the periodic signatures cannot be ruled out, no significant peak could be identified in the window function (see Fig. B1). The complete analysis of these two data sets can be found in Appendix B. Data set #21 also falls in the 'orange square' category with a very wide identified peak, yielding a 10.6 per cent deviation between the retrieved and injected periods, slightly over our 10 per cent threshold.

False positives arose as the highest peak was not the simulated one, which would lead to false identification (if relying solely on DI) for #1 and #17. Regarding #9 and #24, the peaks were barely above the FAP of 0.001 and would not have led to a significant detection.

Although we could not identify a systematic impact, phase is expected to play a role in the injected planets retrieval and we can

see this occurring in the zoomed box in Fig. 6. The only noticeable difference between #13 and #15 is their phases (respectively, $\Phi_{13} = 0.4769$ and $\Phi_{15} = 0.1093$) and yet planet #13 is recovered but not #15.

Studying periodograms for all data sets indicated that planets with periods close to P_{eq} (#12 and #26), $P_{eq/2}$ (#33 and #34), $P_{eq/3}$ (#9 and #24), and $P_{eq/4}$ (#30 and #32) seem to be affected by the activity filtering. The case of #32 has been discussed above. This effect is to some extent expected as DI has the capacity to distort the line profiles, interpreting the rotationally modulated distortions as produced by spots on the brightness maps at harmonics of the rotation period and therefore is likely to absorb part of a planetary signature close to one of these periods.

For RV searches, the LS periodogram has limitations (choosing an FAP limit, interpreting the significance of a result, limitation to sinusoidal signals; see VanderPlas 2018) and we emphasize that dedicated treatment for stellar activity should be performed. We therefore advocate for incorporating a second, complementary method, presented in the following section, allowing both better quantification of the significance of a retrieved signature and a more comprehensive modelling of the activity.

6.2.2 Method 2: GP regression activity modelling

Again, results are detailed in Table 4 and summarized in Fig. 7. We defined successful retrievals (green circles) where the GP strongly favoured the single planet model over the activity-only model with a probability $p > 0.909$ (computed from the marginal likelihood/BF; see Appendix A and Table A1 for further details). We then have substantial evidence (orange

Table 4. Data sets 1 to 6. Each column represents a simulated data set. The first section (rows 1–4) gives the stellar parameters inferred from the DI analysis. The second section (rows 5–7) gives the values of the three parameters used to simulate the injected planet. The third section (rows 8–13) gives the results of the Method #1 DI filtering. RV_0 is an offset; then, we have the three recovered planet parameters, followed by the rms and χ^2 of the residuals. The fourth section (rows 14–22) gives the result of the GP (Method #2) for the no planet (activity-only) model. For the parameters (θ_1 to σ_s), the values are given as: mean $\pm$ std (maximum a posteriori). We then have the rms and χ^2 of the residuals and the resulting natural logarithm of the evidence. The last section (rows 23–36) gives the result of the GP for the single planet model. Again, the parameter (θ_1 to σ_s) values are given as: mean $\pm$ std (maximum a posteriori). We then have the rms and χ^2 of the residuals and the resulting natural logarithm of the evidence. Finally, we give the BF, defined as the ratio between $\mathcal{Z}$ from the single planet model (row 34) and the no planet model (row 22). The last row is the probability in favour of the single planet model associated with the BF value. Only the first six rows (data sets) are shown here; the full version is available as online material.

Data set	#1	#2	#3	#4	#5	#6	...
DI inferred stellar parameters							
$i(°)$	43	40.5	41.5	42	42.5	45	...
$v \sin i$ (km s^{-1})	35.47	35.53	35.75	35.789	35.643	35.862	...
P_{eq} (d)	2.206 15	2.203 45	2.1988	2.194 08	2.197 88	2.201 79	...
$d\Omega$ (rad d^{-1})	0.0769	0.097 96	0.133 33	0.146 94	0.133 33	0.123 08	...
Injected planet parameters							
K (m s^{-1})	60.6	82.9	154.9	267.1	–	532.2	...
P_{orb} (d)	3.6531	3.6531	3.6531	3.6531	–	3.6531	...
Φ [0:1]	0.271	0.303	0.445	0.311	–	0.432	...
Method #1 (DI) mean $\pm$ std							
RV_0 (m s^{-1})	13.1 ± 9	1.7 ± 9	0.5 ± 9	12.2 ± 10	–	-60.0 ± 10	...
K (m s^{-1})	85.5 ± 12	68.7 ± 11	115.6 ± 12	233.0 ± 13	–	430.2 ± 14	...
P_{orb} (d)	2.549 ± 0.054	1.413 ± 0.022	3.538 ± 0.067	3.649 ± 0.041	–	3.640 ± 0.022	...
Φ [0:1]	0.612 ± 0.026	0.209 ± 0.031	0.435 ± 0.018	0.301 ± 0.009	–	0.427 ± 0.006	...
rms (m s^{-1})	81.9	79.6	79.7	84.4	–	89.9	...
χ^2	1.34	1.28	1.28	1.43	–	1.63	...
Method #2 (GP)/no planet model/mean $\pm$ std (MAP)							
θ_1 (m s^{-1})	$314.2 \pm 113.4(217.4)$	$276.8 \pm 89.9(208.2)$	$265.4 \pm 68.0(217.1)$	$320.5 \pm 74.1(277.5)$	$357.6 \pm 109.1(296.6)$	$462.9 \pm 98.0(394.1)$	...
θ_2 (m s^{-1})	$10.2 \pm 6.3(6.9)$	$6.9 \pm 3.8(6.0)$	$3.0 \pm 1.4(4.2)$	$1.7 \pm 0.5(1.7)$	$19.2 \pm 9.3(15.7)$	$1.4 \pm 0.3(1.1)$	...
θ_3 (m s^{-1})	$2.190 \pm 0.036(2.148)$	$2.190 \pm 0.041(2.141)$	$2.205 \pm 0.045(2.164)$	$2.206 \pm 0.049(2.255)$	$2.215 \pm 0.020(2.216)$	$2.207 \pm 0.049(2.268)$	...
θ_4 (m s^{-1})	$0.527 \pm 0.148(0.399)$	$0.468 \pm 0.136(0.347)$	$0.415 \pm 0.137(0.234)$	$0.587 \pm 0.146(0.511)$	$0.628 \pm 0.083(0.597)$	$0.723 \pm 0.151(0.605)$	...
RV_0	$12.5 \pm 156.8(-20.5)$	$19.7 \pm 127.6(16.6)$	$35.6 \pm 97.7(52.1)$	$15.8 \pm 113.3(-0.7)$	$25.3 \pm 194.0(50.9)$	$17.4 \pm 160.3(49.6)$	...
σ_s	$12.3 \pm 9.5(0.4)$	$11.7 \pm 9.1(2.8)$	$12.0 \pm 9.2(9.8)$	$12.0 \pm 9.2(3.0)$	$11.5 \pm 8.8(0.6)$	$12.3 \pm 9.5(0.7)$	...
rms (m s^{-1})	55.9	54.9	54.8	55.8	59.0	56.6	...
χ^2	0.62	0.59	0.59	0.61	0.69	0.63	...
$\ln\mathcal{Z}$	-550.99	-552.15	-560.46	-563.85	-545.03	-571.93	...
Method #2 (GP)/single planet model/mean $\pm$ std (MAP)							
θ_1 (m s^{-1})	$340.1 \pm 117.7(271.1)$	$360.8 \pm 120.8(257.2)$	$400.9 \pm 151.9(301.2)$	$401.0 \pm 138.3(347.7)$	$357.3 \pm 107.9(292.2)$	$473.5 \pm 199.5(338.4)$	...
θ_2 (m s^{-1})	$14.0 \pm 7.8(13.3)$	$17.2 \pm 10.0(17.1)$	$20.9 \pm 13.2(15.3)$	$21.7 \pm 12.3(19.0)$	$19.9 \pm 9.7(15.9)$	$27.7 \pm 16.7(22.3)$	...
θ_3 (m s^{-1})	$2.204 \pm 0.030(2.217)$	$2.213 \pm 0.026(2.233)$	$2.208 \pm 0.024(2.216)$	$2.215 \pm 0.019(2.220)$	$2.216 \pm 0.020(2.228)$	$2.206 \pm 0.020(2.213)$	...
θ_4 (m s^{-1})	$0.595 \pm 0.151(0.602)$	$0.642 \pm 0.154(0.601)$	$0.641 \pm 0.163(0.570)$	$0.677 \pm 0.143(0.645)$	$0.631 \pm 0.081(0.592)$	$0.705 \pm 0.156(0.680)$	...
K (m s^{-1})	$43.1 \pm 23.6(59.8)$	$59.2 \pm 19.7(69.1)$	$130.8 \pm 19.1(138.8)$	$238.7 \pm 14.2(235.6)$	$29.4 \pm 32.1(39.7)$	$474.7 \pm 14.5(473.4)$	...
Period (d)	$3.008 \pm 2.386(3.331)$	$3.223 \pm 1.341(3.420)$	$3.571 \pm 0.438(3.598)$	$3.580 \pm 0.102(3.603)$	$2.481 \pm 2.528(0.172)$	$3.616 \pm 0.035(3.610)$	...
Phase [0:1]	$0.345 \pm 0.210(0.259)$	$0.305 \pm 0.115(0.286)$	$0.444 \pm 0.036(0.448)$	$0.306 \pm 0.013(0.309)$	$0.571 \pm 0.293(0.821)$	$0.432 \pm 0.005(0.430)$	...
RV_0	$38.6 \pm 166.5(-1.6)$	$52.6 \pm 178.1(119.5)$	$27.8 \pm 212.0(81.8)$	$69.8 \pm 211.2(1.2)$	$28.4 \pm 190.2(166.2)$	$39.2 \pm 292.2(121.2)$	...
σ_s	$11.0 \pm 8.3(6.8)$	$11.1 \pm 8.4(3.9)$	$12.4 \pm 9.4(3.2)$	$11.8 \pm 9.1(5.1)$	$10.7 \pm 8.2(6.3)$	$12.7 \pm 9.6(2.8)$	...
rms (m s^{-1})	57.3	62.1	61.3	57.9	60.4	64.4	...
χ^2	0.65	0.77	0.75	0.66	0.72	0.82	...
$\ln\mathcal{Z}$	-551.06	-550.91	-556.77	-556.41	-546.25	-560.60	...
Bayes factor	0.9	3.5	40.0	1702.8	0.3	83283.0	...
$p(\mathcal{M}_1)$	0.48	0.78	0.98	1.00	0.23	1.00	...

squares) of a planet (i.e. $0.75 < p < 0.909$) and non-detections (grey crosses, i.e. $p < 0.75$). We note that most of the injected planets (28/37) were correctly identified by the GP, although not always significant enough to lead to a detection claim.

The GP yielded 16 positive detections, 4 'to be confirmed' findings (i.e. requiring further observations), 17 non-detections, and more importantly no false positives. Again here, all six simulated planets with semi-amplitudes larger than 150 m s^{-1} were correctly found. This drops to half (9/18) between 100 and 150 m s^{-1} (same ratio as DI although not systematically on the same data sets), and increases to 70 per cent (9/13) when removing all data sets close to P_{eq} and its harmonics. Finally, for planets below 100 m s^{-1}, only 1 out of 13 was found (along with two cases requiring further observations, with #2 a correctly identified planet and #27 a missed identification). The GP, compared to DI, is more conservative yet more reliable (i.e. no false positives) due to its accurate measure of the significance for each finding. Fig. 7 shows that, similar to the DI analysis, it is difficult to identify planetary signatures close to P_{eq} or its harmonics.
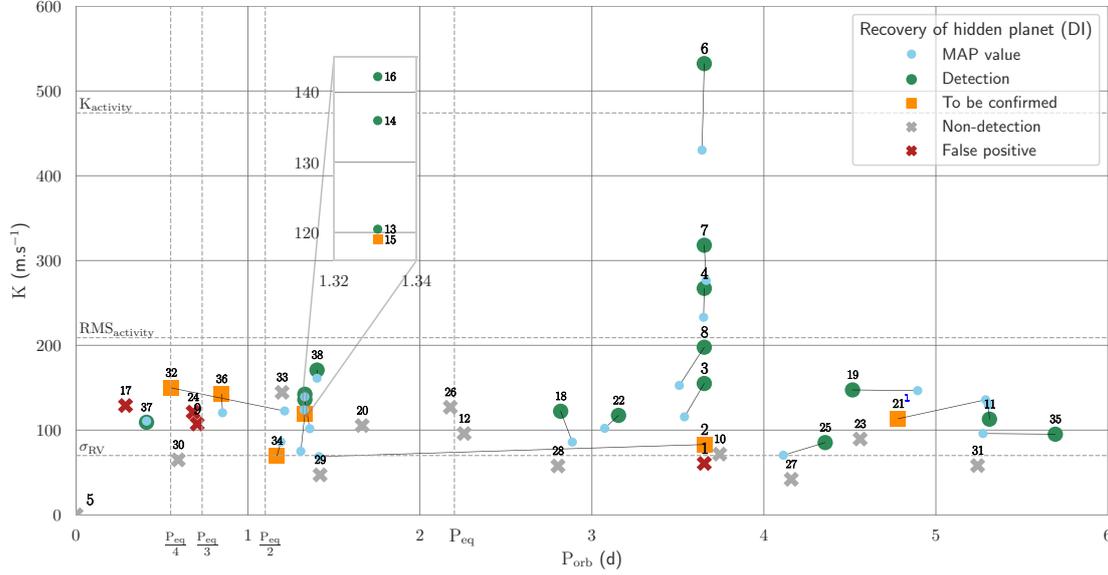
Figure 6. Results using the DI method. Each marker on the plot is a data set containing a simulated injected planet. Orbital period (d) is on the *x*-axis and semi-amplitude (m s^{-1}) is on the *y*-axis. Green circles: identification of the correct planet with a periodogram peak above an FAP of 0.001. Orange square: two peaks of similar height (see Section 6) were found preventing a safe conclusion or when the width of the correctly identified peak yielded a deviation of more than 10 per cent between the retrieved and injected orbital period. Grey crosses: no signature with FAP > 0.001 could be identified. Red crosses: the most significant peak was not matching the injected period peak. Horizontal dashed lines show the stellar activity semi-amplitude ($K_{activity}$) and the error bars on the retrieved RVs (σ_{RV}). Vertical dashed lines show the rotation period of the star and its harmonics. Blue points show, for the 'Detection' and 'To be confirmed' data sets, the P_{eq} and K values corresponding to the highest peak in the periodogram.[1]: For data set #21, the correct peak was found, at an FAP of 1.1×10^{-13}. However, that peak being very broad, it yielded a 10.6 per cent deviation between the retrieved and injected periods, slightly outside our 10 per cent limit.

We finally note that imprecision on the retrieved P_{orb} increases for longer periods (see MAP values indicated on Fig. 7). This is because fewer periods are covered by the data set as we move to the right of Fig. 7.

6.2.3 Consistency between methods

Utilizing two distinct methods serves as cross-validation when a signature is found. However, the GP is the only Bayesian approach, therefore the only one allowing a rigorous quantification of the evidence favouring of a particular model (i.e. presence of a sinusoidal signature in the data or not).

For signatures above 150 m s^{-1} and after removing data sets close to P_{eq} harmonics, both methods yielded systematic detections except for the ambiguity on #36 when using DI. For signatures between 100 and 150 m s^{-1}, the GP showed more consistency than the DI that exhibited three false positives. Out of the 13 data sets below 100 m s^{-1}, we ended up with 1 detection for both GP and DI.

Even though the Bayesian approach using a GP can (i) better handle correlated noise and (ii) more reliably estimate the significance of a detection, the use of the DI filtering method allows an independent validation.

6.2.4 Comparison with previous work

In a study performed in 2014, Jeffers et al. (2014) (J14 hereafter) injected various planets behind simulated activity signatures of

two young G and K stars. Varying parameters were the planet semi-amplitude, orbital period, and $v \sin i$ (shown to be correlated with the activity level). Stellar activity was generated based on DI maps and with different configurations (e.g. adding plages associated with spots, adding extra random spots, etc.; see J14 for more details). The G dwarf was HD 141943, thus making the comparison particularly relevant. Each simulated data set was composed of a single planet in a circular orbit, to which modelled stellar activity and instrumental signatures were added. In that study, the search for injected planets was performed *without* a specific treatment for stellar activity and was considered successful for periodogram peaks below FAP = 0.01 (versus 0.001 in our study).

With 50 observational epochs and for their less complex simulation of activity (only based on the DI maps), J14 were able to retrieve signatures of semi-amplitude $K = 110$ m s^{-1} when $v \sin i = 20$ km s^{-1} and $K = 525$ m s^{-1} when $v \sin i = 50$ km s^{-1}. Regarding their most complex simulation of activity (DI maps + plages + random spots; see J14 for further details), the minimum attainable planetary signature was $K = 525$ m s^{-1} when $v \sin i = 20$ km s^{-1}. In the case of $v \sin i = 50$ km s^{-1}, 200 observational epochs were required to reach the $K = 525$ m s^{-1} detection threshold.

We note that the data sampling is different, which might slightly hinder the comparison.[3] With 23 unevenly spread epochs, the

---

[3]J14 has one datum per night for 50, 100, or 200 consecutive nights, whereas we have 23 epochs over 10 nights.

56
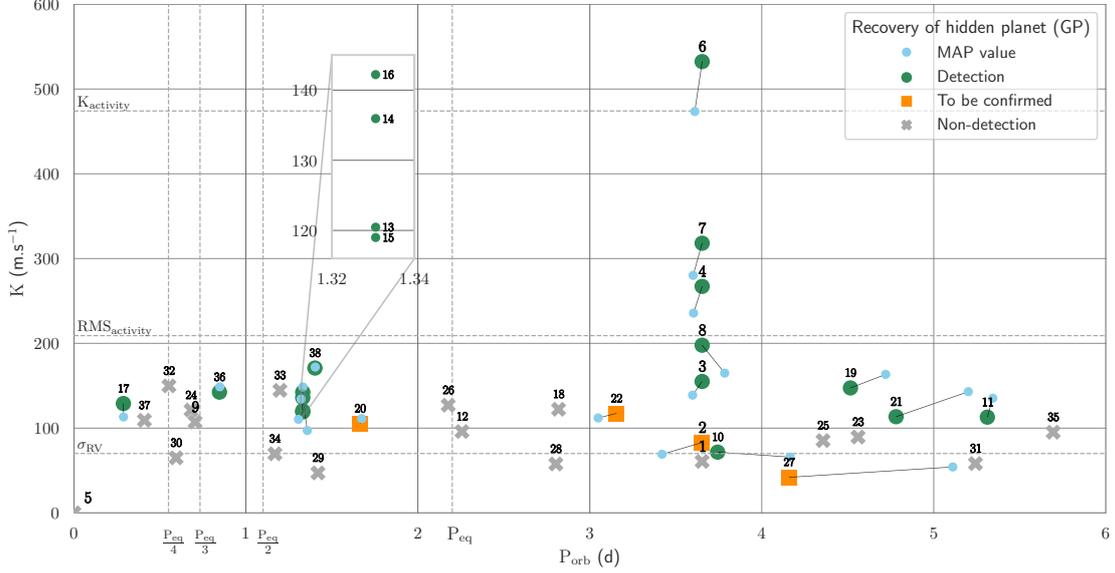
Figure 7. Results using the GP method. Each marker on the plot is a simulated injected planet. Orbital period (d) is on the *x*-axis and semi-amplitude (m s^{-1}) is on the *y*-axis. Green circles: the evidence between the single-planet model and with activity only is $\Delta \ln(\mathcal{Z}) > 10$, associated with at least a strong detection (probability $p > 0.909$). Orange squares: $10 > \Delta \ln(\mathcal{Z}) > 3$ or substantial evidence ($0.95 > p > 0.75$). Grey crosses: $\Delta \ln(\mathcal{Z}) < 3$ ($p < 0.75$). Horizontal dashed lines show the stellar activity semi-amplitude ($K_{\rm activity}$) and the error bars on the retrieved RVs ($\sigma_{\rm RV}$). Vertical dashed lines show the rotation period of the star and its harmonics. Blue points show the Maximum A Posteriori (MAP) values for $P_{\rm eq}$ and K, linked by a line to the corresponding data set.

smallest signature we could reliably detect was $K = 100 \text{ m s}^{-1}$ (down to 70 m s^{-1} for data set #10), emphasizing the benefit granted by our activity filtering approach. Although now systematically applied by the community for planetary searches around active stars, this emphasizes that an independent treatment of stellar activity combined with robust model selection is crucial to improve detection capabilities.

6.2.5 Recovered exoplanets

Here, we translate our results into planetary mass/orbital periods for the case of HD 141943, given $M_\star = 1.3 \, M_\odot$ and $i = 43°$. We consider that the stellar rotation axis is normal to the planet's orbital plan. In this context, our lower detection threshold of 100 m s^{-1} is equivalent to a planet with either

(i) $M_{\rm pla} = 1 \, M_{\rm Jup}$, $P_{\rm orb} = 1.6$ d ($a = 0.03$ au) or
(ii) $M_{\rm pla} = 2 \, M_{\rm Jup}$, $P_{\rm orb} = 12.5$ d ($a = 0.12$ au).

Using M11A's inclination value of $i = 70°$, we get

(i) $M_{\rm pla} = 1 \, M_{\rm Jup}$, $P_{\rm orb} = 5$ d ($a = 0.062$ au);
(ii) $M_{\rm pla} = 2 \, M_{\rm Jup}$, $P_{\rm orb} = 35$ d ($a = 0.23$ au).

In a case of a star of 1 M$_\odot$ with a transiting exoplanet, we can hope to detect a 1 M$_{\rm jup}$ orbiting at up to 10 d, using typical DI non-stabilized observations. This is, of course, given similar conditions in terms of data quality and quantity, observing constraints, and stellar variability level. As more numerous and precise RVs should be easily obtainable, it is fair to expect better results and identify HJs around very young solar-type stars.

6.2.6 Dependence of planet detection to various parameters

In terms of semi-amplitude, our detection threshold of around 100 m s^{-1} corresponds to half of the stellar activity rms and a quarter of its semi-amplitude ($\approx$400–500 m s^{-1} looking at the maximum of the data points, or 357 ± 100 m s^{-1} according to the GP applied to data set #5). Given the scarcity of planets orbiting very young stars discovered solely using RV, comparisons with the literature are limited. When excluding searches in the small activity regimes (i.e. rms$_{\rm activity} < 50 \text{ m s}^{-1}$) only two planets provide a direct comparison, V830 Tau b (Donati et al. 2017) and TAP 26 b (Yu et al. 2017b).

TAP 26 b is thought to have a semi-amplitude of 160 m s^{-1}, or $\frac{1}{8}$ to $\frac{1}{12}$ of the stellar variability semi-amplitude and V830 Tau b ($K \approx 60 \text{ m s}^{-1}$) up to $\frac{1}{20}$. They both exhibit activity levels with a semi-amplitude of $\approx$1200 m s^{-1}. We believe that the difference in performance (detection threshold of $\approx \frac{1}{4}$ of the activity level for this work) can be explained by the fact that (i) both Yu et al. (2017b) and Donati et al. (2017) had more data ($\approx$30 and 60 epochs versus 23 for us), (ii) Donati et al. (2017) had slightly better uncertainties on the RVs [$\sigma_{\rm RV} \approx 50 \text{ m s}^{-1}$ versus 75 for both Yu et al. (2017b) and this study] and most importantly, (iii) both had a longer baseline for the observations: 100 and 35 stellar rotation cycles versus 3 for us. We also used less constraining priors as previous knowledge was not available.

To ensure it was not our method implementation that hindered our capacity to find smaller signatures from our data sets, we ran our code on both Donati et al. (2016)'s and Yu & Donati (2017a)'s data and were able to retrieve the published periodic signatures. We note that with no access to previous knowledge, our prior distributions were less restrictive (i.e. non-Gaussian and/or broader

57

for the concerned parameters), decreasing the evidence and yielding slightly more conservative results. Our limitations can be seen as an upper boundary and that data of better quality and quantity would be able to detect smaller planets.

We find that detecting planets with orbital periods conflicting (i.e. within 0.1 d of) either P_{eq} or its harmonics was unreliable, as illustrated in Figs 6 and 7. Although longer period planets did not seem to be harder to detect, we noticed a significant loss of precision in the orbital period retrieved once we reach periods larger than 40 to 50 per cent of the observing time frame (see MAP values on Fig. 7). This is expected and good practice to sample a few orbital periods at least to get reasonable constraint. A good example of a similar study can be found in Klein et al. (2020), where the authors required 35/50 data points spread over 3 months ($\approx$15 orbital periods) to reliably detect 5/10 m s^{-1} planets behind stellar activity (about 2/3 times the planetary signature level). Finally, as we saw for data sets #32, #36 (see Appendix B for the complete analysis of these two data sets) and to a lesser extent for #2 and #15, spurious periodicity signatures can appear with no relation to harmonics and no obvious relation with the window function, as suggested by Nava et al. (2020).

Regarding orbital phase, the significant difference in peak height between data sets #13 (periodogram peak power = 0.3, Φ = 0.4769) and #15 (periodogram peak power = 0.5, Φ = 0.9183) given their identical period and comparable semi-amplitude suggests that phase impacts the detection capabilities. It is not surprising that particular phases would have an impact on the periodogram as the irregular sampling can yield different phase coverage. That being said, we did not observe a general trend with phase across all data sets.

Finally, data obviously play a huge role in the detection capabilities with crucial aspects being quality, quantity, and sampling. To better characterize the activity, i.e. improve the hyperparameters of the GP, it is important to optimize the sampling (spanning multiple stellar orbits with as dense and as regular sampling as possible). Another successful strategy is to apply a GP to simultaneous photometric data, or at least not too far apart so that there is not too much evolution of the stellar surface features. We also tested (see Appendix D) the improvement brought by the knowledge of the period of the orbiting planet (i.e. characterizing a known transiting planet).

7 CONCLUSIONS

In this paper, we assessed our capabilities to detect exoplanets behind real stellar activity signatures. We used a previously published set of observations gathered with a non-stabilized spectrograph of the young, active G star HD 141943 in which we injected simulated planets. We then utilized two distinct strategies, DI and a GP regression to filter out the stellar activity variability, aiming to recover these injected planets.

Our dedicated treatment of stellar activity allowed significant improvement in the detection capabilities compared to J14, a planet search study done on the same star with no dedicated activity mitigation. As previously shown by Yu & Donati (2017a), these strategies are among the best tools we have to deal with large activity signatures. Although now widely accepted in the community, we further confirmed that dedicated treatment of the activity is crucial, and showed that we can detect short-orbit gas giants even in non-optimally sampled data sets exhibiting a 50–100 m s^{-1} RV precision.

We tested two alternatives to recover the RVs from the LSD line profiles (GND fit and first-order moment), which yielded slightly different RV time series but more importantly different uncertainties. We favoured the FOM approach but other methods such as broadened profiles could be explored.

With a low number of epochs acquired with a non-stabilized spectrograph, the combined use of both GP and ZDI methods enabled us to set a planet detection threshold of around 100 m s^{-1} or $\approx\frac{1}{4}$ of the activity level. Injected planets under this threshold were either non-detections or would require extra observations to confirm. The limitations we faced give a good idea of the upper limit we can hope to achieve for such systems in similar conditions.

Although DI shows less reliability than the GP, it allows us to strengthen the confidence of a finding. This lack of reliability could be explained by the fact that DI does not take into account surface variability due to the active regions' appearance/disappearance. These can evolve quickly, as shown for another G-type star in Petit et al. (2004). We suggest that claiming planets around active star should be done with a dedicated treatment for stellar variability, and preferentially using a Bayesian framework for robustness and to allow a quantification of the evidence of the presence of an orbiting planet.

We attempted to identify some factors that could improve the likelihood to find exoplanets orbiting young stars. Larger and more precise data set is an obvious one. Efficient sampling is also crucial, where dense sampling of the stellar rotation to better constrain the activity should be combined with coverage of multiple planetary orbits.

Orbital periods close to P_{eq} and its harmonics pose serious difficulties, and often lead to non-detections. In our case, it also appears to be a good rule of thumb to sample at least 2, or even 3, orbital periods to constrain P_{orb} with sufficient precision.

Some data sets (#2, #15, #32, and #36) exhibited significant spurious peaks of mysterious origin that compete with the true planetary period, emphasizing the difficulty of RV-only searches.

Detecting young exoplanets that do not transit is difficult but essential if we want to expand the sample of massive short-period exoplanets orbiting very young stars and progress towards settling the long-lasting debate over their origin. This work demonstrates that we can realistically identify potential candidates for follow-up observations and even detect short-orbit gas giant planets in non-optimized data sets exhibiting large activity variability.

8 FUTURE WORK

As follow-up of this work, and to improve precision, producing mean line profiles with either classical approaches (i.e. CCF, shift, and fit) or more recent proposals (Rajpaul, Aigrain & Buchhave 2020 or Collier Cameron et al. 2021), rather than with LSD for the GP analysis, could be explored. Indeed, the strength of LSD is the increase in S/N it provides, at the cost of a poorer estimation of the uncertainties (usually overestimated). It is more relevant to have a better estimation of the uncertainties when it comes to the RV used for the GP rather than a boosted S/N (required for DI).

Now having a better grasp on the capabilities of these activity mitigation strategies, it is possible to study real data of young solar-type stars. Many projects such as the BCool (Marsden et al. 2014) and TOUPIES[4] (Folsom et al. 2016, 2018) surveys, aimed at characterizing star using DI and ZDI, would be good starting points.

Among the overwhelming number of targets observed by the *TESS* mission (Ricker et al. 2015), many are young Solar analogues. Careful planning of follow-up and the availability of photometry for transiting planets would drastically increase the characterization capabilities; see Appendix D. In general, using complementary tools

---

[4]https://ipag.osug.fr/Anr_Toupies//

58

to diagnose the activity, such as activity indicators and photometry, is strongly recommended (e.g. Rajpaul et al. 2015; Jones et al. 2017; Oshagh et al. 2017; Kosiarek & Crossfield 2020).

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author. The full content of Table 4 is available as online material.

REFERENCES

Appenzeller I., Mundt R., 1989, A&AR, 1, 291
Astropy Collaboration, 2013, A&A, 558, A33
Bailer-Jones C., Rybizki J., Fouesneau M., Demleitner M., Andrae R., 2020, Gaia eDR3 Lite Distances Subset Cone Search. VO Resource Provided by the GAVO Data Center. http://dc.zah.uni-heidelberg.de/gedr3dist/q/cone/info, Accessed date: May 15th, 2021
Bailey E., Batygin K., 2018, ApJ, 866, L2
Baluev R. V., 2008, MNRAS, 385, 1279
Baruteau C. et al., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. Univ. Arizona Press, Tucson, AZ, p. 667
Batygin K., Bodenheimer P. H., Laughlin G. P., 2015, ApJ, 829, 1
Beaugé C., Nesvorný D., 2012, ApJ, 751, 119
Boley A. C., Contreras A. P. G., Gladman B., 2015, ApJ, 817, L17
Bouma L. G. et al., 2020, AJ, 160, 239
Brown S. F., Donati J. F., Rees D., Semel M., 1991, A&A, 250, 463
Buchner J. et al., 2014, A&A, 564, A125
Cabot S. H. C. et al., 2021, AJ, 161, 26
Cang T. Q. et al., 2020, A&A, 643, A39
Chatterjee S., Ford E. B., Matsumura S., Rasio F. A., 2008, ApJ, 686, 580
Collier Cameron A., 2021, MNRAS, 505, 1699
Damasso M. et al., 2020, A&A, 642, A133
David T. J. et al., 2016, Nature, 534, 658
David T. J., Petigura E. A., Luger R., Foreman-Mackey D., Livingston J. H., Mamajek E. E., Hillenbrand L. A., 2019, ApJ, 885, L12
Dawson R. I., Johnson J. A., 2018, ARA&A, 56, 175
Donati J. F. et al., 2006, MNRAS, 370, 629
Donati J. F. et al., 2014, MNRAS, 444, 3220
Donati J. F. et al., 2015, MNRAS, 453, 3706
Donati J. F. et al., 2016, Nature, 534, 662
Donati J. F. et al., 2017, MNRAS, 465, 3343
Donati J. F., Brown S. F., 1997, A&A, 326, 1135
Donati J. F., Semel M., Carter B. D., Rees D. E., Cameron A. C., 1997, MNRAS, 291, 658
Donati J.-F. et al., 2020, MNRAS, 491, 5660
Donati J.-F., Collier Cameron A., Petit P., 2003, MNRAS, 345, 1187
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Flagg L., Johns-Krull C. M., Nofi L., Llama J., Prato L., Sullivan K., Jaffe D. T., Mace G., 2019, ApJ, 878, L37
Folsom C. P. et al., 2016, MNRAS, 457, 580
Folsom C. P. et al., 2018, MNRAS, 474, 4956
Foreman-Mackey D., 2016, J. Open Source Softw., 1, 24
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Gaia Collaboration, 2021, A&A, 649, A1
Gaia Collaboration, 2016, A&A, 595, A1
Gaia Collaboration, 2018, A&A, 616, A1
Gregory P. C., 2007, MNRAS, 374, 1321
Hamers A. S., Antonini F., Lithwick Y., Perets H. B., Portegies Zwart S. F., 2017, MNRAS, 464, 688
Harris C. R. et al., 2020, Nature, 585, 357
Haywood R. D. et al., 2018, AJ, 155, 203
Haywood R. D., Cameron A. C., Queloz D., Others, 2014, MNRAS, 443, 2517
Hillenbrand L. A. et al., 2008, ApJ, 677, 630
Hobson M. P., Lasenby A. N., 1998, MNRAS, 298, 905
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jeffers S. V., Barnes J. R., Jones H. R., Reiners A., Pinfield D. J., Marsden S. C., 2014, MNRAS, 438, 2717
Jeffers S. V., Keller C. U., 2009, in Stempels E., ed., AIP Conf. Proc. Vol. 1094, 15th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun. Am. Inst. Phys., New York, p. 664
Jeffreys H., 1961, Theory of Probability/by Harold Jeffreys, 3rd edn. Clarendon Press, Oxford
Johns-Krull C. M. et al., 2016, ApJ, 826, 206
Jones D. E., Stenning D. C., Ford E. B., Wolpert R. L., Loredo T. J., Gilbertson C., Dumusque X., 2017, preprint (arXiv:1711.01318)
Klein B. et al., 2020, MNRAS, 493, L92
Kochukhov O., Makaganiuk V., Piskunov N., 2010, A&A, 524, 1
Kosiarek M. R., Crossfield I. J. M., 2020, AJ, 159, 271
Kupka F. G., 2000, Baltic Astronomy, 9, 590
Lomb N. R., 1976, Ap&SS, 39, 447
Marsden S. C. et al., 2011a, MNRAS, 413, 1922
Marsden S. C. et al., 2011b, MNRAS, 413, 1939
Marsden S. C. et al., 2014, MNRAS, 444, 3517
Nadarajah S., 2005, J. Appl. Stat., 32, 685
Nava C., López-Morales M., Haywood R. D., Giles H. A. C., 2020, AJ, 159, 23
Nelson B. E. et al., 2020, AJ, 159, 73
Newton E. R. et al., 2019, ApJ, 880, L17
Oshagh M. et al., 2017, A&A, 606, A107
Perryman M., 2011, The Exoplanet Handbook. Cambridge University Press, Cambridge
Petit P. et al., 2004, MNRAS, 351, 826
Petit P., Donati J. F., Collier Cameron A., 2002, MNRAS, 334, 374
Petrovich C., 2015, ApJ, 805, 75
Petrovich C., Tremaine S., 2016, ApJ, 829, 132
Plavchan P. et al., 2020, Nature, 582, 497
Rajpaul V. M., Aigrain S., Buchhave L. A., 2020, MNRAS, 27, 3960
Rajpaul V., Aigrain S., Osborne M. A., Reece S., Roberts S., 2015, MNRAS, 452, 2269
Rajpaul V., Aigrain S., Roberts S., 2016, MNRAS, 456, L6
Ricker G. R. et al., 2015, J. Astron. Telesc. Instrum. Syst., 1, 014003
Rizzuto A. C. et al., 2020, AJ, 160, 33
Robert C. P., Chopin N., Rousseau J., 2009, Stat. Sci., 24, 141

Scargle J. D., 1982, ApJ, 263, 835

Semel M., 1989, A&A, 225, 456

Skilling J., Bryan R. K., 1984, MNRAS, 211, 111

Soummer R. et al., 2014, ApJ, 786, 2

VanderPlas J. T., 2018, ApJS, 236, 16

Virtanen P. et al., 2020, Nat. Methods, 17, 261

Wright J. T., Marcy G. W., Howard A. W., Johnson J. A., Morton T. D.,
Fischer D. A., 2012, AJ, 753, 2

Yu L. et al., 2017b, MNRAS, 467, 1342

Yu L. et al., 2019, MNRAS, 489, 5556

Yu L. F., 2017a, in Reylé C., Di Matteo P., Herpin F., Lagadec E., Lançon A.,
Meliani Z., Royer F., eds, SF2A-2017: Proceedings of the Annual meeting
of the French Society of Astronomy and Astrophysics, held 4-7 July, 2017
in Paris, p. 69

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Table 4. Data sets 1 to 6. Each column represents a simulated data
set.

Please note: Oxford University Press is not responsible for the content
or functionality of any supporting materials supplied by the authors.
Any queries (other than missing material) should be directed to the
corresponding author for the article.

APPENDIX A: BAYESIAN MODEL SELECTION

To assess the significance of the presence of periodicity in our
data sets, we had to compare different models: activity only versus
activity + planet(s). From parameter estimation and for a given
model, an MCMC (nested sampling in our case) approach gives
access to the most likely set of parameters for the model along
with the associated maximum likelihood value. This non-normalized
likelihood is, however, not comparable across models. A model
with more parameters will have more flexibility and therefore the
capacity to better fit the data, yielding a greater likelihood value. To
circumvent this, one tries to penalize models with more parameters,
following Ockham's razor (law of parsimony). The notoriously hard
to compute marginal likelihood or evidence (or posterior distribution
normalization constant) naturally applies Ockham's razor and acts
to penalize models with higher number of parameters.

We encourage the reader to refer to Robert, Chopin & Rousseau
(2009) for a thorough description of this Bayesian framework, based
on the work of Jeffreys (1961). We will here give a quick overview.

First let's define the quantities required for this Bayesian frame-
work. With **y** an array containing the data points, θ the set of
parameters composing the model, and $\mathcal{M}_i$ the *i*th model, we define

(i) p($\theta|\mathbf{y}, \mathcal{M}_i$) or $\mathcal{P}(\theta)$: Posterior distribution of the parameters.
(ii) p($\mathbf{y}|\theta, \mathcal{M}_i$) or $\mathcal{L}(\theta)$: Likelihood.
(iii) p($\theta|\mathcal{M}_i$) or $\pi(\theta)$: Prior probability of the parameters.
(iv) p($\mathbf{y}|\mathcal{M}_i$) or $\mathcal{Z}$: Evidence/marginal likelihood/probability of
data given the model/posterior distribution normalization constant.
(v) p($\mathcal{M}_i$) or $\pi(\mathcal{M}_i$): Model's prior.

The first four are linked by Bayes' theorem:

$$\mathcal{P}(\theta) = \frac{\pi(\theta) \times \mathcal{L}(\theta)}{\mathcal{Z}}. \tag{A1}$$

To compare models, we look at the BF, defined for two models
$\mathcal{M}_0$ and $\mathcal{M}_1$ as

$$\text{BF} = \frac{\mathcal{Z}_1}{\mathcal{Z}_0} \frac{\pi(\mathcal{M}_1)}{\pi(\mathcal{M}_0)}. \tag{A2}$$

Table A1. Values of the BF, i.e. the ratio of marginal likelihood between the
single planet model and the 0-planet (stellar activity only) model. Middle
and left columns are the associated probability and level of confidence.

Bayes factor	Probability	Level of confidence
<1	<0.5	None
<3	<0.75	Not worth more than a bare mention
<10	<0.909	Substantial
<30	<0.967	Strong
<100	<0.99	Very strong
>100	>0.99	Decisive

In cases where nothing a priori favours one model over the other
(i.e. $\frac{\pi(\mathcal{M}_1)}{\pi(\mathcal{M}_0)} = 1$), we are left with the ratio of the marginal likelihoods.
To obtain the marginal likelihoods, we need to marginalize (i.e.
integrate) over all parameters:

$$\mathcal{Z} = \int \mathcal{L}(\theta)\pi(\theta)\mathrm{d}^N\theta, \tag{A3}$$

with N the number of parameters. For models with a large number
of parameters, accurate computation of $\mathcal{Z}$ requires integrating over
many dimensions (N) and is therefore often intractable. It quickly
becomes too computationally expensive and needs to be approxi-
mated. Various approaches are used in the literature; see Nelson et al.
(2020) for the most extensive attempt to compare these methods in
the context of exoplanet searches.

Once in possession of the evidence for each model, we can
compute the BF and assess whether the data favour model 1 over
model 0. It is common practice to work with the *natural* logarithm
of the evidence; we then have

$$\text{BF} = \exp(\ln \mathcal{Z}_1 - \ln \mathcal{Z}_0) = \exp(\Delta \ln \mathcal{Z}). \tag{A4}$$

The different level of confidence is then interpreted from the BF,
following Jeffreys (1961), as summarized in Table A1. We emphasize
that this level of evidence is assessing the significance of the model
given the data, and does not take into account how accurately the
data reflect what we wish to observe nor the plausibility of the model
on its own [although that can be added as $\pi(\mathcal{M}_i)$ in equation A2].

A1 Notes on prior probabilities

Prior probabilities are at the core of Bayesian inference, and express
knowledge previously acquired on a particular aspect of the problem,
i.e. a parameter of the model, or on the model itself. One must be
mindful of the choice for these priors. Some are 'uninformative'
(Uniform, Jeffrey's or Modified Jeffrey's priors; see Gregory 2007,
for more details on the last two), meaning that they do not contain
a priori information (they do to some extent as a uniform prior
has limits, but these are rather physical than inferred from previous
analyses). On the other hand, so-called 'informative' priors, typically
a Gaussian prior, can strongly constrain the parameter space to be
explored. This results in a boosted marginal likelihood compared
with the use of an 'uninformative prior'.

Therefore, one has to be extremely cautious when using 'informa-
tive' priors. The previous knowledge yielding that prior has to be sta-
tistically robust, to not mislead the analysis by boosting the evidence.

A2 Note on evidence versus planet detection

Accurate estimation of the evidence grants a statistically robust
framework to measure the significance of *one model relative to
others*. We should keep in mind that models could be wrong and

that drawn conclusions about physical phenomena are only justified as long as these models and their underlying assumptions are reasonable.

Here, our model assesses the likelihood of the presence of a periodic signature in a data set. It does not, however, inform us about the origin of such signatures. Assumptions on the nature of the underlying physical phenomenon or the accuracy of the data collection and reduction are required to go from 'a periodic signal in the data' to 'a planet orbiting the observed star'. For example, a previous exoplanet detection claim was later attributed to the window function by Rajpaul, Aigrain & Roberts (2016). It has been the case again very recently for V830 Tau b (Damasso et al. 2020), where the signature found by Donati et al. (2016, 2017) could not be found in a new data set. Whether this means the planet does or does not exist is a rather challenging question, but it once more highlights the difficulty of RV-only searches.

APPENDIX B: SPURIOUS PERIODOGRAM PEAKS

For the two following data sets:

(i) #32: $K = 149.72$ m s^{-1}, $P_{orb} = 0.5526$ d, and $\Phi = 0.5701$
(ii) #36: $K = 142.36$ m s^{-1}, $P_{orb} = 0.8463$ d, and $\Phi = 0.6225$

the periodogram resulting from the DI analysis showed, in each case and after the activity filtering, the apparition of a spurious peak at periods seemingly unrelated to the stellar rotation harmonics, the orbital period harmonics, or a peak in the window function (see Fig. B1). As shown in Fig. B2, $P_{spurious32} \approx 1.22$ d and $P_{spurious36} \approx 5.7$ d. After removing the identified highest peak from the filtered data (matching with the true period in the case of #36 but not for #32), the competing peak was also filtered out, suggesting an effect of the uneven sampling of the observations.

Figs B3, B4, and B5 show the corner plot of the posterior distribution and the resulting fit to the data for these two data sets following the GP analysis.

For data set #32 (Fig. B3 and top plot in Fig. B5), the GP does not seem to be fooled and clearly identifies the right period. However, the evidence is surprisingly low for such a large planet (BF = 1.67 or $p = 0.62$ in favour of the single model), yielding a non-detection. We strongly suspect that the ambiguity in both methods is due to the orbital period of the injected planet being very close to the third harmonic of the stellar period ($\frac{P_{eq}}{3} = 0.547$ d and $P_{orb} = 0.553$ d).

Regarding data set #36 (Fig. B4 and bottom plot in Fig. B5), once again the correct period is identified by the GP. We also note a slight appearance of the conflicting period around 6 d on the posterior distribution (Fig. B4). This time, however, the evidence is strongly in favour of a detection with a BF of 175.9 ($p = 0.994$).

Although it is not clear where the spurious peak arises from, one of these two cases could be settled by the GP.
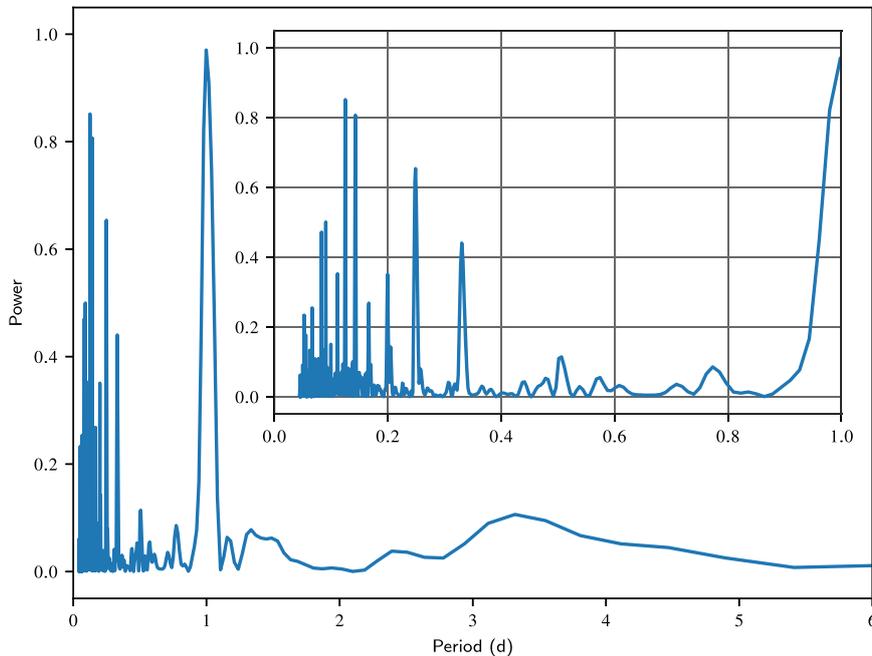
Figure B1. Window function of the observed RVs. The horizontal axis shows the period in days and the vertical axis the periodogram power. The upper right zoomed window shows extra details for periods <1 d.
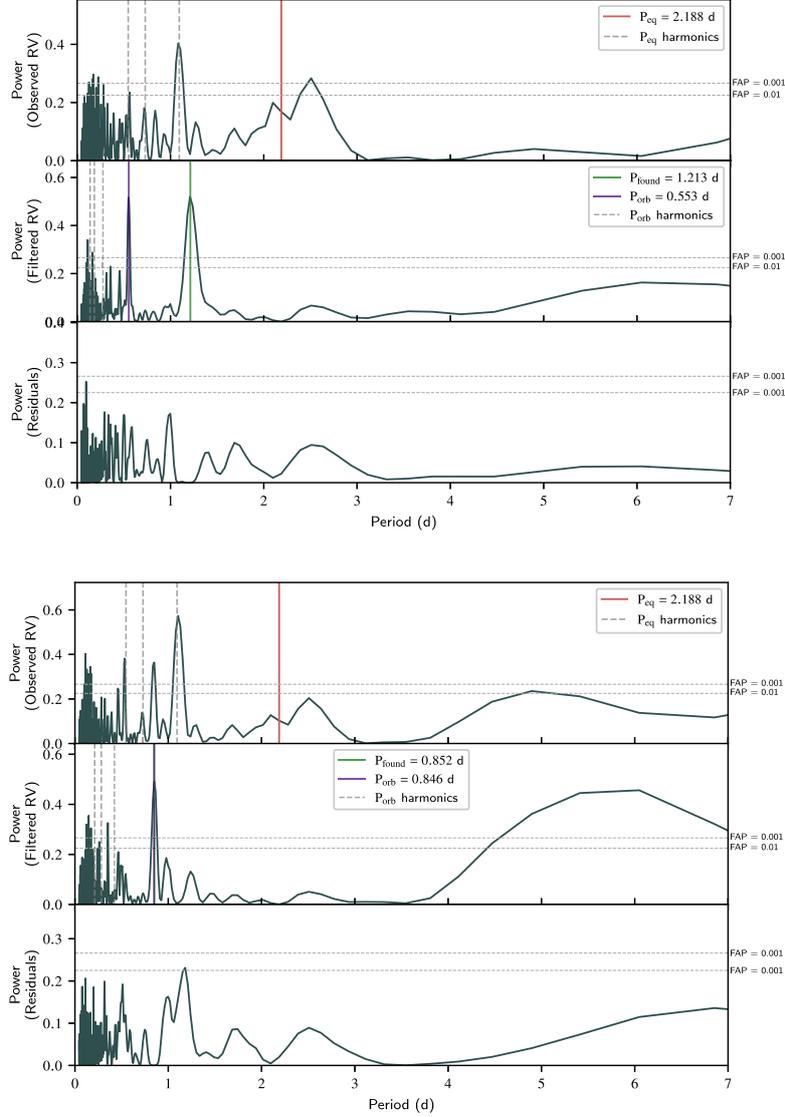
Figure B2. Periodograms of data sets #32 (top) and #36 (bottom). Each plot has three subplots showing the periodograms of the raw observed data (top), DI filtered (middle), i.e. observed – synthetic RVs and residuals (bottom), i.e. filtered – identified periodic signature. Vertical lines on the top subplots mark the stellar rotation period and its harmonics. In the middle subplots, we showed the recovered periodicity (green vertical line) and the true injected period (purple vertical line), along with its harmonics (grey dashed vertical lines).

62
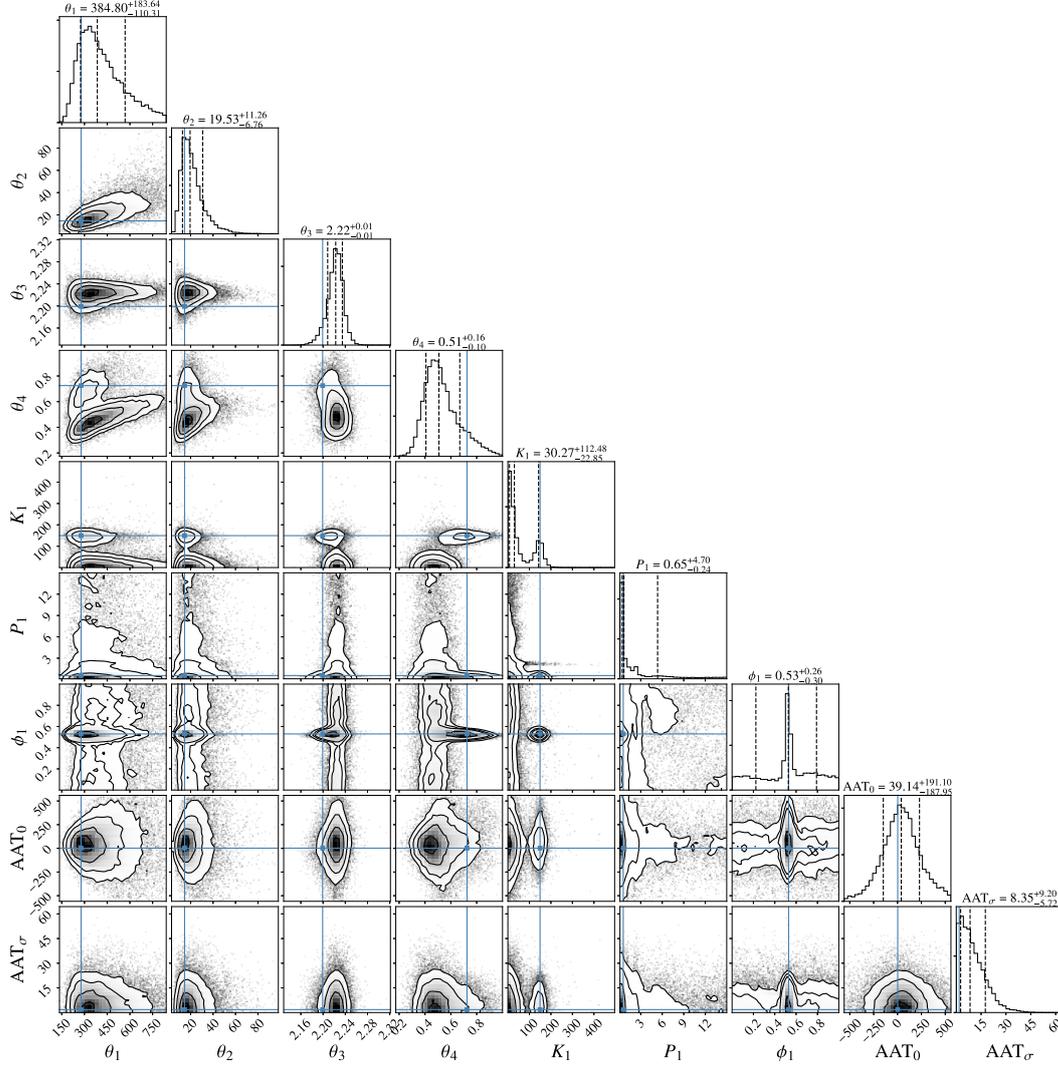
Figure B3. Posterior distribution of the parameters from the GP analysis of data set #32. Prior weighted posterior samples drawn from the PYMULTINEST analysis. Blue lines show the most likely parameters. Dashed vertical lines are 0.16, 0.5, and 0.84 quantiles. Contours are 1σ, 2σ, and 3σ levels (representing, respectively, 39, 86, and 99 per cent for a 2D distribution).

63
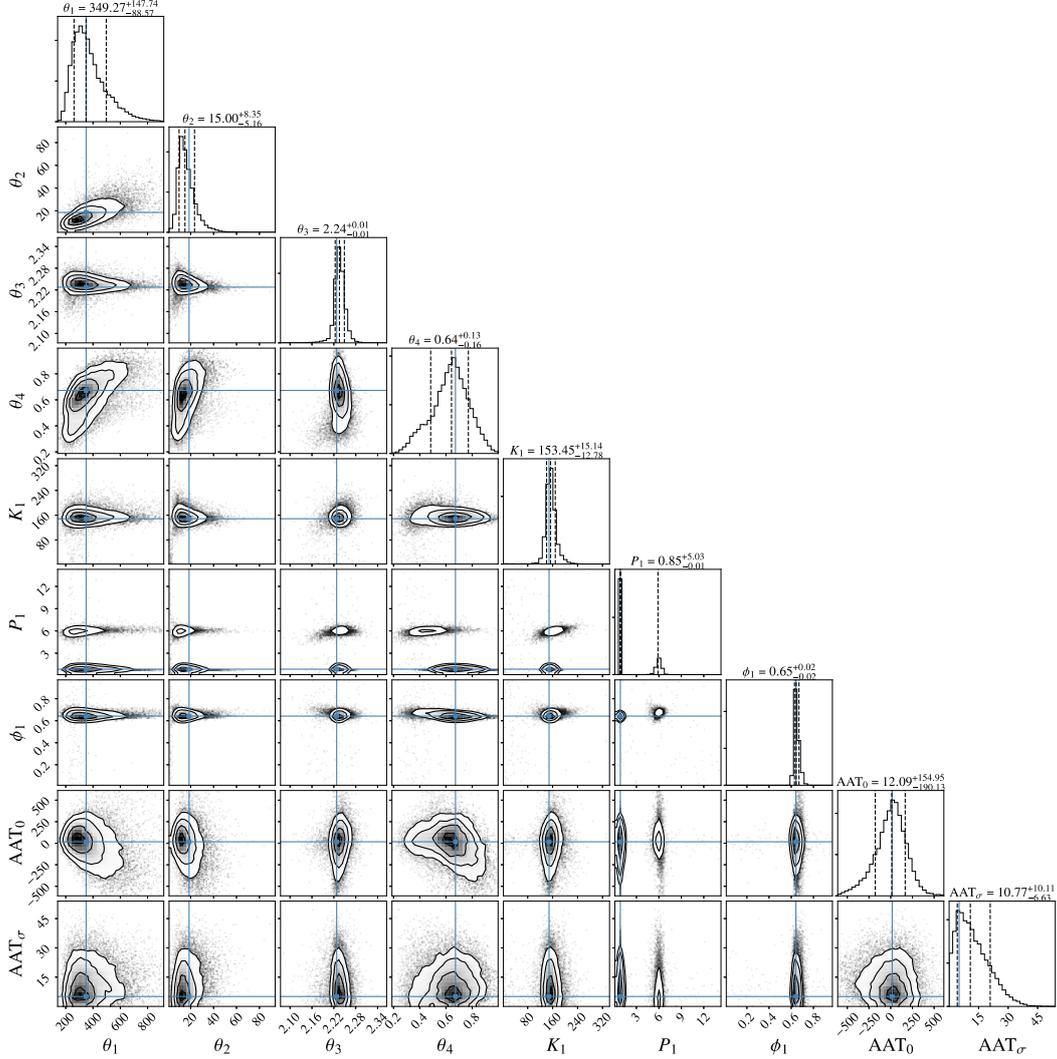
Figure B4. Posterior distribution of the GP analysis of data set #36. Prior weighted posterior samples drawn from the PYMULTINEST analysis. Blue lines show the most likely parameters. Dashed vertical lines are 0.16, 0.5, and 0.84 quantiles. Contours are 1σ, 2σ, and 3σ levels (representing, respectively, 39, 86, and 99 per cent for a 2D distribution).
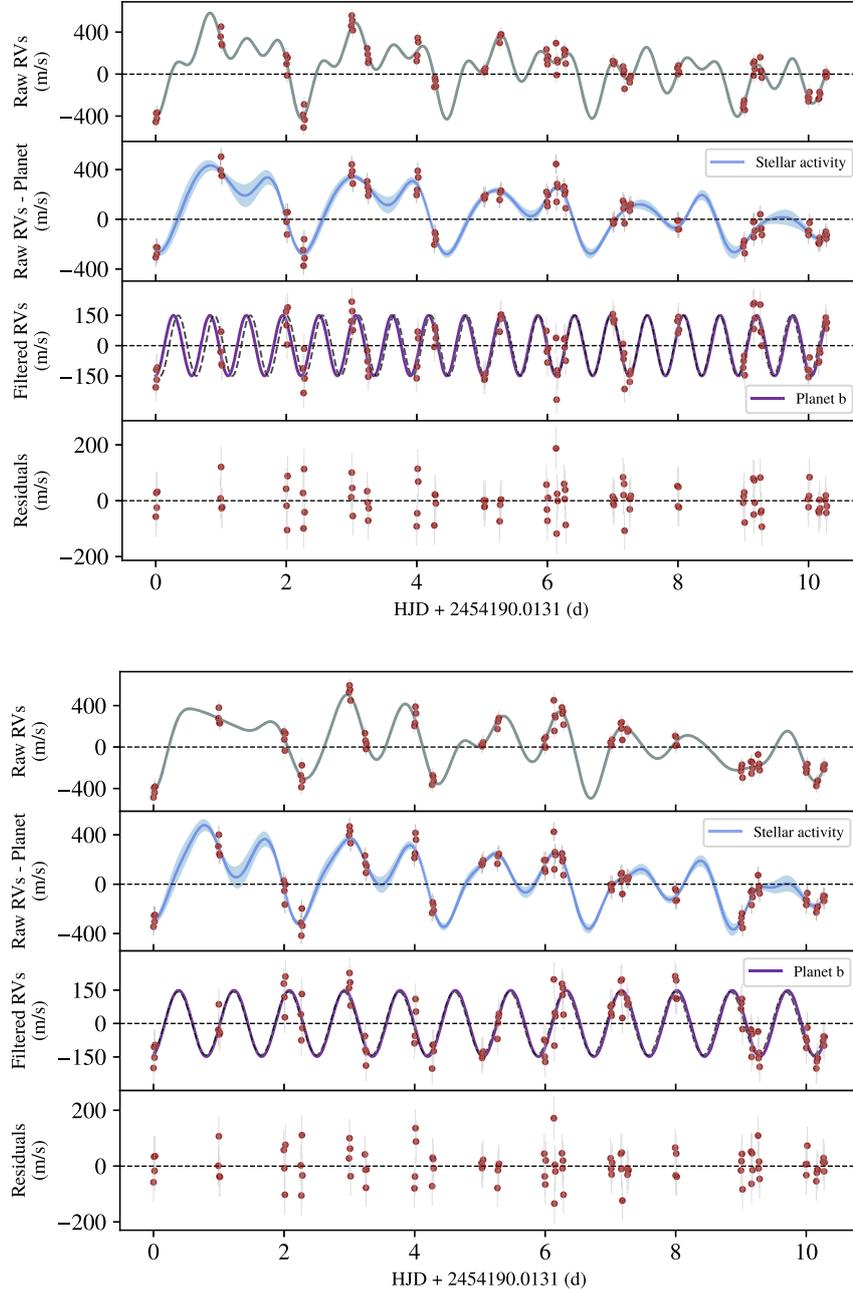
Figure B5. Resulting fits for the GP analysis of data sets #32 (top) and #36 (bottom). Top panel is the total model (grey line) and the observed raw RVs (red points). The second panel is the GP fit with the uncertainty shown by the shadowed area (analytically computed predictive standard deviation of the GP); red points are (raw RVs) – (recovered planet). The third panel is the recovered planet (purple line) on top of the true injected planet (dashed curve). Red points are (raw RVs) – (GP fit). The bottom panel is the residuals, i.e. (raw RVs) – (GP model) – (recovered planet).

65

APPENDIX C: IMPACT OF DIFFERENT DI SOLUTIONS ON DATA SET #22

As we saw in Section 4.2, the stellar parameters found via DI are slightly different from M11A/B. Because the DI filtering uses the synthetic generated line profiles that depend on the stellar parameters, we tested the influence of three different DI solutions on our planet detection capabilities. We re-analysed data set 22, containing an injected planet with $K = 117.28$ m s^{-1}, $P_{orb} = 3.1546$ d, and $\Phi = 0.7077$.

As described in Section 4.2, we derived the alternate DI solutions by fitting (i) only using dark features and (ii) using both dark and bright features but forcing the inclination parameter to 70° (matching M11A/B's value). The optimum parameters are displayed in Table C1.

Regarding the maps, results are similar to data set #5 (no planet), with mainly a difference in contrast. The periodograms of the filtered RVs (synthetic RVs – observed RVs) for each of the three cases are displayed in Fig. C1. In all cases, the features in the periodogram are quite similar. The analysis with dark spots only is the most different. It appears, in that case, that the original peak due to activity was not completely filtered and blends with the peak of the injected period after the filtering process. This acts to slightly hinder the accuracy of the period retrieved.

The fixed inclination case performs slightly better than the dark and bright case, as the retrieved period is 0.03 d away from the injected period for the former versus 0.08 d for the latter. However, the shape of the peak is very similar and we cannot conclude whether the better result is in fact due to a better solution for the DI or not.

Table C1. Test of different DI solutions fitting for dark and bright spots (top row), dark and bright features with the inclination parameter forced to 70° (middle row), and only dark features (bottom row).

Fitting for	$v \sin i$	P_{eq}	dΩ	rad d^{-1}
Dark and bright	35.54	2.2003	0.1231	47
Dark and bright (fixed i)	35.58	2.1982	0.1138	70 (fixed)
Dark	35.38	2.2115	0.0441	44

66
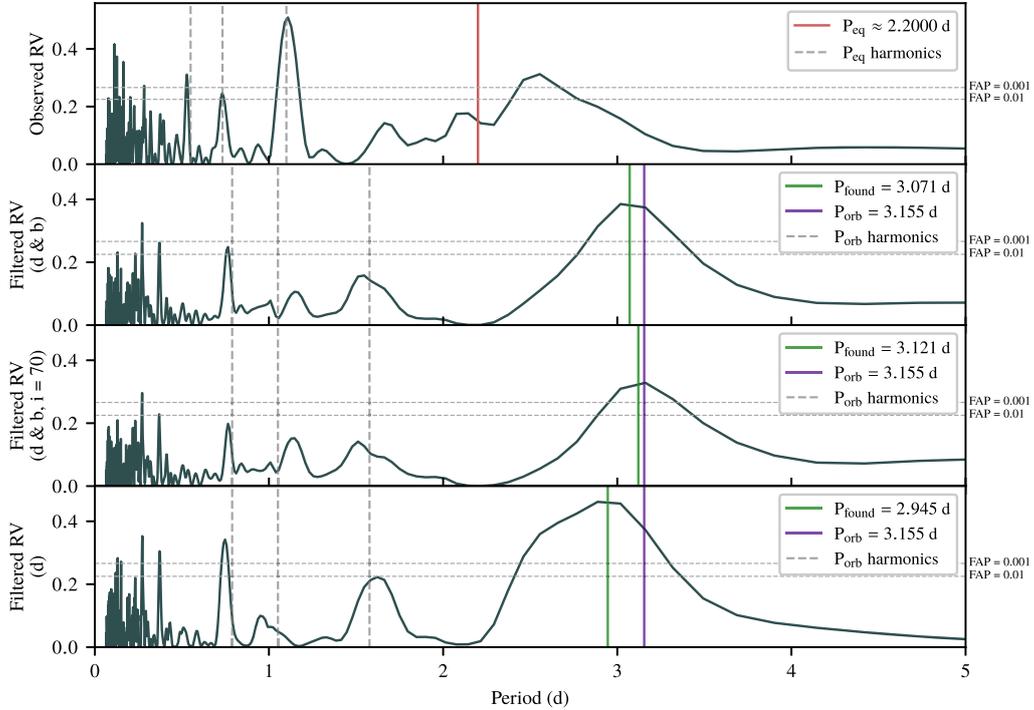
Figure C1. Top: periodogram of the raw data from data set #22. Second to fourth plot: comparison of the periodograms for data set #22 obtained from dark and bright-spot DI analysis (2nd, d and b), from dark and bright-spot DI analysis with imposed inclination (3rd, d and b, $i = 70$), and from dark spots-only DI analysis (4th, d).

APPENDIX D: SIMULATING A TRANSITING PLANET

In order to showcase what access to a photometrically detected planet (i.e. the case of a transiting planet) would add to the retrieval capabilities, we investigated a data set allowing for additional prior constraints. For this, we chose data set #29, which exhibits the smallest semi-amplitude among the data sets that are not close to either P_{eq} or its harmonics. It has a semi-amplitude of 47 m s^{-1}, well below our detection threshold. In Table D1, we compared the pieces of evidence (derived from the GP analysis) for three different cases.

The first case (first line of Table D1) uses the same priors as we did throughout this article, simulating access solely to RV data. The second case (second line of Table D1) simulates the availability of transiting data on the planet. We set Gaussian priors for P_{orb} and

Φ, centred on the true injected value and with a standard deviation of, respectively, 0.0001 d and 0.01. For the third case (third line of Table D1), we fixed the value for P_{orb} and Φ to their true values (simulating the best transit value).

For each case, the BF and probability favouring the single planet model over the activity-only model can be found in the last two columns of Table D1. For the first case, the BF is extremely low (0.4, $p = 0.27$) and leads to a non-detection (see grey cross labelled '29' on Fig. 7). For cases 2 and 3, however (lines 2 and 3 of Table D1), their BFs are comparable and around 9, meaning a 0.9 probability in favour of the one-planet model and therefore a strong evidence for the presence of a planet.

This is expected as our constrained priors act to boost the evidence. It also shows how an inappropriate choice of priors can influence the evidence and bias the claim of a finding, following our discussion in A1. We see once more the difficulty of RV-only searches.

67

Table D1. GP analysis of data set #29 with three different sets of priors. The first column shows the real case configuration leading to the according choice of priors. The second column is the list of parameters with a Gaussian prior. The third column is the list of fixed parameters, i.e. not taking part in the parameter space search process. The last two columns are the BF and the probability favouring the single planet model over the activity-only model resulting from the corresponding analysis. To constrain or fix P_{orb} and Φ, we used the true injected values as they would be available if transits would have been identified.

Analysis	Constrained (i.e. Gaussian prior)	Fixed (not fitted)	BF	Probability
RV only	θ_3	None	0.4	0.27
Transit	$\theta_3, P_{\text{orb}}, \Phi$	None	9.9	0.91
Transit (fixed)	θ_3	P_{orb}, Φ	8.2	0.89

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

68

Chapter 4: The obliquity of HIP 67522 b: a 17 Myr old transiting hot Jupiter-sized planet.

In 2020, Rizzuto et al. (2020) discovered a 17 Myr old transiting hot Jupiter-sized planet in the TESS data, HIP 67522 b. With a 7 day orbit around a bright (V_{mag} = 9.88), fast rotating ($v\sin i \sim 50$ km s^{-1}) host star, it is a prime target for an obliquity measurement. As TESS re-observed HIP 67522 b in May 2021, we simultaneously obtained RVs with the CHIRON spectrograph. In the data, we saw clear signs of the stellar activity that needed to be mitigated in order to confidently recover the spin–orbit alignment. I built from scratch a stellar model with spots on its surface to simulate the star's impact on the data (code shown in appendix C). This was incorporated into the global model comprising stellar activity, photometry and spectroscopy.

Our global model allowed us to unambiguously and precisely recover a spin–orbit angle of $|\lambda|= 5.8°$ $^{+2.8}_{-5.7}$ for HIP 67522 b. From a rough estimate of the stellar inclination $i_\star$, we derived a true obliquity of $\Psi \sim 20.2°$ $^{+10.3}_{-8.7}$, consistent with an aligned orbit. This points towards a smooth mechanism to explain the origin of this planet such as in-situ formation or gas disc mi-

gration. HIP 67522 b is one of only a handful of systems under 150 Myr old with a well constrained age to have an obliquity measurement. Interestingly, all these young planets show well-aligned orbits, and although it is too early to reach any conclusions, additional such measurement will prove to be crucial in our understanding of the early stages of the life of giant planets. HIP 67522 b still has not delivered all its secrets as its mass is currently unknown, and would help to determine if it is more Neptune- or Jupiter-like, which could be used as an input in radius evolution models.

The Obliquity of HIP 67522 b: A 17 Myr Old Transiting Hot, Jupiter-sized Planet

Alexis Heitzmann[1] , George Zhou[1] , Samuel N. Quinn[2] , Stephen C. Marsden[1] , Duncan Wright[1] , Pascal Petit[3] ,
Andrew M. Vanderburg[4] , Luke G. Bouma[5] , Andrew W. Mann[6] , and Aaron C. Rizzuto[7]
[1] Centre for Astrophysics, University of Southern Queensland, West Street, Toowoomba, QLD 4350 Australia; alexis.heitzmann@usq.edu.au
[2] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[3] Institut de Recherche en Astrophysique et Planétologie, Université de Toulouse, CNRS, CNES, 14 avenue Edouard Belin, F-31400 Toulouse, France
[4] Department of Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA
02139, USA
[5] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
[6] Department of Physics and Astronomy, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
[7] Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA

Abstract

HIP 67522 b is a 17 Myr old, close-in ($P_{orb} = 6.96$ days), Jupiter-sized ($R = 10 R_{\oplus}$) transiting planet orbiting a
Sun-like star in the Sco–Cen OB association. We present our measurement of the system's projected orbital
obliquity via two spectroscopic transit observations using the CHIRON spectroscopic facility. We present a global
model that accounts for large surface brightness features typical of such young stars during spectroscopic transit
observations. With a value of $|\lambda| = 5.8^{+2.8}_{-5.7}°$ it is unlikely that this well-aligned system is the result of a high-
eccentricity-driven migration history. By being the youngest planet with a known obliquity, HIP 67522 b holds a
special place in contributing to our understanding of giant planet formation and evolution. Our analysis shows the
feasibility of such measurements for young and very active stars.

1. Introduction

One of the oldest puzzles in the field of exoplanets is the
origin of short-orbit gas giants. With no equivalent in the solar
system, it is far from obvious how these giant planets, orbiting
their host star incredibly close (with orbital periods <10 days),
come to exist.

Among the observational properties easily measurable for
these exoplanets when they transit, their sky-projected stellar
obliquity (λ) angle may help differentiate between the multiple
pathways explaining their origin (Dawson & Johnson 2018).
Unfortunately, on its own, the obliquity angle cannot
unambiguously identify a specific formation pathway. Star–
planet tidal interactions, resulting in angular momentum
exchanges between a host star and its planet, have the ability
to circularize and shrink planetary orbits but also alter the star's
rotation axis alignment. This can erase primordial orbital
characteristics (i.e., misalignments and/or eccentricity), pre-
venting the identification of a specific migration channel.
Because the timescales of such interactions span many orders
of magnitudes (from 10^5 to 10^9 yr), the $\sim 150^{[8]}$ sky-projected
orbital obliquity measurements obtained to date remain difficult
to interpret (Triaud 2018; Albrecht et al. 2021).

Recently, it became possible to compare the obliquity
distribution for very young stars (<100 Myr) against that of
more mature stars. These young planetary systems have not yet
been influenced by star–planet tidal effects, and provide a
glimpse into the primordial orbits of planets post-formation.

The recent efforts to characterize young planets discovered by
the K2 and Transiting Exoplanet Survey Satellite (TESS;
Ricker et al. 2016) missions, such as AU Mic b (Addison et al.
2021; Hirano et al. 2020; Martioli et al. 2020; Palle et al. 2020),
V1298 Tau c (David et al. 2019; Feinstein et al. 2021), DS Tuc
Ab (Montet et al. 2020; Zhou et al. 2020), and TOI 942 b
(Wirth et al. 2021; Zhou et al. 2021), have the potential to
deliver key insights on the formation and migration of close-in
planets. Although these transiting planets orbiting rapidly
rotating young stars are suitable candidates for obliquity
measurements, their host stars' young age imply that strong
intrinsic variability needs to be dealt with in order to recover
the true spin–orbit angles.

In this Letter we present a projected obliquity measurement
for HIP 67522 b. This obliquity measurement is the first for a
hot, Jupiter-sized planet younger than 100 Myr. With an age of
only 17 Myr, HIP 67522 b has a radius of $10.178 \pm 0.440 R_{\oplus}$
(this work), a mass $< 5 M_J$ (Rizzuto et al. 2020), and is orbiting
a bright ($V_{mag} = 9.876 \pm 0.026$) Sun-like star ($T_{eff} = 5675 \pm 75$
K).

We describe both photometric and spectroscopic observa-
tions of HIP 67522 b's transits in Section 2 and present our
combined model used to determine the projected obliquity of
the system in Section 3. Finally, in Section 4, we place this
measurement into context of other planetary systems around
mature main-sequence stars.

2. Observations

2.1. TESS: Photometry

HIP 67522 b was first identified to transit in Sector 11 of the
TESS primary mission (Rizzuto et al. 2020) over the period of
2019 April 22–May 21. The target was subsequently observed
by TESS again during Sector 38 of the extended mission, over

---

[8] TEPCat 2021 July (Southworth 2011).

1
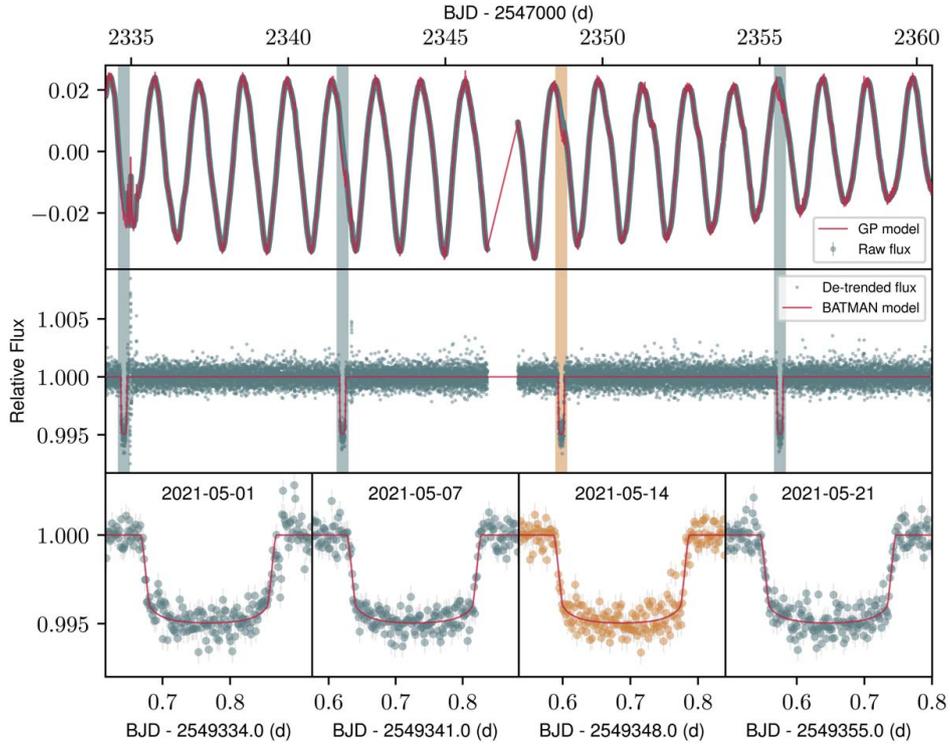
Figure 1. Top: TESS light curve for Sector 38. Data points are shown as gray points. The red line demonstrates the Gaussian process model describing the stellar variability of the TESS light curves. Locations of individual transit events are highlighted by the shaded regions. Middle: the TESS Sector 38 light curve after the Gaussian process model has been removed. Bottom: close-ups of each transit event. The transit event on 2021 May 14 was simultaneously observed by CHIRON, and is marked in orange for clarity.

the period of 2021 April 28–May 26. We make use of the Simple Aperture Photometry (Twicken et al. 2010, 2018; Morris et al. 2020) made available for the target star extracted by the Science Processing Observation Center (SPOC; Jenkins et al. 2016) from the target pixel files, obtained at a 2 minute cadence from both sectors of observations. The light curves were detrended against spacecraft motion via the quaternion detrending technique as per Vanderburg et al. (2019), allowing the recovery of two missing transits in the Sector 11 observations as shown in the discovery paper. Except for these two transits, the light curve was also decorrelated against the PDC band 3 (fast timescale) cotrending basis vectors, and modeled the rotation signal with a basis spline with breakpoints every 0.1 days while excluding points with transits from the systematics correction. This allowed the recovery of a systematic corrected light curve from Sector 38. The light curve from Sector 38 is shown in Figure 1.

2.2. CHIRON/SMARTS: Spectroscopy

We obtained two spectroscopic transits of HIP 67522 b with the CHIRON facility (Tokovinin et al. 2013). CHIRON is a high-resolution echelle spectrograph on the 1.5 m Small and Moderate Aperture Research Telescope System (SMARTS) telescope, located at Cerro Tololo Inter-American Observatory, Chile. CHIRON is fed via a fiber bundle, and has a spectral

resolving power of $\lambda/\Delta\lambda \equiv R \approx 80{,}000$ with a wavelength coverage from 4100 to 8700 Å.

A total of 24 observations, with exposure times of 1200 s each, were obtained on 2021 May 14, capturing the full transit and baselines on pre-ingress and post-egress from 2021 May 14 00:00 to 08:00 UTC. An additional partial transit was observed from 2021 June 17 23:15 to 2021 June 18 05:20 UTC, with a total of 18 observations at 1200 s integration time each. The stellar spectra were extracted via the official CHIRON pipeline (Paredes et al. 2021), with wavelength calibration provided by a set of thorium–argon cathode ray lamp exposures that bracket each transit sequence. To derive line-broadening profiles from each spectrum, we perform a least-squares deconvolution of the observed spectrum against a set of synthetic nonrotating spectral templates (Donati et al. 1997; Collier Cameron et al. 2010), making use of the ATLAS9 atmosphere models (Castelli & Kurucz 2003) computed at the atmospheric parameters of HIP 67522 ($T_{eff} = 5725$ K, $\log g = 4.0$, and $[Fe/H] = 0$). To ensure realistic uncertainties on our measurement of λ, we rebinned these line profiles to the velocity dispersion corresponding to the detector's pixel size. We make use of these for the transit spectroscopic modeling as described in Section 3.2. In addition, we also model each line profile with a kernel that incorporates the effects of rotational, macro-turbulent, and instrumental line broadening. We make use of this model to determine the rotational broadening velocity

$v \sin i_\star$ necessary for the line profile modeling, measuring a rotational broadening velocity of $v \sin i_\star = 50 \pm 3$ km s^{-1}, consistent with that of 54.2 ± 0.7 km s^{-1} reported in the discovery paper from an ensemble of spectroscopic observations.

3. Analysis

The very young age of HIP 67522 goes hand in hand with substantial intrinsic stellar variability in both spectroscopic and photometric observations, with 2%–3% variations seen in both TESS Sector 11 and 38 light curves. The variability seen for HIP 67522 can be mostly attributed to surface brightness features (spots and plages). We develop a model below that incorporates both photometric (Section 3.1) and spectroscopic (Section 3.2) transits as well as spot modeling to deal with the influence of stellar activity on our projected obliquity measurements.

3.1. Transit Photometry

A total of eight transits, four in Sector 11 and four in Sector 38, were observed by TESS. The TESS photometry exhibits significant stellar rotational variability due to the youth of the host star. We apply a Gaussian process model to account for this variability. We model the light curve using a stochastically driven simple harmonic oscillation kernel as is implemented in the CELERITE package (Foreman-Mackey et al. 2017), with free parameters describing its amplitude S_0 and damping coefficient Q. We fixed the frequency $\omega_0 = 1/P_{\rm rot} = 1/1.39$ (see the last paragraph of Section 3.2 for justification). We also simultaneously model the planetary transit using the BATMAN package (Kreidberg 2015), including free parameters describing the transit centroid timing T_c, orbital period $P_{\rm orb}$, radius ratio $R_{\rm p}/R_\star$, normalized semimajor axis $a/R_\star$, and the line-of-sight inclination of the transit i. Following the low eccentricity found in Rizzuto et al. (2020), the orbit is assumed to be circular for this model.

The light-curve and best-fit model for the Sector 38 observations are shown in Figure 1. We subtract the Gaussian process model from the TESS observations, and pass the resulting detrended light curve to the subsequent analysis described in Section 3.3.

3.2. Transit Spectroscopy

Any phenomenon introducing brightness variations on the stellar disk that are asymmetric will leave rotationally modulated imprints on the observed spectroscopic line profiles. As these variations move across the stellar disk, they block/ enhance incoming flux at a wavelength (and corresponding radial velocity) depending on their longitude. This is because light emitted by different parts of the stellar disk experience a different Doppler shift due to the star's rotation. This directly translates into bumps and dips on the rotational line profiles. Doppler tomography consists of monitoring the evolution of line profiles during a planetary transit to catch the distortions induced by the eclipsing body, called the "Doppler shadow". The way this Doppler shadow crosses the line profiles over time yields the angle at which the planet crosses the stellar disk, the projected spin–orbit angle λ. Any other brightness variations (i.e., spot-induced features in a first approximation) will contribute to distort the line profiles in a similar fashion.

Table 1
Global Model Parameters

Parameters	Prior[a]	Results
Transit		
T_c (BJD$_{\rm TDB}$)	$\mathcal{U}$	$1604.023722^{+0.00024}_{-0.00023}$
$P_{\rm orb}$ (days)	$\mathcal{U}$	$6.959471^{+0.0000030}_{-0.0000030}$
$R_{\rm p}$ (R$_\star$)	$\mathcal{U}$	$0.067471^{+0.00019}_{-0.00017}$
a (R$_\star$)	$\mathcal{U}$	$11.685^{+0.112}_{-0.220}$
i ({°})	$\mathcal{U}[86,90]$	$89.23^{+0.37}_{-0.47}$
e	0 (fixed)	...
μ_1	0.148 (fixed)[b]	...
μ_2	0.23 (fixed)[b]	...
Doppler Tomography		
$v \sin i_\star$ (km s^{-1})	$\mathcal{G}[50,1]$	$49.21^{+0.95}_{-0.97}$
λ (°)	$\mathcal{U}$	$-5.8^{+2.8}_{-5.7}$
$v_{\rm macro}$ (km s^{-1})	$\mathcal{U}$	$0.59^{+0.43}_{-0.41}$
ν_1	0.4139 (fixed)[c]	...
ν_2	0.2494 (fixed)[c]	...
#$_{\rm spots}$	1 (fixed)	...
$\theta_{\rm spot1}$ (°)	$\mathcal{U}$	$212.95^{+0.36}_{-0.37}$
$\phi_{\rm spot1}$ (°)	$\mathcal{U}[0,90]$	$57.8^{+2.7}_{-2.2}$
$R_{\rm spot1}$ (°)	$\mathcal{U}$	$4.76^{+0.24}_{-0.71}$
$T_{\rm spot1}$ (K)	$\mathcal{U}[0.6 \times {\rm Teff}, 0.8 \times {\rm Teff}]$	3890^{+288}_{-1550}
$P_{\rm rot}$(days)	1.39 (fixed)	...
Light-curve GP		
$\log S_0$	$\mathcal{U}$	$-1.6^{+0.4}_{-3.7}$
$\log Q$	$\mathcal{U}$	$3.37^{+0.67}_{-0.77}$
$\log \omega_0$	-0.329 (fixed)	...
Derived Parameters		
R_P (R$_\oplus$)	...	10.178 ± 0.440
R_P (R$_J$)	...	0.928 ± 0.040
3D obliquity (°)	...	$20.2^{+10.3}_{-8.7}$
$i_\star$ (°)	...	$>85\ (3\sigma)$

Notes.
[a] $\mathcal{U}$ unconstrained uniform priors; $\mathcal{U}$ [a, b] constrained uniform priors with boundaries a and b; $\mathcal{G}[\mu, \sigma]$ Gaussian priors.
[b] Adopted at the TESS band from Claret (2017).
[c] Adopted at the V band from Claret et al. (2012).

We therefore modeled both contribution (spots and planet) to fit the observed line profiles and recover λ.

The star was modeled as a circular mask on a grid of pixels, each with a value between 0 and 1 representing fractional brightness $f_{\rm b}$. The mask is a combination of (i) a uniform disk ($f_{\rm b} = 1$), (ii) a limb-darkening quadratic law parameterized with a linear ν_1 and a quadratic ν_2 coefficient (given in Table 1), (iii) the spot(s), and (iv) the transiting planet. Spots were modeled as spherical caps defined by their colatitude (complementary angle of latitude, i.e., 0° at the north pole and 180° at the south pole) ϕ, longitude θ, angular radius $R_{\rm spot}$, and temperature $T_{\rm spot}$ conditioning their brightness, as per

$$f_{\rm b} = (T_{\rm spot}/T_{\rm eff})^4. \qquad (1)$$

The projection of each spot on the stellar disk was then computed analytically and its contribution was added to the total mask. Finally, the transiting planet was assumed on a circular orbit and modeled as a disk ($f_b = 0$). Its shadow on the stellar disk was parameterized by T_c, $P_{\rm orb}$, $R_{\rm p}/R_\star$, $a/R_\star$, ω, i, and λ (see the values in Table 1).

3
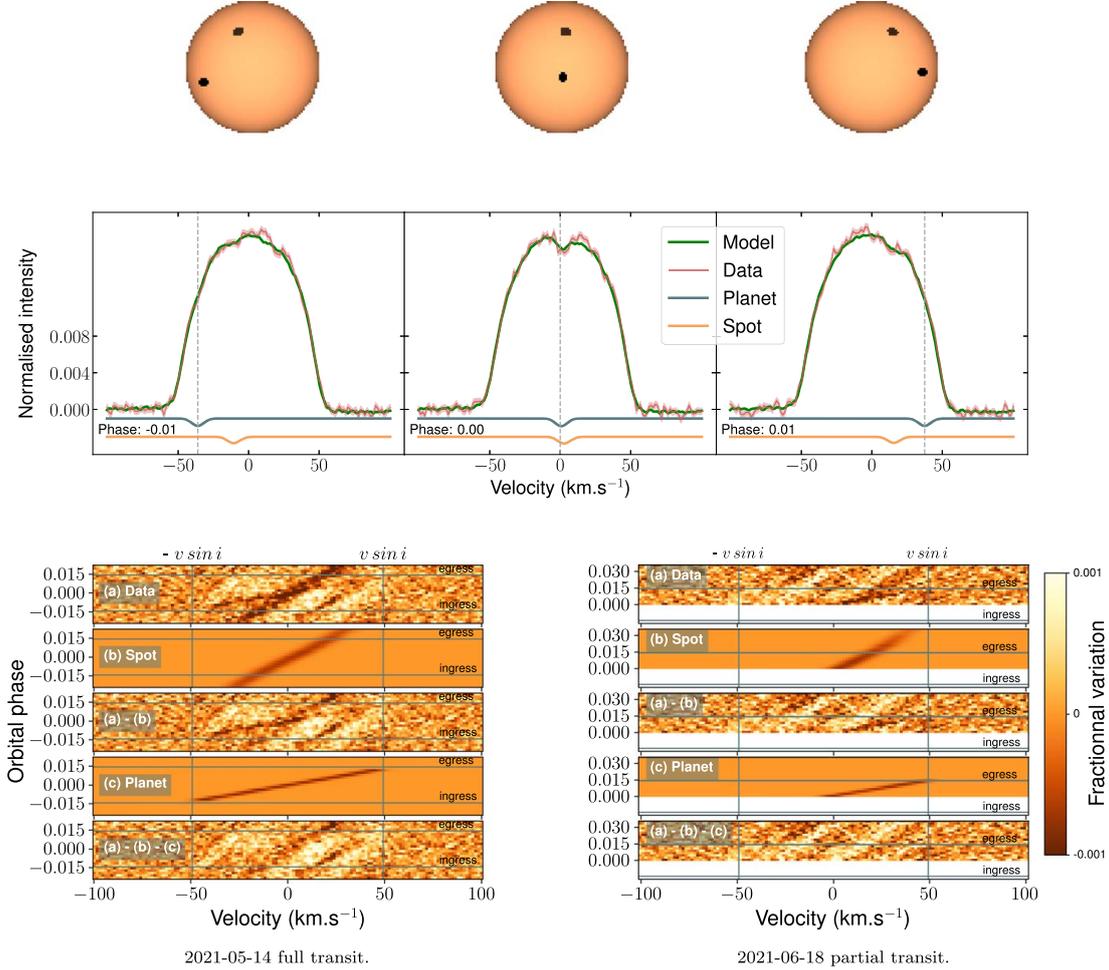
Figure 2. Results from the spectroscopic transit modeling of HIP 67522 b. Top: modeled stellar disk including a single spot and the transiting planet at phase -0.01 (left), 0 (center), and 0.01 (right). Middle: line profiles fit matching the above images. Observed line profiles are shown in red with uncertainties in lighter red. Fit to the data (mean line profile + integrated modeled stellar disk) is shown in green, the planet contribution in blue, and the spot contribution in orange. Gray vertical dotted lines express planet transit phases. Bottom: result plots for the 2021 May 14 full transit (left) and the 2021 June 18 partial transit (right). Each subpanel shows the variation of lines profiles in velocity space (horizontal axis) for different orbital phases of HIP 67522 b (vertical axis). (a) Residuals between each observed line broadening and the mean observed line profile. (b) Spot-only model. (a)–(b) Line profile residuals after subtraction of the spot model; we see that the spot signature has been removed, leaving a clear planetary signature. (c) Modeled planet. (a)–(b)–(c) Residuals after subtraction of both the spot and planet model. Horizontal black lines show the phase of transit ingress (bottom) and egress (top). Vertical black lines are $-v \sin i$ (left) and $+v \sin i$ (right).

To obtain the full disk-integrated line profiles, we summed the pixels along the vertical axis of the stellar disk (i.e., for each radial velocity bin). Finally, we incorporate line broadening in the observations via a convolution of our model with a Gaussian kernel, with the width being the quadrature addition of the instrumental resolution and the macroturbulent velocity (v_{macro}).

A dark spot feature was seen during both spectroscopic transits on 2021 May 14 and 2021 June 18 (Figure 2). We interpret this to be the same spot feature. That it reappears at the same location on the second transit is interpreted to be that the planet orbital period is a multiple of the rotation period, near 5:1 resonance. This is supported by the near-sinusoidal light-curve variability seen during Sector 38 of the TESS observations. We therefore adopt a fixed stellar rotation period of $P_{\mathrm{rot}} = 1.39$ days for both the Gaussian process detrending and our Doppler tomographic analysis, and assume that there is one fixed long-lived spot that does not vary in size or temperature between the two observations. Not taking into account the spot evolution explains the slight residuals left after removal of the main spot feature as seen on the bottom panel of Figure 2 (panels (a)–(b) and (a)–(b)–(c)). For reference, a Lomb–Scargle analysis of the joint Sector 11 and 38 light curves reveals a rotation period of 1.4 ± 0.1 day, in agreement with the value we adopt in this analysis.

4
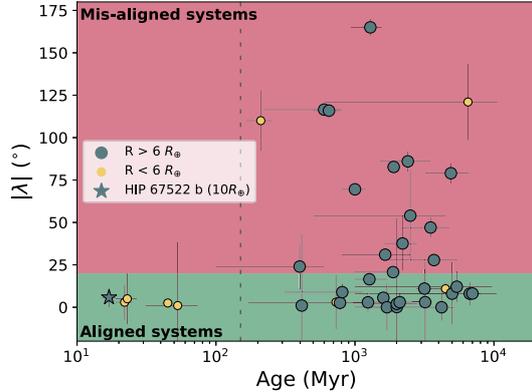
Figure 3. Systems younger then 10 Gyr that have both their age (NASA Exoplanet Archive 2021 July) and obliquity constrained. Yellow smaller circles represent Neptune-sized planets ($R < 6~R_\oplus$), larger blue circles show Jupiter-sized planets ($R > 6~R_\oplus$), and the blue star symbolizes HIP 67522 b. The very young systems (<100 Myr) are, in ascending order of age, HIP 67522, AU Mic, V1298 Tau, DS Tuc A, and TOI 942. The red and green areas show aligned ($|\lambda| < 20°$) and misaligned ($|\lambda| > 20°$) systems, respectively.

3.3. Global Fit

We simultaneously fitted the transit light curves modeled with BATMAN (Section 3.1) to the TESS data and the line profiles obtained through our Doppler tomographic model to the least-squares deconvolution profiles (Section 3.2).

For a given parameter set, we computed the likelihood for both the photometric and spectroscopic data sets. To determine the best-fitting parameter values and resulting posteriors, we made use of a Markov Chain Monte Carlo analysis using the affine-invariant ensemble sampler EMCEE (Foreman-Mackey et al. 2013). The analysis included 240 simultaneous chains with 10000 iterations each. The resulting posteriors are presented in Table 1. We derived a planet radius of $10.178 \pm 0.44~R_\oplus$, agreeing with the previous value of $10.07 \pm 0.47~R_\oplus$ from Rizzuto et al. (2020). We note that the availability of four extra transits did not improve the uncertainty on R_P, limited by the poorer constraint on $R_\star$.

4. Conclusions

We measured the projected obliquity angle of HIP 67522 b to be $|\lambda| = 5.8^{+2.8}_{-5.7}°$. With a stellar inclination estimated following Masuda & Winn (2020), we derived the 3D obliquity to be $\psi = 20.2^{+10.3}_{-8.7}°$. At an age of 17 Myr, HIP 67522 b is the youngest planet to receive such characterization. We demonstrate that a precise measurement of the sky-projected obliquity is possible for such young stars, despite the activity-dominated spectroscopic transit observations. Our single-spot model allows us to unambiguously disentangle the planetary signature from the stellar activity.

Figure 3 places HIP 67522 b into context of other planetary systems that have known obliquities and constrained ages.[9] This particular system joins AU Mic b (Palle et al. 2020; Addison et al. 2021; Hirano et al. 2020; Martioli et al. 2020), V1298 Tau c (David et al. 2019; Feinstein et al. 2021), DS Tuc Ab (Newton et al. 2019; Montet et al. 2020; Zhou et al. 2020),

---

[9] NASA Exoplanet Archive 2021 July.

and TOI 942 b (Wirth et al. 2021) in the group of systems younger than 100 Myr old that have their obliquity measured. Remarkably, all these systems have been found to be on well-aligned orbits, encouraging the pursuit of very young star obliquity measurements to confirm this trend.

With a radius of $10~R_\oplus$, HIP 67522 b is the only hot, Jupiter-sized planet in this very young group of systems. Its future evolution is still unclear, however, and the lack of constraint on its mass prevents us from definitively classifying it as a proto–hot Jupiter. Super-Earths and Neptune-sized planets commonly found around Sun-like stars can have a radii of $\sim 10~R_\oplus$ at the age of HIP 67522. The planet is undergoing Kelvin–Helmholtz contraction and photoevaporation, and its eventual radius depends strongly on its core-envelope makeup (e.g., Lopez & Fortney 2013).

If HIP 67522 b is indeed a proto–hot Jupiter, it is a prime example of one that did not migrate via high-eccentricity pathways. The circularization of HIP 67522 b's orbit is on the Gyr timescale (estimated from Goldreich & Soter 1966), playing against a high-eccentricity-type migration. Classic planet systems such as HD 80606 b (Naef et al. 2001; Wu & Murray 2003; Pont et al. 2009; Hébrard et al. 2010) and HD 17156 b (Barbieri et al. 2007; Cochran et al. 2008; Narita et al. 2008) exhibit highly eccentric and oblique orbits. For most other mature hot Jupiters, though, it is more difficult to decipher their original migration pathways due to tidal synchronization that occurs at the hundreds of millions of years to gigayear timescales (e.g., Lai 2012; Valsecchi & Rasio 2014), erasing evidence of their primordial histories. HIP 67522 b may be the first such example for which the most likely explanation is that the primordial orbit of a close-in Jovian planet is well aligned. Recently, Albrecht et al. (2021) found that hot Jupiters are preferentially found in well-aligned or polar orbits. There are a range of mechanisms that can result in such bimodality in the obliquity distribution, and understanding the age–obliquity distribution can help distinguish between these mechanisms. For example, magnetic warping of the protoplanet disk can result in hot Jupiters that formed in situ being found in oblique orbits (e.g., Lai et al. 2011). With more observations, a prevalence of well-aligned hot Jupiters around young stars may help limit the real-life effectiveness of such pathways.

If HIP 67522 b does become a Neptune-sized planet, it adds to the well-aligned pool of very young systems, contrasting with the often misaligned single-planet Neptune-like systems in close orbit around more mature aged stars. Even with a mass of a few $10~M_\oplus$, the circularization timescale is likely on the few 100 Myr timescale, one order of magnitude older larger than the age of the system, disfavoring high-eccentricity migration. HIP 67522 b would contribute to the growing interest to understand the formation of close-in Neptunes (e.g., Bean et al. 2021) and to make sense of the alignment distribution of these commonly found exoplanets.

Although not seen in Sector 38, HIP 67522 b has a possible nearby exterior transiting companion with an orbital period of > 23 days (Rizzuto et al. 2020). The existence of HIP 67522 c is tentative, but if confirmed, HIP 67522 b would be a more consistent future Neptune-sized planet as hot Jupiters are rarely found with outer companions. Dynamical interactions within closely packed planetary systems can excite mutual inclinations (Hansen & Murray 2013), and often result in the destruction of the interior planetary architecture if the outer companion is

massive (Huang et al. 2017). Systems like HIP 67522 and V1298 Tau (David et al. 2019; Feinstein et al. 2021) are candidates to test planet–planet interactions before planetary systems have settled into their final stable forms.

ORCID iDs

Alexis Heitzmann ⬡ https://orcid.org/0000-0002-8091-7526
George Zhou ⬡ https://orcid.org/0000-0002-4891-3517
Samuel N. Quinn ⬡ https://orcid.org/0000-0002-8964-8377
Stephen C. Marsden ⬡ https://orcid.org/0000-0001-5522-8887
Duncan Wright ⬡ https://orcid.org/0000-0001-7294-5386
Pascal Petit ⬡ https://orcid.org/0000-0001-7624-9222
Andrew M. Vanderburg ⬡ https://orcid.org/0000-0001-7246-5438
Luke G. Bouma ⬡ https://orcid.org/0000-0002-0514-5538
Andrew W. Mann ⬡ https://orcid.org/0000-0003-3654-1602
Aaron C. Rizzuto ⬡ https://orcid.org/0000-0001-9982-1332

References

Addison, B. C., Horner, J., Wittenmyer, R. A., et al. 2021, AJ, 162, 137
Albrecht, S. H., Marcussen, M. L., Winn, J. N., Dawson, R. I., & Knudstrup, E. 2021, ApJL, 916, L1
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Barbieri, M., Alonso, R., Laughlin, G., et al. 2007, A&A, 476, L13
Bean, J. L., Raymond, S. N., & Owen, J. E. 2021, JGRE, 126, e06639
Castelli, F., & Kurucz, R. L. 2003, in IAU Symp. 210, Modelling of Stellar Atmospheres , ed. N. Piskunov et al. (Cambridge: Cambridge Univ. Press)
Claret, A. 2017, A&A, 600, A30
Claret, A., Hauschildt, P. H., & Witte, S. 2012, A&A, 546, A14
Cochran, W. D., Redfield, S., Endl, M., & Cochran, A. L. 2008, ApJL, 683, L59
Collier Cameron, A., Guenther, E., Smalley, B., et al. 2010, MNRAS, 407, 507
David, T. J., Petigura, E. A., Luger, R., et al. 2019, ApJL, 885, L12
Dawson, R. I., & Johnson, J. A. 2018, ARA&A, 56, 175
Donati, J.-F., Semel, M., Carter, B. D., Rees, D. E., & Collier Cameron, A. 1997, MNRAS, 291, 658
Feinstein, A. D., Montet, B. T., Johnson, M. C., et al. 2021, AJ, 162, 213
Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, AJ, 154, 220
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Goldreich, P., & Soter, S. 1966, Icar, 5, 375
Hansen, B. M. S., & Murray, N. 2013, ApJ, 775, 53
Hébrard, G., Désert, J. M., Díaz, R. F., et al. 2010, A&A, 516, A95
Hirano, T., Krishnamurthy, V., Gaidos, E., et al. 2020, ApJL, 899, L13
Huang, C. X., Petrovich, C., & Deibert, E. 2017, AJ, 153, 210
Jenkins, J. M., Twicken, J. D., McCauliff, S., et al. 2016, Proc. SPIE, 9913, 99133E
Kreidberg, L. 2015, PASP, 127, 1161
Lai, D. 2012, MNRAS, 423, 486
Lai, D., Foucart, F., & Lin, D. N. C. 2011, MNRAS, 412, 2790
Lopez, E. D., & Fortney, J. J. 2013, ApJ, 776, 2
Martioli, E., Hébrard, G., Moutou, C., et al. 2020, A&A, 641, L1
Masuda, K., & Winn, J. N. 2020, AJ, 159, 81
Montet, B. T., Feinstein, A. D., Luger, R., et al. 2020, AJ, 159, 112
Morris, R. L., Twicken, J. D., Smith, J. C., et al. 2017, in Kepler Data Processing Handbook: Photometric Analysis (KSCI-19081-003) ed. J. M. Jenkins (Mountain View, CA: NASA Ames Research Center)
Naef, D., Latham, D. W., Mayor, M., et al. 2001, A&A, 375, L27
Narita, N., Sato, B., Ohshima, O., & Winn, J. N. 2008, PASJ, 60, L1
Newton, E. R., Mann, A. W., Tofflemire, B. M., et al. 2019, ApJL, 880, L17
Palle, E., Oshagh, M., Casasayas-Barris, N., et al. 2020, A&A, 643, A25
Paredes, L. A., Henry, T. J., Quinn, S. N., et al. 2021, AJ, 162, 176
Pont, F., Hébrard, G., Irwin, J. M., et al. 2009, A&A, 502, 695
Ricker, G. R., Vanderspek, R., Winn, J., et al. 2016, Proc. SPIE, 9904, 99042B
Rizzuto, A. C., Newton, E. R., Mann, A. W., et al. 2020, AJ, 160, 33
Southworth, J. 2011, MNRAS, 417, 2166
Tokovinin, A., Fischer, D. A., Bonati, M., et al. 2013, PASP, 125, 1336
Triaud, A. H. M. J. 2018, in Handbook of Exoplanets, ed. H. J. Deeg & J. A. Belmonte (Berlin: Springer), 1375
Twicken, J. D., Clarke, B. D., Bryson, S. T., et al. 2010, Proc. SPIE, 7740, 774023
Twicken, Joseph D., Catanzarite, Joseph H., Clarke, Bruce D., et al. 2018, PASP, 130, 064502
Valsecchi, F., & Rasio, F. A. 2014, ApJ, 786, 102
Vanderburg, A., Huang, C. X., Rodriguez, J. E., et al. 2019, ApJL, 881, L19
Wirth, C. P., Zhou, G., Quinn, S. N., et al. 2021, ApJL, 917, L34
Wu, Y., & Murray, N. 2003, ApJ, 589, 605
Zhou, G., Winn, J. N., Newton, E. R., et al. 2020, ApJL, 892, L21
Zhou, G., Quinn, S. N., Irwin, J., et al. 2021, AJ, 161, 2

6

CHAPTER 5: TOI-4562 b: A HIGHLY ECCENTRIC COOL JUPITER ANALOG ORBITING A YOUNG FIELD STAR.

THE LAST PUBLICATION of my candidacy presents the discovery of a new planet, TOI-4562 b. With its 221.5 days period, this young Jupiter-like planet sits at the outer edge of the warm Jupiter population. Its most notable feature is its highly elliptic orbit, with $e = 0.81 \pm 0.05$. Firstly identified in the TESS data, we performed a 2-year (i.e 2 orbits) RV follow-up using the CHIRON spectrograph to constrain its mass and eccentricity. We also used gyrochronology and lithium abundance to constrain its age that we estimate to be 300-400 Myr.

ALTHOUGH REMARKABLY ECCENTRIC, we found that TOI-4562 b's large semi-major axis is inconsistent with a future circularisation of the orbit. This planet is therefore unlikely to become a hot Jupiter, unless invoking exotic mechanisms such as eccentricity oscillations. Rather, we found traces of TTV, up to 20 minutes, which we interpret as the presence of a companion likely as the origin of the high eccentricity. This system has more to offer, and we hope that long term follow-up in RVs, an obliquity measurement

and upcoming Gaia astrometry data will provide, within the next few years, a complete 3D characterisation of the TOI-4562 planetary system, possibly revealing its past history.

TOI-4562 b: A highly eccentric temperate Jupiter analog orbiting a young field star.

Alexis Heitzmann [1], George Zhou [1], Samuel N. Quinn [2], Chelsea X. Huang [1], Jiayin Dong [3, 4, 5, *], L. G. Bouma [6, †], Rebekah I. Dawson [4, 5], Stephen C. Marsden [1], Duncan Wright [1], Pascal Petit [7], Karen A. Collins [2], Khalid Barkaoui [8, 9, 10], Robert A. Wittenmyer [1], Edward Gillen [11, 12, ‡], Rafael Brahm [13, 14], Melissa Hobson [15, 14], Coel Hellier [16], Carl Ziegler [17], César Briceño [18], Nicholas Law [19], Andrew W. Mann [19], Steve B. Howell [20], Crystal L. Gnilka [20], Colin Littlefield [21, 20], David W. Latham [2], Jack J. Lissauer [22], Elisabeth R. Newton [23], Daniel M. Krolikowski [24], Ronan Kerr [24], Rayna Rampalli [23], Stephanie T. Douglas [25], Nora L. Eisner [26], Nathalie Guedj [27], Guoyou Sun [27], Martin Smit [27], Marc Huten [27], Thorsten Eschweiler [27], Lyu Abe [28], Tristan Guillot [28], George Ricker [29], Roland Vanderspek [29], Sara Seager [30, 29, 31], Jon M. Jenkins [20], Eric B. Ting [20], Joshua N. Winn [32], David R. Ciardi [33], Andrew M. Vanderburg [34, 35], Christopher J. Burke [29], David R. Rodriguez [36], and Tansu Daylan [32, 37]

[1] University of Southern Queensland, Centre for Astrophysics, West Street, Toowoomba, QLD 4350 Australia
[2] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[3] Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA
[4] Department of Astronomy & Astrophysics, 525 Davey Laboratory, The Pennsylvania State University, University Park, PA, 16802, USA
[5] Center for Exoplanets and Habitable Worlds, 525 Davey Laboratory, The Pennsylvania State University, University Park, PA, 16802, USA
[6] Cahill Center for Astrophysics, California Institute of Technology, Pasadena, CA 91125, USA
[7] Institut de Recherche en Astrophysique et Planétologie, Université de Toulouse, CNRS, CNES, 14 avenue Edouard Belin, 31400 Toulouse, France
[8] Astrobiology Research Unit, Université de Liège, 19C Allée du 6 Août, 4000 Liège, Belgium
[9] Department of Earth, Atmospheric and Planetary Science, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[10] Instituto de Astrofísica de Canarias (IAC), Calle Vía Láctea s/n, 38200, La Laguna, Tenerife, Spain
[11] Astronomy Unit, Queen Mary University of London, Mile End Road, London E1 4NS, UK
[12] Astrophysics Group, Cavendish Laboratory, J.J. Thomson Avenue, Cambridge CB3 0HE, UK
[13] Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Av. Diagonal las Torres 2640, Peñalolén, Santiago, Chile
[14] Millennium Institute of Astrophysics MAS, Nuncio Monsenor Sotero Sanz 100, Of. 104, Providencia, Santiago, Chile
[15] Max Planck Institute for Astronomy, Königstuhl 17, D-69117—Heidelberg, Germany
[16] Astrophysics Group, Keele University, Staffordshire, ST5 5BG, UK
[17] Department of Physics, Engineering and Astronomy, Stephen F. Austin State University, 1936 North St, Nacogdoches, TX 75962, USA
[18] Cerro Tololo Inter-American Observatory, Casilla 603, La Serena, Chile
[19] Department of Physics and Astronomy, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3255, USA
[20] NASA Ames Research Center, Moffett Field, CA 94035, USA
[21] Bay Area Environmental Research Institute, Moffett Field, CA 94035, USA
[22] Space Science & Astrobiology Division , MS 245-3, NASA Ames Research Center, Moffett Field, CA 94035, USA
[23] Department of Physics and Astronomy, Dartmouth College, Hanover, NH 03755, USA
[24] Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA
[25] Department of Physics, Lafayette College, 730 High St., Easton, PA 18042, USA
[26] Sub-department of Astrophysics, University of Oxford, Keble Rd, Oxford, United Kingdom
[27] Citizen Scientist, Zooniverse c/o University of Oxford, Keble Road, Oxford OX1 3RH, UK
[28] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, CS 34229, F-06304 Nice Cedex 4, France
[29] Department of Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[30] Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[31] Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[32] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

Corresponding author: Alexis Heitzmann
alexis.heitzmann@usq.edu.au

[33]*NASA Exoplanet Science Institute, Caltech/IPAC, 1200 E. California Ave, Pasadena, CA 91125, USA*

[34]*Department of Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

[35]*Department of Astronomy, University of Wisconsin-Madison, Madison, WI 53706, USA*

[36]*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD, 21218, USA*

[37]*LSSTC Catalyst Fellow*

ABSTRACT

We report the discovery of TOI-4562 b (TIC-349576261), a Jovian planet orbiting a young F7V-type star, younger than the Praesepe/Hyades clusters ($\sim$ 300–400 Myr). This planet stands out because of its unusually long orbital period for transiting planets with known masses ($P_{\rm orb} = 225.11757^{+0.00027}_{-0.00025}$ days), and because it has a substantial eccentricity ($e = 0.81^{+0.05}_{-0.05}$). The location of TOI-4562 near the southern continuous viewing zone of *TESS* allowed observations throughout 25 sectors, enabling an unambiguous period measurement from *TESS* alone. Alongside the four available *TESS* transits, we performed follow-up photometry using the South African Astronomical Observatory node of the Las Cumbres Observatory, and spectroscopy with the CHIRON spectrograph on the 1.5 m SMARTS telescope. We measure a radius of $1.072^{+0.044}_{-0.043}$ $R_{\rm J}$, and a mass of $3.29^{+1.88}_{-0.82}$ $M_{\rm J}$ for TOI-4562 b. The radius of the planet is consistent with contraction models describing the early evolution of the size of giant planets. We detect tentative transit timing variations at the $\sim$ 20 min level from five transit events, favouring the presence of a companion that could explain the dynamical history of this system if confirmed by future follow-up observations. With its current orbital configuration, tidal timescales are too long for TOI-4562 b to become a hot-Jupiter via high eccentricity migration, though it is not excluded that interactions with the possible companion could modify TOI-4562 b's eccentricity and trigger circularization. The characterisation of more such young systems is essential to set constraints on models describing giant planet evolution.

1. INTRODUCTION

Planetary systems evolve rapidly within the first hundreds of millions of years of formation. The architectures of the systems evolve before settling into their eventual orbital configuration. Planets with extensive gaseous envelopes are expected to undergo contraction and cooling and experience observable changes in radius within this time frame. Observations of planets around young stars help anchor our understanding of this era of rapid change and help define models of planet formation and evolution. In particular, Jovian planets in distant orbits are less affected by stellar irradiation than close-in hot Jupiters. Transiting cold Jupiters around young stars can therefore provide constraints for cooling and contraction of giant planet evolution models. The orbital properties of these planets can also help to narrow down the timescales of dynamical evolution experienced by many other giant planets discovered to date.

Numerous mechanisms are responsible for the formation and evolution of close-in Jovian planets. These mechanisms vary by the distribution of planets that they

produce and by the timescales at which they operate. We can best assess the prevalence of these multiple formation channels via a census of the gas giant population as a function of time (see Dawson & Johnson 2018). Such a temporal survey of planetary systems can unveil the roles that in-situ formation (review in Chabrier et al. 2014), disk migration (review in Baruteau et al. 2014) and high eccentricity migration (review in Dawson & Johnson 2018) played in shaping our current gas giant population. For example, planets can gravitationally interact with their depleting gas disks, resulting in moderately eccentric final orbits within a few million years (e.g., Nagasawa et al. 2003; Duffell & Chiang 2015; Debras et al. 2021). On the other extreme, excitation via stellar fly-bys can occur on the hundreds of millions of years timescale (e.g., Shara et al. 2016).

Gas giants also undergo significant contraction in the first hundred million years post formation. In models, the rate of contraction is strongly dependent on the initial conditions of the planet post formation, such as their envelope-core mass ratio and initial luminosities (e.g., Fortney et al. 2007; Linder et al. 2019). It is clear, however, that the radius distribution of close-in Jovian planets is shaped by external factors that retard their contraction (e.g., Guillot & Showman 2002; Baraffe et al. 2003; Batygin & Stevenson 2010). Young planets

in distant orbits provide simpler key tests for gas giant evolution.

Missions like Kepler, K2, and the Transiting Exoplanet Survey Satellite (*TESS*; Ricker et al. 2015) have brought forth a growing number of planetary systems about young stars (e.g., Newton et al. 2019; Mann et al. 2020; Plavchan et al. 2020; Bouma et al. 2022a,b; Zhou et al. 2022). However, true young Jovian analogues are rare. Interestingly, Suárez Mascareño et al. (2022) measured the masses of the giant planets in the 22 Myr old V1298 Tau system (David et al. 2019a,b), finding that the two Jovian planets have already settled to their expected final radii, a process that is predicted to take hundreds of millions of years by contraction models. Other close-in Jovian-sized planets have also been found around young stars (Rizzuto et al. 2020; Bouma et al. 2020; Mann et al. 2021), but strong stellar activity has prevented their mass from being measured.

We report the discovery of a young transiting Jovian planet in a distant orbit around a ∼ 300 Myr old star. TOI-4562 hosts a temperate-Jupiter in a 225 day period orbit near the *TESS* continuous viewing zone. Along with additional observations from our ground-based photometric follow-up campaign, five total transits of the planet were obtained, unambiguously identifying the period of the system. Radial velocity monitoring over the following two years provided a mass and eccentricity measurement for the young planet. In addition, data from FEROS helped to constrain the stellar parameters, and high resolution images from Gemini-South and SOAR helped to rule out false positive scenarios, confirming the transit candidate as a true planet. We also constrained the age of TOI-4562 via gyrochronology and lithium. Finally we detect a transit timing variations (TTV) signature, indicative of a perturbing companion in the system. TOI-4562 b is one of the longest period transiting temperate Jupiters discovered by *TESS*, and the youngest amongst such planets. Missions like *TESS* and *PLATO* (Rauer et al. 2014) have the potential to uncover this special population that critically constrains cooling models and migration pathways for Jovian planets.

2. OBSERVATIONS

2.1. TESS*: Photometry*

The transiting planet candidate around TOI-4562 was first identified from observations by *TESS*. TOI-4562 lies in the Southern Continuous Viewing Zone of *TESS*, and therefore received near-uninterrupted photometric monitoring during years 1 and 3 of operations. The target received observations at 30 minute cadence during Sectors 1-8 (2018-07-25 to 2019-02-28) and 10-13 (2019-03-26 to 2019-07-18), and 2 minute target-pixel-stamp observations during sectors 27-39 (2020-07-04 to 2021-06-24). The transit signature of TOI-4562b was detected by the TESS Science Processing Operations Center (SPOC; Jenkins et al. (2016)) at NASA Ames Research Center during a transit search of sectors 27 through 39 with an adaptive, noise-compensating matched filter (Jenkins 2002; Jenkins et al. 2010, 2020). The transit signature passed all the diagnostic tests in the Data Validation report (Twicken et al. 2018) and was fitted with an initial limb-darkened transit model (Li et al. 2019). In particular, the transit signal passed the difference image centroiding test, which localized the source of the transits to within 1.0 ± 2.5 arcsec. The TESS Science Office reviewed the diagnostic information and released an alert to the community for TOI-4562b on 28 October 2021 (Guerrero et al. 2021).

We make use of the MIT Quicklook pipeline (Huang et al. 2020) photometric extraction from the Full Frame Image observations. In addition, where available, we make use of the 2 minute cadence target pixel file observations from the crowding and flux fraction corrected Simple Aperture Photometry (CROWDSAP) light curves (Twicken et al. 2010; Morris et al. 2020) made available by SPOC. Because of the large stellar variability seen in the light curve, we used the SAP light curves rather than the Pre-search Data Conditioning SAP (PDCSAP) flux and performed the detrending using a high order spline interpolation (Vanderburg & Johnson 2014)

The full *TESS* light curve covering all sectors of observations is presented in Figure 1. During the two (nonconsecutive) years of near-continuous observations a total of 4 transits were captured by *TESS*. Figure 2 shows the zoomed in region around each of these transits.

TOI-4562 was first identified as a potential young star due to its strong rotational modulation (Zhou et al. 2021), as part of our program to survey for planets around young field stars. We performed a search for transiting signals around TOI-4562 via a Box-least-squares period search (Kovács et al. 2002) after removal of the stellar modulation signal with the splines. This detrending was not the one used for the transit modeling, described in section 4.1.

2.2. *Follow-up photometry*

We obtained follow-up photometric confirmation of the planetary transit via the Las Cumbres Observatory Global Network (LCOGT; Brown et al. 2013). Transit opportunities for a 225 day period planet are rare from the ground (see Table 3). We captured the full transit of TOI-4562 b on 2022-01-03 UTC from the South African
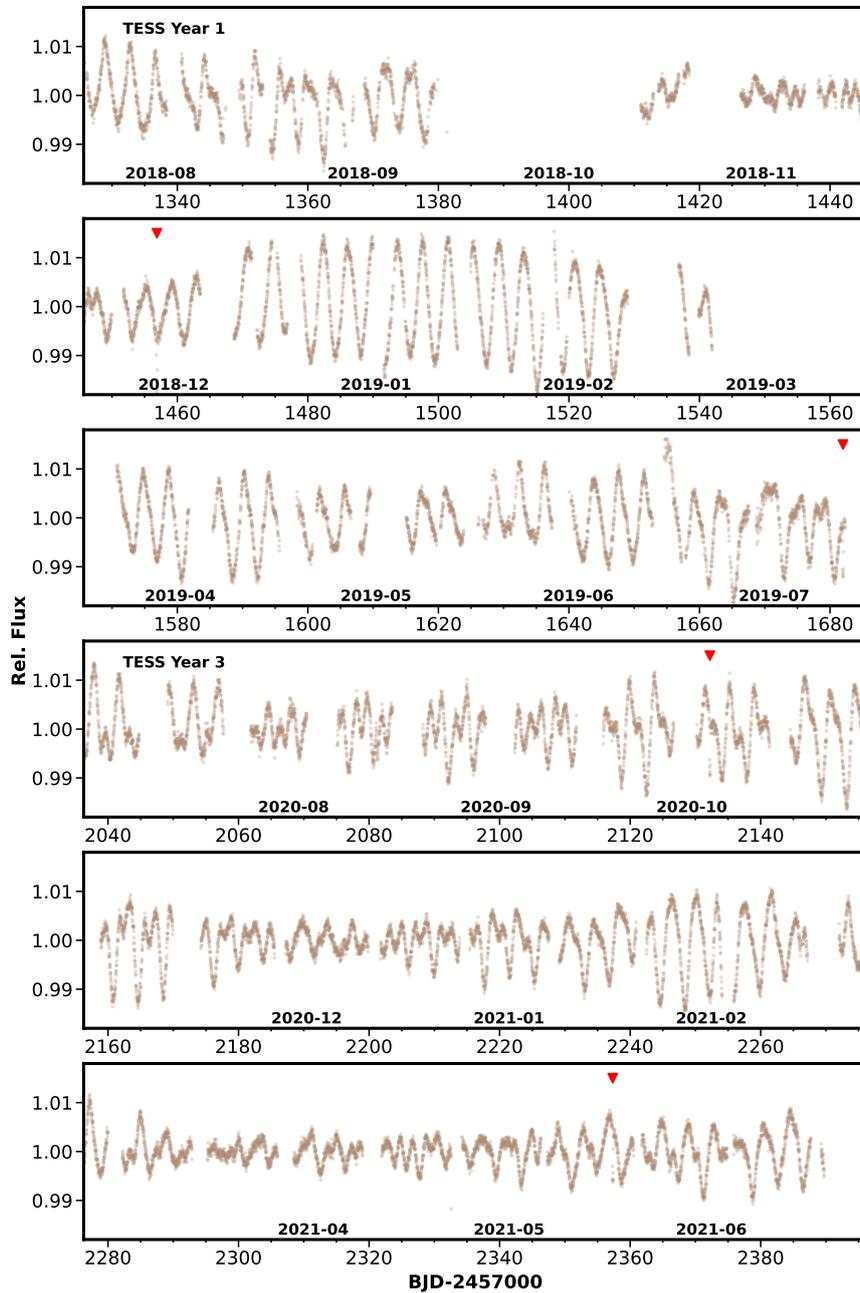
Figure 1. *TESS* lightcurve (brown) of TOI-4562 b from all 25 available sectors. Photometry prior to sector 13 were obtained at 30 minute cadence, while latter observations were obtained at 2 minute cadence, and binned in this figure to 30 min for clarity. The four TOI-4562 b transits from Sectors 5, 13, 30, and 38 are marked by red arrows. The host star exhibits up to ∼3% peak-to-peak stellar rotational modulation due to its youth.
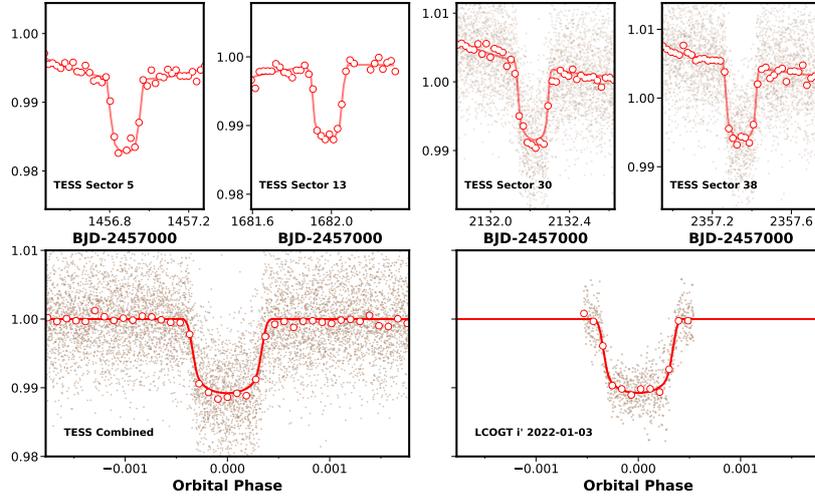
Figure 2. Top Individual transits from *TESS* from Sectors 5 and 13 at 30 minute cadence, and from Sectors 30 and 38 at 2 minute cadence. The red circles indicate the measured data for sectors 5 and 13, and the binned data at 30 minute cadence for Sectors 30 and 38. The 2-minute cadence data for Sectors 30 and 38 are plotted as brown points. The best fit model, incorporating the transit timing variations in Section 4, and out-of-transit trends, are shown by the red lines. **Bottom left** The phase folded *TESS* transit and best fit model. **Bottom right** The combined follow-up LCOGT 1 m observations from 2022-01-03 in i' band.

Astronomical Observatory (SAAO) node of LCOGT via two 1 m telescopes. The observations were obtained with the *Sinistro* 4K×4K cameras in the Sloan i' filter. The observations were calibrated via the *BANZAI* pipeline (McCully et al. 2018), and light curves were extracted via the *AstroImageJ* package (AIJ; Collins et al. 2017) using circular apertures with radius $4\rlap{.}{''}7$, which exclude flux from all known nearby Gaia EDR3 and TESS Input Catalog stars. The combined light curves (after removing systematics) and best fit model are shown in Figure 2.

In addition, a transit on 2022-08-16 was attempted from the SAAO node of LCOGT via one 1 m telescope, as well as the Antarctica Search for Transiting Exo-Planets (ASTEP) facility (Guillot et al. 2015; Mékarnia et al. 2016), located at the East Antarctic plateau. A 25 minute segment was captured out of transit, but no portions of a transit event was recorded, and the dataset not included in the modeling presented below.

2.3. *CHIRON/SMARTS: Spectroscopy*

To characterize the radial velocity orbit of TOI-4562 b and constrain the properties of the host star, we obtained 84 spectroscopic observations of TOI-4562 using the CHIRON facility. To capture the long orbital period of TOI-4562 b, the velocities spanned two observing seasons, from 2020-12-09 to 2022-01-23; the resulting radial velocities are given in Table 4. CHIRON is a fiber-fed high resolution echelle spectrograph on the 1.5 m SMARTS telescope at Cerro Tololo Inter-American Observatory, Chile (Tokovinin et al. 2013). Due to the faintness of the host star, spectral observations were obtained in the 'fiber' mode of CHIRON, yielding a resolving power of $R \sim 28,000$ over the wavelength range of 4100 to 8700 Å, and an average signal-to-noise of ~ 100 per resolution element at the Mg b line wavelength region.

We make use of the extracted spectra from the standard CHIRON pipeline described in Paredes et al. (2021). Radial velocities were derived from the observations via a least-squares deconvolution against a non-rotating ATLAS9 spectral template (Castelli & Kurucz 2004). The resulting broadening profile is fitted via a kernel describing the effects of radial velocity shift, rotational, macroturbulent, and instrumental broadening. The derived velocities are presented in Table 4 and shown in Figure 4.

To estimate the spectroscopic properties of the host star, we matched each spectrum against an observed library of $\sim 10,000$ spectra pre-classified by the Spectroscopic Classification Pipeline (Buchhave et al. 2012). The matching was performed by first training the pre-

classified library via a gradient boosting classifier using SCIKIT-LEARN, and then classifying the observed spectrum. We found that TOI-4562 has an effective temperature of $T_{\mathrm{eff}} = 6096 \pm 50$ K, a surface gravity of $\log g = 4.4 \pm 0.1$ dex, and bulk metallicity of [Fe/H] = 0.1 ± 0.1 dex. Since the CHIRON dataset overwhelms the other datasets we obtained for TOI-4562 in quantity, we adopt these parameters as Gaussian priors in the global analysis of the system described in Section 4. We note a general consensus between the spectral parameters from CHIRON and those presented below in Section 2.4.

We also check for the possibility that the velocity variations we observe are due to a spectroscopically blended companion rather than the host star. We compare the broadening measured from the line profiles against the velocities and find no correlation. If a blended companion is causing the radial velocity offset, then the line profiles should be broadest at the orbital quadratures, and narrowest at conjunctions. We therefore find no evidence that the velocity variations originate from a blended companion.

2.4. *FEROS & GALAH: Spectroscopy*

The FEROS spectrograph, attached to the MPG 2.2 m (Kaufer et al. 1999) telescope at La Silla Observatory, gathered 11 spectra of TOI-4562. Spectra are co-added, with a signal to noise ratio per spectra ranging between 52 and 82, and atmospheric parameters are derived using ZASPE (Brahm et al. 2017b). We find $T_{\mathrm{eff}} = 6280 \pm 100$ K, $\log g = 4.49 \pm 0.10$, [Fe/H] = 0.24 ± 0.05 dex and $v \sin i = 15.67 \pm 0.5$ km s^{-1}. We chose not to include the FEROS data in the RV modelling. All points fall near phases (-0.4, 0.025 and 0.35) where the RV signal is close to 0 and therefore don't meaningfully contribute, while adding one instrument and the associated extra parameters. Using the CERES pipeline (Brahm et al. 2017a), we also recover chromospheric emission indices, tracers of stellar activity. The core emission of the H$_\alpha$ line at 6562.808 Å is H$_\alpha = 0.160 \pm 0.005$ (following Boisse et al. (2009)). Using regions defined by Duncan et al. (1991) and calibrations from Noyes et al. (1984) we measure the core emission of the Ca II H and K lines around 3933 Å and 3968 Å to be $\log R'_{HK} = $-4.503 ± 0.044. This value is consistent with a young active star (Mamajek & Hillenbrand 2008).

Finally, legacy spectra from the GALAH survey (Buder et al. 2021) found $T_{\mathrm{eff}} = 6034 \pm 77$ K, $\log g = 4.36 \pm 0.18$, [Fe/H] = 0.08 ± 0.06 and $v \sin i = 15.6 \pm 2.2$ km s^{-1}.

2.5. Gemini-South and SOAR: High resolution direct Imaging

A first high resolution image of TOI-4562 was obtained on 2022-03-17 with the Zorro Speckle camera on the 8.1 m Gemini-South telescope (Howell & Furlan 2022) and is shown in the top of Figure 3. Simultaneous observations were obtained at 562 and 832 nm respectively. Contrast curves were retrieved following Howell et al. (2011) for both wavelengths and neither shows sign of a companion in the vicinity of TOI-4562 b. A difference in magnitude Δm of 5 is achieved at a separation of ~ 0.1". This allows us to rule out the presence of bright stellar objects in the same TESS pixel as TOI-4562 that would meaningfully impact the transit light curve to a projected distance of ~ 35 au (given TOI-4562's distance of 346.9 ± 3.8 pc).

On 2022-04-19, another high resolution image was acquired with the HRCam instrument on the 4.1 m Southern Astrophysical Research (SOAR) telescope. TOI-4562 was observed as part of the SOAR TESS survey (Ziegler et al. 2020, 2021), and the data was reduced following Tokovinin (2018). The image shown in the bottom panel of Figure 3 and shows a contrast in the I-band of 5 mag within 1" with no sign of a companion, in agreement with the Gemini-South observation.

3. AGE OF TOI-4562

TOI-4562 does not appear in the extensive list of stars with known age and/or belonging to associations and moving groups compiled from the literature in Bouma et al. (2022). Similarily, we do not identify a co-eval population when applying the COMOVE package (Tofflemire et al. 2021) that uses Gaia DR3 astrometric parameters to find whether a given possible young star candidate is co-moving with its visual neighbours.

This lack of evidence of TOI-4562 belonging to any known moving group or open cluster means its age estimation is challenging. The variability seen in both photometry and radial velocity are indicative of the presence of rotationally modulated surface brightness features, likely due to the presence of dark spots and bright plages/faculae. Combined with a fast rotation period ($P_\star = 3.82 \pm 0.05$ days), this strongly suggests that TOI-4562 is a young and active star.

Determining the age of a field star is notoriously difficult (Soderblom 2010). In the following paragraphs, we make use of the rotation and lithium abundance of TOI-4562 to qualitatively assess its youth. We note that though TOI-4562 exhibits signatures of activity and youth indicative of being younger than 1 Gyr, pinpointing its age will remain difficult without placing it within co-moving populations. With increasingly more
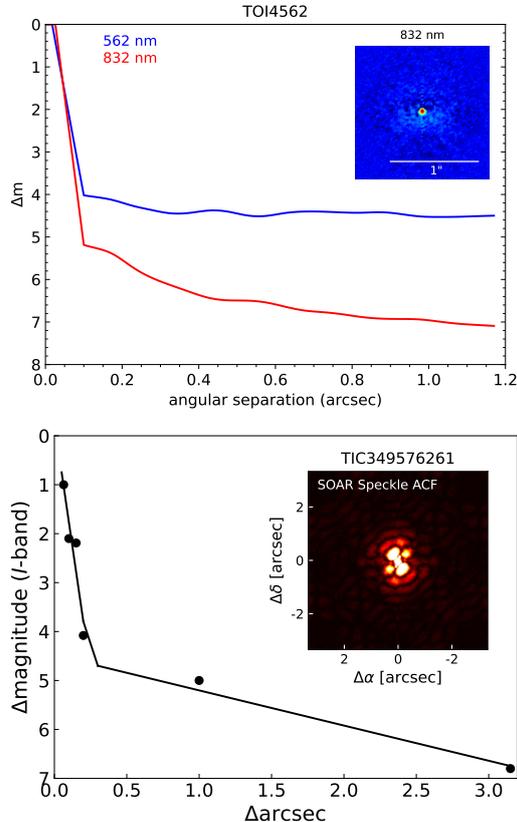
Figure 3. Top High resolution image of TOI-4562 obtained with the Zorro camera attached to the 8.1 m Gemini South telescope. The blue and red curves show difference in magnitude as a function of orbital separation from TOI-4562 obtained at wavelengths of respectively 562 and 832 nm. The inset plot shows the reconstructed image at 832 nm where no companion is detected. **Bottom** SOAR HRCam high resolution imaging of TOI-4562. The difference in magnitude as a function of orbital separation from TOI-4562 is shown by the black line and the autocorrelation function on the inset image. There is no sign of a stellar sized companion to TOI-4562.

sophisticated clustering with updated *Gaia* datasets, we hope that kinematics studies such as Oh et al. (2017), Gagné et al. (2018), Kounkel & Covey (2019) and Ujjwal et al. (2020) can provide improved census of young associations and groups.

3.1. Stellar rotation and Gyrochronology

Table 1. TOI-4562 parameters.

Parameters	Description	Prior	Value	Reference
Name and position				
TOI	*TESS* Object of Interest	-	4562	
TIC	*TESS* Input Catalog	-	349576261	ST18
Gaia	DR2 Source ID	-	5288681857665822080	*Gaia* EDR3
RA	Right ascension (HH:MM:SS, J2000, epoch 2015.5)	-	07:28:02.41	*Gaia* EDR3
DEC	Declination (DD:MM:SS, J2000, epoch 2015.5)	-	-63:31:04	*Gaia* EDR3
μ_{RA}	RA proper motion $(\mathrm{mas\,yr^{-1}})$	-	-5.899 ± 0.015	*Gaia* EDR3
μ_{DEC}	DEC proper motion $(\mathrm{mas\,yr^{-1}})$	-	10.491 ± 0.011	*Gaia* EDR3
Type	Spectral type	-	F7V	
ϖ	Parallax (mas)	$\mathcal{G}[2.881, 0.01]^a$	$2.88356^{+0.0104}_{-0.0099}$	This work
D	Distance (parsec)	-	346.9 ± 3.8	This work
Photospheric parameters				
T_{eff}	Effective temperature (K)	-	$6096 \pm 32\,\mathrm{K}$	This work (CHIRON)
$\log g$	Surface gravity (dex)	-	4.42 ± 0.04	This work
[M/H]	Bulk metallicity (dex)	$\mathcal{G}[0.2, 0.3]$	0.24 ± 0.10	This work
$v \sin i$	Rotational velocity $(\mathrm{km\,s^{-1}})$	-	15.67 ± 0.5	This work (CHIRON)
Physical parameters				
$M_\star$	Mass $(M_\odot)$	$\mathcal{U}[0, 2]$	$1.218^{+0.054}_{-0.048}$	This work
$R_\star$	Radius $(R_\odot)$	-	$1.126^{+0.014}_{-0.012}$	This work
Age	Age (Myr)		300–400	This work
Activity parameters				
$P_\star$	Equatorial rotation period (days)	-	3.82 ± 0.05	This work
Li 6708 EW	Li doublet (~ 6708 Å) equivalent width (Å)	-	0.076 ± 0.022	This work
$\log R'_{HK}$		-	$-4.503^{+0.028}_{-0.052}$	This work (FEROS)
Photometric parameters				
E(B-V)	Interstellar extinction (mag)	$\mathcal{U}[0, 0.1542]^b$	0.077 ± 0.015	This work
T	*TESS* T(mag)	-	11.533 ± 0.006	ST18
V	Johnson V(mag)	-	12.098 ± 0.014	H16
B	Johnson B (mag)	-	12.698 ± 0.025	H16
G	Gaia G(mag)	-	11.948 ± 0.020^c	*Gaia* EDR3
B_{p}	Gaia B_{p}(mag)	-	12.262 ± 0.020^c	*Gaia* EDR3
R_{p}	Gaia R_{p}(mag)	-	11.467 ± 0.020^c	*Gaia* EDR3
J	2MASS J(mag)	-	10.931 ± 0.023	SK06
H	2MASS H(mag)	-	10.693 ± 0.025	SK06
K_{s}	2MASS K_{s}(mag)	-	10.619 ± 0.023	SK06
W_1	WISE W_1(mag)	-	10.578 ± 0.023	W10,C13
W_2	WISE W_2(mag)	-	10.618 ± 0.020	W10,C13
W_3	WISE W_3(mag)	-	10.590 ± 0.061	W10,C13
NUV	GALEX/NUV calibrated AB magnitude (mag)	-	17.133 ± 0.023	B17

[a]Adopted from Gaia EDR3 Gaia Collaboration et al. (2016, 2021)

[b]Adopted from Schlafly & Finkbeiner (2011)

[c]These are inflated uncertainties from the *Gaia* photometric bands, following the convention from Eastman et al. (2013)

Priors: $\mathcal{U}[a, b]$ uniform priors with boundaries a and b; $\mathcal{G}[\mu, \sigma]$ Gaussian priors

References: *Gaia* EDR3 Gaia Collaboration et al. (2016, 2021); ST18 Stassun et al. (2018); SK06 Skrutskie et al. (2006); W10 Wright et al. (2010); C13 Cutri et al. (2021); B17 Bianchi et al. (2017); H16 Henden et al. (2016)
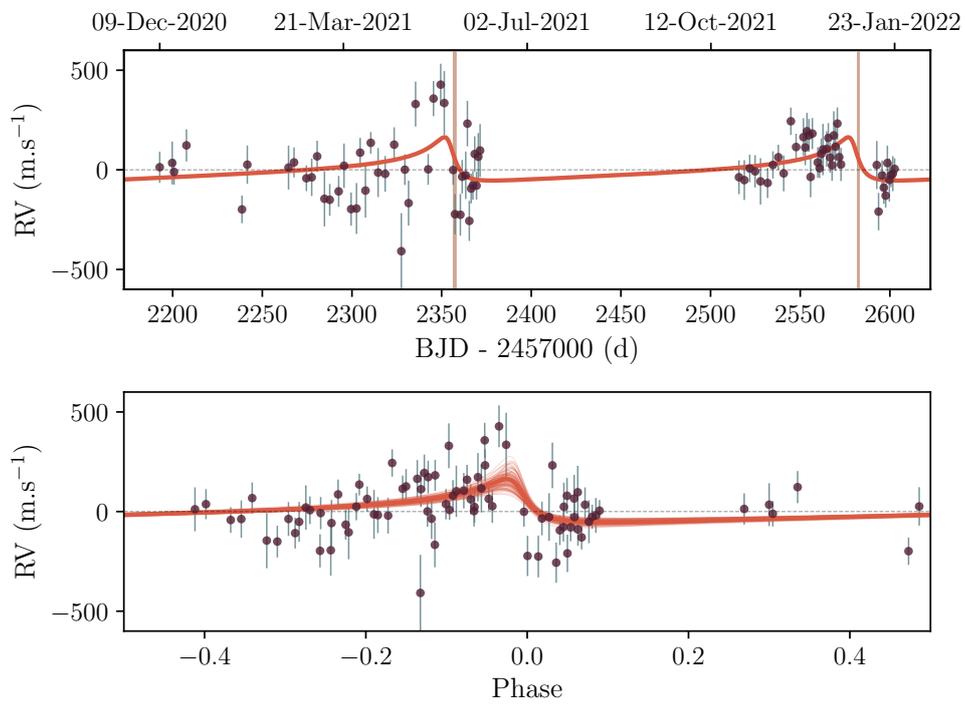
Figure 4. Top: 2-year radial velocity time series of TOI-4562 obtained with the CHIRON spectrograph (brown points) with the associated error bars. The Keplerian orbit fit from our global modelling is shown in red. Transits are highlighted in light brown. **Bottom**: Phase folded RVs (brown) with the Keplerian orbit best fits in red.
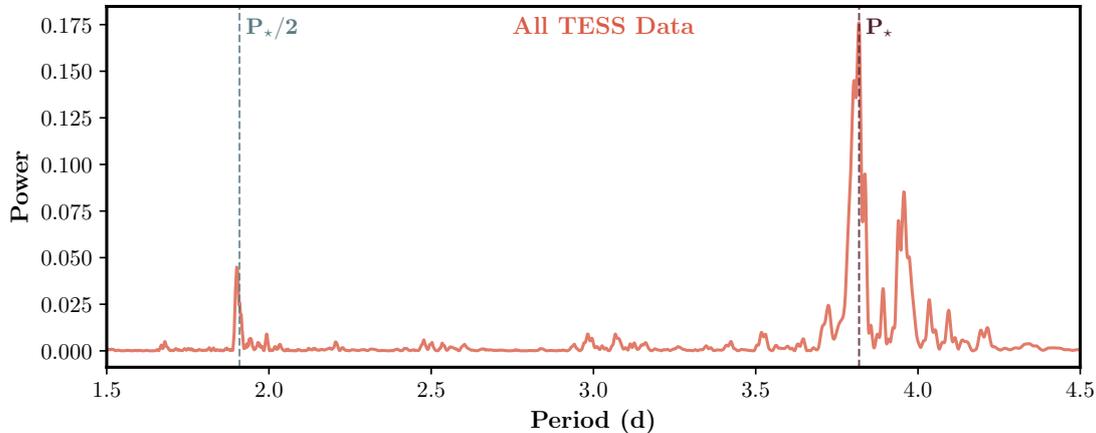
Figure 5. Generalized Lomb Scargle periodogram of TOI-4562's photometry from *TESS* for the 22 available sectors (red). The identified stellar rotation period ($P_\star = 3.82 \pm 0.05$) and the first harmonic ($P_\star/2 \sim 1.91$ days) are shown with the dashed brown and grey lines respectively. The second peak around 3.9 days could be associated with stellar surface features tracing differential rotation.

Young stars on the zero-age main-sequence spin rapidly. Over the course of a few billion years, mass loss from stellar winds spin-down Sun-like stars. The rotation period of Sun-like stars can be a tracer for their age. Rotation–color–age relationships such as those from Barnes (2007) and Mamajek & Hillenbrand (2008) are calibrated against co-eval clusters and associations, and can provide useful metrics to estimate stellar ages. Recent theoretically-motivated models, which are based in wind braking models and can incorporate core-envelope coupling, also provide such relationships (e.g. Spada & Lanzafame 2020).

The 25 sectors of observations gathered by *TESS* provide the means for a good estimation of the rotation period of TOI-4562. As shown in Figure 5, we ran a Lomb-Scargle period analysis (Lomb 1976; Scargle 1982) on the entire dataset and measured a rotation period of $P_\star = 3.82 \pm 0.05$ days.

We note the clear second periodogram peak on Figure 5, close to $P_\star$. This could be showing differential rotation (i.e., the variation of $P_\star$ as a function of stellar latitude). This has been largely observed in Kepler stars (Reinhold et al. 2013). We could suppose that the rotational modulation of two distinct clumps of surface stellar spots evolving at a different latitude would be at origin of the double peak (Lanza et al. 1993).

In addition, TOI-4562 received 4 years of monitoring with the Wide Angle Search for Planets (WASP) Consortium (Pollacco et al. 2006) Southern SuperWASP facility from 2008-2012. WASP-South is located at SAAO, and consists of an array of eight commonly mounted 200 mm f/1.8 Canon telephoto lenses, each with a $2K \times 2K$ detector. A period analysis of the WASP-South light curves reveals a 3.84 day periodicity, in agreement with the *TESS* light curves. The long term stability of the signal helps to confirm it as the correct alias of the rotational modulation signal.

Finally, we run periodograms on the available light curves from the All-Sky Automated Survey for Supernovae (ASAS-SN, Shappee et al. (2014); Jayasinghe et al. (2019)). Sloan g-band data spanning from October 2017 to April 2022 shows very strong peaks in the periodogram around 3.85 days, agreeing with the other photometric datasets. Johnson V-band data was obtained between October 2016 and September 2018. Despite being less extensive and less densely sampled than the g-band photometry, a moderate peak (FAP $\sim 0.2\%$) is found at 3.64 days, close to $P_\star$.

Using the age-rotation relationship from Mamajek & Hillenbrand (2008), we found TOI-4562 b to be 110-490 (3σ) Myr old. We note that age estimates from this relationship assumes that the star lies on the slow-sequence of the age-rotation relationship. Stars are often found to be more rapidly rotating than such sequences for a given age, which has been attributed to binarity in cluster populations (e.g., Douglas et al. 2016; Gillen et al. 2020). Though there is no evidence for TOI-4562 being part of a binary system, caveats still apply for gyrochronology-based age estimates. For a 1.2 $M_\odot$ star with $P_\star = 3.82 \pm 0.05$, the model from Spada & Lanzafame (2020) gives a consistent age estimate of 300-400 Myr.
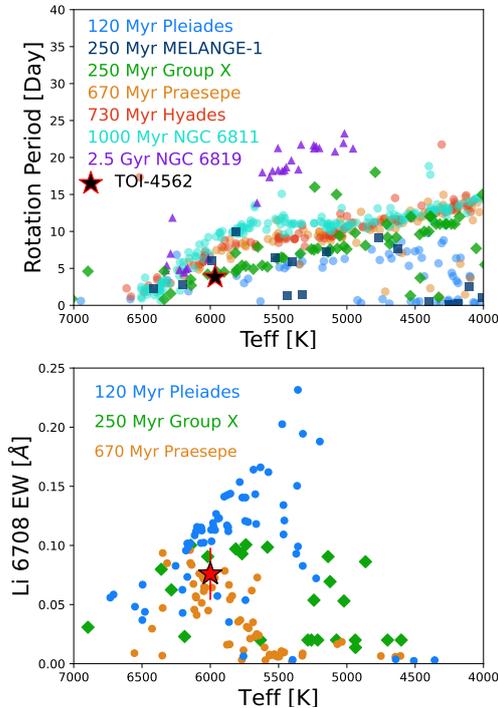
Figure 6. Youth indicators for TOI-4562. **Top:** The rotation period of TOI-4562 compared to the distribution of stars within known associations and clusters, including the Pleiades (Rebull et al. 2016), MELANGE-1 (Tofflemire et al. 2021), Group X (Newton et al. 2022; Messina et al. 2022), Praesepe and Hyades (Douglas et al. 2016, 2019), NGC 6811 (Curtis et al. 2019) and NGC 6819 (Meibom et al. 2015). **Bottom:** Equivalent width of the lithium doublet at 6707.76 and 6707.91 Å for TOI-4562 (red star) and stars in the Praesepe (orange, Cummings et al. 2017), Group X (Newton et al. 2022), and Pleiades (black, Bouvier et al. 2018) clusters. TOI-4562 lies at an age comparable to the Hyades and Praesepe.

The top plot of Figure 6 shows the rotation period of TOI-4562 compared with stars of known nearby clusters and associations. TOI-4562's $P_\star$ is consistent with that of stars belonging to Group X (Newton et al. 2022; Messina et al. 2022), with an estimated age of 300 Myr.

3.2. *Lithium*

The convective envelope of low-mass stars ($M_\star < 1.5$ $M_\odot$) allows efficient transport of lithium to deeper and hotter regions in a star's interior, where it gets destroyed by proton capture. Calibrated with stars in clusters and associations, this lithium depletion can be used as a proxy for stellar age. Using CHIRON spectra (see
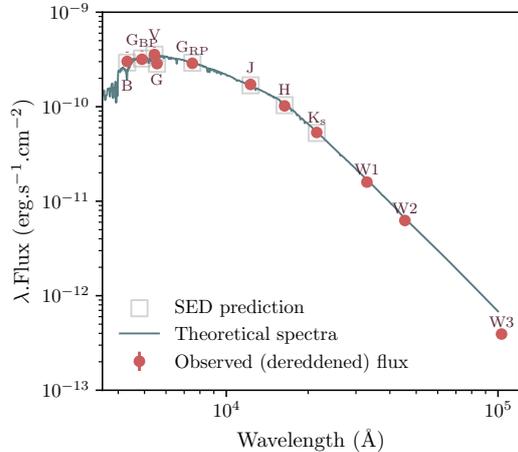
Figure 7. Spectral energy distribution (SED) of TOI-4562 b. Red points are the observed magnitudes in different wavelength bands (labelled with corresponding letters) corrected from interstellar reddening. Predicted magnitudes from the isochrone part of our global model are shown as grey squares. The blue line is a theoretical spectra for a star with $T_{\rm eff}$= 6000 K, log g = 4.5 dex and [M/H] = 0; adopted from Coelho (2014).

section 2.3), we measured the equivalent width of the lithium doublet at 6707.76 and 6707.91 Å. We fit two Gaussian line profiles of the same depth at the respective wavelengths of the lithium doublet and one auxiliary with a different depth to account for the nearby Fe I line at 6707.43 Å, usually blended with the Li doublet. All profiles share the same width as per the rotational broadening of the star. We measure a lithium equivalent width of 0.084 ± 0.007 Å from a median combined spectrum of all our CHIRON observations. These data are displayed in Figure 6. On the same figure, we show the lithium equivalent width as function of effective temperature for stars belonging to clusters with well constrained ages, the Pleiades (~ 125 Myr), Group X (~ 300 Myr) and Praesepe (~ 670 Myr). TOI-4562 exhibits a Li equivalent width shallower than most of the Pleiades stars and of comparable strength to stars from the Praesepe cluster, at an effective temperature of 6000 K. Combined with the Gyrochronology analysis, TOI-4562's age is consistent with a star younger than the Praesepe/Hyades clusters (i.e., $\lesssim 700$ Myr).

4. ANALYSIS AND RESULTS

To best determine the system properties of TOI-4562, we perform a joint modeling of all available photometric and spectroscopic datasets, including stellar isochrone models that constrain the properties of the host star.

The paragraphs below detail individual components of this model.

4.1. Transit modeling

Despite the 225 day orbital period of TOI-4562 b, the extensive observations of TOI-4562 by *TESS* allowed four transits to be observed. Spot modulated variability at the ~3% level is seen on the *TESS* light curve due to the active nature of TOI-4562, as expected given its young age. For the purposes of the transit modeling, we detrend the region around each transit epoch with a fourth-order polynomial. The polynomial is fitted using the out-of-transit regions of the light curve within 0.5 days of the transit center. We model the transits as per Mandel & Agol (2002) via the batman package (Kreidberg 2015). Free parameters that describe the transit model include the transit centre T_c at each transit epoch, radius ratio $R_p/R_\star$, line of sight inclination of the transit i, and the eccentricity parameters $\sqrt{e}\cos\omega$ and $\sqrt{e}\sin\omega$. A quadratic model was used to account for Limb Darkening using coefficients μ_{1TESS} and μ_{2TESS} fixed to those interpolated from Claret (2017) at the atmospheric parameters of TOI-4562 for the *TESS* transits. We note that $a/R_\star$ was not directly sampled but rather computed from the free parameters P_{orb}, $M_\star$, $R_\star$ and planet mass M_p. For the two (same epoch, different telescopes) SAAO LCOGT transits, the Limb Darkening coefficients μ_{1LCO} and μ_{2LCO} are computed for the SDSS i' band from Claret & Bloemen (2011), using the interpolation routine from Eastman et al. (2013) with $T_{eff} = 6000$ K, $\log g = 4.5$ and [M/H] = 0.1, computed with the least square method (LSM). For the SAAO LCOGT data, we also incorporate the effects of instrumental systematic variations that are common to ground-based photometric observations via a simultaneous detrending of the light curve against parameters describing the observation airmass to which we add a linear trend with respect to time. All detrended light curves and the best transit model fits are shown in Figure 2.

4.2. Radial velocity modeling

The radial velocities obtained over the 2 consecutive orbits of TOI-4562 b were modeled using a Keplerian orbit. Some fitted parameters are shared with the transits and stellar isochrone fitting, such as T_c, P_{orb}, $a/R_\star$, $R_\star$, $M_\star$, i, $\sqrt{e}\cos\omega$ and $\sqrt{e}\sin\omega$. To model the velocities, we add the planet mass M_p, a radial velocity offset γ_{rel}, and a white noise term for each year of data, σ_{Y1} and σ_{Y2} to account for the stellar noise being noticeably different from the first year to the next. The semi-amplitude of the planetary signature K_{amp} was computed from the above parameters. The orbital solution and the associated likelihood from the fit to the data are computed from K_{amp}, T_c, P_{orb}, $\sqrt{e}\cos\omega$ and $\sqrt{e}\sin\omega$ via the radvel package (Fulton et al. 2018).

We also try to add a Gaussian Process using a Quasi-Periodic kernel, implemented through radvel to model the stellar noise apparent in the data. The resulting parameter values do not yield a significant difference, therefore not justifying the necessity to use a correlated noise model to account for the stellar intrinsic variability seen in the radial velocities. With one datapoint a day at most, the sampling is too sparse for the Gaussian Process to correctly grasp the ~ 4 days stellar period. Crudely assuming a spot covering 0.6-1.2% (δ_{spot}) of the stellar surface, we can approximate an activity induced radial velocity semi-amplitude K_{act} of $v\sin i \times \delta_{spot} \sim 100 - 200$ m s^{-1}, comparable to the jitter level seen in Figure 4.

We attempted to fit a second longer period circular planet to the radial velocities. We used uniform priors for the period ($\mathcal{U}[300 : 2000]$ days), planet mass ($\mathcal{U}[0.002 : 0.1]$ $M_\odot$) and t_0 ($\mathcal{U}[1398 : 3398]$ TBJD). The posterior distribution are not clearly converging, favouring larger periods and smaller masses. With a K_{amp} of ~ 70 m s^{-1}, the best solution is clearly below the activity level and therefore not trustworthy. Long term data is needed to attempt to constrain a longer period companion.

4.3. Spectral energy distribution model

To constrain the host star parameters $R_\star$, $M_\star$, [Fe/H] and T_{eff} we also model the spectral energy distribution of TOI-4562 simultaneously to the transit and radial velocity models. The stellar parameters are modeled using the MESA Isochrones & Stellar Tracks (Paxton et al. 2011, 2013, 2015; Choi et al. 2016). We interpolate evolution tracks using the minimint package (Koposov 2021) against $M_\star$, age, [Fe/H] and the photometric bands B, V, *Gaia* G, Bp, Rp, 2MASS bands J, H, and K. $R_\star$ is derived from the isochrone predicted values for $\log g$ and $M_\star$. To account for uncertainties in the stellar evolution models, we adopt a 4% uncertainty floor in stellar radius, and 5% floor in stellar mass, where appropriate (Tayar et al. 2022). For the effective temperature T_{eff}, we apply a Gaussian prior such that the predicted T_{eff} interpolated from the isochrone is compared against that measured from the CHIRON spectra as an additional likelihood term. Predicted fluxes from the SED model are corrected for interstellar reddening with the PyAstronomy unred package, that uses the parameterization from Fitzpatrick (1999). Extinction is a free parameter, with a maximum value of $E(B-V) = 0.1542$

mag, as estimated from the Schlafly & Finkbeiner (2011) maps over a 5 arcmin radius[1] around TOI-4562. We also incorporate a Gaussian prior on the distance modulus via the observed *Gaia* parallax to TOI-4562. We offset Gaia DR3's parallax value by -0.023861 mas, the parallax zero-point offset estimated using the routine from Lindegren et al. (2021a)[2] and function of ecliptic latitude, magnitude and colour. At each MCMC jump step, the observed spectral energy distribution is compared against the interpolated MIST model predictions for a given tested stellar parameter.

4.4. *Global model*

The global model includes simultaneous fits of the *TESS* and ground-based photometric datasets (4.1), the CHIRON velocities (4.2), and stellar isochrone model (4.3), as shown in Figure 2, 4 and 7 respectively. We explore the best fit parameters and the posterior distribution via the Affine Invariant Markov chain Monte Carlo Ensemble sampler *emcee* (Foreman-Mackey et al. 2013). The resulting parameters for TOI-4562 b are given in Table 2.

The availability of the radial velocities not only allows us to recover TOI-4562 b's mass, but also helps to break the degeneracy between e and ω. Figure 8 illustrates this by showing the resulting posterior distributions for e and ω from our global model versus a model excluding the radial velocities.

Figure 8. Posterior distributions for TOI-4562 b's orbital eccentricity (e) and argument at periapse (ω). Posterior distributions obtained from photometric data (*TESS* + LCO-SAAO) + SED fit versus the complete model including the radial velocities are shown in black and red respectively. The addition of radial velocities improved the constraints on e and ω.

5. DISCUSSIONS AND CONCLUSIONS

We report the discovery of TOI-4562 b, a temperate gas giant on a highly eccentric orbit around a young Sunlike star. The planet has a mass of $3.29^{+1.88}_{-0.82}$ $M_{\rm J}$ and a radius of $1.072^{+0.044}_{-0.043}$ $R_{\rm J}$. With an orbital period of $225.11757^{+0.00027}_{-0.00025}$ days, it is to date the second longest period planet in the *TESS* sample (after TOI-2180b, Dalba et al. 2022). TOI-4562 b' resides in a highly elliptic orbit ($e = 0.81^{+0.05}_{-0.05}$), and has, based on (Spada & Lanzafame 2020), an age younger than the Praesepe and Hyades clusters. A representation of its orbit alongside the inner Solar System planets is shown in Figure 9.

5.1. *Radius evolution*

At the end of their accretion phase, newly formed gas giants are expected to have radii larger than 1 $R_{\rm J}$. As the planet core radiates its primordial internal heat, Jovian mass planets will cool down via Kelvin-Helmholtz

Figure 9. Two dimensional orbit of TOI-4562 b (yellow solid line) compared with Mercury (red dashed line), Venus (purple dashed line), Earth (blue dashed line) and Mars (orange dashed line).

---

[1] Obtained from the NASA/IPAC Infrared Science Archive
[2] https://gitlab.com/icc-ub/public/gaiadr3_zeropoint

Table 2. TOI-4562 b parameters.

Parameters	Description	Priors	Values
Transit parameters			
$T_{c,1}$ [a]	Transit mid-time (BJD_{TDB})	$\mathcal{U}$	$1456.8761^{+0.0012}_{-0.0011}$
$T_{c,2}$ [a]	Transit mid-time (BJD_{TDB})	$\mathcal{U}$	$1681.9934^{+0.0011}_{-0.0011}$
$T_{c,3}$ [a]	Transit mid-time (BJD_{TDB})	$\mathcal{U}$	$2132.2280^{+0.0010}_{-0.0010}$
$T_{c,4}$ [a]	Transit mid-time (BJD_{TDB})	$\mathcal{U}$	$2357.34522^{+0.00083}_{-0.00086}$
$T_{c,5}$ [a]	Transit mid-time (BJD_{TDB})	$\mathcal{U}$	$2582.46250^{+0.00028}_{-0.00024}$
T_c [a]	Derived linear ephemeris	$\mathcal{U}[1456.83, 1456.93]$	$1456.874636^{+0.001172}_{-0.001249}$
$P_{\rm orb}$ [a]	Orbital period (days)	Derived linear ephemeris	$225.11757^{+0.00027}_{-0.00025}$
T_{14}	Transit total duration (hours)	-	4.23 ± 0.14
$R_{\rm p}/R_\star$	Radius ratio	$\mathcal{U}[0, 0.2]$	$0.09450^{+0.00050}_{-0.00042}$
$a/R_\star$	Normalised Semi-major axis, derived from $[M_\star, R_\star, P_{\rm orb}, M_{\rm p}]$	-	$147.1^{+6.6}_{-6.1}$
b	Impact parameter ($a \cos i$ ($R_\star$))	-	$0.29^{+0.24}_{-0.32}$
δ	Transit depth (ppm)	-	9066^{+225}_{-243}
$\sqrt{e} \cos \omega$ [a]	Reparameterization of e and ω	$\mathcal{U}[-1, 1]$	$0.37^{+0.11}_{-0.12}$
$\sqrt{e} \sin \omega$ [a]	Reparameterization of e and ω	$\mathcal{U}[-1, 1]$	$0.808^{+0.046}_{-0.052}$
i	Planet orbit inclination (°)	$\mathcal{U}[84, 90]$	$89.72^{+0.17}_{-0.15}$
$\mu_{1_{TESS}}$ [b]	Quadratic Limb Darkening law coefficient 1 (TESS)	Fixed	0.28
$\mu_{2_{TESS}}$ [b]	Quadratic Limb Darkening law coefficient 2 (TESS)	Fixed	0.29
$\mu_{1_{LCO}}$ [c]	Quadratic Limb Darkening law coefficient 1 (LCO)	Fixed	0.28
$\mu_{2_{LCO}}$ [c]	Quadratic Limb Darkening law coefficient 2 (LCO)	Fixed	0.29
Radial velocities parameters			
K	RV Semi-amplitude ($\rm m\,s^{-1}$)	-	353^{+73}_{-88}
$M_{\rm p}$	Mass ($M_\odot$)	$\mathcal{U}[0, 0.2]$	$1.218^{+0.054}_{-0.048}$
γ_{CHIRON}	RV offset ($\rm m\,s^{-1}$)	$\mathcal{U}[5200, 5400]$	5352 ± 13
σ_{Y1}	RV jitter, first orbit ($\rm m\,s^{-1}$)	$\mathcal{U}[0, 600]$	73^{+15}_{-15}
σ_{Y2}	RV jitter, second orbit ($\rm m\,s^{-1}$)	$\mathcal{U}[0, 600]$	$28.5^{+18.7}_{-20.9}$
Planetary parameters			
R_p	Radius ($R_\oplus$)	-	11.72 ± 0.28
	Radius ($R_{\rm J}$)	-	$1.072^{+0.044}_{-0.043}$
M_p	Mass ($M_\oplus$)	-	768^{+170}_{-165}
	Mass ($M_{\rm J}$)	-	$3.29^{+1.88}_{-0.82}$
e	Eccentricity	-	$0.81^{+0.05}_{-0.05}$
w	Argument at periapse (°)	-	63 ± 5
ρ_p	Density (g cm^{-3})	-	$2.63^{+0.58}_{-0.56}$
a	Semi-major axis (AU)	-	$0.771^{+0.013}_{-0.013}$
$\langle T_{\rm eq} \rangle$	Temporal average equilibrium temperature (K) [d]	-	349 ± 10
$T_{\rm peri}$	Equilibrium temperature at periapsis (K)	-	731^{+55}_{-34}
$T_{\rm apo}$	Equilibrium temperature at apoapsis (K)	-	261^{+2}_{-3}

Priors: $\mathcal{U}\,[a, b]$ uniform priors with boundaries a and b

[a] Parameters common to the Transit and RV models

[b] Adopted at the *TESS* band from Claret (2017), using ATLAS model with $T_{\rm eff} = 6000$K, $\log g = 4.5$ and [M/H] = 0.1 and computed with the least square method (LSM)

[c] Computed for the SDSS i' band from Claret & Bloemen (2011), using the interpolation routine from Eastman et al. (2013) with $T_{\rm eff} = 6000$K, $\log g = 4.5$ and [M/H] = 0.1 and computed with the least square method (LSM)

[d] Computed for an elliptical orbit from Méndez & Rivera-Valentín (2017), using an albedo of $A = 0.4$, $\epsilon = 1$ and $\beta = 0.74$.
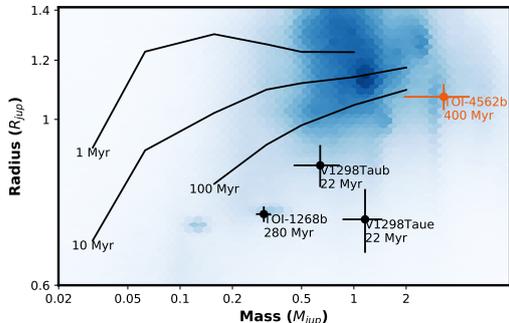
Figure 10. Young gas giants can help constrain the cooling and contraction models. To date, TOI-4562 b is only the fourth Jovian planet younger than 500 Myr to have both its mass and radius measured. The mass-radius of TOI-4562 b is plotted in orange alongside the planets in the V1298 Tau (David et al. 2019a; Suárez Mascareño et al. 2022) and TOI-1268 (Dong et al. 2022; Šubjak et al. 2022) systems. Unlike others, TOI-4562 b sits along the isochrone tracks that model the contraction of young planets (Linder et al. 2019). The mass-radius distribution of other known planets are shown with a density plot in blue in the background.

contraction to $\sim 1\,R_{\rm J}$. Only Hot Jupiters, orbiting extremely close to their parent star are expected to remain inflated due to their increased irradiation. According to cooling models (Baraffe et al. 2003; Fortney et al. 2007; Baraffe et al. 2008; Linder et al. 2019), shown on Figure 10, the most drastic changes in radius occur at the earliest ages. Measuring radii of young gas giants like TOI-4562 b is therefore essential to set constraints on these such models, as emphasized in Fortney et al. (2007).

The current picture is unclear as the recently measured mass of V1298 Tau b & e (Suárez Mascareño et al. 2022) yield much denser planets than predicted at 20 Myr old and require dramatic heavy element enrichment to somewhat reconcile with cooling models (see Figure 10). Conversely, TOI-4562 b's radius is as expected for its age. At the closest approach to its host star (~ 0.18 AU), it receives stellar irradiation of $\sim 9.3 \times 10^{4}\,{\rm W\,m^{-2}}$, or ~ 68 times that of Earth. Although above the $\sim 1.6 \times 10^{4}\,{\rm W\,m^{-2}}$ threshold to trigger inflation, given by Sestovic et al. (2018) for planets more massive than 2.5 $M_{\rm J}$, TOI-4562 b's orbital eccentricity means this level of irradiation affects the planet for a very short fraction of the orbit, not sufficient to trigger radius inflation.

5.2. Dynamical history of TOI-4562 b and benefits of additional follow-up

In its current observed state, TOI-4562 b's semi major axis and eccentricity (see Figure 11) are not in favour of a high eccentricity migration scenario as a circularization of its orbit would take orders of magnitudes longer than the age of the universe ($\tau_{circ} \sim 1 \times 10^{7}$ Gyr, Goldreich & Soter 1966). It is possible, however, that the planet is experiencing ongoing eccentricity cycles and we happen to be observing it at a lower eccentricity. Reduction of the star-planet distance at periastron at the eccentricity peak of such cycles might allow the circularization process to be triggered as described in Dong et al. (2014). Disk-planet interactions can in principle excite the eccentricity of the orbit (Duffell & Chiang 2015) but this is restricted to low ($e \lesssim 0.2$) values, as shown with the red area on Figure 11. Debras et al. (2021) proposed that migration inside wide gaps carved in protoplanetary disks could result in gas giants with eccentricities up to 0.4. This is still insufficient to explain the very high eccentricity from TOI-4562 b's orbit.

Another possible scenario to account for TOI-4562 b's very high eccentricity is in-situ formation (or alternatively, smooth disk migration), followed by excitation from a companion. This can occur via secular interactions, or slow angular momentum exchanges with another body located further out, either periodically through e.g., von Zeipel-Lidov-Kozai cycles (von Zeipel 1910; Lidov 1962; Kozai 1962; Naoz 2016; Nagasawa et al. 2008) or chaotically in secular chaos (Wu & Lithwick 2011; Hamers et al. 2017). High eccentricity can also be triggered sporadically in planet-planet scattering (Weidenschilling & Marzari 1996; Rasio & Ford 1996; Ford & Rasio 2006; Chatterjee et al. 2008), or stellar flybys (Shara et al. 2016; Rodet et al. 2021). Planet-planet scattering could have happened quickly and potentially early if triggered by the dissipation of the gas disk or if the planets were initially closely spaced. Constraints on an outer companion (if not ejected as a result of scattering) could provide crucial insights on dynamical evolution timescales give the young age of the system.

The five transits of TOI-4562 b show modest deviation from a linear ephemeris fit on the 5 – 20 min level (see Figure 12). This potential detection of a transit timing variation signal suggests the presence of a companion in the system, to which TOI-4562 b probably owes its high eccentricity. The existing data are not sufficient to set meaningful constraints on the companion and most configurations for period (i.e., inner or outer companion), eccentricity and mutual inclination remain possible. TOI-4562 b will be observed by *TESS* again in its second extended mission in 2023. In Table 3, we
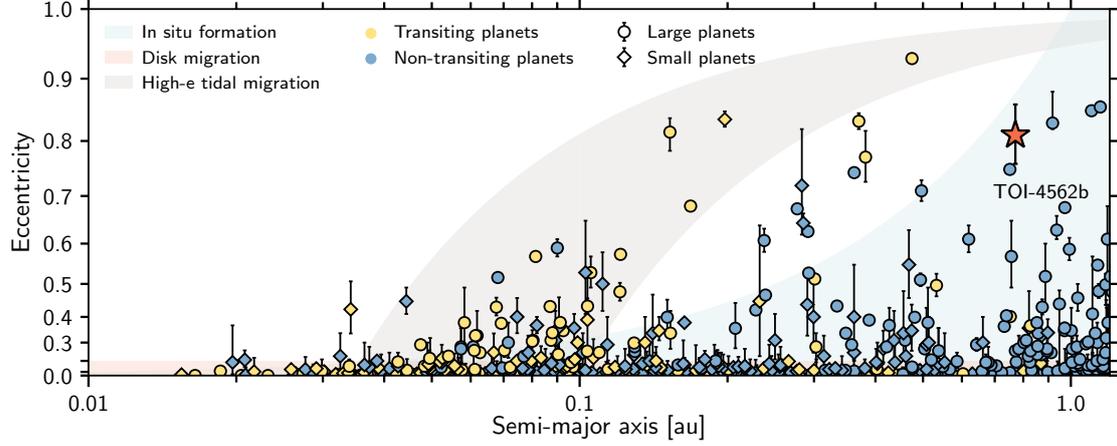
Figure 11. Eccentricity versus semi-major axis for all confirmed planets (Obtained from the NASA exoplanet archive 13 Feb. 2022) with $M_p < 13\,M_J$. The vertical coordinate is scaled to e^2 to emphasize non-circular planets. Shaded areas highlight different formation scenarios. Planets in the grey region are on the path of high-eccentricity migration, with a final semi major axis between 0.034 and 0.1 au. The upper and lower bounds of this region are set by the Roche limit and the circularization timescale respectively. Disk migration, expected to only marginally excite orbital eccentricity is shown as the red shaded region. Finally, in situ formation, with eccentricity excited by, e.g., planet-planet scattering is shown in blue. Transiting versus non-transiting planets are labeled in yellow and blue respectively. Circles are representing larger ($R_p > 6\,R_\oplus$ and/or $M_p > 100\,M_\oplus$) planets and diamonds smaller planets ($R_p < 6\,R_\oplus$ and/or $M_p < 100\,M_\oplus$). Only planets with $e \geq 0.2$ and with uncertainties on e smaller than 50% of the measured e or planets with $e < 0.2$ and with uncertainties less than 0.2 are shown.

Table 3. Next 10 transit opportunities for TOI-4562b.

Transit mid-time (BJD)	Transit date	Visible from (Partial (P) or Full (F))	TESS simultaneity
2460032.6977	29-Mar-2023	Paranal (P)	Y
2460257.8142	9-Nov-2023	Paranal (P)	TBD
2460482.9307	21-Jun-2024	MKO (P) & **ASTEP (F)**	TBD
2460708.0473	1-Feb-2025	**MKO (F)**	TBD
2460933.1638	14-Sep-2025	MKO (P) & **ASTEP (F)**	TBD
2461158.2803	27-Apr-2026	SAAO (P) & RUN (P) & **ASTEP (F)**	TBD
2461383.3968	8-Dec-2026	SAAO (P) & **RUN (F)**	TBD
2461608.5134	27-Jul-2027	ASTEP (P)	TBD
2461833.6299	3-Mar-2028	SAAO (P) & **Paranal (F)**	TBD

Locations SAAO: South African Astronomical Observatory, South Africa (latitude = -32.379444, longitude = -339.189306), Paranal: European Southern Observatory at Paranal, Chile (latitude = -24.625, longitude = -70.403333), MKO: Mt. Kent Observatory, Australia (latitude = -27.797861, longitude = 151.855417), RUN: Observatoire astronomique des Makes, Reunion Island (latitude = -21.199359, longitude = 55.409464), ASTEP: Antarctic Search for Transiting ExoPlanets, Dome C, Antarctica (latitude = -75.09978, longitude = 123.332196)
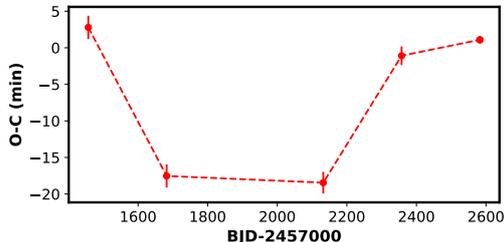
Figure 12. Observed - Calculated mid-transit time for the 5 transits of TOI-4562 b, in minutes. The second and third transits (from TESS Sectors 13 and 30) show a ~ 20 min mid-transit time difference with the other transits, suggesting the presence of a third body in the system.

show future opportunities to continue monitoring transits of TOI-4562 b in the years to come. Combining these with long-term radial velocity follow-up might enable us to unravel the 3-D architecture and dynamical history of this system, as has been successfully performed for Kepler-419 b & c (Dawson et al. 2012, 2014).

The orbital astrometric motion of an outer companion could be retrieved from *Gaia* in the upcoming release of astrometric solutions for ~ 1.3 billion stars (Lindegren et al. 2021b). When archival Hipparchos and *Gaia* observations have been analysed jointly for previous brighter systems (e.g., Venner et al. 2021), astrometric accelerations have often yielded constraints for outer stellar massed companions to key exoplanet systems. Additional *Gaia* observations over the next ~ 10 years will allow us to achieve similar constraints for TOI-4562. Combined with the diffraction limited adaptive optics observations estimated to reach ~ 35 au (see section 2.5), these constraints can inform the presence of exterior stellar companions and provide means to distinguish between evolution scenarios.

Another candidate tracer for dynamical history is the angle between the star's rotation axis and the planet's orbital axis, or (sky projected) obliquity. From $P_\star$, $R_\star$ and $v \sin i$, we estimate the stellar inclination with respect to the line of sight to have a 3σ lower bound of 70° as per Masuda & Winn (2020), consistent with being well aligned. Similarly to other planetary characteristics, the young (< 1 Gyr) end of the obliquity distribution is under sampled. Recent measurements resulting from *TESS* discoveries reveal a remarkable systematic alignment of young systems, including the Jupiter-sized planet HIP 67522 b (Rizzuto et al. 2020; Heitzmann et al. 2021), as well as a number of smaller planets (e.g., AU Mic b & c; Plavchan et al. (2020); Palle et al. (2020); Martioli et al. (2020); Hirano et al. (2020);

Addison et al. (2021), DS Tuc Ab; Newton et al. (2019); Zhou et al. (2020); Montet et al. (2020), TOI 942 b & c Wirth et al. (2021), and TOI 251 Zhou et al. (2021)). The estimated amplitude of the Rossiter McLaughlin effect (Rossiter 1924; McLaughlin 1924) for TOI-4562 b is $\Delta V \sim 70$–150 m s^{-1}. Given the ~ 4 hours transit duration, combined with a brightness of $V = 12.098$ and a rotational broadening of $v \sin i = 17.5$ km s^{-1}, this is well within the grasp of a 4m-class telescope and such an eccentric system would provide a precious addition to the age-obliquity distribution. It is important to note that the long orbital period remains a major obstacle to transit spectroscopy for ground-based facilities.

In the coming years, we aim to conduct extensive follow-ups of the TOI-4562 system to unravel the full architecture of the system and potentially provide insights into the processes shaping the current gas giant planet distribution. Such follow-up will include radial velocities, ground and space based photometry, astrometry and transit spectroscopy for obliquity measurements and/or atmospheric characterisation.

Heitzmann et al.

Table 4. CHIRON radial velocities for TOI-4562. The two left columns cover TOI-4562 b's first orbit (late 2020 to mid 2021) and the two rights columns the second orbit (late 2021 to early 2022).

BJD	RV ($\mathrm{m\,s^{-1}}$)	BJD	RV ($\mathrm{m\,s^{-1}}$)
2457919.279641	5362.7 ± 78.4	2457951.584780	5312.4 ± 84.9
2457919.975844	5384.0 ± 107.5	2457951.882422	5298.0 ± 99.8
2457920.077478	5338.8 ± 64.0	2457952.183280	5356.5 ± 69.2
2457920.770002	5472.3 ± 80.6	2457952.483116	5341.9 ± 81.1
2457923.866595	5150.9 ± 69.1	2457952.785734	5291.7 ± 116.4
2457924.163829	5375.4 ± 96.4	2457953.181198	5283.6 ± 75.4
2457926.463031	5361.0 ± 110.4	2457953.478271	5374.3 ± 68.4
2457926.769474	5386.8 ± 75.5	2457953.779370	5412.8 ± 62.9
2457927.459646	5307.2 ± 64.6	2457954.078933	5331.6 ± 90.0
2457927.755236	5311.9 ± 93.8	2457954.477875	5593.4 ± 68.6
2457928.059467	5417.3 ± 78.6	2457954.770163	5464.1 ± 66.4
2457928.466230	5204.1 ± 138.8	2457955.179212	5513.8 ± 94.8
2457928.761200	5199.1 ± 80.0	2457955.280537	5461.7 ± 89.2
2457929.259193	5241.1 ± 77.4	2457955.375845	5543.3 ± 92.6
2457929.558164	5370.1 ± 111.4	2457955.482781	5522.4 ± 82.3
2457929.953730	5152.4 ± 83.3	2457955.578297	5313.6 ± 101.5
2457930.253505	5155.4 ± 127.6	2457955.677910	5531.5 ± 77.8
2457930.457509	5435.7 ± 74.4	2457955.978882	5386.7 ± 76.3
2457930.752931	5245.4 ± 135.4	2457956.078722	5357.4 ± 49.4
2457931.049893	5485.1 ± 54.3	2457956.174936	5429.0 ± 78.6
2457931.459114	5335.7 ± 122.9	2457956.268873	5451.4 ± 128.8
2457931.852818	5329.3 ± 89.8	2457956.477261	5454.5 ± 88.7
2457932.354462	5475.9 ± 87.3	2457956.567821	5509.4 ± 74.6
2457932.753588	4941.1 ± 191.1	2457956.671976	5411.0 ± 75.6
2457932.950353	5350.4 ± 77.1	2457956.776594	5372.3 ± 78.1
2457933.157726	5182.6 ± 112.6	2457956.872101	5522.2 ± 121.3
2457933.548574	5679.6 ± 112.6	2457956.968008	5465.6 ± 120.6
2457934.251269	5352.0 ± 77.5	2457957.070850	5581.7 ± 80.9
2457934.548758	5707.1 ± 88.4	2457957.173735	5413.1 ± 59.1
2457934.949392	5777.3 ± 104.8	2457957.273394	5376.5 ± 83.0
2457935.148356	5684.9 ± 160.8	2457959.269741	5373.7 ± 120.7
2457935.644737	5348.6 ± 74.5	2457959.369361	5139.8 ± 94.1
2457935.744986	5126.8 ± 101.0	2457959.565821	5321.7 ± 74.9
2457936.045174	5124.1 ± 104.9	2457959.662741	5260.0 ± 81.5
2457936.146598	5314.9 ± 84.3	2457959.765341	5220.0 ± 60.7
2457936.345446	5322.3 ± 119.3	2457959.861135	5383.7 ± 87.1
2457936.438155	5581.4 ± 114.3	2457959.969637	5298.7 ± 108.3
2457936.545021	5093.0 ± 100.9	2457960.060741	5323.5 ± 60.7
2457936.646301	5255.5 ± 74.4	2457960.162519	5330.1 ± 87.7
2457936.744909	5271.4 ± 71.9	2457960.265488	5354.1 ± 53.1
2457936.846528	5428.8 ± 89.1	-	-
2457936.947656	5270.0 ± 69.6	-	-
2457937.046070	5415.1 ± 119.0	-	-
2457937.145298	5446.5 ± 132.6	-	-

Facility: TESS, Exoplanet Archive, CTIO 1.5 m, LCOGT, Gemini:Zorro, CTIO SOAR, ESO 2.2 m

Software: AstroImageJ (Collins et al. 2017), astropy (Astropy Collaboration et al. 2013a, 2018a), batman (Kreidberg 2015), emcee (Foreman-Mackey et al. 2013), pyastronomy (Czesla et al. 2019), comove (https://github.com/adamkraus/Comove), pyphot (https://mfouesneau.github.io/pyphot/), radvel (Fulton et al. 2018), scikit-learn (Pedregosa et al. 2011), minimint (https://zenodo.org/record/4900576), numpy (Harris et al. 2020), matplotlib (Hunter 2007), astropy (Astropy Collaboration et al. 2013b, 2018b), unred (https://github.com/pbrus/unredden-stars), pandas (pandas development team 2020), corner (Foreman-Mackey 2016)

REFERENCES

Addison, B. C., Horner, J., Wittenmyer, R. A., et al. 2021, AJ, 162, 137, doi: 10.3847/1538-3881/ac1685

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013a, A&A, 558, A33, doi: 10.1051/0004-6361/201322068

—. 2013b, A&A, 558, A33, doi: 10.1051/0004-6361/201322068

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018a, AJ, 156, 123, doi: 10.3847/1538-3881/aabc4f

Astropy Collaboration, Price-Whelan, A. M., SipHocz, B. M., et al. 2018b, aj, 156, 123, doi: 10.3847/1538-3881/aabc4f

Baraffe, I., Chabrier, G., & Barman, T. 2008, A&A, 482, 315, doi: 10.1051/0004-6361:20079321

Baraffe, I., Chabrier, G., Barman, T. S., Allard, F., & Hauschildt, P. H. 2003, A&A, 402, 701, doi: 10.1051/0004-6361:20030252

Barnes, S. A. 2007, ApJ, 669, 1167, doi: 10.1086/519295

Baruteau, C., Crida, A., Paardekooper, S. J., et al. 2014, in Protostars and Planets VI, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 667, doi: 10.2458/azu_uapress_9780816531240-ch029

Batygin, K., & Stevenson, D. J. 2010, ApJL, 714, L238, doi: 10.1088/2041-8205/714/2/L238

Bianchi, L., Shiao, B., & Thilker, D. 2017, ApJS, 230, 24, doi: 10.3847/1538-4365/aa7053

Boisse, I., Moutou, C., Vidal-Madjar, A., et al. 2009, A&A, 495, 959, doi: 10.1051/0004-6361:200810648

Bouma, L. G., Hartman, J. D., Brahm, R., et al. 2020, AJ, 160, 239, doi: 10.3847/1538-3881/abb9ab

Bouma, L. G., Curtis, J. L., Masuda, K., et al. 2022a, AJ, 163, 121, doi: 10.3847/1538-3881/ac4966

Bouma, L. G., Kerr, R., Curtis, J. L., et al. 2022b, arXiv e-prints, arXiv:2205.01112. https://arxiv.org/abs/2205.01112

Bouma, L. G., Curtis, J. L., Masuda, K., et al. 2022, The Astronomical Journal, 163, 121, doi: 10.3847/1538-3881/ac4966

Bouvier, J., Barrado, D., Moraux, E., et al. 2018, A&A, 613, A63, doi: 10.1051/0004-6361/201731881

Brahm, R., Jordán, A., & Espinoza, N. 2017a, PASP, 129, 034002, doi: 10.1088/1538-3873/aa5455

Brahm, R., Jordán, A., Hartman, J., & Bakos, G. 2017b, MNRAS, 467, 971, doi: 10.1093/mnras/stx144

Brown, T. M., Baliber, N., Bianco, F. B., et al. 2013, PASP, 125, 1031, doi: 10.1086/673168

Buchhave, L. A., Latham, D. W., Johansen, A., et al. 2012, Nature, 486, 375, doi: 10.1038/nature11121

Buder, S., Sharma, S., Kos, J., et al. 2021, MNRAS, 506, 150, doi: 10.1093/mnras/stab1242

Castelli, F., & Kurucz, R. L. 2004, ArXiv Astrophysics e-prints

Chabrier, G., Johansen, A., Janson, M., & Rafikov, R. 2014, in Protostars and Planets VI, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 619, doi: 10.2458/azu_uapress_9780816531240-ch027

Chatterjee, S., Ford, E. B., Matsumura, S., & Rasio, F. A. 2008, ApJ, 686, 580, doi: 10.1086/590227

Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102, doi: 10.3847/0004-637X/823/2/102

Claret, A. 2017, A&A, 600, A30, doi: 10.1051/0004-6361/201629705

Claret, A., & Bloemen, S. 2011, A&A, 529, A75, doi: 10.1051/0004-6361/201116451

Coelho, P. R. T. 2014, MNRAS, 440, 1027, doi: 10.1093/mnras/stu365

Collins, K. A., Kielkopf, J. F., Stassun, K. G., & Hessman, F. V. 2017, AJ, 153, 77, doi: 10.3847/1538-3881/153/2/77

Cummings, J. D., Deliyannis, C. P., Maderak, R. M., & Steinhauer, A. 2017, AJ, 153, 128, doi: 10.3847/1538-3881/aa5b86

Curtis, J. L., Agüeros, M. A., Douglas, S. T., & Meibom, S. 2019, ApJ, 879, 49, doi: 10.3847/1538-4357/ab2393

Cutri, R. M., Wright, E. L., Conrow, T., et al. 2021, VizieR Online Data Catalog, II/328

Czesla, S., Schröter, S., Schneider, C. P., et al. 2019, PyA: Python astronomy-related packages. http://ascl.net/1906.010

Dalba, P. A., Kane, S. R., Dragomir, D., et al. 2022, AJ, 163, 61, doi: 10.3847/1538-3881/ac415b

David, T. J., Petigura, E. A., Luger, R., et al. 2019a, ApJL, 885, L12, doi: 10.3847/2041-8213/ab4c99

David, T. J., Cody, A. M., Hedges, C. L., et al. 2019b, AJ, 158, 79, doi: 10.3847/1538-3881/ab290f

Dawson, R. I., & Johnson, J. A. 2018, ARA&A, 56, 175, doi: 10.1146/annurev-astro-081817-051853

Dawson, R. I., Johnson, J. A., Morton, T. D., et al. 2012, ApJ, 761, 163, doi: 10.1088/0004-637X/761/2/163

Dawson, R. I., Johnson, J. A., Fabrycky, D. C., et al. 2014, ApJ, 791, 89, doi: 10.1088/0004-637X/791/2/89

Debras, F., Baruteau, C., & Donati, J.-F. 2021, MNRAS, 500, 1621, doi: 10.1093/mnras/staa3397

Dong, J., Huang, C. X., Zhou, G., et al. 2022, ApJL, 926, L7, doi: 10.3847/2041-8213/ac4da0

Dong, S., Katz, B., & Socrates, A. 2014, ApJL, 781, L5, doi: 10.1088/2041-8205/781/1/L5

Douglas, S. T., Agüeros, M. A., Covey, K. R., et al. 2016, ApJ, 822, 47, doi: 10.3847/0004-637X/822/1/47

Douglas, S. T., Curtis, J. L., Agüeros, M. A., et al. 2019, ApJ, 879, 100, doi: 10.3847/1538-4357/ab2468

Duffell, P. C., & Chiang, E. 2015, ApJ, 812, 94, doi: 10.1088/0004-637X/812/2/94

Duncan, D. K., Vaughan, A. H., Wilson, O. C., et al. 1991, ApJS, 76, 383, doi: 10.1086/191572

Eastman, J., Gaudi, B. S., & Agol, E. 2013, PASP, 125, 83, doi: 10.1086/669497

Fitzpatrick, E. L. 1999, PASP, 111, 63, doi: 10.1086/316293

Ford, E. B., & Rasio, F. A. 2006, ApJL, 638, L45, doi: 10.1086/500734

Foreman-Mackey, D. 2016, JOSS, 1, 24, doi: 10.21105/joss.00024

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: 10.1086/670067

Fortney, J. J., Marley, M. S., & Barnes, J. W. 2007, ApJ, 659, 1661, doi: 10.1086/512120

Fulton, B. J., Petigura, E. A., Blunt, S., & Sinukoff, E. 2018, PASP, 130, 044504, doi: 10.1088/1538-3873/aaaaa8

Gagné, J., Mamajek, E. E., Malo, L., et al. 2018, ApJ, 856, 23, doi: 10.3847/1538-4357/aaae09

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1, doi: 10.1051/0004-6361/201629272

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, A&A, 649, A1, doi: 10.1051/0004-6361/202039657

Gillen, E., Briegal, J. T., Hodgkin, S. T., et al. 2020, MNRAS, 492, 1008, doi: 10.1093/mnras/stz3251

Goldreich, P., & Soter, S. 1966, Icarus, 5, 375, doi: 10.1016/0019-1035(66)90051-0

Guerrero, N. M., Seager, S., Huang, C. X., et al. 2021, ApJS, 254, 39, doi: 10.3847/1538-4365/abefe1

Guillot, T., & Showman, A. P. 2002, A&A, 385, 156, doi: 10.1051/0004-6361:20011624

Guillot, T., Abe, L., Agabi, A., et al. 2015, Astronomische Nachrichten, 336, 638, doi: 10.1002/asna.201512174

Hamers, A. S., Antonini, F., Lithwick, Y., Perets, H. B., & Portegies Zwart, S. F. 2017, MNRAS, 464, 688, doi: 10.1093/mnras/stw2370

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357, doi: 10.1038/s41586-020-2649-2

Heitzmann, A., Zhou, G., Quinn, S. N., et al. 2021, ApJL, 922, L1, doi: 10.3847/2041-8213/ac3485

Henden, A. A., Templeton, M., Terrell, D., et al. 2016, VizieR Online Data Catalog, II/336

Hirano, T., Krishnamurthy, V., Gaidos, E., et al. 2020, ApJL, 899, L13, doi: 10.3847/2041-8213/aba6eb

Howell, S. B., Everett, M. E., Sherry, W., Horch, E., & Ciardi, D. R. 2011, AJ, 142, 19, doi: 10.1088/0004-6256/142/1/19

Howell, S. B., & Furlan, E. 2022, Frontiers in Astronomy and Space Sciences, 9, 871163, doi: 10.3389/fspas.2022.871163

Huang, C. X., Vanderburg, A., Pál, A., et al. 2020, Research Notes of the American Astronomical Society, 4, 204, doi: 10.3847/2515-5172/abca2e

Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90, doi: 10.1109/MCSE.2007.55

Jayasinghe, T., Stanek, K. Z., Kochanek, C. S., et al. 2019, MNRAS, 485, 961, doi: 10.1093/mnras/stz444

Jenkins, J. M. 2002, ApJ, 575, 493, doi: 10.1086/341136

Jenkins, J. M., Tenenbaum, P., Seader, S., et al. 2020, Kepler Data Processing Handbook: Transiting Planet Search, Kepler Science Document KSCI-19081-003

Jenkins, J. M., Chandrasekaran, H., McCauliff, S. D., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7740, Software and Cyberinfrastructure for Astronomy, ed. N. M. Radziwill & A. Bridger, 77400D, doi: 10.1117/12.856764

Jenkins, J. M., Twicken, J. D., McCauliff, S., et al. 2016, in Proc. SPIE, Vol. 9913, Software and Cyberinfrastructure for Astronomy IV, 99133E, doi: 10.1117/12.2233418

Kaufer, A., Stahl, O., Tubbesing, S., et al. 1999, The Messenger, 95, 8

Koposov, S. 2021, segasai/minimint: Minimint 0.3.0, v0.3.0, Zenodo, doi: 10.5281/zenodo.5610692

Kounkel, M., & Covey, K. 2019, AJ, 158, 122, doi: 10.3847/1538-3881/ab339a

Kovács, G., Zucker, S., & Mazeh, T. 2002, A&A, 391, 369, doi: 10.1051/0004-6361:20020802

Kozai, Y. 1962, AJ, 67, 591, doi: 10.1086/108790

Kreidberg, L. 2015, PASP, 127, 1161, doi: 10.1086/683602

Lanza, A. F., Rodono, M., & Zappala, R. A. 1993, A&A, 269, 351

Li, J., Tenenbaum, P., Twicken, J. D., et al. 2019, PASP, 131, 024506, doi: 10.1088/1538-3873/aaf44d

Lidov, M. L. 1962, Planet. Space Sci., 9, 719, doi: 10.1016/0032-0633(62)90129-0

Lindegren, L., Bastian, U., Biermann, M., et al. 2021a, A&A, 649, A4, doi: 10.1051/0004-6361/202039653

Lindegren, L., Klioner, S. A., Hernández, J., et al. 2021b, A&A, 649, A2, doi: 10.1051/0004-6361/202039709

Linder, E. F., Mordasini, C., Mollière, P., et al. 2019, A&A, 623, A85, doi: 10.1051/0004-6361/201833873

Lomb, N. R. 1976, Ap&SS, 39, 447, doi: 10.1007/BF00648343

Mamajek, E. E., & Hillenbrand, L. A. 2008, ApJ, 687, 1264, doi: 10.1086/591785

Mandel, K., & Agol, E. 2002, ApJL, 580, L171, doi: 10.1086/345520

Mann, A. W., Johnson, M. C., Vanderburg, A., et al. 2020, AJ, 160, 179, doi: 10.3847/1538-3881/abae64

Mann, A. W., Wood, M. L., Schmidt, S. P., et al. 2021, arXiv e-prints, arXiv:2110.09531. https://arxiv.org/abs/2110.09531

Martioli, E., Hébrard, G., Moutou, C., et al. 2020, A&A, 641, L1, doi: 10.1051/0004-6361/202038695

Masuda, K., & Winn, J. N. 2020, AJ, 159, 81, doi: 10.3847/1538-3881/ab65be

McCully, C., Volgenau, N. H., Harbeck, D.-R., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10707, Proc. SPIE, 107070K, doi: 10.1117/12.2314340

McLaughlin, D. B. 1924, ApJ, 60, 22, doi: 10.1086/142826

Meibom, S., Barnes, S. A., Platais, I., et al. 2015, Nature, 517, 589, doi: 10.1038/nature14118

Mékarnia, D., Guillot, T., Rivet, J. P., et al. 2016, MNRAS, 463, 45, doi: 10.1093/mnras/stw1934

Méndez, A., & Rivera-Valentín, E. G. 2017, ApJL, 837, L1, doi: 10.3847/2041-8213/aa5f13

Messina, S., Nardiello, D., Desidera, S., et al. 2022, A&A, 657, L3, doi: 10.1051/0004-6361/202142276

Montet, B. T., Feinstein, A. D., Luger, R., et al. 2020, AJ, 159, 112, doi: 10.3847/1538-3881/ab6d6d

Morris, R. L., Twicken, J. D., Smith, J. C., et al. 2020, Kepler Data Processing Handbook: Photometric Analysis, Kepler Data Processing Handbook (KSCI-19081-003)

Nagasawa, M., Ida, S., & Bessho, T. 2008, ApJ, 678, 498, doi: 10.1086/529369

Nagasawa, M., Lin, D. N. C., & Ida, S. 2003, ApJ, 586, 1374, doi: 10.1086/367884

Naoz, S. 2016, ARA&A, 54, 441, doi: 10.1146/annurev-astro-081915-023315

Newton, E. R., Mann, A. W., Tofflemire, B. M., et al. 2019, ApJL, 880, L17, doi: 10.3847/2041-8213/ab2988

Newton, E. R., Rampalli, R., Kraus, A. L., et al. 2022, arXiv e-prints, arXiv:2206.06254. https://arxiv.org/abs/2206.06254

Noyes, R. W., Hartmann, L. W., Baliunas, S. L., Duncan, D. K., & Vaughan, A. H. 1984, ApJ, 279, 763, doi: 10.1086/161945

Oh, S., Price-Whelan, A. M., Hogg, D. W., Morton, T. D., & Spergel, D. N. 2017, AJ, 153, 257, doi: 10.3847/1538-3881/aa6ffd

Palle, E., Oshagh, M., Casasayas-Barris, N., et al. 2020, A&A, 643, A25, doi: 10.1051/0004-6361/202038583

pandas development team, T. 2020, pandas-dev/pandas: Pandas, latest, Zenodo, doi: 10.5281/zenodo.3509134

Paredes, L. A., Henry, T. J., Quinn, S. N., et al. 2021, AJ, 162, 176, doi: 10.3847/1538-3881/ac082a

Paxton, B., Bildsten, L., Dotter, A., et al. 2011, ApJS, 192, 3, doi: 10.1088/0067-0049/192/1/3

Paxton, B., Cantiello, M., Arras, P., et al. 2013, ApJS, 208, 4, doi: 10.1088/0067-0049/208/1/4

Paxton, B., Marchant, P., Schwab, J., et al. 2015, ApJS, 220, 15, doi: 10.1088/0067-0049/220/1/15

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Plavchan, P., Barclay, T., Gagné, J., et al. 2020, Nature, 582, 497, doi: 10.1038/s41586-020-2400-z

Pollacco, D. L., Skillen, I., Collier Cameron, A., et al. 2006, PASP, 118, 1407, doi: 10.1086/508556

Rasio, F. A., & Ford, E. B. 1996, Science, 274, 954, doi: 10.1126/science.274.5289.954

Rauer, H., Catala, C., Aerts, C., et al. 2014, Experimental Astronomy, 38, 249, doi: 10.1007/s10686-014-9383-4

Rebull, L. M., Stauffer, J. R., Bouvier, J., et al. 2016, AJ, 152, 113, doi: 10.3847/0004-6256/152/5/113

Reinhold, T., Reiners, A., & Basri, G. 2013, A&A, 560, A4, doi: 10.1051/0004-6361/201321970

Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003, doi: 10.1117/1.JATIS.1.1.014003

Rizzuto, A. C., Newton, E. R., Mann, A. W., et al. 2020, AJ, 160, 33, doi: 10.3847/1538-3881/ab94b7

Rodet, L., Su, Y., & Lai, D. 2021, ApJ, 913, 104, doi: 10.3847/1538-4357/abf8a7

Rossiter, R. A. 1924, ApJ, 60, 15, doi: 10.1086/142825

Scargle, J. D. 1982, ApJ, 263, 835, doi: 10.1086/160554

Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103, doi: 10.1088/0004-637X/737/2/103

Sestovic, M., Demory, B.-O., & Queloz, D. 2018, A&A, 616, A76, doi: 10.1051/0004-6361/201731454

Shappee, B., Prieto, J., Stanek, K. Z., et al. 2014, in American Astronomical Society Meeting Abstracts, Vol. 223, American Astronomical Society Meeting Abstracts #223, 236.03

Shara, M. M., Hurley, J. R., & Mardling, R. A. 2016, ApJ, 816, 59, doi: 10.3847/0004-637X/816/2/59

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163, doi: 10.1086/498708

Soderblom, D. R. 2010, ARA&A, 48, 581, doi: 10.1146/annurev-astro-081309-130806

Spada, F., & Lanzafame, A. C. 2020, A&A, 636, A76, doi: 10.1051/0004-6361/201936384

Stassun, K. G., Oelkers, R. J., Pepper, J., et al. 2018, AJ, 156, 102, doi: 10.3847/1538-3881/aad050

Suárez Mascareño, A., Damasso, M., Lodieu, N., et al. 2022, Nature Astronomy, 6, 232, doi: 10.1038/s41550-021-01533-7

Tayar, J., Claytor, Z. R., Huber, D., & van Saders, J. 2022, ApJ, 927, 31, doi: 10.3847/1538-4357/ac4bbc

Tofflemire, B. M., Rizzuto, A. C., Newton, E. R., et al. 2021, AJ, 161, 171, doi: 10.3847/1538-3881/abdf53

Tokovinin, A. 2018, PASP, 130, 035002, doi: 10.1088/1538-3873/aaa7d9

Tokovinin, A., Fischer, D. A., Bonati, M., et al. 2013, PASP, 125, 1336, doi: 10.1086/674012

Twicken, J. D., Clarke, B. D., Bryson, S. T., et al. 2010, in Proc. SPIE, Vol. 7740, Software and Cyberinfrastructure for Astronomy, 774023, doi: 10.1117/12.856790

Twicken, J. D., Catanzarite, J. H., Clarke, B. D., et al. 2018, PASP, 130, 064502, doi: 10.1088/1538-3873/aab694

Ujjwal, K., Kartha, S. S., Mathew, B., Manoj, P., & Narang, M. 2020, AJ, 159, 166, doi: 10.3847/1538-3881/ab76d6

Vanderburg, A., & Johnson, J. A. 2014, PASP, 126, 948, doi: 10.1086/678764

Venner, A., Vanderburg, A., & Pearce, L. A. 2021, AJ, 162, 12, doi: 10.3847/1538-3881/abf932

von Zeipel, H. 1910, Astronomische Nachrichten, 183, 345, doi: 10.1002/asna.19091832202

Šubjak, J., Endl, M., Chaturvedi, P., et al. 2022, arXiv e-prints, arXiv:2201.13341. https://arxiv.org/abs/2201.13341

Weidenschilling, S. J., & Marzari, F. 1996, Nature, 384, 619, doi: 10.1038/384619a0

Wirth, C. P., Zhou, G., Quinn, S. N., et al. 2021, ApJL, 917, L34, doi: 10.3847/2041-8213/ac13a9

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868, doi: 10.1088/0004-6256/140/6/1868

Wu, Y., & Lithwick, Y. 2011, ApJ, 735, 109, doi: 10.1088/0004-637X/735/2/109

Zhou, G., Winn, J. N., Newton, E. R., et al. 2020, ApJL, 892, L21, doi: 10.3847/2041-8213/ab7d3c

Zhou, G., Quinn, S. N., Irwin, J., et al. 2021, AJ, 161, 2, doi: 10.3847/1538-3881/abba22

Zhou, G., Wirth, C. P., Huang, C. X., et al. 2022, AJ, 163, 289, doi: 10.3847/1538-3881/ac69e3

Ziegler, C., Tokovinin, A., Briceño, C., et al. 2020, AJ, 159, 19, doi: 10.3847/1538-3881/ab55e9

Ziegler, C., Tokovinin, A., Latiolais, M., et al. 2021, AJ, 162, 192, doi: 10.3847/1538-3881/ac17f6

Chapter 6: Discussion and Conclusions

This Ph. D. aimed to advance our understanding of planetary formation and evolution through the lens of young planetary systems. It started with the realisation that the large sample of known exoplanets contained only a very small fraction of youngsters. Not only are such planets rare, but their detection is a major challenge due to the strong natural variability of their host star.

For my first paper, I chose 2 promising strategies, Doppler Imaging and Gaussian Processes, to bench-mark our capability to discover gas giants in short-orbit around young and active stars using radial velocity legacy datasets.

Still focusing on young stars, I took a slight turn for my second publication by measuring the spin–orbit angle of the 17 Myr HIP 67522 b. My knowledge of stellar activity allowed me to build a model that yielded a unambiguous and precise measurement.

Finally, the last publication of this Ph. D. is the discovery of a new young planet, at the edge of the warm Jupiter population and on a highly eccentric orbit. Extensive follow-up combined with a global model allowed strong constraints on its parameters and hinted at the presence of an outer companion rendering this system unique and attractive for further characterisation.

6.1 Gas giants around active young stars: The need for reliable datasets and strong activity mitigation strategies.

Surveys such as BCool (Marsden et al., 2014) or TOUPIES (Folsom et al., 2016, 2018) systematically obtain spectra of fast rotating stars to model their magnetic evolution. Using DI and Zeeman DI, they map the surface brightness and magnetic field distribution of these stars. Although these spectroscopic datasets were not intended for planet searches, they are amenable to produce RVs. In the work presented in chapter 3, we assessed the viability of using such legacy datasets to search for planets and at the same time, give insights on the good practices to increase chances to discover planets when doing 'blind' RV planet searches targeting young stars.

We generated 37 different datasets, each containing a different Jupiter-like planets in short-orbit buried in a real dataset of the young (17–32 Myr) and very active ($\sigma_{RV_{activity}} \sim 400 - 600\,\mathrm{m\,s^{-1}}$) HD 141943. One part of the challenge came from the dataset itself, non-optimized for planet search and with a RV precision of 50–75 $\mathrm{m\,s^{-1}}$, comparable to the RV amplitude from the lightest/furthest planets simulated. The other challenge, which is the reason of being for this work, is stellar variability. We decided upon two strategies, DI and GP, inspired from previous works (e.g. Donati et al., 2016, 2017; Yu et al., 2017, 2019; Klein et al., 2020), to mitigate the stellar activity and try to recover the hidden planets.

We found that DI was able to remove some contribution of the activity and allowed to recover planets RV signatures down to 2.5 times below the activity level. The main drawback of this strategy to be its reliance on a classical pe-

riodogram search post-activity filtering and a False Alarm Probability value, making it difficult to assess the statistical significance of suspected planetary signatures.

The GP proved to be more efficient to remove the stellar contribution. It is also part of a Bayesian framework allowing a more rigorous estimation of the likelihood of the presence of a planet. Although widely used, one of its main difficulty is the estimation of the evidence, a quantity allowing model comparison, i.e., statistically vetting between a model containing only stellar activity versus a combination of stellar activity and one (or more) planet(s). Extremely expensive computationally to evaluate, the evidence can be estimated. Following the results from the community survey on Bayesian evidence estimation (Nelson et al., 2020), we used a Nested sampling approach (using PYMULTINEST, Buchner et al. 2014) to explore our parameter space and sample the posterior distribution. This second approach allowed us to dig up planets down to ~ 4 times below the stellar activity level with confident detection probabilities.

We conclude that finding planets orbiting young active stars solely relying on these type of RV legacy datasets is very challenging, and only the largest planet will be detectable. Adding the fact the young planets are rare (see section 2.4.1) gave us some perspective on the weaknesses of the datasets, which are all addressable with current practices. Even for fast rotating stars, current state-of-the-art instruments can easily provide RV precision one order of magnitude higher than the dataset used in our study. Improvement are also required on the quantitative side, with more datapoints, both periodically densely sampling the stellar rotation period and regularly covering multiple planet periods. This would allow an efficient modelling of the stellar activ-

ity and persistence of the planetary signature. We briefly mention the case of transiting planets, adding very strong prior knowledge on the presence of a planet and granting access to P_{orb}, allowing more efficient planning of observations to observe the planet in quadrature (i.e., at phases 0.25 and 0.75) when the RV is the strongest.

6.2 Gas giants origin: Insights from exoplanets

In papers II & III, we performed observations of young exoplanets. The measurement of the obliquity and HIP 67522 b and the discovery of the very eccentric TOI-4562 b are important contributor to our understanding of the origin of planetary systems.

6.2.1 The aligned orbit of HIP 67522 b

In paper II (Chapter 4), we performed two observing campaigns of the young, hot and Jupiter sized HIP 67522 b (discovered by Rizzuto et al., 2020) using the CHIRON high-resolution echelle spectrograph to measure the spin–orbit alignment of the system. Sky-projected obliquity (λ) measurements are usually done using the Rossiter–McLaughlin (RM) effect (Rossiter, 1924; McLaughlin, 1924), where the transiting planet creates an anomaly in the RV curve at the time of transit. As we now know, stellar activity, strongly exhibited by HIP 67522 due to its 17 Myr of age, deforms the line profiles that are used to infer the RVs. For this analysis, we chose to work directly on the mean line profiles rather than the RVs to retain the information about the stellar activity and filter it more efficiently.

Our global model comprised the photometry to fit TESS data and a model of the line profile made of both a transiting planet (to infer λ) and the model

of a rotating star with surface spots. I designed and implemented the latter model from scratch (see appendix C) which successfully removed part of the stellar noise that was partially hiding the signature of the planet.

The result is a precise recovery of HIP 67522 b's spin–orbit angle of $|\lambda| = 5.8° ^{+2.8}_{-5.7}$. Given its young age and alignment, it seems very unlikely that HIP 67522 b experienced high-eccentricity migration as the circularisation timescale of the orbit (second stage of high-eccentricity migration) is 2-3 orders of magnitudes longer than the age of the system. Also, we do not expect high-eccentricity migration to favour well-aligned orbits. Tidal realignment is also unlikely to have had time to occur. In conclusion, HIP 67522 b might be a prime example of a planet that arrived to its current location by a smooth mechanism, i.e., in-situ or gas disc migration.

Further characterisation of this system, such as a mass measurement (being currently performed from priv. conversation with the investigator of this campaign), search for outer companion and atmospheric characterisation (the science case of the GO 2498[1] programme for cycle 1 of JWST) will bring additional information and constrain HIP 67522 b's past history.

HIP 67522 b belongs to an intriguing population of very young aligned systems. Because measuring young planets obliquities is just a few years old, only 5 systems below 150 Myr of age have their λ constrained. Techniques such as the one presented in paper II will yield a larger sample of young system obliquities needed to validate and investigate this seemingly systematic alignment among young planets. This, in turn, would inform us on the prevalence of originally misaligned disc and formation scenarios.

---

[1] https://www.stsci.edu/jwst/science-execution/program-information.html?id=2498

6.2.2 The highly eccentric TOI-4562 b and its companion

In the last piece of work of this thesis, paper III (Chapter 5), we present a new exoplanet, TOI-4562 b. This gas giant with a 225.1 day had four transits identified in the TESS data. On 2022 January 3rd, we caught a fifth transit from the South African Astronomical Observatory node of the Las Cumbres Observatory Global Network. The addition of 2 years RV follow-up using CHIRON allowed a precise recovery of the mass and eccentricity of the planet.

The youth of the system, informed by the variations seen in both the light curve and the RVs, was constrained using gyrochronology and lithium abundance as TOI-4562 does not appear to belong to any known open cluster or young association. We infer an age of 300–400 Myr.

Our global model, simultaneously accounting for the photometry, the spectroscopy and the stellar energy distribution revealed a planet of $3.29^{+1.88}_{-0.82}$ M_J, $1.072^{+0.044}_{-0.043}$ R_J on an highly eccentric ($e = 0.81 \pm 0.05$), relatively long orbit ($P_{orb} = 225.11757^{+0.00027}_{-0.00025}$ days) around a Sun-like star. With the second longest orbit from TESS to date, TOI-4562 b's extreme eccentricity is intriguing. Despite the large eccentricity, the planet does not orbit close enough to its host star to experience orbit circularisation, therefore discarding the high-eccentricity migration scenario, unless significant oscillation of the eccentricity are at play, occasionally bringing the planet close enough to the star to trigger circularisation. The latter explanation is very challenging to test and deemed less likely than other migration scenarios.

By analysing each transits mid-time, we found that the second and third observed transits occur ~20 minutes earlier than expected. This Transit Timing Variation (TTV) is non-negligible and characteristic of the gravitational influence of an additional planet on TOI-4562 b. Constraining the orbit of

the companion is currently out of reach with only 5 available transits, however this is compelling evidence that the large eccentricity of TOI-4562 b has been excited by this companion.

Similarly to HIP 67522 b, further follow-up of TOI-4562 b will reveal the full extent of this system and provide important clues on its history, as it has been the case for extensively characterised system like Kepler-419 (Dawson et al., 2014). Such follow-ups include: long term RVs to search for a trend left by the potential outer companion, further transits observations to constrain the TTV signal (we note that the long period of TOI-4562 b renders full transits observable from the ground especially rare, so transit follow-up campaigns need to take leverage of space mission such as TESS or CHEOPS), obliquity measurement (although also suffering from the rarity of transit) or atmospheric characterisation. Last but not least, data from the third release form ESA's Gaia mission has the potential to bring (i) strong astrometric constraints on the possible outer companion and (ii) find new or extensions of known stellar young association to which TOI-4562 might belong, providing a much more precise age for this system.

6.3 Prospective and future work

This Ph. D. project is part of a large community effort aimed to find rare young planets and circumvent the barrier resulting from the stellar activity and we started the discussion by asking the two following questions:

1. **Can we reliably detect gas giants in short orbit around young and very active stars using radial velocities?**

   The work from chapter 3 propose a conditional answer. We conclude that although stellar activity will remain a major challenge and stellar

mitigation strategies need to be perfected, high quality datasets can provide means to detect gas giants in short orbit, even around very active stars.

It is essential to improve detection capabilities when targeting young stars to reveal of the true young exoplanet population. A large enough sample is needed to compare very young gas giants in short orbit to their well-studied more mature counterparts. Resulting differences (or similarities) in orbital characteristics or occurrence rates could help to set strong constraints on the timescales of formation scenarios and assess their prevalence. This work suggests that this will soon be possible for such planets thanks to: new generation high resolution spectrographs, the community effort to improve stellar activity mitigation and the expanded young star searching ground allowed by the upcoming discoveries of new young associations by the Gaia mission.

2. **Can the characterisation of young planets provide meaningful insights into the processes shaping the observed variety in the gas giants exoplanet population?**

Paper II & III clearly reinforce the growing consensus in the exoplanet community that the origin of gas giants in short-orbit can not be accounted for by a single mechanism. Although some extremely eccentric planets such as HD 80606 b or TOI 3362 b seem to be strong evidence of the high-eccentricity migration scenario, our obliquity measurement for HIP 67522 b and the localisation in the e vs a parameter space for TOI-4562 b, combined with the likely presence of a companion to the latter, does not support the high eccentricity migration scenario. We believe that a more precise description of HIP 67522 b and TOI-4562 b's

past history are possible with further follow-up in the coming years. As for the bigger picture, the extent to which each scenario contribute to shaping the observed distribution of exoplanets remains to be determined but would be greatly informed by expanding the young planets sample.

To conclude, we describe below some promising avenues that have the potential to bring further insights to the above questions.

IMPROVING STELLAR MITIGATION strategies by integrating other activity diagnostics. It is possible to add complementary datasets such as photometry in different bands, near infrared spectroscopy or activity indicators to the GP model (e.g. Barragán et al., 2022). For DI, photometry (e.g. from TESS) can be used as integrated photometry to constrain the stellar brightness distributions and the parameter exploration is currently still a grid search, and could be improved with a Monte Carlo type framework.

FURTHER CHARACTERISE HIP 67522 b and TOI-4562 b. As previously discussed, our knowledge of both systems remain incomplete, i.e., mass and composition for HIP 67522 b and obliquity, composition and nature of the companion for TOI-4562 b. For HIP 67522 b, these are attainable with high precision extensive RVs observing campaigns combined with efficient stellar activity mitigation strategies. For TOI-4562 b, Gaia astrometric data will be key, along with space based transit follow-up and ground RV long-term monitoring. Both target atmospheric composition (e.g. using JWST) could also provide constrains on e.g. gas giant radii evolution models.

FIND MORE YOUNG PLANETS, and populate the young end of various planetary characteristic distributions. Because of the stellar activity signatures observed in light curves, transits from young planets transits can be missed from automatic light curve transit searches such as the TESS pipeline. Performing careful searches in noisy light curves and developing adapted algorithms will yield to discoveries of young planets (e.g. Fernandes et al. 2022).

FIND MORE YOUNG STARS, by identifying new stellar associations and open clusters and mapping the full extent of the known ones. The 3rd data release from Gaia will yield tangential velocities for half a billion stars and RVs for ~7 Million of them. This unprecedented dataset, will allow to significantly increase the number of stars belonging to an association or cluster with a well constrained age.

References

Note that the references presented here are for chapters 1, 2 and 6 and appendices A to C. References for the papers included in chapters 3 to 5 are included in the references sections of the papers.

Addison, B., Wright, D. J., Wittenmyer, R. A., et al. (2019). Minerva-Australis. I. Design, Commissioning, and First Photometric Results. PASP, 131(1005), 115003.

Addison, B. C., Horner, J., Wittenmyer, R. A., et al. (2021). The youngest planet to have a spin-orbit alignment measurement AU mic b. *The Astronomical Journal*, 162(4), 137.

Addison, B. C., Wright, D. J., Nicholson, B. A., et al. (2021). TOI-257b (HD 19916b): a warm sub-saturn orbiting an evolved F-type star. MNRAS, 502(3), 3704–3722.

Al Moulla, K., Dumusque, X., Cretignier, M., et al. (2022). Measuring precise radial velocities on individual spectral lines. III. Dependence of stellar activity signal on line formation temperature. *arXiv e-prints*, (pp. arXiv:2205.07047).

Albrecht, S., Winn, J. N., Johnson, J. A., et al. (2012). Obliquities of Hot Jupiter Host Stars: Evidence for Tidal Interactions and Primordial Misalignments. ApJ, 757(1), 18.

Albrecht, S. H., Dawson, R. I., & Winn, J. N. (2022). Stellar Obliquities in Exoplanetary Systems. PASP, 134(1038), 082001.

Anderson, K. R., Lai, D., & Pu, B. (2020). In situ scattering of warm Jupiters and implications for dynamical histories. MNRAS, 491(1), 1369–1383.

Anderson, K. R., Storch, N. I., & Lai, D. (2016). Formation and stellar spin-orbit misalignment of hot Jupiters from Lidov-Kozai oscillations in stellar binaries. MNRAS, 456(4), 3671–3701.

Angelo, I., Naoz, S., Petigura, E., et al. (2022). Kepler-1656b's Extreme Eccentricity: Signature of a Gentle Giant. AJ, 163(5), 227.

Antonini, F., Hamers, A. S., & Lithwick, Y. (2016). Dynamical Constraints on the Origin of Hot and Warm Jupiters with Close Friends. AJ, 152(6), 174.

Astropy Collaboration, Price-Whelan, A. M., SipHocz, B. M., et al. (2018). The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. *aj*, 156(3), 123.

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. (2013). Astropy: A community Python package for astronomy. A&A, 558, A33.

Barnes, S. A. (2007). Ages for Illustrative Field Stars Using Gyrochronology: Viability, Limitations, and Errors. ApJ, 669(2), 1167–1189.

Barragán, O., Aigrain, S., Rajpaul, V. M., & Zicher, N. (2022). PYANETI - II. A multidimensional Gaussian process approach to analysing spectroscopic time-series. MNRAS, 509(1), 866–883.

Baruteau, C., Crida, A., Paardekooper, S. J., et al. (2014). Planet-Disk Interactions and Early Evolution of Planetary Systems. In H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning (Eds.), *Protostars and Planets VI* (pp. 667).

Batygin, K. (2012). A primordial origin for misalignments between stellar spin axes and planetary orbits. Nature, 491(7424), 418–420.

Batygin, K., Bodenheimer, P. H., & Laughlin, G. P. (2016). In Situ Formation and Dynamical Evolution of Hot Jupiter Systems. ApJ, 829(2), 114.

Bellotti, S., Petit, P., Morin, J., et al. (2022). Mitigating stellar activity jitter with different line lists for least-squares deconvolution. Analysis of a parametric and a randomised line selection. A&A, 657, A107.

Berdyugina, S. V. (2005). Starspots: A Key to the Stellar Dynamo. *Living Reviews in Solar Physics*, 2(1), 8.

Boisse, I., Bouchy, F., Hébrard, G., et al. (2011). Disentangling between stellar activity and planetary signals. A&A, 528, A4.

Boisse, I., Moutou, C., Vidal-Madjar, A., et al. (2009). Stellar activity of planetary host star HD 189 733. A&A, 495(3), 959–966.

Bouma, L. G., Curtis, J. L., Masuda, K., et al. (2022a). A 38 Million Year Old Neptune-sized Planet in the Kepler Field. AJ, 163(3), 121.

Bouma, L. G., Kerr, R., Curtis, J. L., et al. (2022b). Kepler and the Behemoth: Three Mini-Neptunes in a 40 Million Year Old Association. *arXiv e-prints*, (pp. arXiv:2205.01112).

Brown, D. J. A., Cameron, A. C., Anderson, D. R., et al. (2012). Rossiter-McLaughlin effect measurements for WASP-16, WASP-25 and WASP-31. MNRAS, 423(2), 1503–1520.

Brown, S. F., Donati, J. F., Rees, D., & Semel, M. (1991). Zeeman-Doppler Imaging of active stars IV. Maximum entropy reconstruction of 2D magnetic topologies. *Astronomy & Astrophysics*, (250), 463–474.

Buchner, J., Georgakakis, A., Nandra, K., et al. (2014). X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue. A&A, 564, A125.

Cale, B. L., Reefe, M., Plavchan, P., et al. (2021). Diving Beneath the Sea of Stellar Activity: Chromatic Radial Velocities of the Young AU Mic Planetary System. AJ, 162(6), 295.

Carlos, M., Meléndez, J., Spina, L., et al. (2019). The Li-age correlation: the Sun is unusually Li deficient for its age. MNRAS, 485(3), 4052–4059.

Cegla, H. (2019). The Impact of Stellar Surface Magnetoconvection and Oscillations on the Detection of Temperate, Earth-Mass Planets Around Sun-Like Stars. *Geosciences*, 9(3), 114.

Chabrier, G., Johansen, A., Janson, M., & Rafikov, R. (2014). Giant Planet and Brown Dwarf Formation. In H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning (Eds.), *Protostars and Planets VI* (pp. 619).

Chang, S. H., Gu, P. G., & Bodenheimer, P. H. (2010). Tidal and Magnetic Interactions Between a Hot Jupiter and its Host Star in the Magnetospheric Cavity of a Protoplanetary Disk. ApJ, 708(2), 1692–1702.

Chatterjee, S., Ford, E. B., Matsumura, S., & Rasio, F. A. (2008). Dynamical Outcomes of Planet-Planet Scattering. ApJ, 686(1), 580–602.

Coleman, G. A. L. & Nelson, R. P. (2016). Giant planet formation in radially structured protoplanetary discs. MNRAS, 460(3), 2779–2795.

Collier Cameron, A., Duncan, D. K., Ehrenfreund, P., et al. (1990). Fast spectroscopic variations on rapidly rotating, cool dwarfs III. Masses of circumstellar absorbing clouds on AB Doradus. MNRAS, 247, 415.

Cretignier, M., Dumusque, X., Allart, R., et al. (2020). Measuring precise radial velocities on individual spectral lines. II. Dependance of stellar activity signal on line depth. A&A, 633, A76.

Damasso, M., Lanza, A. F., Benatti, S., et al. (2020). The GAPS Programme at TNG. XXVII. Reassessment of a young planetary system with HARPS-N: is the hot Jupiter V830 Tau b really there? A&A, 642, A133.

David, T. J., Petigura, E. A., Luger, R., et al. (2019). Four Newborn Planets Transiting the Young Solar Analog V1298 Tau. ApJ, 885(1), L12.

Dawson, R. I. & Johnson, J. A. (2018). Origins of Hot Jupiters. ARA&A, 56, 175–221.

Dawson, R. I., Johnson, J. A., Fabrycky, D. C., et al. (2014). Large Eccentricity, Low Mutual Inclination: The Three-dimensional Architecture of a Hierarchical System of Giant Planets. ApJ, 791(2), 89.

Debras, F., Baruteau, C., & Donati, J.-F. (2021). Revisiting migration in a disc cavity to explain the high eccentricities of warm Jupiters. MNRAS, 500(2), 1621–1632.

Desort, M., Lagrange, A. M., Galland, F., et al. (2007). Search for exoplanets with the radial-velocity technique: quantitative diagnostics of stellar activity. A&A, 473(3), 983–993.

Donati, J. F., Bouvier, J., Alencar, S. H., et al. (2020). The magnetic field and accretion regime of CI Tau. MNRAS, 491(4), 5660–5670.

Donati, J. F. & Brown, S. F. (1997). Zeeman-Doppler imaging of active stars. V. Sensitivity of maximum entropy magnetic maps to field orientation. A&A, 326, 1135–1142.

Donati, J. F., Hébrard, E., Hussain, G., et al. (2014). Modelling the magnetic activity and filtering radial velocity curves of young suns: The weak-line T tauri star LkCa 4. *Monthly Notices of the Royal Astronomical Society*, 444(4), 3220–3229.

Donati, J. F., Hébrard, E., Hussain, G., et al. (2014). Modelling the magnetic activity and filtering radial velocity curves of young Suns : the weak-line T Tauri star LkCa 4. MNRAS, 444(4), 3220–3229.

Donati, J. F., Moutou, C., Malo, L., et al. (2016). A hot Jupiter orbiting a 2-million-year-old solar-mass T Tauri star. Nature, 534(7609), 662–666.

Donati, J. F., Yu, L., Moutou, C., et al. (2017). The hot Jupiter of the magnetically active weak-line T Tauri star V830 Tau. MNRAS, 465(3), 3343–3360.

Dong, J., Huang, C. X., Zhou, G., et al. (2021). TOI-3362b: A Proto Hot Jupiter Undergoing High-eccentricity Tidal Migration. ApJ, 920(1), L16.

Dravins, D. (1987). Stellar granulation. II. Stellar photospheric line asymmetries. A&A, 172, 211–224.

Dumusque, X. (2018). Measuring precise radial velocities on individual spectral lines. I. Validation of the method and application to mitigate stellar activity. A&A, 620, A47.

Dumusque, X., Santos, N. C., Udry, S., et al. (2011a). Planetary detection limits taking into account stellar noise. II. Effect of stellar spot groups on radial-velocities. A&A, 527, A82.

Dumusque, X., Udry, S., Lovis, C., et al. (2011b). Planetary detection limits taking into account stellar noise. I. Observational strategies to reduce stellar oscillation and granulation effects. A&A, 525, A140.

Eggleton, P. P., Kiseleva, L. G., & Hut, P. (1998). The Equilibrium Tide Model for Tidal Friction. ApJ, 499(2), 853–870.

Feinstein, A. D., Montet, B. T., Johnson, M. C., et al. (2021). H-alpha and Ca II Infrared Triplet Variations During a Transit of the 23 Myr Planet V1298 Tau c. *arXiv e-prints*, (pp. arXiv:2107.01213).

Fernandes, R. B., Mulders, G. D., Pascucci, I., et al. (2022). Pterodactyls: A Tool to Uniformly Search and Vet for Young Transiting Planets In TESS Primary Mission Photometry. *arXiv e-prints*, (pp. arXiv:2206.03989).

Feroz, F. & Hobson, M. P. (2014). Bayesian analysis of radial velocity data of GJ667C with correlated noise: evidence for only two planets. MNRAS, 437(4), 3540–3549.

Feroz, F., Hobson, M. P., & Bridges, M. (2009). MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics. MNRAS, 398(4), 1601–1614.

Fischer, D. A., Vogt, S. S., Marcy, G. W., et al. (2007). Five Intermediate-Period Planets from the N2K Sample. ApJ, 669(2), 1336–1344.

Flagg, L., Johns-Krull, C. M., Nofi, L., et al. (2019). CO Detected in CI Tau b: Hot Start Implied by Planet Mass and M_K. ApJ, 878(2), L37.

Folsom, C. P., Bouvier, J., Petit, P., et al. (2018). The evolution of surface magnetic fields in young solar-type stars II: The early main sequence (250-650 Myr). *MNRAS*, 474(4), 4956–4987.

Folsom, C. P., Petit, P., Bouvier, J., et al. (2016). The evolution of surface magnetic fields in young solar-type stars - I. The first 250 Myr. *MNRAS*, 457(1), 580–607.

Ford, E. B. & Rasio, F. A. (2006). On the Relation between Hot Jupiters and the Roche Limit. ApJ, 638(1), L45–L48.

Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. (2017). Fast and Scalable Gaussian Process Modeling with Applications to Astronomical Time Series. AJ, 154(6), 220.

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. (2013). emcee: The MCMC Hammer. PASP, 125(925), 306.

Fulton, B. J., Petigura, E. A., Howard, A. W., et al. (2017). The California-Kepler Survey. III. A Gap in the Radius Distribution of Small Planets. AJ, 154(3), 109.

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. (2021). Gaia Early Data Release 3. Summary of the contents and survey properties. A&A, 649, A1.

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. (2016). The Gaia mission. A&A, 595, A1.

Gaidos, E., Hirano, T., Beichman, C., et al. (2021). Zodiacal Exoplanets in Time (ZEIT) XIII: Planet Orbits and Atmospheres in the V1298 Tau System, a Keystone in Studies of Early Planetary Evolution. *arXiv e-prints*, (pp. arXiv:2110.10689).

Gomes da Silva, J., Santos, N. C., Bonfils, X., et al. (2012). Long-term magnetic activity of a sample of M-dwarf stars from the HARPS program . II. Activity and radial velocity. A&A, 541, A9.

Hamers, A. S., Antonini, F., Lithwick, Y., et al. (2017). Secular dynamics of multiplanet systems: implications for the formation of hot and warm Jupiters via high-eccentricity migration. MNRAS, 464(1), 688–701.

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362.

Haywood, R. D., Cameron, A. C., Queloz, D., & Others (2014). Planets and stellar activity: hide and seek in the {CoRoT}-7 system. *Monthly notices of the royal astronomical society*, 443(3), 2517–2531.

Heitzmann, A., Marsden, S. C., Petit, P., et al. (2021a). Planets around young active solar-type stars: assessing detection capabilities from a non-stabilized spectrograph. MNRAS, 505(4), 4989–5011.

Heitzmann, A., Zhou, G., Quinn, S. N., et al. (2022). TOI-4562 b: A highly eccentric temperate Jupiter analog orbiting a young field star. *arXiv e-prints*, (pp. arXiv:2208.10854).

Heitzmann, A., Zhou, G., Quinn, S. N., et al. (2021b). The Obliquity of HIP 67522 b: A 17 Myr Old Transiting Hot, Jupiter-sized Planet. ApJ, 922(1), L1.

Hirano, T., Krishnamurthy, V., Gaidos, E., et al. (2020). Limits on the Spin-Orbit Angle and Atmospheric Escape for the 22 Myr Old Planet AU Mic b. ApJ, 899(1), L13.

Huang, C., Wu, Y., & Triaud, A. H. M. J. (2016). Warm Jupiters Are Less Lonely than Hot Jupiters: Close Neighbors. ApJ, 825(2), 98.

Jeffers, S. V., Barnes, J. R., Jones, H. R. A., et al. (2014a). Is it possible to detect planets around young active G and K dwarfs? MNRAS, 438(4), 2717–2731.

Jeffers, S. V., Mengel, M., Moutou, C., et al. (2018). The relation between stellar magnetic field geometry and chromospheric activity cycles - II The rapid 120-day magnetic cycle of τ Bootis. MNRAS, 479(4), 5266–5271.

Jeffers, S. V., Petit, P., Marsden, S. C., et al. (2014b). ε Eridani: an active K dwarf and a planet hosting star?. The variability of its large-scale magnetic field topology. A&A, 569, A79.

Johansen, A. & Lambrechts, M. (2017). Forming Planets via Pebble Accretion. *Annual Review of Earth and Planetary Sciences*, 45(1), 359–387.

Johnson, M. C., Cochran, W. D., Collier Cameron, A., & Bayliss, D. (2015). Measurement of the Nodal Precession of WASP-33 b via Doppler Tomography. ApJ, 810(2), L23.

Johnson, M. C., David, T. J., Petigura, E. A., et al. (2021). An Aligned Orbit for the Young Planet V1298 Tau b. *arXiv e-prints*, (pp. arXiv:2110.10707).

Jones, H. R. A., Butler, R. P., Tinney, C. G., et al. (2003). An exoplanet in orbit around τ^1 Gruis. MNRAS, 341(3), 948–952.

Jurić, M. & Tremaine, S. (2008). Dynamical Origin of Extrasolar Planet Eccentricity Distribution. ApJ, 686(1), 603–620.

Klein, B., Donati, J.-F., Moutou, C., et al. (2020). Investigating the young AU Mic system with SPIRou: large-scale stellar magnetic field and close-in planet mass. MNRAS.

Kozai, Y. (1962). Secular perturbations of asteroids with high inclination and eccentricity. AJ, 67, 591–598.

Kraus, S., Le Bouquin, J.-B., Kreplin, A., et al. (2020). Spin-Orbit Alignment of the β Pictoris Planetary System. ApJ, 897(1), L8.

Kreidberg, L. (2015). batman: BAsic Transit Model cAlculatioN in Python. PASP, 127(957), 1161.

Lagrange, A. M., Meunier, N., Rubini, P., et al. (2019). Evidence for an additional planet in the β Pictoris system. *Nature Astronomy*, 3, 1135–1142.

Lecar, M., Podolak, M., Sasselov, D., & Chiang, E. (2006). On the Location of the Snow Line in a Protoplanetary Disk. ApJ, 640(2), 1115–1118.

Lee, E. J. & Connors, N. J. (2021). Primordial Radius Gap and Potentially Broad Core Mass Distributions of Super-Earths and Sub-Neptunes. ApJ, 908(1), 32.

Lidov, M. L. (1962). The evolution of orbits of artificial satellites of planets under the action of gravitational perturbations of external bodies. Planet. Space Sci., 9(10), 719–759.

Lin, D. N. C., Bodenheimer, P., & Richardson, D. C. (1996). Orbital migration of the planetary companion of 51 Pegasi to its present location. Nature, 380(6575), 606–607.

Mamajek, E. E. & Hillenbrand, L. A. (2008). Improved Age Estimation for Solar-Type Dwarfs Using Activity-Rotation Diagnostics. ApJ, 687(2), 1264–1293.

Mancini, L., Esposito, M., Covino, E., et al. (2022). The GAPS Programme at TNG: XXXVI. Measurement of the Rossiter-McLaughlin effect and revising the physical and orbital parameters of the HAT-P-15, HAT-P-17, HAT-P-21, HAT-P-26, HAT-P-29 eccentric planetary systems. *arXiv e-prints*, (pp. arXiv:2205.10549).

Marsden, S. C., Petit, P., Jeffers, S. V., et al. (2014). A BCool magnetic snapshot survey of solar-type stars. *MNRAS*, 444(4), 3517–3536.

Martin, R. G. & Livio, M. (2012). On the evolution of the snow line in protoplanetary discs. MNRAS, 425(1), L6–L9.

Martioli, E., Hébrard, G., Moutou, C., et al. (2020). Spin-orbit alignment and magnetic activity in the young planetary system AU Mic. A&A, 641, L1.

Marzari, F. & Nelson, A. F. (2009). Interaction of a Giant Planet in an Inclined Orbit with a Circumstellar Disk. ApJ, 705(2), 1575–1583.

Masuda, K. (2017). Eccentric Companions to Kepler-448b and Kepler-693b: Clues to the Formation of Warm Jupiters. AJ, 154(2), 64.

Masuda, K. & Winn, J. N. (2020). On the Inference of a Star's Inclination Angle from its Rotation Velocity and Projected Rotation Velocity. AJ, 159(3), 81.

Mayor, M. & Queloz, D. (1995). A Jupiter-mass companion to a solar-type star. Nature, 378(6555), 355–359.

McLaughlin, D. B. (1924). Some results of a spectrographic study of the Algol system. ApJ, 60, 22–31.

Mengel, M. W., Fares, R., Marsden, S. C., et al. (2016). The evolving magnetic topology of τ Boötis. MNRAS, 459(4), 4325–4342.

Meunier, N. (2021). Stellar variability in radial velocity. *arXiv e-prints*, (pp. arXiv:2104.06072).

Meunier, N., Lagrange, A. M., & Borgniet, S. (2017). A new method of correcting radial velocity time series for inhomogeneous convection. A&A, 607, A6.

Michel, E., Baglin, A., Auvergne, M., et al. (2008). CoRoT Measures Solar-Like Oscillations and Granulation in Stars Hotter Than the Sun. *Science*, 322(5901), 558.

Montet, B. T., Feinstein, A. D., Luger, R., et al. (2020). The Young Planet DS Tuc Ab Has a Low Obliquity. AJ, 159(3), 112.

Naef, D., Latham, D. W., Mayor, M., et al. (2001). HD 80606 b, a planet on an extremely elongated orbit. A&A, 375, L27–L30.

Naoz, S. (2016). The Eccentric Kozai-Lidov Effect and Its Applications. ARA&A, 54, 441–489.

Nelson, B. E., Ford, E. B., Buchner, J., et al. (2020). Quantifying the Bayesian Evidence for a Planet in Radial Velocity Data. AJ, 159(2), 73.

Newton, E. R., Mann, A. W., Tofflemire, B. M., et al. (2019). TESS Hunt for Young and Maturing Exoplanets (THYME): A Planet in the 45 Myr Tucana-Horologium Association. ApJ, 880(1), L17.

O'Toole, S. J., Tinney, C. G., & Jones, H. R. A. (2008). The impact of stellar oscillations on doppler velocity planet searches. MNRAS, 386(1), 516–520.

Owen, J. E. & Wu, Y. (2013). Kepler Planets: A Tale of Evaporation. ApJ, 775(2), 105.

Palle, E., Oshagh, M., Casasayas-Barris, N., et al. (2020). Transmission spectroscopy and Rossiter-McLaughlin measurements of the young Neptune orbiting AU Mic. A&A, 643, A25.

Palumbo, Michael L., I., Ford, E. B., Wright, J. T., et al. (2022). GRASS: Distinguishing Planet-induced Doppler Signatures from Granulation with a Synthetic Spectra Generator. AJ, 163(1), 11.

Petigura, E. A., Marcy, G. W., Winn, J. N., et al. (2018). The California-Kepler Survey. IV. Metal-rich Stars Host a Greater Diversity of Planets. AJ, 155(2), 89.

Petrovich, C. (2015). Steady-state Planet Migration by the Kozai-Lidov Mechanism in Stellar Binaries. ApJ, 799(1), 27.

Petrovich, C. & Tremaine, S. (2016). Warm Jupiters from Secular Planet-Planet Interactions. ApJ, 829(2), 132.

Piskunov, N. E. (1991). The Art of Surface Imaging. In I. Tuominen, D. Moss, & G. Rüdiger (Eds.), *IAU Colloq. 130: The Sun and Cool Stars. Activity, Magnetism, Dynamos*, volume 380 (pp. 309).

Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. (1996). Formation of the Giant Planets by Concurrent Accretion of Solids and Gas. Icarus, 124(1), 62–85.

Psaridi, A., Bouchy, F., Lendl, M., et al. (2022). Three new brown dwarfs and a massive hot Jupiter revealed by TESS around early-type stars. *arXiv e-prints*, (pp. arXiv:2205.10854).

Rajpaul, V., Aigrain, S., Osborne, M. A., et al. (2015). A Gaussian process framework for modelling stellar activity signals in radial velocity data. MNRAS, 452(3), 2269–2291.

Rasio, F. A. & Ford, E. B. (1996). Dynamical instabilities and the formation of extrasolar planetary systems. *Science*, 274, 954–956.

Reiners, A., Bean, J. L., Huber, K. F., et al. (2010). Detecting Planets Around Very Low Mass Stars with the Radial Velocity Method. ApJ, 710(1), 432–443.

Rice, M., Wang, S., & Laughlin, G. (2022). Origins of Hot Jupiters from the Stellar Obliquity Distribution. ApJ, 926(2), L17.

Rice, W. K. M., Armitage, P. J., & Hogg, D. F. (2008). Why are there so few hot Jupiters? MNRAS, 384(3), 1242–1248.

Rizzuto, A. C., Newton, E. R., Mann, A. W., et al. (2020). TESS Hunt for Young and Maturing Exoplanets (THYME). II. A 17 Myr Old Transiting Hot Jupiter in the Sco-Cen Association. AJ, 160(1), 33.

Rodet, L., Su, Y., & Lai, D. (2021). On the Correlation between Hot Jupiters and Stellar Clustering: High-eccentricity Migration Induced by Stellar Flybys. ApJ, 913(2), 104.

Roettenbacher, R. M., Monnier, J. D., Korhonen, H., et al. (2016). No Sun-like dynamo on the active star ζ Andromedae from starspot asymmetry. Nature, 533(7602), 217–220.

Rogers, T. M., Lin, D. N. C., & Lau, H. H. B. (2012). Internal Gravity Waves Modulate the Apparent Misalignment of Exoplanets around Hot Stars. ApJ, 758(1), L6.

Rossiter, R. A. (1924). On the detection of an effect of rotation during eclipse in the velocity of the brigher component of beta Lyrae, and on the constancy of velocity of this system. ApJ, 60, 15–21.

Saar, S. H. & Baliunas, S. L. (1992). The Magnetic Cycle of Kappa Ceti. In K. L. Harvey (Ed.), *The Solar Cycle*, volume 27 of *Astronomical Society of the Pacific Conference Series* (pp. 197–202).

Santerne, A., Moutou, C., Tsantaki, M., et al. (2016). SOPHIE velocimetry of Kepler transit candidates. XVII. The physical properties of giant exoplanets within 400 days of period. A&A, 587, A64.

Schlaufman, K. C. (2010). Evidence of Possible Spin-orbit Misalignment Along the Line of Sight in Transiting Exoplanet Systems. ApJ, 719(1), 602–611.

Setiawan, J., Weise, P., Henning, T., et al. (2007). Evidence for a Planetary Companion around a Nearby Young Star. ApJ, 660(2), L145–L148.

Shara, M. M., Hurley, J. R., & Mardling, R. A. (2016). Dynamical Interactions Make Hot Jupiters in Open Star Clusters. ApJ, 816(2), 59.

Simola, U., Dumusque, X., & Cisewski-Kehe, J. (2019). Measuring precise radial velocities and cross-correlation function line-profile variations using a Skew Normal density. A&A, 622, A131.

Socrates, A., Katz, B., Dong, S., & Tremaine, S. (2012). Super-eccentric Migrating Jupiters. ApJ, 750(2), 106.

Spalding, C. & Batygin, K. (2015). Magnetic Origins of the Stellar Mass-Obliquity Correlation in Planetary Systems. ApJ, 811(2), 82.

Strassmeier, K. G. (2009). Starspots. A&A Rev., 17(3), 251–308.

Trilling, D. E., Benz, W., Guillot, T., et al. (1998). Orbital Evolution and Migration of Giant Planets: Modeling Extrasolar Planets. ApJ, 500(1), 428–439.

Ujjwal, K., Kartha, S. S., Mathew, B., et al. (2020). Analysis of Membership Probability in Nearby Young Moving Groups with Gaia DR2. AJ, 159(4), 166.

Valsecchi, F., Rappaport, S., Rasio, F. A., et al. (2015). Tidally-driven Roche-lobe Overflow of Hot Jupiters with MESA. ApJ, 813(2), 101.

Virtanen, P., Gommers, R., Oliphant, T. E., et al. (2020). Scipy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*.

Vogt, S. S. & Penrod, G. D. (1983). Doppler imaging of spotted stars : application to the RS Canum Venaticorum star HR 1099. PASP, 95, 565–576.

Vogt, S. S., Penrod, G. D., & Hatzes, A. P. (1987). Doppler Images of Rotating Stars Using Maximum Entropy Image Reconstruction. ApJ, 321, 496.

von Zeipel, H. (1910). Sur l'application des séries de M. Lindstedt à l'étude du mouvement des comètes périodiques. *Astronomische Nachrichten*, 183(22), 345.

Weidenschilling, S. J. & Marzari, F. (1996). Gravitational scattering as a possible origin for giant planets at small stellar distances. Nature, 384(6610), 619–621.

Winn, J. N., Fabrycky, D., Albrecht, S., & Johnson, J. A. (2010). Hot Stars with Hot Jupiters Have High Obliquities. ApJ, 718(2), L145–L149.

Wirth, C. P., Zhou, G., Quinn, S. N., et al. (2021). TOI-942b: A Prograde Neptune in a 60 Myr Old Multi-transiting System. ApJ, 917(2), L34.

Wittenmyer, R. A., O'Toole, S. J., Jones, H. R. A., et al. (2010). The Frequency of Low-mass Exoplanets. II. The "Period Valley". ApJ, 722(2), 1854–1863.

Wolszczan, A. & Frail, D. A. (1992). A planetary system around the millisecond pulsar PSR1257 + 12. Nature, 355(6356), 145–147.

Wong, I., Shporer, A., Zhou, G., et al. (2021). TOI-2109: An Ultrahot Gas Giant on a 16 hr Orbit. AJ, 162(6), 256.

Wright, J. T., Marcy, G. W., Howard, A. W., et al. (2012). The Frequency of Hot Jupiters Orbiting nearby Solar-type Stars. ApJ, 753(2), 160.

Wu, Y. & Lithwick, Y. (2011). Secular Chaos and the Production of Hot Jupiters. ApJ, 735(2), 109.

Young, J. S., Baldwin, J. E., Boysen, R. C., et al. (2000). New views of Betelgeuse: multi-wavelength surface imaging and implications for models of hotspot generation. MNRAS, 315(3), 635–645.

Yu, L., Donati, J. F., Grankin, K., et al. (2019). Magnetic field, activity, and companions of V410 Tau. MNRAS, 489(4), 5556–5572.

Yu, L., Donati, J. F., Hébrard, E. M., et al. (2017). A hot Jupiter around the very active weak-line T Tauri star TAP 26. MNRAS, 467(2), 1342–1359.

Zechmeister, M., Reiners, A., Amado, P. J., et al. (2018). Spectrum radial velocity analyser (SERVAL). High-precision radial velocities and two alternative spectral indicators. A&A, 609, A12.

Zhao, L. L., Fischer, D. A., Ford, E. B., et al. (2022). The EXPRES Stellar Signals Project II. State of the Field in Disentangling Photospheric Velocities. AJ, 163(4), 171.

Zhou, G., Huang, C. X., Bakos, G. Á., et al. (2019). Two New HATNet Hot Jupiters around A Stars and the First Glimpse at the Occurrence Rate of Hot Jupiters from TESS. AJ, 158(4), 141.

Zhou, G., Quinn, S. N., Irwin, J., et al. (2021). Two Young Planetary Systems around Field Stars with Ages between 20 and 320 Myr from TESS. AJ, 161(1), 2.

Zhou, G., Winn, J. N., Newton, E. R., et al. (2020). A Well-aligned Orbit for the 45 Myr-old Transiting Neptune DS Tuc Ab. ApJ, 892(2), L21.

Zhou, G., Wirth, C. P., Huang, C. X., et al. (2022). A Mini-Neptune from TESS and CHEOPS Around the 120 Myr Old AB Dor Member HIP 94235. AJ, 163(6), 289.

Appendix A: Detecting exoplanets: The radial velocity method.

The radial velocity detection method has yielded 930 exoplanets, 18% of all the discoveries to date, making it the second most successful method behind the transit method.

Two gravitational bound bodies orbit around their common centre of mass. In the case of a planet orbiting a star (i.e. $M_{\rm p} \ll M_{\star}$), the system's centre of mass is located close, but not exactly at the stellar barycenter. As a consequence, the star orbits around the system's centre of mass and wobbles. This orbital motion will yield a periodic acceleration of the star in the radial (i.e. along the line of sight) and tangential (i.e. in the plane of the sky) directions. The radial component depends on the mass of the two bodies, their separation and the angle of the orbit relative to the line of sight, usually referred as the stellar inclination $i_{\star}$.

Stellar radial acceleration, or change in radial velocity induced by an orbiting planet is detectable through the Doppler effect (illustrated on Figure A.1), where light emitted from a receding/approaching source will be observed with a longer/shorter wavelength (i.e. redshifted/blueshifted). In the case of stars, elements and chemical compounds present in their atmosphere will absorb

some of the light originating from deeper layers and create spectral absorption lines. These lines are clearly identifiable in stellar spectra and located at a very specific wavelength, λ_{em}, corresponding to the atomic/molecular absorption transition. The stellar reflex motion due to an orbiting planet will yield a Doppler shift of the spectral lines in wavelength space, from λ_{em} to their observed, shifted values λ_{obs}. The stellar radial velocity at the considered epoch v_r is given by:

$$
v_r = \left(\frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}} \right) c \tag{A.1}
$$

with c the speed of light. By monitoring the radial velocity over time, we can recover K, the semi-amplitude of the planet induced stellar RV, and the orbital period of the planet P_{orb}. Considering circular orbits (with the eccentricity $e = 0$), K and P_{orb} grant access to the relative masses of the star ($M_\star$) and orbiting planet (M_{p}):

$$
K = \left(\frac{2\pi G}{P_{orb}} \right) \frac{M_p \sin i_\star}{(M_\star + M_p)^{2/3}} \tag{A.2}
$$

with G the gravitational constant. Because shifts induced by planets on star can be very small (tens of $\mathrm{m\,s^{-1}}$ for a Sun-like star due to a hot Jupiter down a few $\mathrm{cm\,s^{-1}}$ for the Sun reflex motion due to the Earth), it is essential to precisely measure λ_{obs}. Precisely measuring the location of one single spectral line in an observed spectra is limited by instrumental resolution. The most widely used approach is called the cross-correlation function (CCF). A template spectra containing lines of precisely known position (obtained from laboratory measurement, λ_{em}), or line mask, is cross-correlated with the observed spectra. The result is a single spectral line that can be understood as

a mean line of all the lines present in the template, with a drastically higher signal to noise ratio. Fitting a single line model to the CCF and recovering its radial velocity (converted from wavelength value using Equation (A.1)) is possible to the sub m s^{-1} level with state of the art, highly stabilised spectrographs (e.g. ESPRESSOS, HARPS or CARMENES). An alternative mean line profile generation technique is Least Square Deconvolution (LSD Donati & Brown 1997). A thourough description of the LSD is beyond the scope of this thesis, but we redirect the reader to Chapter 3 and references therein for complementary information.
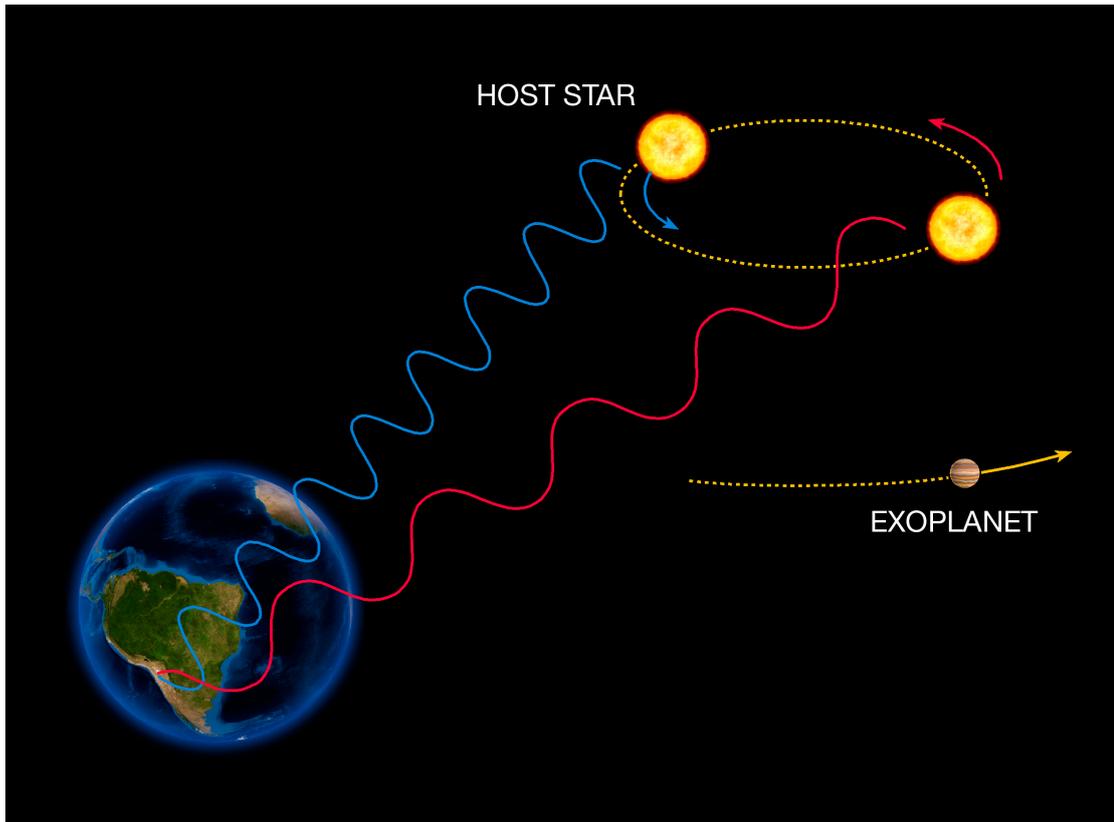
Figure A.1: Illustration of the radial velocity method. The orbiting planet induces a reflex motion on the host star which periodically moves towards/away from the observer resulting in a blue/red shift of its emitted light. Credits: ESO, obtained from https://www.eso.org/public/images/eso0722e/

.

Appendix B: Involvement with MINERVA-Australis.

The MINERVA-Australis array (Addison et al., 2019) is made of four, 0.7m telescopes and is located at USQ's Mount Kent Observatory in regional Queensland, Australia. The telescopes work simultaneously and feed a high resolution (R > 80,000) optical (480 to 630 nm) spectrograph and are capable of ≈ 2 m s^{-1} RV precision on bright stable stars. Since the end of 2019, I have, in parallel of my Ph. D., been heavily involved in the development of the array, mostly software-wise.

The array is now fully robotized and autonomous. The interface between the telescopes, spectrograph and weather station was written using Node-red, a visual programming framework well suited to automate processes. We also implemented a front-end user interface and live status updates on a Discord server.

My involvement started with performing manual observation prior to the completion of the automated system. I have contributed in many aspect of the development of the observatory's automation: set up of daily calibrations, metadata collection, implementing solutions to compute the flux weighted mean time of observations. I have also been tasked with the design and im-

plementation of the automated scheduler of the observatory. Finally, I developed a software solution to enable guiding on difficult targets that have nearby bright stars.

My current role is to maintain the system, add features (e.g. target priority assessment, improvement of the telescopes automated procedure, automation of a focusing routine, provide metadata for monitoring instrument systematics, optimisation of the scheduler) and handle the organisational aspect of the development of the observatory. My ability to qualitatively judge the quality of spectra and reduce line profile/RVs allows me to take part in the decision making regarding targets follow-up. I also plan and am in charge of handling the observations campaigns allocated to the NASA/NSF NN-Explore program.

APPENDIX C: PYTHON CODE DEVELOPED FOR PAPER II

This Appendix contains the python code that I developed and that was used to analyse the data from paper II (see chapter 4). The star spot model to which a transiting planet shadow was added was built from scratch. The rest of the model made use of python packages developed by the community: NUMPY (Harris et al., 2020), SCIPY (Virtanen et al., 2020), ASTROPY (Astropy Collaboration et al., 2013, 2018), BATMAN (Kreidberg, 2015), CELERITE (Foreman-Mackey et al., 2017) and EMCEE (Foreman-Mackey et al., 2013).

C.1 MAIN CODE

Below is the main function, which read the data, set up the parameters, and runs the MCMC (Monte Carlo Markov Chain) to explore the parameter space.

```python
import numpy as np
import emcee
import pickle
import time
from matplotlib import rc
from pylab import *
from scipy.stats import binned_statistic
```

```python
from LC_model_emcee import run_batman, lnlike_LC, compute_ingress_egress
from DT_model_emcee import compute_DT_model, lnlike_DT


def separate_params(params):

    """
    From the dictionnary of parameters, separate the fitted from the fixed
 →  ones.
    Inputs:
    - params (dict): Dictionnary of all parameters
    Outputs:
    - fittedParams (dict): Dictionnary containging parameters part of the MCMC
 →  fit
    - fixedParams (dict): Dictionnary containing fixed parameters, not part
 →  of the MCMC fit
    """

    fixedParams = {
        'name':[],
        'type':[],
        'val1':[],
        'val2':[],
        'init':[],
        'initvar':[]
    }
    fittedParams = {
        'name':[],
        'type':[],
        'val1':[],
        'val2':[],
        'init':[],
        'initvar':[]
    }

    for i in range(len(params['type'])):
        if params['type'][i] == 'F':
            for key in params:
                fixedParams[key].append(params[key][i])
        else:
```

```
44                for key in params:
45                    fittedParams[key].append(params[key][i])
46
47        return fittedParams, fixedParams
48
49  def merge_params(fittedOrFixed, fittedParams, fixedParams):
50
51        """
52        Merges the fitted and fixed parameter dictionaries.
53        Inputs:
54        - fittedOrfixed (list of stings): list mapping each parameters with 'F'
    →   indicating fixed ones and 'U' or 'G'for fitted ones.
55        - fittedParams (dict): Dictionary of fitted parameters.
56        - fixedParams (dict): Dictionary of fixed parameters.
57        Outputs:
58        - mergedParams (dict): Merged dictionary of fitted and fixed parameters.
59
60        """
61        mergedParams = []
62        i = 0
63        j = 0
64
65        for value in fittedOrFixed:
66            if value == 'F':
67                mergedParams.append(fixedParams[i])
68                i += 1
69            else:
70                mergedParams.append(fittedParams[j])
71                j += 1
72
73        return mergedParams
74
75  def lnlike(x, fixedParams, params, dataPhoto, dataSpectro, nbSpots, perStar,
    →   Teff, limbDarkLaw, instResolution, do_mcmc = True):
76
77        """
78        Compute the likelihood of the model given a set of parameters.
79        1) check that the current drawn parameters are within the prior and
    →   compute the prior likelihood.
```

```
80    2) compute the likelihood from the transit model
81    3) compute the likelihood from the Doppler Tomography model
82    Inputs:
83    - x (array): array of parameter values
84    - fixedParams (dict): Dictionary of fixed, non-fitted parameters
85    - params (dict): Dictionary of all parameters
86    - dataPhoto (dict): Dictionary containing the photometric data
87    - dataSpectro (dict): Dictionary containing the spectroscopic data
88    - nbSpots (int): Number of spots for the stellar DT model
89    - perStar (float): Stellar rotation period
90    - Teff (float): Stellar effective temperature
91    - limbDarkLaw (str): law used to model the Limb Darkening
92    - instResolution (float): Spectral resolution of the spectrograph
93    - do_mcmc (bool): If true, returns only the likelihood. If False, returns
   →  other parameters for plots and visualisation.
94    Outputs:
95    - lnlike (float): Total likelihood value.
96    """
97    fitOrfixed = params['type']
98    currentParams = x.copy()
99    fixedParamsValues = fixedParams.copy()
100
101   parameterValues = merge_params(fitOrfixed, currentParams,
      →  fixedParamsValues)
102
103   ## Set up priors and boundaries ##
104   lnlikePriors = 0
105   for i in range(len(fitOrfixed)):
106       if fitOrfixed[i] == 'F': # Fixed parameter
107           pass
108       elif fitOrfixed[i] == 'G': #Gaussian prior
109           lnlikePriors +=
              →  -0.5*(parameterValues[i]-params['val1'][i])**2/(2*params['val2
110           '][i]**2)
111       else: # Uniform prior
112           if np.abs(parameterValues[i]-params['val1'][i]) >
              →  params['val2'][i]:
113               lnlike = -1*np.inf
114               return lnlike
```

```
115
116     #### Transit model ####
117
118     t0  = parameterValues[0]
119     per = parameterValues[1]
120     rp  = parameterValues[2]
121     sma = parameterValues[3]
122     inc = parameterValues[4]
123     ecc = parameterValues[5]
124     w   = parameterValues[6]
125     u11 = parameterValues[7]
126     u12 = parameterValues[8]
127
128     paramsTr = [t0,per,rp,sma,inc,ecc,w,u11,u12]
129
130     lnlikeTransits = 0
131
132     # timeGP = []
133     # meanGP = []
134     # varGP = []
135     transitModel = []
136     nbSectors = len(dataPhoto['tbjd'])
137     # Detrend and fit transits for each sector to get a likelihood for each
    →  (use same fitting parameters for all sector for both GP and transits)
138     for i in range(nbSectors): # for each sector
139
140         model = run_batman(paramsTr, dataPhoto['tbjd'][i], limbDarkLaw)
141
142         transitModel.append(model)
143
144         #Remove the transit model
145         #noTransitFlux = dataPhoto['flux'][i]-model
146
147         # #### GP detrend LC ####
148         # ln_S0 = parameterValues[9]
149         # ln_Q  = parameterValues[10]
150         # ln_w0 = parameterValues[11]
151
152         # paramsGP = [ln_S0, ln_Q, ln_w0]
```

```python
153
154         ### if lc model is flat -- ie if it predicts no transit
155         if np.std(model)<0.000001:
156             lnlikeTransit = -np.inf
157         else:
158             # # Compute the ln(likelihood) for the transit (detrend woth GP
                 →  then substract data - model)
159             # if do_mcmc == False:
160             #     returnGP = True
161             #     lnlikeTransit, t, GPmean, GPvar = lnlike_LC(paramsGP,
                 →  dataPhoto['tbjd'][i], noTransitFlux, dataPhoto['err'][i],
                 →  return_gp_model=returnGP)
162             #     timeGP.append(t)
163             #     meanGP.append(GPmean)
164             #     varGP.append(GPvar)
165             # else:
166             #     returnGP = False
167             #     lnlikeTransit = lnlike_LC(paramsGP, dataPhoto['tbjd'][i],
                 →  noTransitFlux, dataPhoto['err'][i], return_gp_model=returnGP)
168             lnlikeTransit =
                 →  -0.5*np.sum((dataPhoto['flux'][i]-model)**2/(dataPhoto['err']
169         [i]**2))
170
171         lnlikeTransits += lnlikeTransit
172
173     if do_mcmc == False:
174         ingress, egress, _ = compute_ingress_egress(rp, sma, inc)
175
176
177     #### DT model ####
178
179     vsini   = parameterValues[9]
180     lam     = parameterValues[10]
181     nonRotV = parameterValues[11]
182     u21  = parameterValues[12]
183     u22  = parameterValues[13]
184
185     colatitudes = []
186     longitudes  = []
```

```
187        spotSizes   = []
188        spotTemps   = []
189
190        for i in range (nbSpots): # In case we have multiple spots
191            increment = 4*i
192            colatitudes.append(parameterValues[14+increment])
193            longitudes.append(parameterValues[15+increment])
194            spotSizes.append(parameterValues[16+increment])
195            spotTemps.append(parameterValues[17+increment])
196
197        spots = {
198            'colatitude' : colatitudes,
199            'longitude' : longitudes,
200            'spotSize' : spotSizes,
201            'spotTemp' : spotTemps
202        }
203
204        allMasks = []
205        allProfPlanet = []
206        allProfTotal = []
207        allProfPureNorm = []
208        allprofSpots = []
209        dist = []
210        planetPhases = []
211
212        nbEpochsPerDatasets = []
213
214        lnlikeDT = 0
215        nbTransitsSpectro = len(dataSpectro['velLSD'])
216
217        for j in range (nbTransitsSpectro):
218
219            nbEpochs = len(dataSpectro['epochsLSD'][j])
220
221            # Phases of LSD observations
222            nbOrbitsSincet0 = np.floor((dataSpectro['epochsLSD'][j][-1]-t0)/per)
223
224            plaPhases = (dataSpectro['epochsLSD'][j] - t0 -
                ↪   nbOrbitsSincet0*per)/per
```

```
225
226         paramsDT = [vsini, nbSpots, spots, Teff, limbDarkLaw, u21, u22,
             → plaPhases, nonRotV]
227
228         paramsPla = [rp, sma, inc, lam]
229
230         if do_mcmc == False:
231             allMasksTemp, allProfPlanetTemp, allProfTotalTemp,
                 → allProfPureNormTemp, _, allprofSpotsTemp, distTemp =
                 → compute_DT_model(dataSpectro['velLSD'][j],
                 → dataSpectro['meanLSD'][j], instResolution, paramsDT,
                 → paramsPla, dataSpectro['epochsLSD'][j], perStar, paramSearch =
                 → do_mcmc, addNoise = False)
232             allMasks.append(allMasksTemp)
233             allProfPlanet.append(allProfPlanetTemp)
234             allProfTotal.append(allProfTotalTemp)
235             allProfPureNorm.append(allProfPureNormTemp)
236             allprofSpots.append(allprofSpotsTemp)
237             dist.append(distTemp)
238
239             planetPhases.append(plaPhases)
240             nbEpochsPerDatasets.append(nbEpochs)
241
242         else:
243             _, allProfTotalTemp, _ =
                 → compute_DT_model(dataSpectro['velLSD'][j],
                 → dataSpectro['meanLSD'][j], instResolution, paramsDT,
                 → paramsPla, dataSpectro['epochsLSD'][j], perStar, paramSearch =
                 → do_mcmc, addNoise = False)
244
245         for n in range(nbEpochs):
246             lnlikeDT +=
                 → lnlike_DT(allProfTotalTemp[n],dataSpectro['intLSD'][j][n],
                 → dataSpectro['errLSD'][j][n])
247
248     lnlike = lnlikePriors + lnlikeTransits + lnlikeDT
249
250
251     if do_mcmc == False:
```

```python
252              otherParams = [planetPhases, vsini, ingress, egress,
              →  nbEpochsPerDatasets]
253          return lnlike, transitModel, allMasks, allProfPlanet, allProfTotal,
              →  allProfPureNorm, _, allprofSpots, dist, otherParams
254      else:
255          return lnlike
256
257  def sec_to_hours(sec):
258      # Convert seconds to hours
259      time_hours = time.gmtime(sec)
260      return time.strftime('%HH%MM%SS',time_hours)
261
262  def run_mcmc(fittedParams, fixedParams, params,  dataPhoto, dataSpectro,
      →  nwalkers, niter, nbSpots, perStar, Teff, limbDarkLaw, instResolution,
      →  nCPUs, constants):
263
264      """
265      Main fitting loop. Perform the MCMC
266      Inputs:
267      - fixedParams (dict): Dictionary of fixed, non-fitted parameters
268      - params (dict): Dictionary of all parameters
269      - dataPhoto (dict): Dictionary containing the photometric data
270      - dataSpectro (dict): Dictionary containing the spectroscopic data
271      - nwalkers (float): Number of walkers (chains) for the MCMC parameter
      →  exploration
272      - niter (float): Number of MCMC iterations
273      - nbSpots (int): Number of spots for the stellar DT model
274      - perStar (float): Stellar rotation period
275      - Teff (float): Stellar effective temperature
276      - limbDarkLaw (str): law used to model the Limb Darkening
277      - instResolution (float): Spectral resolution of the spectrograph
278      - nCPUs (int): Number of CPUs to use for the MCMC
279      - constants (list): Constants.
280      """
281      x0 = []
282
283      # Start time
284      start_time = time.time()
285
```

```python
        ### Setup walkers
        for i in range(nwalkers):
            ### tryout a walk
            goodwalker = False
            while not goodwalker:
                x0i = []
                for j in range(len(fittedParams['name'])):
                    value = np.random.normal(fittedParams['init'][j],fittedParams
                    ['initvar'][j])
                    x0i.append(value)

                if abs(lnlike(x0i, fixedParams['val1'], params, dataPhoto,
                → dataSpectro, nbSpots, perStar, Teff, limbDarkLaw,
                → instResolution, do_mcmc = True)) < np.inf:
                    x0.append(x0i)
                    goodwalker=True

        ndim = len(x0[0])

        sampler = emcee.EnsembleSampler(nwalkers, ndim, lnlike, args =
        → [fixedParams['val1'], params, dataPhoto, dataSpectro, nbSpots,
        → perStar, Teff, limbDarkLaw, instResolution], threads=nCPUs)

        # We'll track how the average autocorrelation time estimate changes
        index = 0
        autocorr = np.empty(max_iter)

        # This will be useful to testing convergence
        old_tau = np.inf

        # Now we'll sample for up to max_n steps
        for sample in sampler.sample(x0, iterations=max_iter, progress=True):
            # Only check convergence every 100 steps
            if sampler.iteration % 100:
                continue

            # Compute the autocorrelation time so far
            # Using tol=0 means that we'll always get an estimate even if it isn't
            → trustworthy
```

```python
        tau = sampler.get_autocorr_time(tol=0)
        meanTau = np.mean(tau)
        if np.isnan(meanTau):
            meanTau = 0
        autocorr[index] = meanTau
        index += 1

        # Check convergence
        converged = np.all(tau * 50 < sampler.iteration)
        converged &= np.all(np.abs(old_tau - tau) / tau < 0.01)
        if converged:
            break
        old_tau = tau

    #pos, prob, state = sampler.run_mcmc(x0, niter, progress=True)
    #chain = sampler.chain.reshape((-1,ndim))

    # Data for export
    data_to_pickle = [constants, dataPhoto, dataSpectro, params, fittedParams,
    →  fixedParams, sampler, autocorr, index]

    # End time
    end_time = time.time()

    # Process time
    # print(end_time-start_time)
    process_time_sec = (end_time-start_time)
    #total_time_sec = process_time_sec/nCPUs

    process_time = sec_to_hours(process_time_sec)
    #total_time = sec_to_hours(total_time_sec)

    with open('results_' + str(nbSpots) + 'spots_' + str(nwalkers) +
    →  'walkers_' + str(niter) + 'iter_' + str(process_time), 'wb') as fh:
        pickle.dump(data_to_pickle, fh)

if __name__ == '__main__':

    ###############################
```

```
357    ######### Constants ############
358    ###############################
359
360    MSun = ^^I1.9885e30
361    RSun = 6.957e8
362    G = 6.67e-11
363    c = 3e5
364
365    ## Instrument ##
366    instResolution = 80000 # Chiron
367
368    ##  HIP67522 ##
369    Teff = 5675                    # K
370    perStar = 1.3918809138374308    # days (deduced with assuming exactly 25
       →  rotation cycles between our 2 datasets and Porb = 6.959503)
371    #perStar = 1.413508 # from LS periodogram
372    MStar = 1.22 * MSun          # kg
373    RStar = 1.38 * RSun          # m
374    # logg = np.log10(MStar) - 2* np.log10(RStar) + 4.437
375
376    ## DT Simulation ##
377    limbDarkLaw ='quadratic'
378    nbSpots = 1
379    constants = [MSun, RSun, G, c, instResolution, Teff, perStar, MStar,
       →  RStar, limbDarkLaw]
380
381    ##########################
382    ######### Data ##########
383    ##########################
384
385    pathFolder = 'HIP67522/'
386
387    ## Photometry ##
388    filesPhoto =
       →  ['sector_11_GP_detrended_Andrew','sector_38_GP_detrended_Andrew']
389    NbSectors = len(filesPhoto)
390
391    dataPhoto = {
392        'tbjd': [],
```

```python
            'flux': [],
            'err': [],
            'rawFlux': [],
            'GPModel': [],
        }

    for i in range (NbSectors):
        dataFilePhoto = open(pathFolder + filesPhoto[i],'rb')
        dataP = pickle.load(dataFilePhoto)
        dataFilePhoto.close()

        dataPhoto['tbjd'].append(dataP['tbjd'])
        dataPhoto['flux'].append(dataP['GPModel']-dataP['rawFlux'])
        dataPhoto['err'].append(dataP['err'])
        dataPhoto['rawFlux'].append(dataP['rawFlux'])
        dataPhoto['GPModel'].append(dataP['GPModel'])

    ## Spectroscopy ##

    # with open('lc_s38_detrended.txt', 'w') as out:
    #     for i in range(len(dataPhoto['tbjd'][1])):
    #         out.write('%f %f
    →  %f\n'%(dataPhoto['tbjd'][1][i],dataPhoto['flux'][1][i],dataPhoto
    ['err'][1][i]))

    filesSpectro = ['LSD_data_14_05_2021','LSD_data_18_06_2021']
    #filesSpectro = ['LSD_data_14_05_2021']
    #filesSpectro = ['LSD_data_18_06_2021']

    NbTransitsSpectro = len(filesSpectro)

    dataSpectro = {
        'epochsLSD': [],
        'nbEpochs': [],
        'velLSD': [],
        'intLSD': [],
        'errLSD': [],
        'meanLSD': [],
    }
```

```python
    for i in range (NbTransitsSpectro):

        filename = filesSpectro[i]
        dataFileSpectro = open(pathFolder + filename,'rb')
        dataS = pickle.load(dataFileSpectro)
        dataFileSpectro.close()

        dataSpectro['epochsLSD'].append(dataS['epochsLSD'])
        dataSpectro['nbEpochs'].append(dataS['nbEpochs'])
        dataSpectro['velLSD'].append(dataS['velLSD'])
        dataSpectro['intLSD'].append(dataS['intLSD'])
        dataSpectro['errLSD'].append(dataS['errLSD'])
        dataSpectro['meanLSD'].append(dataS['meanLSD'])

    ###################
    #### Parameters ####
    ###################

    paramsNoSpots = {
        'name'    : ['t0'     ,'per'   ,'rp'   ,'a'   ,'inc' ,'ecc','w'
        → ,'u11' ,'u12','vsini','lam','nonrotv','u21'  ,'u22' ],
        'type'    : ['U'       ,'U'     ,'U'    ,'U'    ,'U'   ,'F' ,'F' ,'F'
        → ,'F' ,'G'    ,'U'  ,'U'       ,'F'    ,'F'   ], # U for uniform,
        → G for Gaussian, F if fixed
        'val1'    : [1604.0237, 6.9595 , 0.0687, 11.55, 88   , 0   , 0   ,
        → 0.148, 0.23, 50    , 0   , 3       , 0.4139, 0.2494], # mean if U,
        → mean if G, fixed value of F
        'val2'    : [0.1      , 0.02   , 0.01  , 2.55 , 2    , None, None,
        → None , None, 3     , 180 , 3       , None  , None ], # (+-) if U,
        → std if G, none if F
        'init'    : [1604.0237, 6.9595 , 0.0667, 11.7 , 89.34, None, None,
        → None , None, 50    , 0   , 2       , None  , None ], # (+-) if U,
        → std if G, none if F
        'initvar' : [0.01     , 0.002  , 0.001 , 0.2  , 0.2  , None, None,
        → None , None, 1     , 18  , 1       , None  , None ], # (+-) if U,
        → std if G, none if F
    }

```

```python
    paramsSpot1 = {
        'name'   : ['clatSpot','lonSpot','sizeSpot','tempSpot'],
        'type'   : ['U'        ,'U'      ,'U'       ,'U'        ], # U for
        →  uniform, G for Gaussian, F if fixed
        'val1'   : [ 45        , 180     , 25       , 0.6*Teff ], # mean if U,
        →  mean if G, fixed value of F
        'val2'   : [ 45        , 180     , 25       , 0.4*Teff ], # (+-) if U,
        →  std if G, none if F
        'init'   : [ 50        , 210     , 5        , 0.7*Teff ], # (+-) if U,
        →  std if G, none if F
        'initvar' : [ 5        , 18      , 2.5      , 0.1*Teff ], # (+-) if U,
        →  std if G, none if F
    }

    paramsSpot2 = {
        'name'   : ['clatSpot2','lonSpot2','sizeSpot2','tempSpot2'],
        'type'   : ['U'        ,'U'      ,'U'       ,'U'        ], # U for
        →  uniform, G for Gaussian, F if fixed
        'val1'   : [ 45        , 180     , 25       , 0.6*Teff ], # mean if U,
        →  mean if G, fixed value of F
        'val2'   : [ 45        , 175     , 25       , 0.4*Teff ], # (+-) if U,
        →  std if G, none if F
        'init'   : [ 50        , 180     , 5        , 0.7*Teff ], # (+-) if U,
        →  std if G, none if F
        'initvar' : [ 5        , 175     , 2.5      , 0.1*Teff ], # (+-) if U,
        →  std if G, none if F
    }


    params = paramsNoSpots
    for key in paramsNoSpots:
        if nbSpots == 1:
            for i in range(4):
                params[key].append(paramsSpot1[key][i])
        elif nbSpots ==2:
            for i in range(4):
                params[key].append(paramsSpot1[key][i])
            for i in range(4):
                params[key].append(paramsSpot2[key][i])
```

```
488          else:
489              continue
490
491
492     # params = paramsNoSpots
493     # for key in paramsNoSpots:
494     #     for i in range(4*nbSpots):
495     #         params[key].append(paramsSpots[key][i%4])
496
497
498     ## With GP ##
499
500     # params = {
501     #     'name'    : ['t0'     ,'per'  ,'rp'   ,'a'    ,'inc','ecc','w'
       → ,'u11' ,'u12','ln_S0' ,'ln_Q' ,'ln_w0','vsini','lam','nonrotv','u21'
       → ,'u22'  ,'clatSpot','lonSpot','sizeSpot','tempSpot'],
502     #     'type'    : ['U'      ,'U'    ,'U'    ,'U'    ,'U' ,'F'  ,'F'  ,'F'
       → ,'F' ,'U'    ,'U'   ,'G'    ,'G'    ,'U' ,'U'      ,'F'    ,'F'
       → ,'U'       ,'U'     ,'U'      ,'U'       ], # U for uniform, G for
       → Gaussian, F if fixed
503     #     'val1'    : [1604.0237, 6.9595, 0.0687, 11.55, 88 , 0   , 0   ,
       → 0.148, 0.23, -2.5   , 3.1  , -0.33 , 50    , 0  , 3        , 0.4139,
       → 0.2494, 45       , 180     , 25       , 0.6*Teff ], # mean if U, mean
       → if G, fixed value of F
504     #     'val2'    : [0.1      , 0.02 , 0.01 , 2.55 , 2   , None, None,
       → None , None, 2.5   , 2    , 0.001 , 3     , 180 , 3        , None  ,
       → None , 45       , 180     , 25       , 0.4*Teff ], # (+-) if U, std
       → if G, none if F
505     # }
506     #     'val1'    : [1604.0217, 6.9395, 0.0557, 9   , 86 , 0   , 0   ,
       → 0.148, 0.23, 0   , 3  , 1.4, 45     , -180, 0        , 0.4139, 0.2494, 1
       → , 0         , 0        , 1        , 0.2*Teff ], # mean if U, mean if G,
       → fixed value of F
507     #     'val2'    : [1604.0257, 6.9795, 0.0787, 14.1, 90  , None, None, None
       → , None, 0.05, 200, 0.1, 60    , 180 , 6        , None  , None   , None
       → , 90       , 360     , 50       , Teff     ], # (+-) if U, std if G,
       → none if F
508     # }
509
```

149

```python
510    # Separate fitted from fixed parameters
511    fittedParams, fixedParams = separate_params(params)
512
513
514    ###################
515    #### Run Emcee ####
516    ###################
517
518    ## EMCEE ##
519
520    nwalkers = 40
521    max_iter = 1000
522    nCPUs = 120
523
524    total_time = 0
525    process_time = 0
526
527    x0 = run_mcmc(fittedParams, fixedParams, params, dataPhoto, dataSpectro,
       →   nwalkers, max_iter, nbSpots, perStar, Teff, limbDarkLaw,
       →   instResolution, nCPUs, constants)
```

C.2 Transit model

The functions presented below generate the model for the photometric transits applied on the TESS data.

```python
import numpy as np
from scipy import interpolate
from scipy.stats import norm
import batman
import copy
import celerite
from scipy import interpolate

def run_batman(parameters, tbjd, limbDarkLaw):

    """
    phase is tbjd
    u1 u2 -- quadratic limb darkening parameters
    vsini -- free parameter constrained by a Gaussian prior
    per -- orbital period in days
    rp -- Rp/R*
    sma -- a/R*
    inc -- inclination in deg
    lam -- projected obliquity in deg
    nonrotv -- non-rotational broadening component in km/s, constrained by
    →  Gaussian prior if possible
    resolution -- spectral resolution of instrument
    """

    ### use batman to calculate the photometric transit
    params = batman.TransitParams()
    #t0 = parameters[0]
    params.t0  = parameters[0]                      #time of inferior conjunction
    params.per = parameters[1]                          #orbital period
    params.rp  = parameters[2]                          #planet radius (in units
    →  of stellar radii)
```

```python
        params.a   = parameters[3]                          #semi-major axis (in
            → units of stellar radii)
        params.inc = parameters[4]                          #orbital inclination (in
            → degrees)
        params.ecc = parameters[5]                          #eccentricity
        params.w   = parameters[6]                          #longitude of periastron
            → (in degrees)
        params.u   = [parameters[7],parameters[8]]  #limb darkening coefficients
            → [u1, u2] EDIT THIS for your own star!!
        params.limb_dark = limbDarkLaw                      #limb darkening model

        phase = tbjd
        model = batman.TransitModel(params, phase)
        return model.light_curve(params)        #calculates light curve


def compute_ingress_egress(RPlanet, semiMajorAxis, inclination):

    """
    Compute the transit duration + ingress and egress times.
    Inputs:
    - RPlanet (float): Planet radius [RSun]
    - semiMajorAxis (float): semi major axis [RSun]
    - inclination (float): orbit inclination in degrees
    Outputs:
    - ingress (float): ingress time in days relative to t0
    - egress (float): egress time in days relative to t0
    - TD (float): Transit duration in days
    """

    Rp = copy.deepcopy(RPlanet)
    a = copy.deepcopy(semiMajorAxis)
    inc = copy.deepcopy(inclination)


    inc *= np.pi/180


    # impact parameter

    b = a*np.cos(inc)

```

```python
65      # Transit duration in days
66      TD = np.arcsin(np.sqrt((1/a)**2 + 2*(1/a)**2 * Rp + (Rp/a)**2 -
   →       b**2/a**2))/np.pi
67
68      ingress = -TD/2
69      egress = TD/2
70
71      return ingress, egress, TD
72
73  def lnlike_LC(params,t,flux,err,return_gp_model=False):
74
75      """"
76      Compute the likelihood of a Gaussian process model on a lightcurve using
   →  celerite.
77
78      Inputs:
79      - params (list of floats): the three GP parameters (Semi Harmonic
   →  Oscillator)
80      - t (array): time vector
81      - flux (array): flux vector
82      - err (array): error on flux vector
83      - return_gp_model (bool): If false return only the likelihood, if True
   →  also returns the prediction from the gp model
84      Outputs:
85      -  lnlike (float): the likelohood of gp_model - data
86      if return_gp_model is True
87      - t (array): time vector
88      - predMean (array): predicted mean of the gp (size matching t)
89      - predVar (array): predicted variance of the gp (size matching t)
90
91      """"
92
93
94      ln_S0,ln_Q,ln_w0 = params # Should already be in log form
95
96      # Apply SHOT GP
97      kernel = celerite.terms.SHOTerm(log_S0=ln_S0, log_Q=ln_Q,
   →   log_omega0=ln_w0)
98      kernel.freeze_parameter("log_Q")
```

```python
99         kernel.freeze_parameter("log_S0")
100        kernel.freeze_parameter("log_omega0")
101
102        gp = celerite.GP(kernel, mean=np.mean(flux))
103
104        try:
105            gp.compute(t,err)
106            lnlike = gp.log_likelihood(flux)
107        except:
108            lnlike = 0
109
110        predMean = np.array([])
111        predVar = np.array([])
112
113        if return_gp_model:
114            print('predicting...')
115            #t2 = np.linspace(min(t),max(t),1000) ### compute for 1000 points to
                → save some time
116            ## to the prediction in chunks to save memory
117            batchSize = 1000
118            if len(t)%batchSize == 0:
119                nbBatches = int(len(t)/batchSize)
120            else:
121                nbBatches = int(len(t)/batchSize) + 1
122
123            for i in range (nbBatches):
124                if i == (nbBatches-1): # last iteration
125                    gp.compute(t[i*batchSize:],err[i*batchSize:])
126                    predMeanBatch, predVarBatch = gp.predict(flux[i*batchSize:],
                        → t[i*batchSize:], return_var=True)
127                else:
128                gp.compute(t[i*batchSize:(i+1
129                )*batchSize],err[i*batchSize:(i+1)*batchSize])
130                predMeanBatch, predVarBatch =
                    → gp.predict(flux[i*batchSize:(i+1)*batchSize],
                    → t[i*batchSize:(i+1)*batchSize], return_var=True)
131
132                predMean = np.append(predMean,predMeanBatch)
133                predVar = np.append(predVar,predVarBatch)
```

```python
134
135            # pred_mean = interpolate.splrep(t2,pred_mean)
136            # pred_mean = interpolate.splev(t2,pred_mean)
137            return lnlike, t, predMean, predVar
138
139        return lnlike
```

C.3 Spectroscopic model

The functions presented below generate the line profile resulting from the modelled stellar disk with a spot and a transiting planet.

```python
from math import floor
from pylab import *
import copy
import numpy as np
from scipy.stats import norm
import os
#warnings.filterwarnings("ignore", category=RuntimeWarning)


## Constants ##
MSun = 1.9885e30
RSun = 6.957e8
G = 6.67e-11
c = 3e5

def read_dt_data(path, centervel, truncate):

    """
    Read the LSD files and create 4 arrays:
    - epochsLSD  # Epochs of LSD
    - velLSD     # LSD velocities
    - intLSD     # LSD intensity
    - errLSD     # LSD error

    Inputs:
    - path (str): path of the folder containing the LSD files
    - centervel (bool): if True, center all the LSD profiles on the mean value
 →  of all profiles
    - truncate (float): Truncate the profile at + and _ the given value.
    """

```

```python
filenamesDT = os.listdir(path)

epochsLSD = np.array([])      # Epochs of LSD
velLSD = []                   # LSD velocities
intLSD = []                   # LSD intensity
errLSD = []                   # LSD error

# Iterate through files
for filename in filenamesDT:

    with open(path + filename, 'r') as LSDprof:
        # get velocities, intensities and errors from current file
        velCurrent = np.array([])
        IntCurrent = np.array([])
        errCurrent = np.array([])

        i = 0
        for line in LSDprof:
            if i == 0:
                # get epoch
                epochCurrent = float(line.split(' ')[1])
            else:
                velCurrent = np.append(velCurrent,float(line.split('
                →  ')[0]))
                IntCurrent = np.append(IntCurrent,float(line.split('
                →  ')[1]))
                errCurrent = np.append(errCurrent,float(line.split('
                →  ')[2]))
            i += 1

        # Add info from current files to the whole arrays/lists containing
        →  all the LSD data
        epochsLSD = np.append(epochsLSD, epochCurrent)
        velLSD.append(velCurrent)
        intLSD.append(IntCurrent)
        errLSD.append(errCurrent)


nbEpochs = len(epochsLSD)
```

```python
67
68      if centervel: # center the LSD profiles, using the meanvel computed from
    →     the mean of all maximums
69          maxs = np.array([])
70          for i in range (nbEpochs):
71              maximum = intLSD[i].max()
72              indexmax = np.where(intLSD[i] == maximum)
73              maxs = np.append(indexmax,maxs)
74
75          meanvelindex = maxs.mean()
76          meanvel = velLSD[0][int(np.round(meanvelindex))]
77
78          for j in range(nbEpochs):
79              velLSD[j] -= meanvel
80
81      if truncate != -1:
82          indexmaxvel = np.where(velLSD[0] >= truncate)[0][0]
83          indexminvel = np.where(velLSD[0] <= -truncate)[0][-1]
84
85          for k in range(nbEpochs):
86              velLSD[k] = velLSD[k][indexminvel:indexmaxvel+1]
87              intLSD[k] = intLSD[k][indexminvel:indexmaxvel+1]
88              errLSD[k] = errLSD[k][indexminvel:indexmaxvel+1]
89
90      return epochsLSD, nbEpochs, velLSD, intLSD, errLSD
91
92  def limb_darkening(limbLaw, coefs, dist):
93
94      """
95      Compute the mask due to limb darkening (LD).
96
97      Inputs:
98      - limbLaw (str): the LD law to be applied
99      (['uniform','linear','quadratic','sqrt','exp' or 'non-linear'],
100     definitions are taken from the batman package.
101     - coefs (list of floats): coefficients for the LD law
102     - dist (meshgrid): grid of pixel with value of 0 for pixels outside the
    →   unit circle
103     and value 1 for the pixels inside the unit circle
```

```python
104        Outputs: grid of pixel with fraction of brightness due to the LD (function
    →  of distance from center)
105        """
106
107        ## Taken from the batman package
108        dist[dist > 1] = 1
109        mu = np.sqrt(1-dist**2) # radial distance to the center of the star (pixel
            →  grid)
110
111        try:
112            # Uniform
113            if limbLaw == 'uniform':
114                return 1
115            # Linear
116            elif limbLaw == 'linear':
117                return (1 - coefs[0]*(1-mu))
118            # Quadratic
119            elif limbLaw == 'quadratic':
120                return (1 - coefs[0]*(1-mu) - coefs[1]*(1-mu)**2)
121            # Square-root
122            elif limbLaw == 'sqrt':
123                return (1 - coefs[0]*(1-mu) - coefs[1]*(1-np.sqrt(mu)))
124            # log
125            elif limbLaw == 'log':
126                return (1 - coefs[0]*(1-mu) - coefs[1]*mu*np.log(mu))
127            # exp
128            elif limbLaw == 'exp':
129                return (1 - coefs[0]*(1-mu) - coefs[1]/(1-np.exp(mu)))
130            # non-linear
131            elif limbLaw == 'non-linear':
132                return (1 - coefs[0]*(1-np.power(mu,1/2)) - coefs[1]*(1-mu) -
                    →  coefs[2]*(1-np.power(mu,3/2)) - coefs[3]*(1-np.power(mu,2)))
133        except:
134            print('unknown Limb-darkening law')
135
136    def gaussian (x, amp, cen, wid):
137        """
138        Gaussian function.
139        Inputs:
```

```python
140        - x: input values
141        - amp: height of the curve's peak
142        - cen: mean
143        - wid: standard deviation
144        """
145        return amp * np.exp(-(x-cen)**2/(2 * wid**2))
146
147    def coord_rot(x,y,angle):
148
149        """
150        Rotate the coordinate system '(x,y)' by the 'angle'
151        Inputs:
152        - x (array)
153        - y (array)
154        - angle (float)
155        """
156        angle *= np.pi/180
157
158        xRot =    x * np.cos(angle) - y * np.sin(angle)
159        yRot =    x * np.sin(angle) + y * np.cos(angle)
160
161        return xRot,yRot
162
163    def make_spot (clat, lon, size, inc = None):
164
165        """
166        Generates the location of the center, the semi-minor and semi-major axis
167        of the ellpise representing a projected spot on the stellar surface grid.
168        Inputs:
169        - clat (float): co-latitude of spot in degrees
170        - lon (float): longitude of spot in degrees
171        - size (float): Angular size of spot in degrees
172        - inc (float): stellar inclination relative to the line of sight in
    →   degrees
173        Outputs:
174        - xc (float): x coordinate of the spot center on the pixel grid
175        - yc (float): y coordinate of the spot center on the pixel grid
176        - ax (float): semi minor axis of projected spot (ellipse)
177        - ay (float): semi major axis of projected spot (ellipse)
```

```python
      """

      # phi (0:180) : colatitude, from north pole downwards
      # theta (0:360) : longitude, starting at (x,y) = (0,0) on the disk plane
      phi = copy.deepcopy(clat)
      theta = copy.deepcopy(lon)
      angularSize = copy.deepcopy(size)

      # given in deg, convert to red
      theta *= np.pi/180
      phi *= np.pi/180
      angularSize *= np.pi/180

      # center of projected spot
      xc = np.cos(angularSize) * np.sin(theta) * np.sin(phi)
      yc = np.cos(angularSize) * np.cos(phi)

      # radius of spot:
      r = np.sin(angularSize)

      # projected ellipse surface:
      ax = r * np.cos(np.pi/2 -phi) * np.cos(theta)
      ay = r

      return xc, yc, ax, ay

def spot_mask (clatSpot, lonSpot, angularSizeSpot, Teff, TSpot, X, Y):

      """
      Generate the mask to be applied on the stellar surface for one spot. Based
      on the location and sizes of the projected ellipe computed on make_spot,
      we buid a mask for the spot. The projection can be the shape of an ellipse
      if the entire spot is visible or either an ellipse+lune shape or a lune
  →   shape
      when it is partially behinf the visible hemisphere. See details in
      'Stratified sampling of projected spherical caps, Carlos Urena 2018
      in Eurographics Symposium on Rendering 2018'.

      Inputs:
```

```
216        - clatSpot (float): co-latitude of spot in degrees
217        - lonSpot (float): longitude of spot in degrees
218        - angularSizeSpot (float): Angular size of spot in degrees
219        - Teff (float): stellar effective temperature [K]
220        - TSpot (float): Spot temperature
221        - X (nd array): x coordinates of the meshgrid
222        - Y (nd array): y coordinates of the meshgrid
223        Output:
224        - maskSpot: gris with values 1 outside the spot projected area and
225        value 'spotbrightness' (depending on spot temperature) inside the
226         spot projected area.
227        """
228
229        # Get spot projected ellipse position and size
230        xSpot, ySpot, minAxis, majAxis =
           →  make_spot(clatSpot,lonSpot,angularSizeSpot)
231
232        alpha = angularSizeSpot
233
234        # Compute angle to rotate the surface area
235        distCentre = np.sqrt(xSpot**2 + ySpot**2)
236
237        if lonSpot < 90 or lonSpot > 270: # Spot on the invisible hemisphere
238            #beta  = 90 - lonSpot
239            #beta = np.arccos(majAxis**2-minAxis**2)
240            invisibleHemisphere = True
241            beta = np.arcsin(minAxis/majAxis)/np.pi*180
242        else: # Spot on the visible hemisphere
243            #beta  = lonSpot - 270
244            #beta = - np.arccos(majAxis**2-minAxis**2)
245            invisibleHemisphere = False
246            beta = np.arcsin(minAxis/majAxis)/np.pi*180
247
248        if alpha >= 0 and beta >= alpha:  # Spot entirely visible
249            projectedAreashape = 'Ellipse'
250        elif beta >= 0 and alpha > beta:  # Spot partially visible
251            projectedAreashape = 'EllipseAndLune'
252        elif -beta > 0 and alpha > -beta:  # Spot partially visible
253            projectedAreashape = 'Lune'
```

```python
254        else: # alpha >= 0 and -beta >= alpha:  # Spot entirely behind the horizon
255            projectedAreashape = 'None'
256
257        if ySpot > 0:
258            rotAngle = - np.arccos(xSpot/distCentre)
259        else:
260            rotAngle = np.arccos(xSpot/distCentre)
261
262        # Spot temperature
263        spotBrightness = TSpot**4/(Teff**4)
264
265        #### Draw spot projection
266
267        ## Ellipse
268        # Rotate axes with the center of the ellipse as origin
269        XRotSpot,YRotSpot = coord_rot(X-xSpot,Y-ySpot,rotAngle*180/np.pi)
270
271        #maskEllipse = np.sqrt(((X-xSpot)/minAxis)**2 + ((Y+ySpot)/majAxis)**2)
272
273        maskEllipse = np.sqrt((XRotSpot/minAxis)**2 + (YRotSpot/majAxis)**2)
274        # Draw
275        outsideEllipse = maskEllipse > 1
276        insideEllipse = maskEllipse <= 1
277        maskEllipse[insideEllipse] = spotBrightness
278        maskEllipse[outsideEllipse] = 1
279
280        #if projectedAreashape == 'EllipseAndLune' or projectedAreashape ==
           ↪  'Lune':
281        ## Lune
282        # Rotate axes
283        XRot,YRot = coord_rot(X,Y,rotAngle*180/np.pi)
284
285        if projectedAreashape == 'Ellipse' and not invisibleHemisphere:
286            maskSpot = X*Y
287            maskSpot[:] = 1
288            return maskSpot
289
290        if projectedAreashape in ['EllipseAndLune','Lune']:
291
```

```python
292          # union of C1 and C2, with C1 our stellar disk and C2 a intersecting
             ↪  circle
293          # from equation 11 in 'Stratified sampling of projected spherical
             ↪  caps,
294          # Carlos Urena 2018 in Eurographics Symposium on Rendering 2018'
             ↪  (Carlos et al. hereafter)
295          xl = np.cos(alpha*np.pi/180)/np.cos(beta*np.pi/180)
296          #(majAxis**2*distCentre)/(majAxis**2-minAxis**2)
297
298          yl = np.sqrt(1-xl**2)
299
300          centerToOuterCircle = (distCentre + np.abs(minAxis))
301
302          # Compute the radius of the second (outer) circle C2
303          H = centerToOuterCircle - xl # see figure 2 in Carlos et al. with
             ↪  xSpot = x_e and minAxis = a_x)=
304          W = 2*yl
305          radiusOuterCircle = (H/2) + (W**2/(8*H))
306          distancebetweenCircles = radiusOuterCircle - centerToOuterCircle
307          innerCircle = np.sqrt(XRot**2 + YRot**2)
308          outerCircle = np.sqrt(((XRot +
             ↪  distancebetweenCircles)/radiusOuterCircle)**2 +
             ↪  ((YRot)/radiusOuterCircle)**2)
309          #Draw
310          outsideInnerCircle = innerCircle > 1
311          insideInnerCircle = innerCircle <= 1
312          innerCircle[insideInnerCircle] = 1
313          innerCircle[outsideInnerCircle] = 0
314
315          outsideOuterCircle = outerCircle > 1
316          insideOuterCircle = outerCircle <= 1
317          outerCircle[insideOuterCircle] = 1
318          outerCircle[outsideOuterCircle] = spotBrightness
319
320          maskStar = np.sqrt(X**2 + Y**2)
321          # Draw
322          outsideStar = maskStar > 1
323          insideStar = maskStar <= 1
324          maskStar[insideStar] = 0
```

```python
325            maskStar[outsideStar] = 1

326

327            maskLune = innerCircle * outerCircle

328

329        if projectedAreashape == 'Ellipse':
330            maskSpot = maskEllipse
331        elif projectedAreashape == 'EllipseAndLune':
332            maskSpot = maskEllipse*maskLune
333        elif projectedAreashape == 'Lune':
334            maskSpot = maskLune
335        else: # alpha >= 0 and -beta >= alpha:  # Spot entirely above the horizon
336            maskSpot = X*Y
337            maskSpot[:] = 1

338

339        return maskSpot

340

341    def planet_mask(X, Y, phase, rp, sma, inc, lam):

342

343        """"
344        Generate the mask to be applied on the stellar surface for the passing
    →   planet
345        at a given phase. Based on planet and ster parameters.

346

347        Inputs:
348        - X (nd array): x coordinates of the meshgrid
349        - Y (nd array): y coordinates of the meshgrid
350        - phase (float): phase of the planetary orbit
351        - rp (float): Stellar radius [R*]
352        - sma (float): orbit semi major axis [R*]
353        - inc (float): orbit inclination in deg
354        - lam (float): projected obliquity in deg
355        Output:
356        - maskSpot: gris with values 1 outside the spot projected area and
357        value 0 inside the planet projected area.
358        """"
359        # Planet coordinates on the plane of the stellar disc, in stellar radius
360        xp = sma*np.sin(2*np.pi*phase) #(+omega)
361        yp = -1*sma*np.cos(2*np.pi*phase)*np.cos(inc*np.pi/180.)

362
```

```python
363      ### Rotate the coordinates by the orbital obliquity lambda
364      XRot , YRot = coord_rot(X, Y, lam)
365
366      # Draw circle (eclipsing planet)
367      maskPlanet = np.sqrt((XRot - xp)**2 + (YRot-yp)**2)
368      outside = maskPlanet > rp
369      inside = maskPlanet <= rp
370      # For presentation plot
371      # maskPlanet[inside] = 0.01
372      # maskPlanet[outside] = 1
373      maskPlanet[inside] = 0
374      maskPlanet[outside] = 1
375
376      return maskPlanet
377
378  def compute_DT_model(vel, meanLSD, instResolution, paramsDopplerTomography,
     →  paramsPlanet, epochsLSD, perStar, paramSearch, addNoise = False):
379
380      """"""
381      Main function of the Doppler Tomography. We generate grid of pixel, in
     →  which a unit circle is drawn
382      with values of 1 inside and 0 outside representing the star. Each pixel
     →  value is a brightness value
383      (from 0 dark to 1 max brightness). For each phase that we wish to compute
     →  an image and associated
384      simulated line profile we add to the grid of pixel:
385      Limb Darkening (with the limb_darkening function), spot masks (with the
     →  spot_mask function) and planet mask (with
386      the planet_mask function). We then sum the brightness contribution
     →  vertically to generate the line profiles to which we
387      add the instrmental and macroturbulence contribution.
388      Inputs:
389      - vel: velocity array of the LSD profiles (same for all LSD profiles)
390      - meanLSD: mean of all the LSD profiles
391      - instResolution: instrument resolution
392      - paramsDopplerTomography: list of DT parameters, see below in the
     →  function
393      - paramsPlanet: planet parameters (see below)
394      - epochsLSD: epochs of the LSD profiles
```

```
395        - perStar: Stellar rotation period
396        - paramSearch (bool): Set to True when doing the parameter search,
    →    otherwise
397        will generate extra things useful for plotting but slowing things for
    →    parameter
398        retrievial
399        - addNoise (bool): add noise to the generated line profiles.
400
401        Outputs:
402        - allMasks: Generated masks containing spots, planet and Limb Darkening
403        - allProfPlanet: Generated line profiles from the planet
404        - allProfTotal: Generated complete line profiles (generaed from allMasks)
405        - allProfRaw: Generated masks for profiles only with Limb Darkening
406        - MasksStackTot: Mask which is the product of all independant masks
407        - allProfSpots: Generated line profiles from the spots
408        - dist: grid of pixel with value of 0 for pixels outside the unit circle
409        and value 1 for the pixels inside the unit circle
410
411        """"
412
413        paramsDT = copy.deepcopy(paramsDopplerTomography)
414        paramsPla = copy.deepcopy(paramsPlanet)
415
416        vsini            = paramsDT[0] # vsini
417        nbSpots          = paramsDT[1] # number of spots
418        spots            = paramsDT[2] # spot parameters, dictionnary with
    →    {'colatitude' : colatitudes, 'longitude' : longitudes, 'spotSize' :
    →    spotSizes, 'spotTemp' : spotTemps}
419        Teff             = paramsDT[3] # effective temperature
420        limbDarkLaw      = paramsDT[4] # law for limb darkening
421        limbDarkCoeffs   = [paramsDT[5],paramsDT[6]] # coefficient for LD law
422        phase            = paramsDT[7] # planetary phases
423        nonRotV          = paramsDT[8] # macroturbulence
424
425        rp               = paramsPla[0] # planet radius
426        sma              = paramsPla[1] # semi major axis
427        inc              = paramsPla[2] # orbit inclination
428        lam              = paramsPla[3] # projected obliquity
429
```

```python
## Model initialization ##
## Setup grid ##
profileExtent = vel[-1]/vsini
nbBinsVel = len(vel)
x = np.linspace(-profileExtent,profileExtent,nbBinsVel)
y = np.linspace(profileExtent,-profileExtent,nbBinsVel)

X, Y = np.meshgrid(x, y, sparse = True)

## Model stellar disk ##
maskStar = np.zeros((len(x),len(y)))
dist = np.sqrt(X**2 + Y**2)
# For presentation plot
# maskStar[dist <= 1] = 1
# maskStar[dist > 1] = 2
maskStar[dist <= 1] = 1
maskStar[dist > 1] = 0

## Stellar brightness normalisation for line profile ##
sumStar = np.sum(maskStar, axis = 0) # sum along the vertical (y) axis
#sumStar[sumStar == 0] = 0 # replace zeros with ones to avoid dividing by
    →  0

## (2) Instrument and macroturbulence profile ##
gauMean = 0
gauStd = np.sqrt((c/instResolution)**2 + nonRotV**2) # in km/s
profInst = norm.pdf(vel, gauMean, gauStd)

## Randomly generate spots lon, clat and angular size ##
allMasks = []
allProfRaw = []
allProfSpots = []
allProfPlanet = []
allProfTotal = []
MasksStackTot = np.ones((len(x),len(y)))

nbEpochs = len(epochsLSD)

# Express phases in unit of perStar
```

```python
468      #phaseStar = (epochsLSD - epochsLSD[0]) / perStar # Old way, taking the
    →    first epoch of the current dataset as origin
469      phaseStar = (epochsLSD / perStar) - np.floor(epochsLSD / perStar) # Taking
    →    0 as origin of time, which is 2457000 (TBJD)
470
471      # Compute the associated increment in longitude
472      lonIncrement = phaseStar * 360
473
474      ## Iterate through the phases ##
475      for i in range(nbEpochs):
476
477          # Initialise spots mask:
478          maskSpots = np.ones((len(x),len(y)))
479          # For presentation plot
480          # maskSpots[dist <= 1] = 1
481          # maskSpots[dist > 1] = 2
482          maskSpots[dist <= 1] = 1
483          maskSpots[dist > 1] = 0
484          # Initialise total mask:
485          maskTotal = np.ones((len(x),len(y)))
486
487          ## Iterate through the spots ##
488          for j in range(nbSpots):
489
490              # Current spot parameters
491              clatSpot = spots['colatitude'][j]
492              if j == 0:
493                  lonSpot = spots['longitude'][j]
494              else:
495                  lonSpot = spots['longitude'][0] + spots['longitude'][j]
496              # Separate spots in longitude: # to be reworked if used
497              # lonSpot = spots['longitude'][j] * 360 * ( (j+1)/nbSpots +
    →            j/nbSpots) / nbSpots + (j * 360 / nbSpots)
498              angularSizeSpot = spots['spotSize'][j]
499              # Add the phase to rotate the spots
500              lonSpot = (lonSpot + lonIncrement[i]) % 360
501              TSpot = spots['spotTemp'][j]
502
503              # Make and compute spot mask
```

```python
            maskSingleSpot = spot_mask(clatSpot, lonSpot, angularSizeSpot,
            →  Teff, TSpot, X, Y)
            # Add the now spot to the mask
            maskSpots *= maskSingleSpot


        # LimbDarkening
        maskLimbDarkening = limb_darkening(limbDarkLaw, limbDarkCoeffs, dist)


        # Mask raw star
        maskRaw = maskStar*maskLimbDarkening


        # Compute the normalisation constant on the first iteration
        if i == 0:
            sumRaw = np.sum(maskRaw,axis = 0)/np.max(sumStar) # unspotted pure
            →  mask
            # Convolve with the instrumental profile
            profRaw = np.convolve(sumRaw,profInst, mode = 'same') # unspotted
            →  pure profile (uniform disk + limb darkening)
            normConstant = np.sum(profRaw) # Normalisation constant
            profRaw /= normConstant # Normalise


        # Mask raw with spots
        maskSpotted = maskRaw*maskSpots


        # Planet
        maskPlanet = planet_mask(X, Y, phase[i], rp, sma, inc, lam)


        # Mask total (uniform disk + spots)
        maskTotal = maskSpotted*maskPlanet


        # Total profile, sum along the vertical axis
        sumTotal = np.sum(maskTotal,axis = 0)/np.max(sumStar) # sum of
        →  vertical slice of width 'resolution'


        ## Convolve with the instrumental profile
        profTotal = np.convolve(sumTotal,profInst, mode = 'same')


        # New way
```

```
538         profTotal /= normConstant # Normalise using the sum of the raw
      →   unspotted profiles
539         profTotalNorm = meanLSD  + (profTotal - profRaw) # Add to the mean of
      →   LSD profiles the variation from the unspotted normalized profile
540         profTotal = profTotalNorm
541
542         # Old way, use when we don't have LSD prof? i.e. for simulations
543         # profTotal /= normConstant
544
545         # Normalise (old, already done above)
546         # normConstant = np.max(meanLSD)   # normalisation constant to scale
      →   to max of mean LSD prof
547         # normConstant = np.max(profTotal) # normalisation constant to scale
      →   up to one
548         # normConstant = np.sum(profTotal) # normalisation constant to scale
      →   to a sum of 1
549
550
551
552         ## Raw profile
553         if not paramSearch: # to not pollute computations as useless when
      →   doing the main fitting
554             # Raw profile
555             allProfRaw.append(profRaw)
556
557             # Spotted profile (no planet)
558             profSpot = np.sum(maskSpots*maskStar,axis = 0)/np.max(sumStar)
559             profSpot -= (sumStar/np.max(sumStar))
560             profSpot = np.convolve(profSpot,profInst, mode = 'same')
561             profSpot /= normConstant
562             allProfSpots.append(profSpot)
563
564             ## Planet contribution
565             profPlanet = np.sum(maskPlanet*maskStar,axis = 0)/np.max(sumStar)
566             profPlanet -= (sumStar/np.max(sumStar))
567             profPlanet = np.convolve(profPlanet,profInst, mode = 'same')
568             profPlanet /= normConstant
569             #profPlanet[np.where(profPlanet == 0)] = np.max(profPlanet)
570             #profPlanet -= (np.max(profPlanet) + np.max(profTotal)/5)
```

```python
            #if np.isnan(profPlanet[0]):
            #    print('nan')
            allProfPlanet.append(profPlanet)


            # Stacked masks
            maskStack = maskStar*maskSpots*maskPlanet
            MasksStackTot *= maskStack


        ## Add noise
        if addNoise:
            noise = np.zeros((len(profTotal)))
            noise = np.random.normal(noise, 0.03)
            profTotal = profTotal+noise


        allMasks.append(maskTotal)
        allProfTotal.append(profTotalNorm)


    if not paramSearch: # to not pollute computations as useless when doing
    →   the main fitting
        #allMasksStack  = np.prod(allMasksStack, axis = 0)
        MasksStackTot *= maskLimbDarkening
        return allMasks, allProfPlanet, allProfTotal, allProfRaw,
        →   MasksStackTot, allProfSpots, dist
    else:
        return allMasks, allProfTotal, dist


def compute_RVs(allProfTotal, vel):

    """
    Compute the resulting RVs from each line profiles
    Inputs:
    - allProfTotal: All generated line profiles
    - vel: velocity array
    Outputs:
    - RV: array of radial velocities for each profiles.
    """

    RV = np.array([])

```

```python
608      for prof in allProfTotal:
609          RVCurrent = np.sum(prof*vel)
610          RV = np.append(RV,RVCurrent)
611
612      return RV
613
614  def lnlike_DT(profTotal,intLSD, errLSD):
615      # Compute the ln(likelihood) for the DT fit:
616      return -0.5*np.sum((intLSD-profTotal)**2/(errLSD**2))
```

This page intentionally left blank.