Stellar activity in Radial Velocity surveys



Haochuan Yu

St Cross College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2025

DECLARATION

I carried out the work presented in this thesis as a student at the University of Oxford between December 2021 and July 2025, under the supervision and guidance of Professor Suzanne Aigrain. My research was funded by the European Research Council under the European Union's Horizon 2020 research and innovation programme (title: GPRV, grant agreement No 865624, PI: S. Aigrain). I hereby declare that no part of this thesis has been submitted in support of another degree, diploma or other qualification at the University of Oxford or other higher learning institute. Except where otherwise stated or where reference is made to the work of others, the work in this thesis is entirely my own.

The work presented in Chapter 2 is based on the peer-reviewed paper Yu et al. (2024b) *Modelling stellar variability in archival HARPS data: I - Rotation and activity properties with multidimensional Gaussian processes*, for which I am the lead author, contributed the majority of the work and wrote the manuscript.

The work in Chapter 3 is based on the peer-reviewed paper Yu et al. (2024a) *A Gaussian process model for stellar activity in 2D line profile time-series*, where I am the lead author, proposed the original idea, conducted the majority of the research, and wrote the manuscript.

The work in Chapter 4 is based on the peer-reviewed paper Yu et al. (2025) *A possible misaligned orbit for the young planet AU Mic c.* I am the lead author and principal investigator (PI) of the ESPRESSO observations. I led the analysis, conducted the joint photometric and spectroscopic analysis, and wrote the majority of the manuscript. The ESPRESSO data reduction and spectral-level flare correction were carried out by Michael Cretignier. The CHEOPS observations, data reduction, and a separate analysis—while not included in the final study—were highly valuable in shaping the final version of the analysis. These were conducted by Zoltán Garai, Gyula Szabó, Davide Gandolfi, and members of the CHEOPS consortium. The NGTS observations, data reduction and analysis in the appendix were performed by Edward Bryant and members of the NGTS consortium. The dynamical stability analysis was conducted by Alexandre C. M. Correia.

Acknowledgements

As my time as a DPhil student at Oxford draws to a close, I look back on this journey with deep appreciation – it has been profoundly rewarding. I am immensely grateful to all the people who have supported me along the way. None of this would have been possible without their generosity, encouragement, and guidance.

First and foremost, I would like to express my deep gratitude to my supervisor, Suzanne Aigrain. From the very beginning, she introduced me to the fascinating world of exoplanets and has been unwavering in her support throughout. Suzanne is an exceptionally talented scientist with sharp insight and deep expertise in the field, yet she remains remarkably humble and caring. Thank you, Suzanne, for your trust, your thoughtful guidance, and the freedom you gave me to explore and grow. Your consistently detailed and constructive feedback at every stage helped me build confidence as an independent researcher. I feel incredibly fortunate for the opportunity to learn and develop under your mentorship.

I would also like to extend my gratitude to the Oxford Exoplanet Group, including Michael Crétignier, Baptiste Klein, Oscar Barragán, Geert Jan Talens, Annabella Meech, Niamh O'Sullivan, and many others. In particular, thank you, Michael, for the many valuable discussions we have had. I have learned a great deal from you, especially about instruments and data pipelines. Some of my work would not have been possible without your generous support. I also really enjoyed our time playing tennis and table tennis together!

More broadly, I am grateful to the Department of Physics at the University of Oxford for providing such a stimulating and supportive research environment. In particular, I would like to thank Jayne Birkby and Aris Karastergiou, who served as assessors for my annual progress reviews. Your thoughtful discussions and feedback played a meaningful role in shaping the direction of my research.

I am also grateful to the many people I have met in the broader exoplanet community. Thank you for supporting observational projects, discussing ideas, collaborating, and kindly hosting my visits. In particular, my sincere thanks go to Davide Gandolfi, Fei Dai, Yan Liang, Andrew Howard, Jennifer Burt, Annelies Mortier, and Samantha Thompson for their support and encouragement.

I have been incredibly fortunate to meet many great mentors throughout my scientific journey. Especially, I would like to thank Jiao He and Harold Linnartz, who

first introduced me to astrophysics research and placed trust in me at a very early stage. Your support and belief in my potential played a pivotal role in setting me on this path. I was deeply saddened by Harold's passing in 2023, and I will always remain grateful for his guidance, encouragement, and the opportunities he gave me.

I am deeply grateful to have met Richard Teague and Karin Öberg, who mentored me during my research on protoplanetary disks and guided me through my first first-authored paper. Your support laid a strong foundation for both my DPhil work and ongoing research. Thank you for your continued encouragement, for hosting my visits to MIT and Harvard, and for the many inspiring conversations about science and proposal ideas.

I am sincerely thankful to Joshua Winn for hosting me at Princeton during a three-month exchange in the final year of my DPhil. I truly enjoyed our scientific discussions and the opportunity to sketch out ideas for future projects. Those conversations broadened my perspective and deepened my understanding of the field. Thank you for your generosity, openness, and inspiring mentorship.

I am especially grateful to Stephen Roberts for his generous support and mentorship – not only in the statistical machine learning aspects of my project, but also in broadening my perspective on research, interdisciplinarity, and lasting scientific impact. His guidance has significantly shaped my outlook and opened doors for future opportunities, inspiring me to think boldly and with greater vision.

To my parents, thank you, always, for your unconditional love and support, which have shaped me into who I am today. I wish words could fully express my gratitude, but they never will. Thank you for believing in me and encouraging my creativity from a young age.

And finally, to Chuqi, thank you for your unwavering support and love throughout this journey, especially for believing in me during the hardest times. You have always stood by me when I chose to take risks, and your encouragement has meant the world. Thank you for pushing me forward, and for lifting me up when I needed it most. I am so excited for all the adventures that lie ahead for us.

Abstract

Although instruments for measuring the radial velocities (RVs) of stars now routinely reach sub-meter per second accuracy, the detection of low-mass planets is still very challenging, and Earth analogues are still beyond reach with current methods. The rotational modulation and evolution of spots and/or faculae can distort the shape of the spectral lines, so as to induce variations in the measured RV time-series.

To overcome this, one approach is to use a multi-dimensional Gaussian Process (GP) framework to model the stellar activity signal using spectroscopic activity indicators together with the RVs. A recently published computationally efficient implementation of this framework using semi-separable matrices enables the rapid and robust analysis of large samples of targets with sizable data sets. I introduce the framework and apply it to HARPS observations of 268 well-observed targets with precisely determined stellar parameters. I focus on the activity indicators, namely S-index and Bisector Inverse Slope, and use them to a) measure rotation periods for 49 slow rotators in my sample, b) explore the impact of these results on the spin-down of middle-aged late F, G & K stars, and c) explore indirectly how the spot to facular ratio varies across my sample. The results should provide valuable clues for planning future RV planet surveys such as the Terra Hunting test (THE) or the PLATO ground-based follow-up observations program, and help fine-tune current stellar structure and evolution models.

An alternative approach is to model the stellar activity signals directly in the 2-D time-series of cross-correlation functions (CCFs). I present a novel activity model for CCFs using Gaussian processes. Compared to previous GP models using 1D time-series (e.g., RV, BIS, FWHM), this uses more of the information available in the spectra, and limits the potential degeneracy between the planet and stellar activity signals. Compared to methods which model the line-shape variations without taking the time dimension into account, the new approach exploits the time correlation in the variations of the CCF. I tested this method for both synthetic and real datasets down to Signal-To-Noise Ratio (SNR) $\sim$ 100, and showed it can disentangle activity from planet signals even if the periods of the planet and activity-induced signals are identical. Furthermore, injection-recovery tests on the HARPS-N solar data show that the framework can blindly recover planet signals down to 0.3 m/s during the high-activity part of the Sun's cycle.

Finally, I discuss stellar activity mitigation in the context of stellar spin–orbit alignment measurements with the case of the AU Microscopii (AU Mic) system, a young (24 Myr old) star known for its intense stellar activity. Here, I present the first measurement of the Rossiter-McLaughlin effect for the warm sub-Neptune AU Mic c, using two transits observed simultaneously with the European Southern Observatory's (ESO's) Very Large Telescope (VLT)/Echelle SPectrograph for Rocky Exoplanets and Stable Spectroscopic Observations (ESPRESSO), CHaracterising ExOPlanet Satellite (CHEOPS), and Next-Generation Transit Survey (NGTS). After correcting for flares and for the magnetic activity of the host star, and accounting for transit-timing variations, we find the sky-projected spin-orbit angle of planet c to be in the range $\lambda_c = 67.8^{+31.7}_{-49.0}$ degrees (1-σ), in contrast to the aligned orbit of the inner planet b ($\lambda_b = -2.96^{+10.44}_{-10.30}$ degrees). We examine the possibility that planet c is misaligned with respect to the orbit of the inner planet b, and the equatorial plane of the host star, and discuss scenarios that could explain both this and the planet c's high density, including secular interactions with other bodies in the system or a giant impact. We note that a significantly misaligned orbit for planet c is in some degree of tension with the dynamical stability of the system, and with the fact that we see both planets in transit, though these arguments alone do not preclude such an orbit. Further observations would be highly desirable to constrain the spin-orbit angle of planet c more precisely.

This thesis advances the field of exoplanet detection by addressing a major limitation of the RV method: stellar activity. The techniques developed here improve planetary signal recovery and contribute to the broader goal of detecting Earth-like exoplanets in the presence of stellar variability.

Contents

1 Introduction **1**

1.1 Context and motivation 1

1.2 Astrophysical background 3

1.2.1 Exoplanets 3

1.2.2 The radial velocity (RV) method 5

1.2.3 Rossiter–McLaughlin effect 10

1.3 Stellar activity 14

1.3.1 Types of stellar activity 14

1.3.2 Methods for mitigating stellar activity 18

1.4 Statistical methods for modelling stellar activity 21

1.4.1 A brief introduction to Gaussian process 21

1.4.2 Multi-dimensional Gaussian process framework 24

1.5 Planets around young stars 27

1.6 Thesis outline 30

2 Modelling stellar variability in archival HARPS data: Rotation and activity properties with multi-dimensional Gaussian Processes **33**

2.1 Introduction 33

2.2 The HARPS Sample 34

2.2.1 Sample selection 34

2.2.2 Data reduction and pre-processing 35

2.2.3 Basic properties of the sample 37

2.3 Model for stellar activity 38

2.4 Stellar rotation 42

2.4.1 Rotation period measurement 42

2.4.2 Rossby number 46

2.4.3 Implications for spin-down of middle-aged stars 47

2.5 Faculae-to-spots ratio 51

2.6 Conclusions 55

3 A Gaussian process model for stellar activity in 2-D line profile time-series 59
3.1 Introduction . . . 59
3.2 A physics-based model for line profile variations . . . 61
3.2.1 Preliminaries . . . 61
3.2.2 Extension to line-profile variations . . . 61
3.2.3 Disentangling activity signals from planets . . . 67
3.2.4 Model inputs . . . 69
3.3 Tests and applications . . . 71
3.3.1 Simulated data from SOAP 2.0 . . . 71
3.3.2 Injection and recovery test on HARPS-N solar data . . . 77
3.3.3 Application to real observations . . . 81
3.4 Discussion . . . 83
3.5 Conclusions . . . 84

4 A possible misaligned orbit for the young planet AU Mic c 87
4.1 Introduction . . . 88
4.2 Observations and data reduction . . . 92
4.2.1 CHEOPS photometry . . . 92
4.2.2 NGTS photometry . . . 94
4.2.3 ESPRESSO spectroscopy . . . 95
4.3 Flare mitigation . . . 97
4.3.1 Identifying and masking flares in the CHEOPS lightcurve . . . 97
4.3.2 Extracting a flare proxy from the ESPRESSO spectra . . . 97
4.3.3 Extracting flare-resistant CCFs and RVs . . . 99
4.4 Joint modelling of the photometry and RVs . . . 100
4.4.1 The model . . . 100
4.4.2 Results . . . 101
4.5 Discussion and conclusions . . . 102
4.5.1 Limitations of our measurement . . . 107
4.5.2 The importance of joint modelling . . . 107
4.5.3 Dynamic stability and transit probability . . . 108
4.5.4 Possible explanations for a misaligned orbit . . . 110

5 Conclusions and Future Work 113
5.1 Thesis summary . . . 113
5.2 Ongoing work . . . 117
5.2.1 2-D spot model . . . 117
5.2.2 Preliminary study of the obliquity of HD 3167 b . . . 120
5.3 Future work . . . 128

Appendices

A Rotation and activity in HARPS data **135**
A.1 Confirmed rotation period measurements 135
A.2 Basic properties of the sample 138

B Obliquity measurement of AU Mic c **151**
B.1 Table for derived values of all free parameters and their posterior distributions 151
B.2 NGTS light curve analysis 154

References **157**

1 Introduction

Contents

1.1 Context and motivation 1
1.2 Astrophysical background 3
1.2.1 Exoplanets 3
1.2.2 The radial velocity (RV) method 5
1.2.3 Rossiter–McLaughlin effect 10
1.3 Stellar activity 14
1.3.1 Types of stellar activity 14
1.3.2 Methods for mitigating stellar activity 18
1.4 Statistical methods for modelling stellar activity . . . 21
1.4.1 A brief introduction to Gaussian process 21
1.4.2 Multi-dimensional Gaussian process framework 24
1.5 Planets around young stars 27
1.6 Thesis outline 30

1.1 Context and motivation

Since the groundbreaking detection of the first exoplanet orbiting a Sun-like star, 51 Peg b (Mayor and Queloz, 1995), using the radial velocity (RV) method, numerous techniques have been developed and campaigns initiated to discover exoplanets with the ultimate goal of finding another habitable world. The RV method continues to be one of the most promising techniques for detecting Earth-like planets, thanks to the development of optical ultra-stable high-resolution échelle spectrographs such

as High Accuracy Radial Velocity Planet Searcher (HARPS; Mayor et al., 2003), HARPS-N (Cosentino et al., 2012a), ESPRESSO (Pepe et al., 2021a), EXPRES (Jurgenson et al., 2016), NEID (Schwab et al., 2018), the KPF (Gibson et al., 2016) and the upcoming HARPS-3 (Thompson et al., 2016). These new instruments can achieve extreme precision down to 0.3 ms^{-1} (e.g., Suárez Mascareño et al., 2020; Faria et al., 2022) and even beyond.

Despite the remarkable sub-meter-per-second precision achieved by modern radial velocity (RV) instruments, stellar activity remains a fundamental barrier to detecting Earth-like exoplanets. The challenge arises because activity-induced RV signals can significantly exceed the planetary signals researchers aim to measure. For instance, in our own solar system, Earth's RV signal has an amplitude of only 0.1 m/s, yet solar activity can induce RV variations exceeding several m/s (e.g., Meunier et al., 2010a; Haywood et al., 2022). This discrepancy underscores the difficulty of distinguishing the weak signatures of small, temperate exoplanets from the much larger stellar noise.

Stellar activity manifests in various forms, including starspots, faculae, convective granulation, and oscillations, all of which introduce complex, time-dependent distortions into RV measurements. Without effective mitigation techniques, these activity-driven variations can obscure or even mimic planetary signals, leading to false detections or missed planets. Consequently, robust methods for modeling and correcting stellar activity are essential for advancing the field of exoplanet characterization. The development of improved statistical frameworks, data-driven models, and physics-informed techniques plays a crucial role in overcoming these challenges, enabling the detection of increasingly Earth-like exoplanets in the presence of stellar variability.

1.2 Astrophysical background

1.2.1 Exoplanets

The discovery and characterisation of exoplanets — planets orbiting stars beyond our solar system — revolutionised our understanding of planetary systems and their diversity. Over the past few decades, astronomers have developed various techniques to detect exoplanets, each with its own strengths and limitations. Figure 1.1 shows a mass–period diagram of the planets detected to date, with colours indicating the different detection methods. The figure was generated using the online tool provided by the NASA Exoplanet Archive website[1] (Christiansen et al., 2025). Among them, the transit method and radial velocity (RV) method are the most widely used, providing crucial insights into planetary properties such as mass, radius, composition, and orbital dynamics.

An Overview of Detection Methods

The **transit** method identifies exoplanets by detecting periodic dips in a star's brightness when a planet passes in front of it. This technique allows for the measurement of planetary radii and, when combined with RV observations, enables the determination of the stellar and planetary bulk densities and surface gravities, providing insights into the planet's compositions and offering essential constraints on its atmospheric scale height. One of its major advantages is the ability to monitor thousands of stars simultaneously, making it highly efficient for large-scale surveys. Planets discovered through transits require follow-up RV measurements to constrain their masses and confirm their planetary nature (e.g., Spiegel and Burrows, 2012; Owen, 2020).

The **radial velocity** method detects exoplanets by measuring the gravitational wobble they induce on their host stars. This technique provides direct constraints on planetary masses, making it a powerful complement to transit observations. Unlike the transit method, which is predominantly employed by space missions

[1] https://exoplanetarchive.ipac.caltech.edu/

such as Kepler (Koch et al., 1996) and TESS (Ricker et al., 2014), RV observations are primarily conducted from ground-based observatories. The RV method is slightly more sensitive to planets with longer orbital periods than those typically detected by transits, though its efficiency is constrained by observational cadence and instrumental precision.

Beyond the two dominant techniques, several other methods have been developed to detect and characterise exoplanets.

The **transit timing variation (TTV)** method detects gravitational interactions between planets by measuring deviations from strictly periodic transit times, enabling the identification of previously unknown planetary candidates. When multiple planets transit the same star, mutual gravitational perturbations can cause measurable variations in their transit timings, which can be used to confirm their planetary nature. Moreover, timing anomalies in a known transiting planet can reveal the presence of additional, non-transiting companions. A compelling early example is the discovery of Kepler-19c, the first planet inferred solely from its gravitational effect on the transit timing of another planet (Ballard et al., 2011). Since then, numerous systems have been characterised using the TTV technique (e.g., Gillon et al., 2017; Hébrard et al., 2020). There are similar techniques, such as the **eclipse timing variation (ETV)** method, which infers the presence of planets from variations in the eclipse timings of binary systems. In addition, a few planets have been detected using the **pulsation timing variation (PTV)** method, where slight, regular changes in the timing of pulsar signals indicate that the pulsar is moving back and forth as it orbits the system's center of mass with one or more planetary companions.

The **astrometry** method detects planets by measuring small changes in a star's position, caused by the gravitational influence of an orbiting companion (e.g., Benedict et al., 2002). Although its current sensitivity limits it primarily to the detection of massive planets, ongoing and future missions such as Gaia (Gaia Collaboration et al., 2016) are expected to significantly enhance its precision. This

is especially promising for identifying planets around young stars, where stellar activity often hampers the effectiveness of RV measurements.

The **microlensing** method enables the detection of exoplanets by leveraging the gravitational bending of light, as predicted by general relativity. When a foreground star hosting a planet (the lens) passes closely in front of a more distant background star (the source), the lens's gravitational field distorts spacetime and bends the light from the source. This causes a temporary magnification of the source's brightness, producing a distinctive light curve. If a planet orbits the lens star, its additional gravitational influence can induce brief anomalies in the light curve, thereby revealing its presence. To date, more than 200 exoplanets have been discovered using this technique (e.g., Beaulieu et al., 2006; Gaudi et al., 2008).

Direct imaging, also known as high-contrast imaging, aims to directly capture the photons emitted or reflected by exoplanets (e.g., Marois et al., 2008; Lagrange et al., 2009). While this technique offers the most direct evidence of planetary existence, it requires planets to be both sufficiently bright and widely separated from their stars, making it currently feasible only for young, massive exoplanets with long orbital periods (>1000 days). Nevertheless, direct imaging provides unique insights into planetary atmospheres and evolution, complementing indirect methods.

1.2.2 The radial velocity (RV) method

The radial velocity method measures the line-of-sight velocity of the star by observing the Doppler shift of its spectral lines. If the star has a planetary system, the time-dependent variation of its line-of-sight velocity will reflect how its motion is influenced by its planetary system gravitationally. The bulk properties of planets, especially the masses of the planets, can be characterised by modelling such variations properly.

Here, I consider a simplified case, where the planets are much less massive than their host star, and we can ignore the gravitational interaction between planets. The star-planet system can be simplified to a two-body system. The orbit of

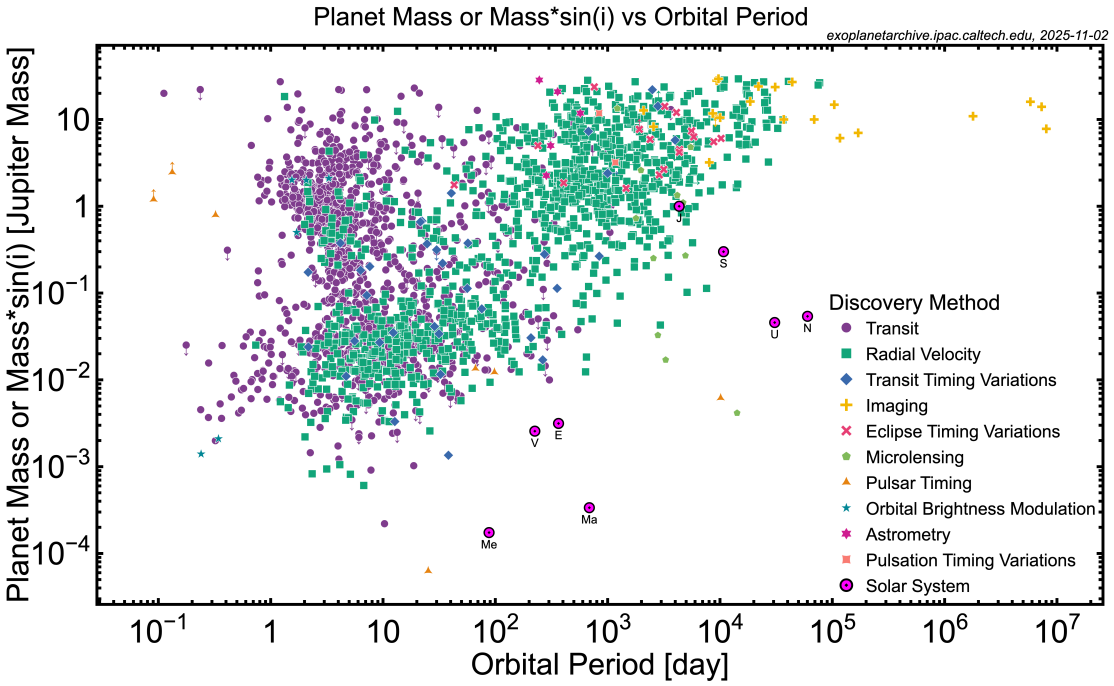


Figure 1.1: A mass-period diagram shows the planets that have been detected to date. The colour indicates different detection methods. The figure was generated using the online tool provided by the NASA Exoplanet Archive website.

the planet can then be described as

$$r(t) = \frac{a\,(1-e^2)}{1+e\cos\theta(t)}, \tag{1.1}$$

where $r(t)$ is the distance between the planet and the centre of mass, $\theta(t)$ is the true anomaly, e is the orbital eccentricity, and a is the semi-major axis of the planetary orbit. Similarly, the star's orbital position $r(t)$ and velocity $v(t)$ can be derived and projected along the line of sight. The RV of the star induced by the planet, following Murray and Dermott (1999), can then be expressed as

$$\mathrm{RV}_{\mathrm{planet}} = K\left[\cos\left(\theta+\omega\right)+e\cos\omega\right], \tag{1.2}$$

in which

$$\begin{aligned}
\theta(t) &= 2\arctan\left[\sqrt{\frac{1+e}{1-e}}\tan\left(\frac{E(t)}{2}\right)\right] \\
A(t) &= E(t) - e\sin[E(t)] \\
A(t) &= \frac{2\pi}{P}\left(t-T_{\mathrm{p}}\right),
\end{aligned} \tag{1.3}$$

where K is the semi-amplitude of the RV variation, P is the orbital period, $A(t)$ is the mean anomaly, and $E(t)$ is the eccentric anomaly. ω is the argument of periastron of the star, which is related to the true anomaly at the time of mid-transit through $\theta_{\rm tr} = \pi/2 - \omega$. $T_{\rm p}$ is the epoch of periastron. For transiting planets, it can be derived from the time of inferior conjunction (or transit midpoint) T_0. $E(t)$ is often calculated by solving the bottom two equations numerically. Figure 1.2 from Lucas Snyder [2] illustrates the different angles used to describe the orbital motion of the planet. The mass of the planet $M_{\rm p}$ can be estimated from the semi-amplitude K through

$$K = \left(\frac{2\pi \mathrm{G}}{P}\right)^{1/3} \frac{M_{\rm p} \sin i}{(M_{\rm p} + M_\star)^{2/3}} \frac{1}{(1-e^2)^{1/2}}, \tag{1.4}$$

where $M_\star$ is the mass of the star and i is the inclination of the planet's orbital plane. Based on this relation, we see that the RV method yields $M_p \sin i$, not the true planetary mass M_p. Resolving the $\sin i$ degeneracy typically requires complementary techniques, such as the transit method. Additionally, an accurate estimate of $M_\star$ is essential in order to infer M_p.

To summarize, the planet-induced RV variation, $\mathrm{RV}_{\rm planet}$, can be modeled parametrically as

$$\mathrm{RV}_{\rm planet}(t) = f(\{T_{\rm p}, P, e, \omega, K\}), \tag{1.5}$$

where $T_{\rm p}$ is the epoch of periastron, P is the orbital period, e is the eccentricity, ω is the argument of periastron, and K is the RV semi-amplitude. Note that if the planet transits, P and T_0 are typically well constrained by the transit observations, which helps to better constrain the remaining parameters when fitting the model to RV data. Figure 1.3 illustrates a toy model illustrating a typical RV signal and how planetary properties can be inferred from such measurements.

Several instrumental and astrophysical factors can affect the accuracy of RV planet detections.

[2] http://web.archive.org/web/20120210201315/http://lasunncty.110mb.com/

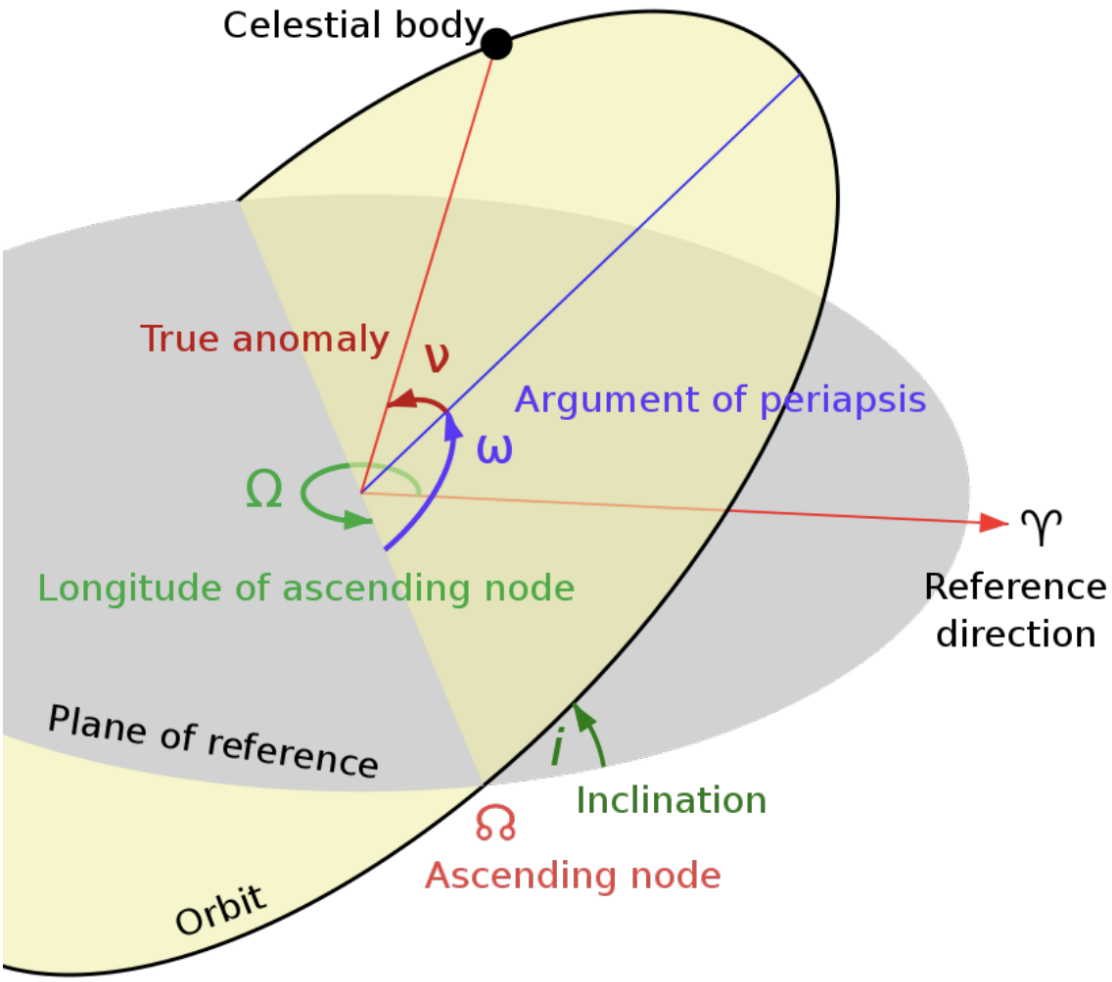


Figure 1.2: The figure illustrates different angles used to describe the orbital motion of the planet. Figure by Lucas Snyder.

Regarding instrumentation, though RV measurements do not always need large telescopes, especially when the target is relatively bright, they require extreme precision and stability of the spectrograph. Considering the working wavelength (e.g., optical) and the resolving power (e.g., R $\sim$ 100,000) of most RV spectrographs, measuring any RV signals below 1 km/s would require a precision within 1 pixel, which is around 10 μm on the CCD (Wright, 2018). To achieve such a super-resolution, i.e., a precision below 1 pixel, a cross-correlation technique (e.g., Mayor, 1985; Mayor et al., 2003), which cross-correlates the stellar spectrum with a binary mask to produce a cross-correlation function (CCF), is widely implemented. The CCF represents the averaged line profile. An accurate RV can be measured by

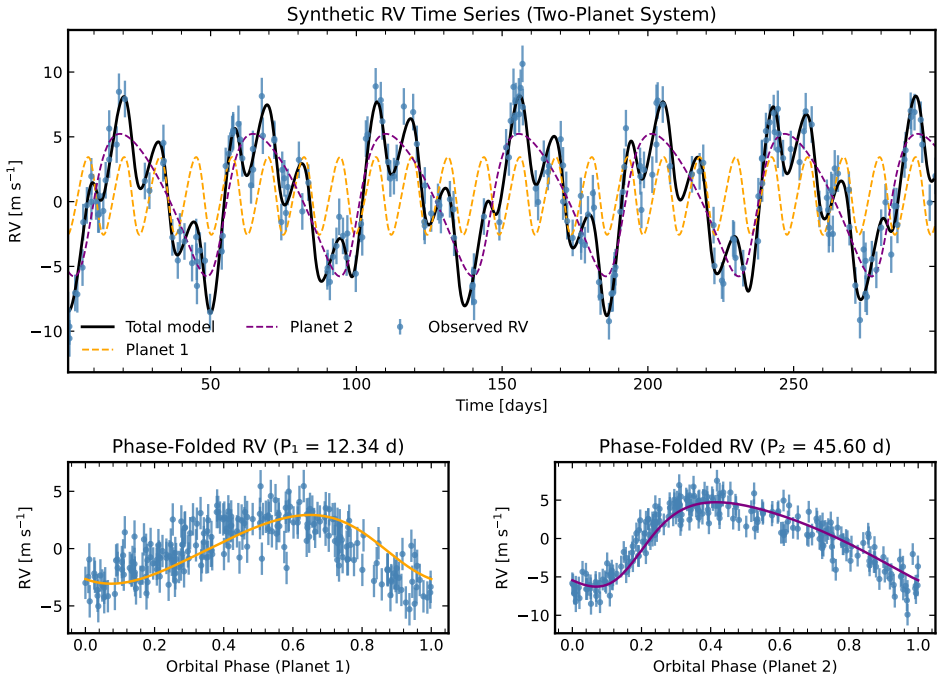


Figure 1.3: Illustration of a typical RV time series and phase-folded signals generated from a two-planet toy model. The upper panel shows simulated RV measurements (blue points with uncertainties) together with the combined model (black line) and individual planetary contributions (orange and purple dashed lines). The lower panels show the data phase-folded on the orbital periods of each planet after subtracting the contribution of the other.

fitting a parametric model, e.g., a Gaussian function, to the CCF. Moreover, any instrument-induced RV shift needs to be avoided or calibrated carefully. This can be done by, e.g., placing the whole spectrograph in a vacuum chamber as well as stabilising the temperature to mK level within a night, and conducting simultaneous calibration from a spectral reference in real time (e.g., Mayor et al., 2003; Pepe et al., 2010; Pepe et al., 2021a).

Regarding astrophysical factors, one of the major influences on RV measurements is stellar activity. Figure 1.5, generated from the `exoplanet.eu` website[3], shows the RV semi-amplitudes of detected exoplanets over time. While RV instrumental precision has steadily improved (indicated by the blue line), the detection limit for planetary signals (orange line) appears to have plateaued in recent years. This

[3] http://exoplanet.eu/

stagnation is not solely due to instrumental limitations; rather, it also reflects the presence of astrophysical noise—primarily arising from stellar activity. I will discuss stellar activity in greater detail in the following part of the thesis.

Additionally, there is the challenge of identifying an unknown number of planets, which correspond to periodic signals embedded in irregularly sampled data. When the number of periodic signals is not known a priori, the model must either be capable of transitioning between different numbers of planets, for example, using trans-dimensional methods such as those in Faria et al. (2018), or one must set up models with varying numbers of planets and perform Bayesian model comparison to infer the most likely configuration. Figure 1.4, adapted from Figure 1 in Faria et al. (2018), illustrates how the trans-dimensional approach operates. The `kima` framework is capable of treating the number of planets, N_p, as a free parameter, sampling directly from its posterior distribution using a trans-dimensional sampling scheme.

Irregular sampling further complicates the analysis. Although in principle we aim for regular observational cadence, real-world constraints such as target visibility, weather conditions, and scheduling pressure at major observatories often lead to non-uniform sampling. As a result, traditional frequency-domain techniques like the Fast Fourier Transform (FFT) become unsuitable due to issues such as spectral leakage and aliasing. These effects can introduce spurious signals or obscure true periodicities. To address this, methods such as the Lomb-Scargle periodogram (Lomb, 1976; Scargle, 1982), which is specifically designed for unevenly spaced data, are commonly used. Alternatively, modern time-domain approaches offer a flexible and robust framework for modelling RV data under these realistic conditions, e.g.,(Barragán et al., 2022).

1.2.3 Rossiter–McLaughlin effect

When a planet transits its host star, it blocks part of the rotating stellar surface, removing light that is either blue-shifted or red-shifted depending on its position along the transit chord. This asymmetric obscuration distorts the stellar absorption lines, producing a time-variable anomaly in the measured radial velocity (RV)

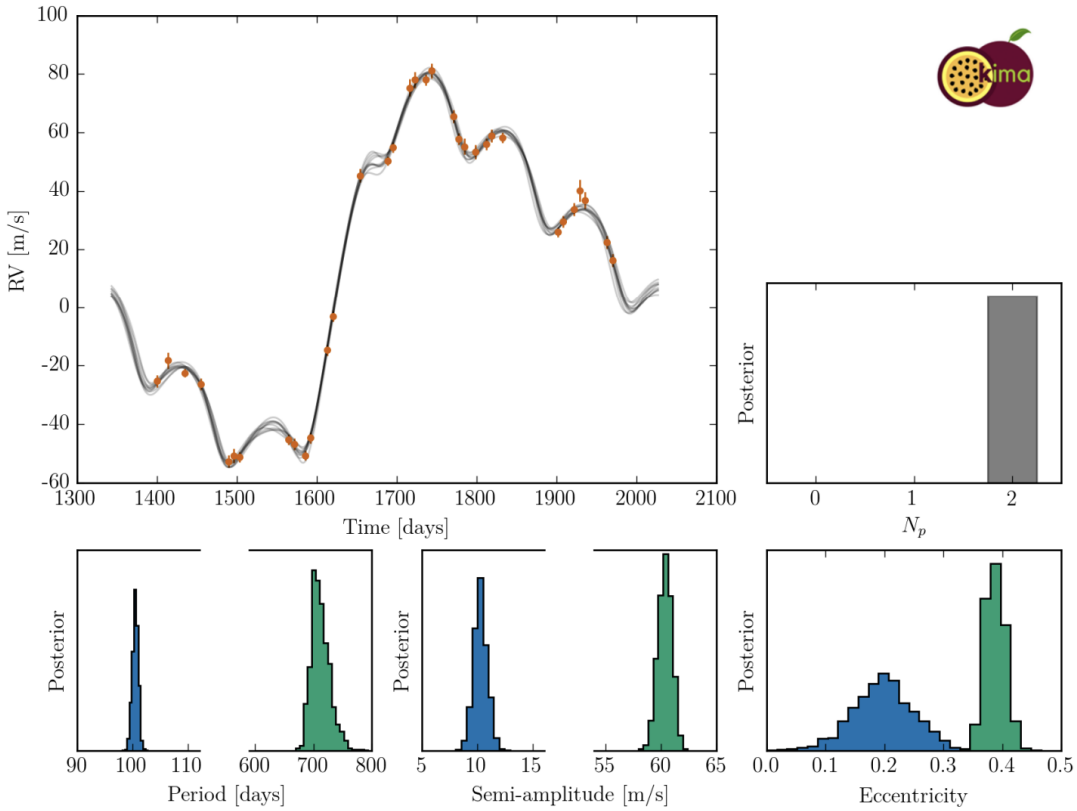


Figure 1.4: Example of a typical `kima` analysis. The top panel presents a simulated RV dataset with two injected planetary signals. The black curves correspond to ten random samples drawn from the posterior distribution. In the upper-right panel, the posterior distribution for the number of planets, N_p, strongly supports the presence of two planets (this parameter was assigned a uniform prior between 0 and 2). The lower panels display the posterior distributions for the orbital periods, semi-amplitudes, and eccentricities associated with the two detected signals. Figure and captions adapted from Figure 1 in Faria et al. (2018).

curve—an effect known as the Rossiter–McLaughlin (RM) effect (Queloz et al., 2000; Winn et al., 2005). The RM effect allows for the measurement of the sky-projected spin–orbit angle, λ, between the planet's orbital axis and the stellar rotational axis.

Since its first detection by Queloz et al. (2000), the RM effect has become a key technique for characterising exoplanet orbital architectures. Observations have uncovered a wide range of spin–orbit alignments. While many systems show alignment between the planetary orbital plane and the stellar equator, a substantial fraction, particularly among hot Jupiters, exhibit significant misalignments, including retrograde orbits (e.g., Winn et al., 2009; Collier Cameron et al., 2010b;

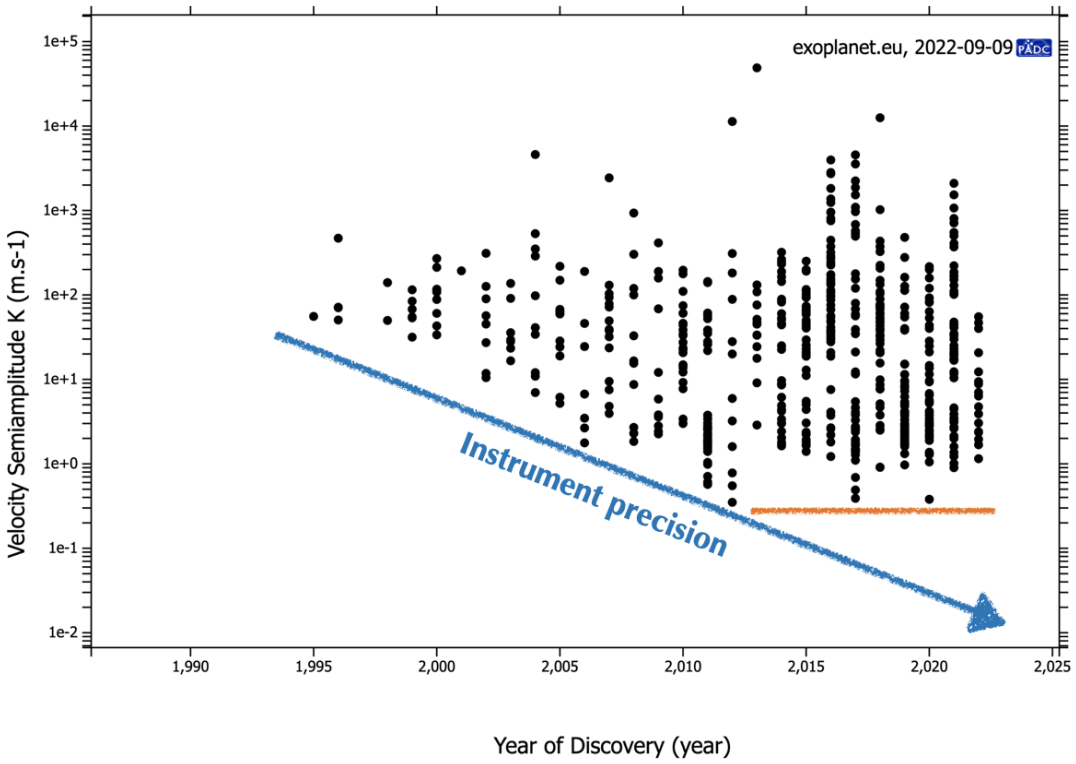


Figure 1.5: The semi-amplitude of the RV signals of the planets versus the years of detection. The blue and orange lines indicate the increasing instrument precision and the bottleneck of the detection limit, respectively. The figure is generated from exoplanet.eu website.

Esposito et al., 2014). These findings have challenged traditional models of planet formation and migration, suggesting the involvement of dynamical processes such as planet–planet scattering or Kozai–Lidov oscillations.

Under simplified assumptions, the RM-induced RV anomaly can be modelled analytically. One common approximation expresses the anomaly as:

$$\Delta V_{\mathrm{RM}}(t) \approx -\left(\frac{R_p^2}{R_\star^2}\right) v \sin i_\star \left[\frac{x(t)}{R_\star}\right], \tag{1.6}$$

where R_p and $R_\star$ are the planetary and stellar radii, $v \sin i_\star$ is the projected stellar rotational velocity, and $x(t)$ is the projected position of the planet across the stellar disk. More detailed models incorporate factors such as stellar limb darkening, and a direct modelling approach applied to the CCFs, also known as Doppler tomography, will be discussed later in this thesis.

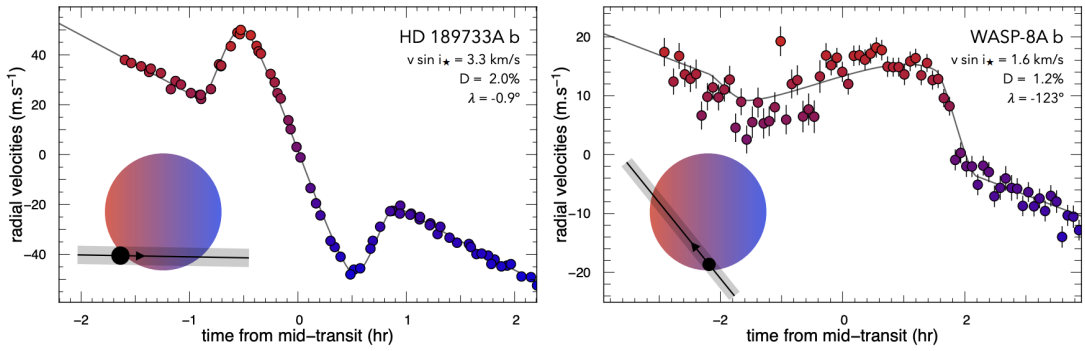


Figure 1.6: Examples of the RM effect observed in two transiting hot Jupiter systems, as measured with the HARPS spectrograph. *Left:* HD 189733Ab. *Right:* WASP-8Ab. Data points are colour-coded by Doppler content. The underlying RV trend reflects the stellar reflex motion induced by the planet. Figure from Triaud (2018).

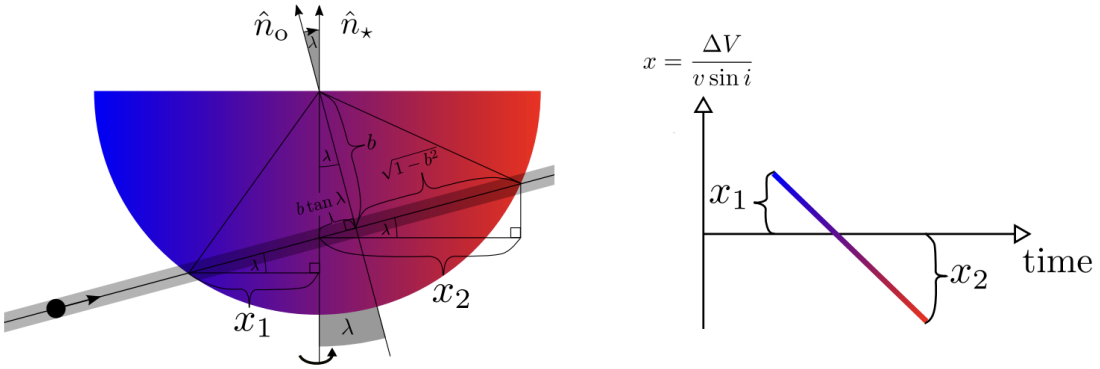


Figure 1.7: Schematic geometry of the RM effect. The left panel illustrates a transiting planet with orbital inclination $i_o < 90°$ moving across the stellar disk from left to right. With a coordinate system defined such that the x-axis aligns with the stellar equator, and assuming solid-body rotation, the radial velocity at the sub-planet point is approximately $(v \sin i_\star)\, x$. Figure from Albrecht et al. (2022a).

Importantly, while the RM effect yields the projected spin–orbit angle λ, the true three-dimensional obliquity can be inferred if the stellar inclination is known (e.g., Fabrycky and Winn, 2009a; Winn et al., 2009), for instance, by combining measurements of the stellar rotation period ($P_{\rm rot}$) and projected rotational velocity ($v \sin i_\star$).

RM observations remain a powerful diagnostic of planetary system architectures, offering crucial insights into their angular momentum evolution and migration histories. These topics are discussed in detail in Chapter 4.

1.3 Stellar activity

Beyond the limitations of instrumental precision, stellar activity remains the primary obstacle to detecting low-mass exoplanets via the RV method, especially those similar to Earth. In our own Solar System, the Earth induces an RV signal with an amplitude of less than 0.1 $\mathrm{m\,s^{-1}}$, whereas solar activity can produce RV variations of several $\mathrm{m\,s^{-1}}$ (Haywood et al., 2016). As shown in the bottom panel of Figure 1.8, HARPS-N RV observations of the Sun, treated as a distant unresolved star, reveal RV fluctuations (grey points) dominated by stellar variability, with the expected RV signal of an Earth analogue overlaid in green for comparison (Collier Cameron et al., 2019).

This stark contrast illustrates that the RV signature of a small, temperate planet like Earth is often buried within the noise induced by stellar activity. Detecting such planets therefore requires precise modelling and mitigation of stellar signals. The top panels of Figure 1.8 also summarise the key physical processes contributing to stellar RV variability, colour-coded by their relative importance to extreme-precision RV (EPRV) efforts.

1.3.1 Types of stellar activity

Different types of stellar activity can induce variations in RV time-series at divergent timescales. Most of them are the results of the magnetic field of the star interplay with the convection. Here I discuss a few of them which affect RV planet detection most severely.

Pressure-mode oscillation.

On short timescales, surface granulation and magnetic activity can excite stellar oscillations at characteristic frequencies, often referred to as acoustic or pressure modes (p-modes) (e.g., Kjeldsen and Bedding, 1995; Arentoft et al., 2008). In Sun-like stars, p-mode oscillations dominate, typically with periods of several minutes and RV amplitudes ranging from 0.1 to 1 $\mathrm{m\,s^{-1}}$. For the Sun, these oscillations have a period of approximately 5 minutes and induce RV variations of around 0.5 $\mathrm{m\,s^{-1}}$

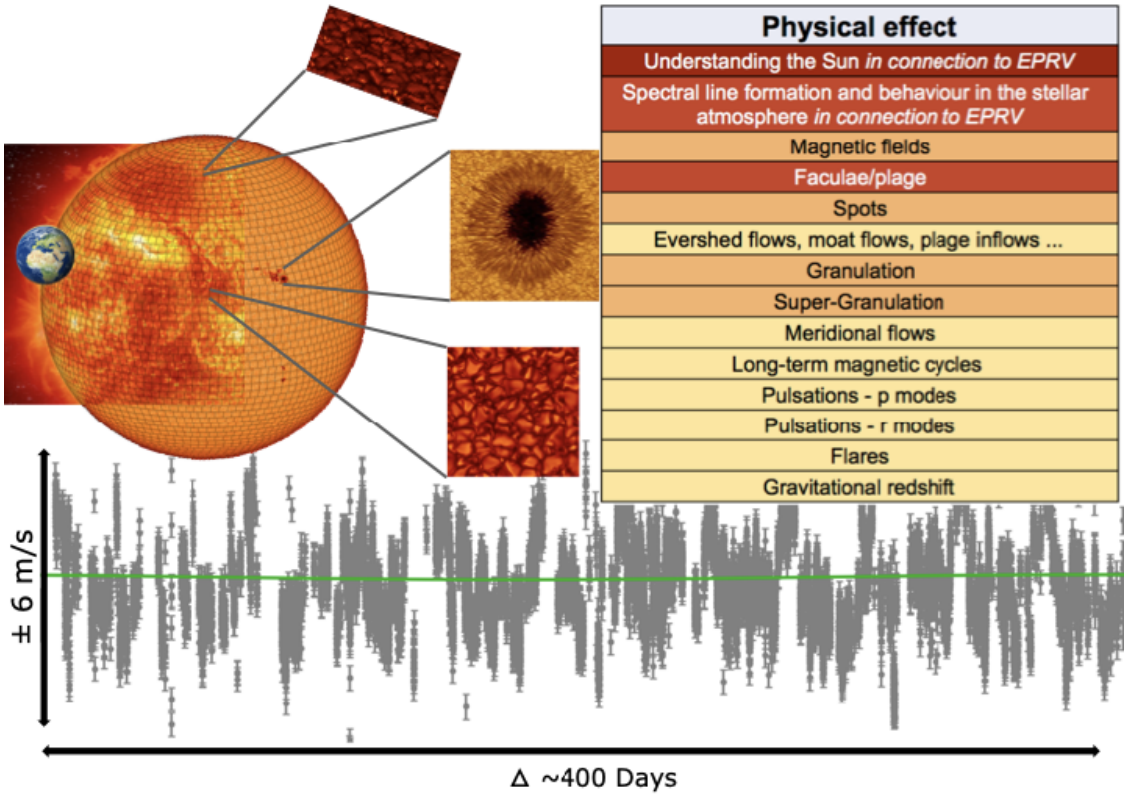


Figure 1.8: Top left: Conceptual illustration highlighting the challenge that stellar variability poses for extreme precision radial velocity (EPRV) measurements. Top right: Overview of physical processes contributing to stellar RV variability, with colour coding indicating their relative impact on EPRV—from red (high priority) to yellow (long-term significance). Bottom panel: HARPS-N solar RV measurements of the Sun observed as a distant star (grey points; Collier Cameron et al. 2019), with the expected RV signal from an Earth analogue shown in green for comparison. Figure from Crass et al. (2021a).

(Strassmeier et al., 2018; Cegla, 2019). In the context of exoplanet searches, Chaplin et al. (2019) have shown that such oscillatory signals can be effectively mitigated by carefully tuning the exposure time of RV observations.

Granulation.

The surface of FGKM-type stars is composed of convective cells where hot material rises and cooler material descends, a phenomenon known as granulation. In each granule, rising hot plasma causes blue-shifted spectral lines, while sinking cooler plasma leads to redshifts. Because the hotter, ascending regions are brighter than the cooler, descending ones, the integrated stellar spectrum exhibits an asymmetric

line profile, resulting in a net blueshift when measuring line centroids. As these granules evolve stochastically across the stellar surface, they induce RV variations at the level of several $\mathrm{m\,s^{-1}}$, with characteristic timescales ranging from a few minutes to several hours (e.g., Dumusque et al., 2011a; Meunier et al., 2015a; Meunier and Lagrange, 2019a; Cegla et al., 2013; Cegla et al., 2018; Cegla et al., 2019).

Active regions.

On moderate timescales of several days, effects caused by active regions on the stellar surface become significant. These regions are shaped by magnetic fields, which can interfere with the convection on the stellar surface in two ways. First, when the magnetic field suppresses the local heat transport, it results in dark regions known as spots. Second, if the magnetic field is not strong enough to entirely suppress the heat transport, it alters local opacity, which typically gives rise to brighter regions, as seen on the Sun, commonly referred to as faculae or plages. It should be noted, however, that faculae are not always brighter than their surroundings; their apparent brightness depends on the opacity structure and the three-dimensional appearance of the flux tubes. The enhanced magnetic fields in plages reduce local convection, and hence suppress convective blueshift.

These active regions can induce RV variations through two primary mechanisms: the photometric effect, marked by a localised flux alteration, and the inhibition of the convective blueshift effect, characterised by a reduced blueshift (or a net redshift) on the local line profile. Therefore, disk-integrated line profiles exhibit distortion due to these localised changes, manifesting as shifts in the RV. As these active regions rotate with the surface of the star, the induced RV signal is modulated by such rotation, resulting in a signal with quasi-periodic variations. Meunier et al. (2010b) found that, for the Sun, the amplitude of such variation in the RV is around 0.4 to 1.4 m/s and is dominated by the convective blueshift effect. The challenge lies in the fact that the rotation periods of Sun-like stars can be of the same order of magnitude as the orbital period of exoplanets. This similarity led to a few early claims being subsequently refuted (e.g., Queloz et al., 2001; Rajpaul et al., 2016).

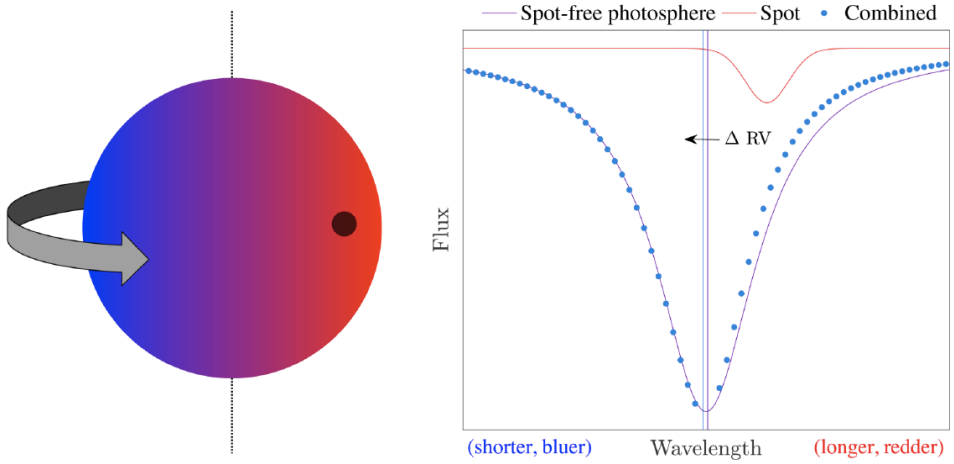


Figure 1.9: A figure from Rajpaul (2017) demonstrates how spots can induce variations in the RV.

Magnetic activity cycles.

On longer timescales of several years, the influence of magnetic activity cycles becomes prominent. Solar-type stars exhibit cycles of magnetic activity spanning several years, which can potentially affect all the above activity signals. Such cycles have been linked to long-term RV fluctuations over the years (Dumusque et al., 2011b). This means the RV signals attributed to these cycles can be easily confused with the RV signals from long-period planets. Meunier et al. (2019) demonstrated that such effects can be substantially reduced by decorrelating the RV from chromospheric emissions.

In addition to the long-term baseline variations, magnetic activity cycles can alter the characteristic period of signals induced by active regions over time. This is because the distribution of active regions across latitudes changes throughout the magnetic activity cycle (e.g., the butterfly diagram of the Sun) (e.g., Foing, 1988). If the surface rotation rate varies from equator to pole (differential rotation), the rotation period inferred from the modulation of surface heterogeneities (whether in photometry or spectroscopic indicators) will vary along the magnetic cycle. For a Sun-like differential rotation and butterfly pattern, the active regions move from higher to lower latitudes as the cycle progresses, and their rotation period decreases.

Although various models have been proposed to explain the origin of stellar magnetic activity cycles, the underlying physical mechanisms remain incompletely understood and are the subject of ongoing research.

1.3.2 Methods for mitigating stellar activity

Most of the methods proposed so far can be grouped into three broad categories: a) time-domain methods, which model activity signals in RV time-series and, in some cases, a small number of additional activity indicators (e.g. Boisse et al., 2011; Haywood et al., 2014; Rajpaul et al., 2015; Barragán et al., 2022; Delisle et al., 2022), b) velocity(wavelength)-domain methods, which typically work by separating pure Doppler shifts from line-shape variations in the Cross-Correlation Function (CCF) or full spectrum (e.g. Collier Cameron et al., 2021; Zhao et al., 2022; de Beurs et al., 2022; Liang et al., 2024; Zhao et al., 2024; Klein et al., 2024), c) line-by-line methods, which exploit the fact that different spectral lines respond to activity in a different manner, while Doppler shifts affect all lines in the same way (e.g. Dumusque, 2018; Cretignier et al., 2020a; Cretignier et al., 2022; Shahaf and Zackay, 2023; Lienhard et al., 2023). Each of these approaches has shown promising results, but also suffers from limitations.

Time-domain methods require good sampling of the stellar rotation, active region evolution, and planetary orbit timescales, and do not fully utilise the extensive information contained in the line-shape variations (and *a fortiori* how these variations vary from line to line). Velocity-domain and Line-By-Line (LBL) methods, on the other hand, typically do not account for the time-dependence of the activity signals in general. Later in Chapter 3 of this thesis, I introduce a hybrid method, which models both the time- and the velocity-dependence of activity signals explicitly, in time-series of line-profiles (or CCFs).

To place both existing methods and the new approach that I present later in the thesis, it is helpful to recall the standard workflow for stellar RV measurements for ultra-stable spectrographs, which is illustrated in Figure 1.10. The workflow

begins with a series of reduced, wavelength-calibrated, continuum-normalised high-dispersion spectra. In the most commonly used pipelines, such as the Data Reduction Software (DRS) used for the HARPS, HARPS-N, and ESPRESSO instruments (Pepe et al., 2021b), each spectrum is cross-correlated with a template (usually, a digital binary mask constructed from a line list, Baranne et al. 1996; Pepe et al. 2002), resulting in a CCF for each epoch. A simple function (usually a Gaussian) is then fit to the CCF, to derive the RV, together with the amplitude (contrast) and Full-Width at Half Maximum (FWHM) of the CCF. Additionally, a measurement of the CCF asymmetry, such as the Bisector Inverse Slope (BIS) is usually extracted at the same time. Furthermore, activity indicators such as the Mount Wilson S-index (or $\log R'_{\rm HK}$ computed from S-index) (Noyes et al., 1984a), which quantify the amount of chromospheric emission in the cores of the Ca II H & K lines, are frequently extracted from the spectra alongside the RVs and line-shape parameters.

Alternatives to this workflow exist, particularly the LBL approach mentioned above, which bypasses the cross-correlation step and instead involves measuring separate RVs for thousands of individual spectral lines (the line lists used in LBL RVs are closely related to those used to construct the cross-correlation templates), or the template-free RV extraction method of Rajpaul et al. (2020), which works by pairwise comparison of the spectra from different epochs, within individual orders or segments thereof. The RVs for individual lines or segments are of course much less precise than the global RVs, but they can then be combined carefully to minimise sensitivity to instrumental effects and stellar signals.

This workflow, in all its various incarnations, can be seen as a dimensionality reduction procedure: each spectrum (consisting of $\sim 10^5$ pixels) is reduced to a set of a few 10^3 LBL or segment RVs, or to a CCF (consisting of a few tens to hundreds of velocity bins.), and then finally to a handful of numbers (one RV, plus a few line-shape and activity indicators). Stellar activity can be modelled and mitigated at any point in this process. Doing so at an earlier stage, when the dimensionality of the signal is larger, can be advantageous, as more information is retained, and the degeneracy between instrumental and stellar signals and pure Doppler shifts induced

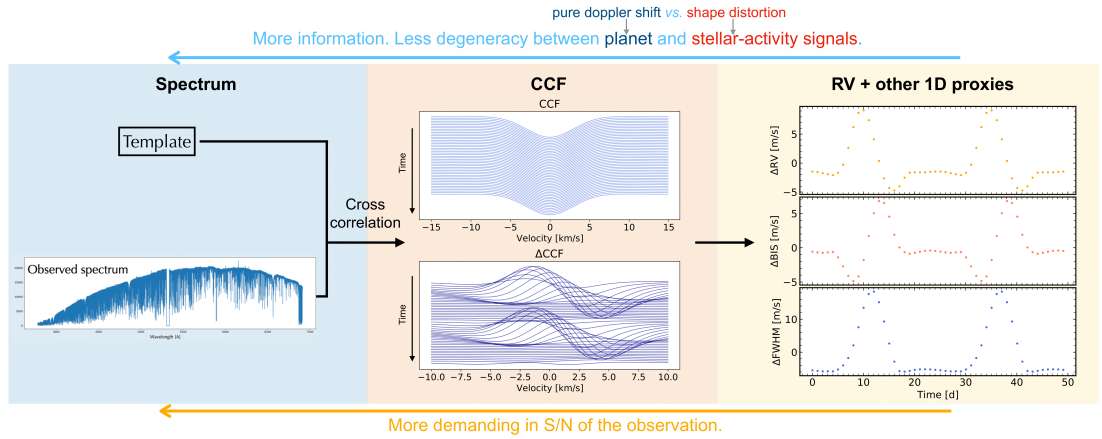


Figure 1.10: A typical workflow for RV extraction.

by companions is lower. However, this also requires that every spectrum has an excellent Signal-to-Noise Ratio (SNR), which is very expensive in terms of telescope time. Furthermore, these methods tend not to take the time-dependence of activity signals into account. On the other hand, the time-domain methods, which start from the lowest-dimension, end-result of the workflow, benefit from much more precise inputs (and hence have lower SNR requirements on the input spectra) and do use the time-domain information, but at the expense of discarding information about how activity signals affect the line shape and vary from line to line.

One subset of the existing approaches to model activity signals in time-series of line profiles does explicitly model their time-dependence: Doppler imaging (e.g., Donati et al., 2014; Petit et al., 2015; Yu et al., 2017; Klein et al., 2021; Klein et al., 2022) seeks to reconstruct a map of the inhomogeneities on the stellar surface that reproduces the line-profile variations. Limitations of the approach include that it does not explicitly model the effects of convective blueshift suppression in active regions and does not account for the potential evolution of active regions over time. A recently developed `SpotCCF` framework (Di Maio et al., 2024) uses similar principles as Doppler imaging but restricts the number of spots in the reconstructed map of the stellar surface. This approach does allow flexibility for the spots to evolve, though it does not specifically model the effects of convective blueshift suppression as well. The new method presented in Chapter 3 addresses these limitations.

Efforts have been made to systematically compare different stellar activity mitigation techniques, such as those presented in the ESSP framework by Zhao et al. (2024). For a broader methodological perspective, Hara and Ford (2023) provides a comprehensive review of statistical techniques for exoplanet detection.

1.4 Statistical methods for modelling stellar activity

Stellar activity introduces complex, time-correlated variability in RV and other spectroscopic time series. Due to its stochastic nature and the diversity of underlying physical processes, activity-induced signals are difficult to capture using deterministic models alone. Consequently, recent studies have increasingly adopted stochastic time-series methods to model this variability. Among these, Gaussian Processes (GPs) have become a widely used and versatile tool. They offer a non-parametric framework for modelling correlated noise by directly parameterising the covariance structure of the data, and allow key properties of stellar activity, such as the stellar rotation period and the evolution timescale, to be inferred from observations without requiring an explicit physical model.

1.4.1 A brief introduction to Gaussian process

In the definition of GP, the joint distribution of a finite set of observations $\mathbf{y} = \{y_i\}_{i=1,\dots,N}$ follows a multivariate normal distribution:

$$p(\mathbf{y}) = \mathcal{N}(\mathbf{m}, \mathbf{K}), \tag{1.7}$$

where $\mathbf{m}$ is the mean vector and $\mathbf{K}$ is the covariance matrix, both of which can be specified through parametric models. In particular, the covariance matrix is constructed using a kernel function, which encodes assumptions about the structure of the data, such as its smoothness, periodicity, or timescale of variability. The kernel function is defined by a set of hyperparameters that govern these properties. While hyperparameters do not dictate the precise form of the data, they constrain the behaviour of the modelled process.

A commonly used choice for modelling stellar activity is the Quasi-Periodic (QP) kernel, which captures both the periodic modulation due to stellar rotation and the evolutionary timescale of active regions.

$$k_{QP}\left(t, t'\right) = A \exp\left[-\Gamma \sin^2\left(\frac{\pi\left(t - t'\right)}{P}\right) - \frac{\left(t - t'\right)^2}{2l^2}\right], \tag{1.8}$$

where A is the amplitude representing the overall scale of the variation from the mean function, P is the characteristic period of the variation, Γ is the harmonic complexity, which in simpler terms indicates how much the variation strays from a pure sine oscillation within a given period. l is the 'evolution timescale', which scales with the maximum distance of two data points that are strongly correlated. Figure 1.11 from Barragán et al. (2022) demonstrates how the modelled curves varies with different values of hyper-parameters. In this example, it compares curves with low and high harmonic complexities in the QP kernel.

With a known mean $\mathbf{m}$ and covariance matrix $\mathbf{K}$, the likelihood takes the standard form:

$$\log \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\phi}) = -\frac{N_{\text{obs}}}{2} \log(2\pi) - \frac{1}{2} \log |\boldsymbol{K}| - \frac{1}{2} \boldsymbol{r}^{\mathrm{T}} \boldsymbol{K}^{-1} \boldsymbol{r}, \tag{1.9}$$

where $\boldsymbol{\theta}$ are the parameters of the mean function and $\boldsymbol{\phi}$ are the hyperparameters of the kernel function. $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{\mu}$ contains the difference between the observations $\boldsymbol{y}$ and the mean function $\boldsymbol{\mu}$.

When applying GP regression (GPR), we aim to infer the posterior distribution over the model parameters and kernel hyperparameters given the data. The posterior can be written using Bayes' theorem:

$$p(\boldsymbol{\theta}, \boldsymbol{\phi} \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\phi})\, p(\boldsymbol{\theta}, \boldsymbol{\phi})}{p(\mathbf{y})}, \tag{1.10}$$

where $p(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\phi})$ is the likelihood, $p(\boldsymbol{\theta}, \boldsymbol{\phi})$ is the prior, and $p(\mathbf{y})$ is the marginal likelihood (also called the evidence).

A key feature of GPR is its ability to produce not just point predictions, but full predictive distributions for new input locations. Suppose we have a set of training observations $\mathbf{y}$ at input times $\mathbf{t}$, and we wish to predict values at new

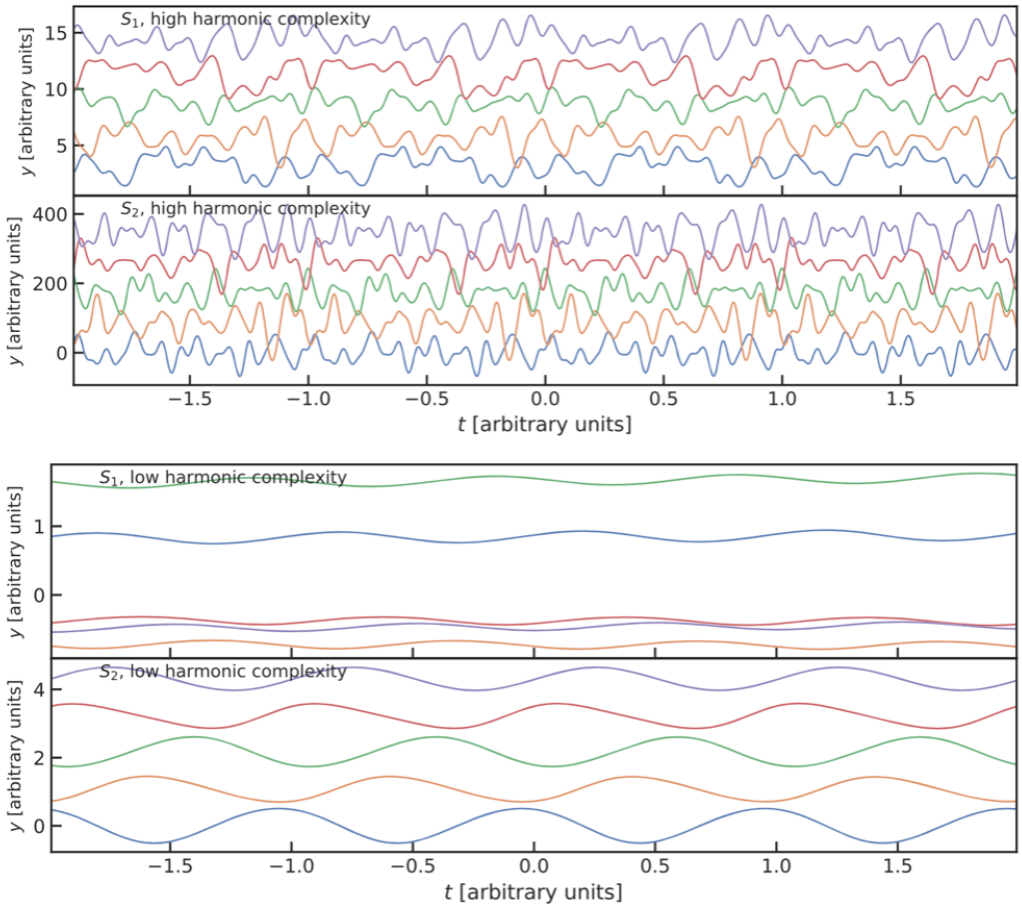


Figure 1.11: This is a figure from Barragán et al. (2022) demonstrating how the modelled curves vary with different values of hyperparameters. In this example, it compares curves with low and high harmonic complexities in the QP kernel.

times $\mathbf{t}_\star$. The goal is to determine the distribution $p(\mathbf{f}_\star \mid \mathbf{y})$, where $\mathbf{f}_\star$ represents the predicted values. Under the GP formalism, this predictive distribution is also Gaussian, and its mean and covariance are given by:

$$\mathbf{f}_\star = \mathbf{m}_\star + \mathbf{K}_\star^{\mathrm{T}}\mathbf{K}^{-1}(\mathbf{y} - \mathbf{m}), \qquad \mathbf{C}_\star = \mathbf{K}_{\star\star} - \mathbf{K}_\star^{\mathrm{T}}\mathbf{K}^{-1}\mathbf{K}_\star, \tag{1.11}$$

where $\mathbf{m}$ denotes the mean function evaluated at the training inputs $\mathbf{t}$, while $\mathbf{m}_\star$ is the mean function evaluated at the new test inputs $\mathbf{t}_\star$. The matrix $\mathbf{K}$ represents the covariance among the training points, $\mathbf{K}_\star$ is the cross-covariance between the training and test inputs, and $\mathbf{K}_{\star\star}$ is the covariance matrix evaluated at the test inputs themselves.

The elements of the kernel matrices $\mathbf{K}_\star$ and $\mathbf{K}_{\star\star}$ are computed from the kernel function,

$$[\mathbf{K}_\star]_{ij} = k(t_i, t_{\star,j}; \boldsymbol{\phi}), \qquad [\mathbf{K}_{\star\star}]_{ij} = k(t_{\star,i}, t_{\star,j}; \boldsymbol{\phi}). \tag{1.12}$$

This formulation enables both interpolation and extrapolation of time-series data, along with estimation of uncertainty.

For comprehensive references on Gaussian processes, I refer the readers to Rasmussen and Williams (2006) and Roberts et al. (2017). A great domain-specific overview of GP applied to astronomical time series is provided in the review by Aigrain and Foreman-Mackey (2023).

1.4.2 Multi-dimensional Gaussian process framework

The multi-dimensional Gaussian process (multi-GP) framework first introduced by Rajpaul et al. (2015) has, over the past decade, proven to be a powerful and effective technique for disentangling planetary signals from stellar activity in RV data. This approach forms the basis of several components of my work and is therefore introduced in detail here. The multi-GP framework can be viewed as an extension of the FF' framework introduced by Aigrain et al. (2012), which uses simple geometric arguments to describe the relationship between stellar activity signals observed in photometry and those measured in RVs. I begin by reviewing the key elements of the FF' framework before presenting the structure of the multi-GP model.

The FF' framework

In the FF' framework, the quantity F represents the relative flux variation due to active regions on the visible stellar hemisphere. F is negative when the observed flux decreases, typically due to dark spots, though it can also be modulated by bright faculae. As the star rotates, the projected area of these active regions changes, leading to time variations in $F(t)$. Additional contributions to the time dependence of $F(t)$ may arise from the intrinsic evolution of the regions (growth and decay) and differences in limb darkening, but these are generally treated as secondary effects.

The photometric contribution of active regions to the total RV signal scales as the product of $F(t)$ and $V_{\rm loc}(t)$, the line-of-sight velocity at the region's location. In addition to these photometric effects, magnetically active regions also reduce the local convective blueshift (CB), resulting in a net redshift. This inhibition of convective blueshift (ICB) is expected to scale with the projected area of the active region, $F(t)$, and the line-of-sight component of the convective upflow velocity (e.g., Gray, 2008).

The key insight from Aigrain et al. (2012) is that $V_{\rm loc}(t)$ approximately corresponds to the time derivative $F'(t)$, while the line-of-sight convective component scales with $F(t)$. As a result, the total RV variation can be approximated by two terms, one proportional to $F(t)F'(t)$ and the other to $F^2(t)$. If we donote $G(t) = F^2(t)$, then $\dot{G}(t) = 2F(t)F'(t)$. This relates to the GP model, which is described in detail in the next subsection.

Model structure

Aigrain et al. (2012) proposed modelling $F(t)$ as a GP with a quasi-periodic (QP) kernel to capture the temporal evolution of stellar activity. Building on this, Rajpaul et al. (2015) extended the formalism to jointly model RVs and spectroscopic activity indicators, such as the Mt Wilson S-index and line-shape diagnostics like the bisector inverse slope (BIS), extracted from the same set of spectra.

Following the framework of Aigrain et al. (2012), Rajpaul et al. (2015), Barragán et al. (2022), and Delisle et al. (2022), we define a set of N observable time series, $Y_{i=1,\ldots,N}(t)$, where each $Y_i(t)$ may correspond to either an RV measurement or an activity indicator. Each time series is modelled as:

$$Y_i(t) = f_i(t) + \Big[a_i G(t) + b_i \dot{G}(t)\Big] + \epsilon_i(t), \tag{1.13}$$

where $f_i(t)$ is the deterministic part of the model. For RV, it is where the mean function and planet-induced RV are incorporated. For activity indicators, it is simply the mean function. $\epsilon_i(t)$ is the measurement 'white' noise, which normally includes photon noise, calibration noise, etc. Every $Y_i(t)$ has a shared entity, $G(t)$. We interpret this as a latent GP variable, approximating the proportion of the

visible stellar disc blanketed by active regions. a and b are free parameters to be determined by the model fit.

Commonly, a GP with a quasi-periodic (QP) kernel is the modelling choice for this variable. $\dot{G}(t)$ is the first temporal derivative of $G(t)$ and remains a GP, roughly representing how the active regions evolve in time on the stellar disc. Coefficients a_i and b_i are free parameters that harmonise the interplay between the latent GP variables, G and $\dot{G}$, and various observable time-series, Y_i.

From this model, a fully parameterised covariance matrix $\mathbf{K}$ for the multi-output vector $\mathbf{y}$ of observed time series can be constructed. Bayesian inference techniques, such as Markov Chain Monte Carlo (MCMC) or nested sampling, can then be employed to infer both model parameters and GP hyperparameters. For a more detailed methodological discussion, I refer the reader to Rajpaul et al. (2015) and Barragán et al. (2022).

Modelling $G(t)$ with the traditional QP kernel (Equation 1.8) becomes computationally expensive for large datasets, as its complexity scales cubically with the number of data points. To mitigate this, we can adapt the Matern 3/2 Exponential Periodic (MEP) kernel as implemented in the `S+LEAF` package[4] (Delisle et al., 2022), which offers a close approximation to the QP kernel while scaling linearly with data size. This computational efficiency is crucial when analysing large numbers of targets with extensive time series. The MEP kernel is described as

$$k(\Delta t) = \sigma^2 \frac{k_{3/2}(\Delta t) + f k_{\text{SHO, fund.}}\,(\Delta t) + \frac{f^2}{4} k_{\text{SHO, harm.}}\,(\Delta t)}{1 + f + \frac{f^2}{4}}, \tag{1.14}$$

where

$$\begin{aligned} k_{3/2}(\Delta t) &= \exp\left(-\frac{\sqrt{3}\Delta t}{l}\right)\left(1 + \frac{\sqrt{3}\Delta t}{l}\right), \\ k_{\text{SHO, fund.}}\,(\Delta t) &= \exp\left(-\frac{\Delta t}{l}\right)\left(\cos(v\Delta t) + \frac{1}{vl}\sin(v\Delta t)\right), \\ k_{\text{SHO, harm.}}\,(\Delta t) &= \exp\left(-\frac{\Delta t}{l}\right)\left(\cos(2v\Delta t) + \frac{1}{2vl}\sin(2v\Delta t)\right), \end{aligned} \tag{1.15}$$

[4] https://gitlab.unige.ch/Jean-Baptiste.Delisle/spleaf

with $v = 2\pi/P$. Similar to QP kernel, σ^2 is the amplitude representing the overall scale of the variation from the mean function, P is the characteristic period of the variation, f is the harmonic complexity and l is the evolution timescale.

1.5 Planets around young stars

Transiting planets in nearby young stellar associations ($t < 1$ Gyr) are crucial to bridging the gap between planet formation and early evolution theories, and the mature population of exoplanets around field stars. Processes such as orbital migration, tidal circularisation, atmospheric escape, and planetary contraction—which shape the final architectures and physical properties of planetary systems—are still ongoing at these early stages. Importantly, the precise ages of young stellar groups provide critical constraints on the timescales of these mechanisms, enabling discrimination between competing theoretical models (Owen and Wu, 2017; Gupta and Schlichting, 2020).

The first ∼100 Myr represents a particularly dynamic epoch in planetary evolution. Bulk planet properties, such as density and composition, change rapidly due to cooling, contraction, and atmospheric mass loss. Measuring masses and radii during this phase helps constrain key quantities like the initial entropy and core-to-envelope mass fraction, both of which are highly uncertain (Spiegel and Burrows, 2012; Owen, 2020). When the host star is nearby and sufficiently bright, young planets can be characterised in detail via transit and RV measurements, enabling inferences about their internal structures. Transmission and emission spectroscopy further allow the study of atmospheric composition and dynamics.

Despite their importance, fewer than 50 transiting planets have been detected around stars in nearby young associations, and just ten of these have published mass measurements. This scarcity is largely due to the observational challenges posed by young stars, which are typically fast rotators with strong magnetic activity. Active regions evolving and rotating on their surfaces induce significant RV variability that can dwarf the planetary signal. Also, the high occurenece of flare also affect the . Nonetheless, recent advances in modelling stellar activity, such as through

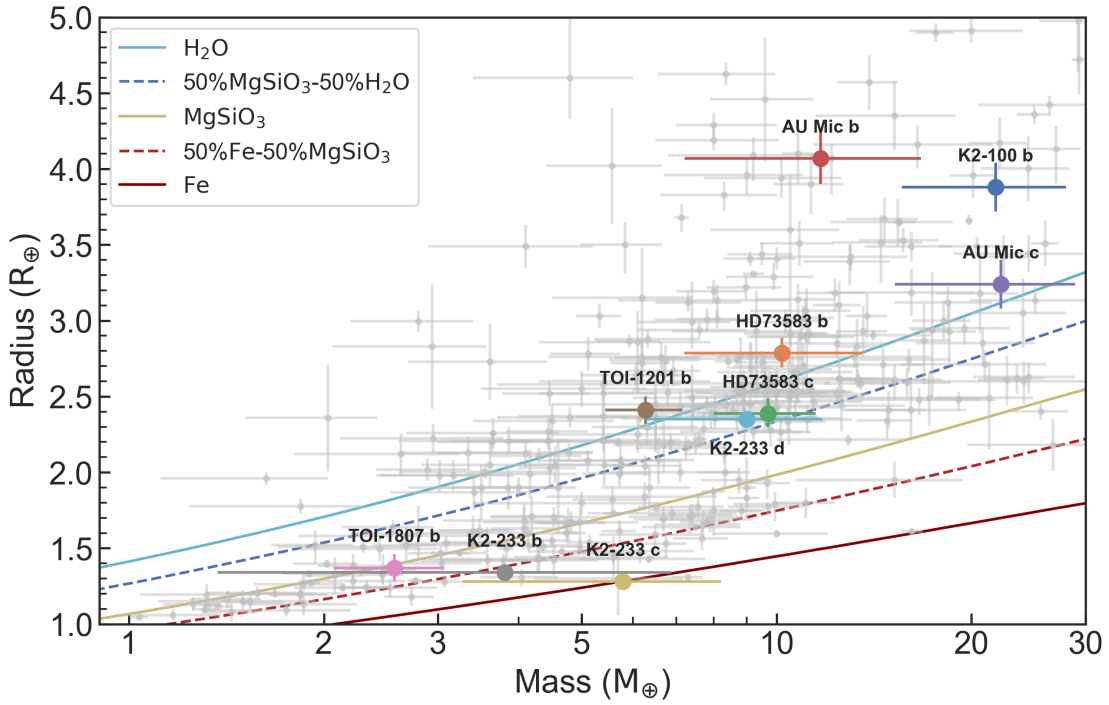


Figure 1.12: Mass–radius diagram for exoplanets. The grey points with error bars represent planets with mass and radius measurements accurate to better than 50%, based on data from the NASA Exoplanet Archive (accessed on 21 July 2025; Akeson et al. 2013). Highlighted in colour are young exoplanets (younger than 1 Gyr) with well-constrained masses and radii. The solid lines indicate two-layer core composition models from Zeng et al. (2016), while the non-solid lines denote models with additional hydrogen-rich envelopes, as labelled in the figure. The figure was generated using `fancy-massradius-plot` (Barragán et al., 2018).

multi-GP frameworks have enabled several mass measurements, including those for the AU Mic system (Rajpaul et al., 2015; Zicher et al., 2022a).

In Figure 1.12, I present a mass–radius diagram for exoplanets, highlighting a subset of young systems (<1 Gyr) with precise mass and radius measurements. For comparison, older planets with <50% uncertainties are shown in grey. Theoretical models from Zeng et al. (2016) are also included to illustrate internal compositions.

Moreover, the orbital orientation of young planets encodes their early dynamical histories. Planets formed within protoplanetary disks may acquire high obliquities through planet-planet scattering, giant impacts, or secular perturbations (Chatterjee et al., 2008; Wu and Murray, 2003), while tidal dissipation may dampen their obliquities (Bolmont and Mathis, 2016). Measuring the sky-projected obliquity

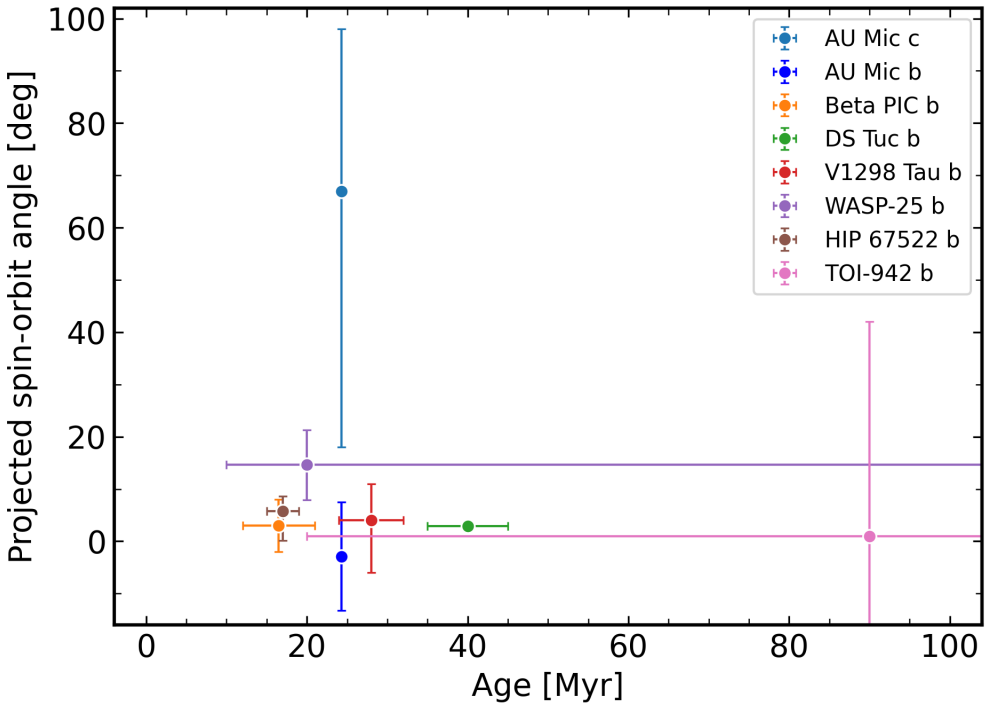


Figure 1.13: Projected spin-orbit angle of planets in systems younger than 100 Myr. The measurement of AU Mic c obtained in the work is shown in light blue. The measurements of other systems are from Albrecht et al. (2022b).

of young planets with known ages provides direct constraints on the timescales of these processes. However, only eight such measurements exist to date due to the difficulty of performing high-precision spectroscopic observations on active, rapidly rotating stars. In Figure 1.13, I show the projected spin-orbit angle versus the age of these planets, together with the measurement of AU Mic c obtained Chapter 4 of this thesis.

One such young system is AU Mic, which lies just 10 pc from the Sun and belongs to the 22 ± 3Myr-old Beta Pictoris moving group. It is one of the brightest M-dwarfs in the sky and hosts an edge-on debris disk with fast-moving substructures, as well as strong magnetic activity that makes it both a challenging and highly valuable target for planet characterisation. Two close-in transiting Neptunes, AU Mic b and c, were discovered in 2020–2021 with orbital periods of 8.5 and 18.9 days. Their masses—$11.7 \pm 5.0 M_\oplus$ and $22.2 \pm 6.7 M_\oplus$, respectively—were derived from

intensive HARPS radial velocity monitoring using a multi-GP model to account for stellar activity. Interestingly, the outer planet, AUMic c, is denser than the inner one, contrary to standard formation expectations, suggesting either divergent formation environments or the aftermath of a giant impact. With its proximity, youth, and well-characterised planets, AU Mic serves as a benchmark for studying early planet formation, atmospheric evolution, and dynamical history. A detailed investigation of this system will be presented in Chapter 4.

1.6 Thesis outline

This thesis focuses on understanding stellar activity, developing techniques to mitigate its effects, and demonstrating the scientific advancements enabled by such mitigation. The structure of the thesis is as follows.

Chapter 2 focuses on rotation and activity properties of stars. This chapter investigates stellar activity through a study of 268 well-observed targets from HARPS. A multi-GP framework was applied to model activity signals in spectroscopic indicators (S-index and BIS). The study measured rotation periods for 49 slow rotators and analysed the variation in spot-to-facular ratios across the sample. These results provided valuable insights for radial velocity (RV) planet surveys such as the Terra Hunting Experiment and PLATO ground-based follow-up, as well as refinements to stellar evolution models. The findings were published in Yu et al. (2024b).

Chapter 3 focuses on a novel physics-driven Gaussian process model for stellar activity. Building on the previous study, this chapter presents the development of CCF-GP, a physics-informed Gaussian process framework that models activity signals directly in time-series of line profiles. Unlike traditional methods, CCF-GP exploits time correlations in velocity bins, enabling superior separation of planetary and activity signals. Tests on synthetic and real HARPS-N datasets demonstrated its ability to recover planetary signals down to 0.3 m/s, even during high solar activity. The framework outperformed existing state-of-the-art approaches and is detailed in Yu et al. (2024a).

Chapter 4 focuses on investigating the spin-orbit alignment in the AU Mic System. This chapter explores the architecture of the young AU Mic system using spectroscopic and photometric data. Utilising ESPRESSO and CHEOPS observations, the first obliquity measurement for the young Neptune AU Mic c was obtained. Despite strong stellar activity, the analysis suggested that AU Mic c is misaligned relative to the inner planet AU Mic b, indicating a dynamically complex history. If confirmed, this would be the first discovery of a misaligned planet younger than 100 Myr. The findings were published in Yu et al. (2025).

Chapter 5 provides a synthesis of the key findings and discusses ongoing and future efforts to enhance stellar activity mitigation techniques and their applications in exoplanet detection. As part of ongoing work, I present two projects. The first involves a two-dimensional spot model designed to correct spot-induced distortions in CCFs during planetary transits. The second is a preliminary study of the spin-orbit alignment of planets in the HD 3167 system, conducted in collaboration with Professor Joshua Winn during my three-month visit to Princeton University as a Visiting Student Research Collaborator (VSRC).

2

Modelling stellar variability in archival HARPS data: Rotation and activity properties with multi-dimensional Gaussian Processes

Contents

2.1 Introduction 33
2.2 The HARPS Sample 34
2.2.1 Sample selection 34
2.2.2 Data reduction and pre-processing 35
2.2.3 Basic properties of the sample 37
2.3 Model for stellar activity 38
2.4 Stellar rotation 42
2.4.1 Rotation period measurement 42
2.4.2 Rossby number 46
2.4.3 Implications for spin-down of middle-aged stars 47
2.5 Faculae-to-spots ratio 51
2.6 Conclusions 55

2.1 Introduction

Since 2003, the high-resolution spectrograph HARPS has been observing thousands of stars. The rich archival, high-quality data from HARPS provides a unique

opportunity to test and apply state-of-the-art methods for activity modelling and mitigation. In this chapter, we apply the multi-dimensional Gaussian Process framework to HARPS observations of 268 well-observed targets. We discuss sample selection, data reduction and pre-processing, as well as the basic properties of the sample in section 2.2. We then discuss the implementation of the multi-dimensional GP framework in section 2.3. We discuss results regarding stellar rotation in section 2.4, and facular to spot ratio in section 2.5. We summarise our key conclusions in section 2.6.

2.2 The HARPS Sample

2.2.1 Sample selection

We utilised archived data from the HARPS, accessible via the ESO science archive [1].

We retrieved all publicly available HARPS spectra using the ESO's `astroquery` [2] (Ginsburg et al., 2019) package. For every observation and each target, we cross-matched the stellar name and position with their respective counterparts in the `simbad` database (Wenger et al., 2000). This step ensured the removal of potential misclassifications from the ESO archive and close-in binary systems from our dataset. Polarimetric data, which only constitute a minor portion of the HARPS archive, have been discarded as they were not reduced automatically by the standard pipeline. We did not implement cuts based on the signal-to-noise ratio (SNR) of the spectra, or cuts based on whether the simultaneous calibration has been applied. We anticipated that these factors would manifest themselves through the uncertainties of the pipeline products, which would be taken into account statistically by our Bayesian framework introduced later. Our preliminary target set included a total of 1438 targets with spectral types spanning from F to M. Their effective temperatures ($T_{\rm eff}$) approximately ranged between 2800 K and 7500 K.

To guarantee a sufficient observation duration for our targets, we excluded those that had fewer than 40 daily-binned observations. In pursuit of homogeneity

[1] http://archive.eso.org/scienceportal/home
[2] https://github.com/astropy/astroquery

in the determination of the stellar parameters, e.g., effective temperature $T_{\rm eff}$, surface gravity log g, we restricted our sample to stars listed in the catalogue published in Gomes da Silva et al. (2021), and used the parameters listed in that catalogue throughout this chapter.

The Gomes da Silva et al. (2021) catalogue provides a homogeneous set of stellar atmospheric parameters for 1674 FGK-type stars, derived from high-resolution HARPS spectra. Line equivalent widths were measured using the ARES code (Sousa et al., 2007; Sousa et al., 2015), and elemental abundances were determined using the MOOG spectral synthesis code (Sneden, 1973). Stellar parameters were then derived by fitting to Kurucz ATLAS9 model atmospheres (Kurucz, 1993). We note that the catalogue is limited to FGK stars; hence, M-type stars are not included in our sample.

Our final sample contains 268 targets, with spectral types spanning from F to K. This sample includes targets that satisfy both criteria: having more than 40 daily-binned observations and being listed in the Gomes da Silva et al. (2021) catalogue. Figure 2.1 shows our sample selection in an Hertzsprung-Russell (HR) diagram, with log g versus $T_{\rm eff}$ from GAIA DR2 (Gaia Collaboration et al., 2016; Gaia Collaboration et al., 2018b). The grey points show the 1438 targets on our initial list, while the selected 268 targets in our final sample are highlighted in orange.

2.2.2 Data reduction and pre-processing

The raw data underwent automatic reduction using version 3.5 and/or 3.8 of the HARPS Data Reduction Software (DRS). The DRS first processes each raw image through a series of calibration steps, including bias and dark subtraction, cosmic ray removal, and flat-field correction. This is followed by order identification and extraction of the echelle spectra, wavelength calibration, and blaze function correction. The extracted spectral orders are then merged into a continuous one-dimensional (1D) spectrum. While some methods operate directly on individual spectral orders, in our work, we use the merged 1D spectrum for cross-correlation.

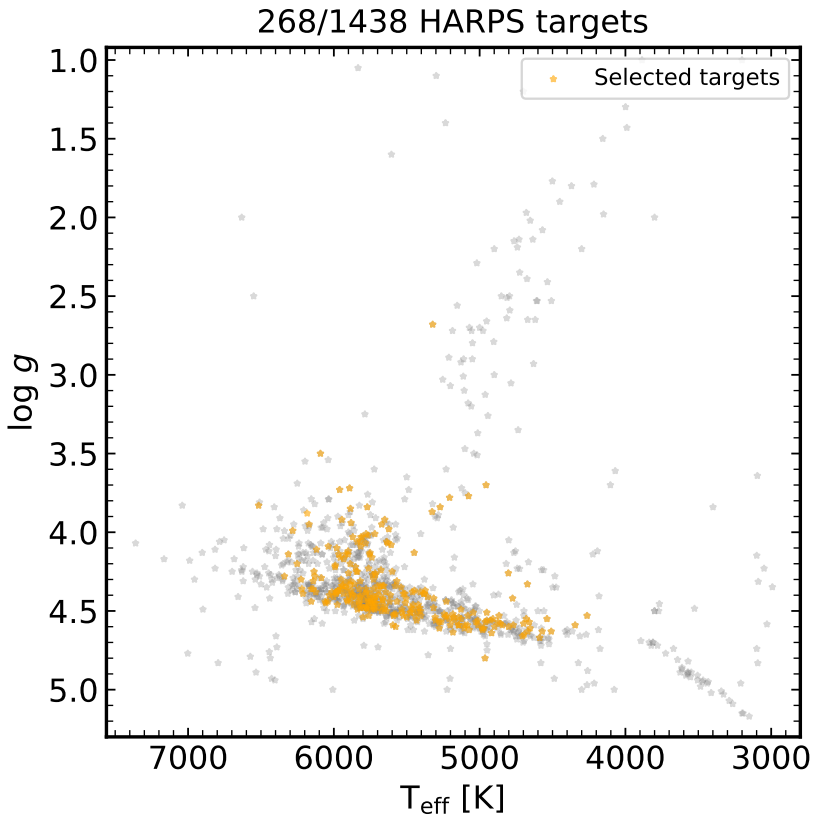


Figure 2.1: Our sample selection in an HR diagram, with log g versus $T_{\rm eff}$ from GAIA DR2 (Gaia Collaboration et al., 2016; Gaia Collaboration et al., 2018b). The grey points show the 1438 targets on our initial list. The selected 268 targets in our final sample are highlighted in orange.

For details about the DRS, we refer readers to (Dumusque et al., 2021a) and the DRS official manual for details.

For each target, cross-correlation functions (CCFs) were computed by correlating the reduced spectrum with a numerical mask corresponding to the star's spectral type. CCF proxies, namely RV and bisector inverse slopes (BIS), were then measured from the CCFs within the DRS. These products are included in the data provided by the ESO Science Archive.

Additionally, we computed a chromospheric emission metric, S-index, from the Ca II H and K lines (resp. 3968.47 and 3933.66 Å) in the S1D spectra, employing the `ACTIN`[3] (Gomes da Silva et al., 2018) package.

[3] https://github.com/gomesdasilva/ACTIN

For each time-series, we calculated the median absolute deviation (MAD) metric, which is considered a more reliable measure of statistical dispersion than the often-employed standard deviation, especially in the presence of outliers. Data points exceeding three times the MAD value were identified as outliers and subsequently excluded. Any time-series extending beyond BJD=2457161.5 was partitioned into two sections for the purpose of outlier removal, given the expectation of two separate baselines resulting from the HARPS fibre upgrade. Finally, we binned each time-series to a maximum of one data point per day.

2.2.3 Basic properties of the sample

We showcase the basic properties of the 268 targets in our sample using HR diagrams in Figure 2.2. The colours of the points in the left, middle, and right panels represent the number of data points in the pre-processed time-series, the mean SNR per pixel (in échelle order 50) of the acquired spectra, and the mean $\log R'_{\rm HK}$ of the star, using values from the catalogue provided by Gomes da Silva et al. (2021), respectively. We can see that the numbers of data points in our sample vary, ranging from 40 to over 400, though the majority fall between 40 and 100. The mean SNR spans from 50 to above 300. The mean $\log R'_{\rm HK}$ for the majority of the targets are below -4.4, suggesting that they are predominantly inactive stars. This aligns with expectations, given that the primary goal of HARPS is exoplanet detection, leading the survey to prioritise low-activity stars.

In Figure 2.3, we show the root mean square (RMS) of the time series for both the S-index and BIS, normalised by their respective maximum values (set to 1), plotted against the mean $\log R'_{\rm HK}$. The RMS values serve as direct metrics of variability in the activity indicators, and as such, they are anticipated to correlate more intimately with the challenges of activity mitigation. We find, however, no clear correlation between the overall RMS values and the mean $\log R'_{\rm HK}$, even though there appears to be a tentative positive correlation between the RMS's upper envelope and the mean $\log R'_{\rm HK}$. For instance, at a mean $\log R'_{\rm HK}$ value of approximately -4.9, the corresponding RMS for either S-index or BIS can fluctuate by an order of magnitude.

Therefore, we deduce that relying solely on the mean $\log R'_{\rm HK}$ is inadequate as an activity metric within the context of activity mitigation. We recommend that future survey target selections consider additional metrics alongside mean $\log R'_{\rm HK}$.

2.3 Model for stellar activity

Different processes of stellar activity, e.g., p-mode oscillations, granulation, active regions, and magnetic activity cycles, can induce variations in the time-series of RV and spectroscopic activity indicators at various timescales. In this section, we focus on modelling the variations induced by the active regions — notably the effects of faculae/plages and spots. This is primarily because such activity signals is modulated by the rotation of the star, and the associated timescales (i.e., tens of days) are closest to the orbital periods of the planets.

We employed the multi-dimensional GP framework introduced in Chapter 1 to model the activity-induced signals in the activity indicators. Although a joint model of the activity indicators and RVs is a natural next step, it remains non-trivial due to the presence of residual systematics in the RV time series and the added complexity of unknown planetary signals. Unlike activity indicators, RVs require accounting for one or more planetary components whose number and properties are not known a priori. This significantly enlarges the parameter space and increases the difficulty of achieving a robust and interpretable fit.

As we initially focused on the activity indicators in the present work, the exact form of the framework we used can be outlined as follows

$$\begin{aligned} \mathrm{BIS}(t) &= f_{\rm BIS}(t) + a_{\rm BIS} G(t) + b_{\rm BIS} \dot{G}(t) + \epsilon_{\rm BIS}(t), \\ S\mathrm{-index}(t) &= f_{S-\rm index}(t) + a_{S-\rm index} G(t) + \epsilon_{S-\rm index}(t), \end{aligned} \tag{2.1}$$

where

$$\begin{aligned} f_{\rm BIS}(t) &= \overline{\mathrm{BIS}(t)}, \\ f_{S-\rm index}(t) &= \overline{S\mathrm{-index}(t)}, \end{aligned} \tag{2.2}$$

and $\epsilon(t)$ are white noise models containing measurement noise and jitter terms. BIS is proved to be interlinked with both $G(t)$ and its derivative $\dot{G}(t)$, while S-index is exclusively related to $G(t)$ (e.g., Isaacson and Fischer, 2010; Dumusque et al.,

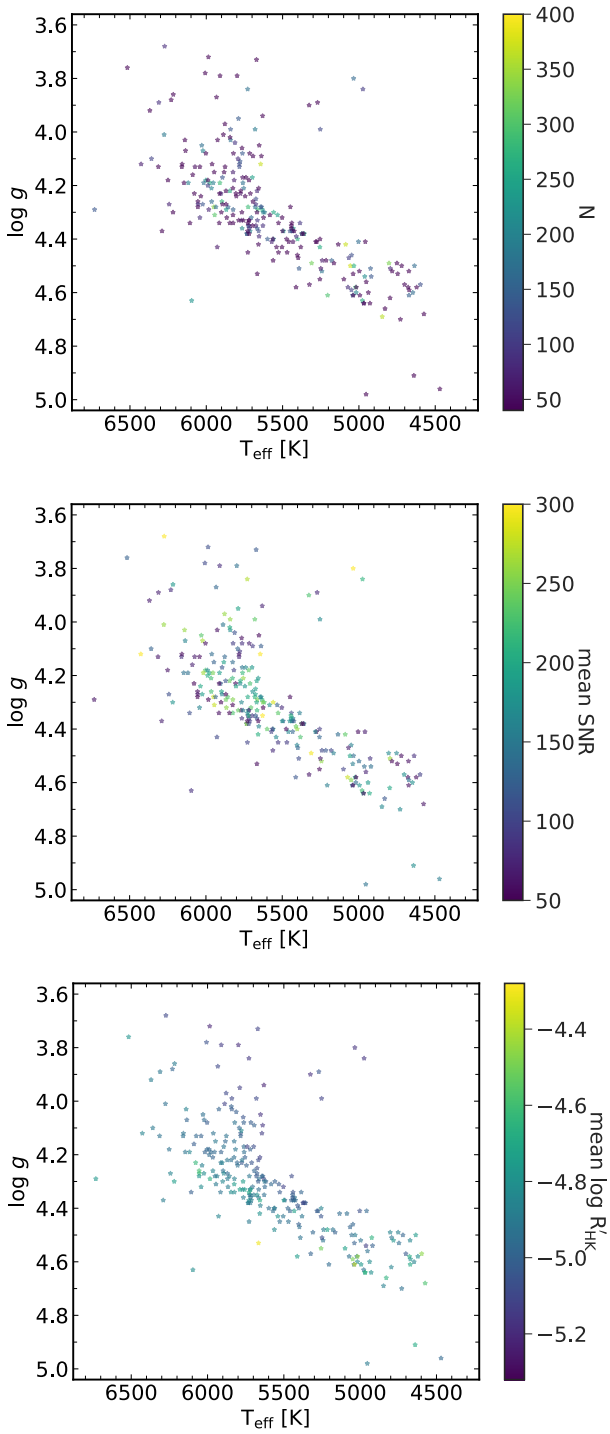


Figure 2.2: The figure shows the basic properties of the 268 targets in our sample. The targets are displayed in HR diagrams, and the colours of the points in the upper, middle, lower panels represent the number of data points in the pre-processed time-series, the mean signal-to-noise ratio (SNR) of the acquired spectra, and the mean $\log R'_{\rm HK}$ of the star, respectively.

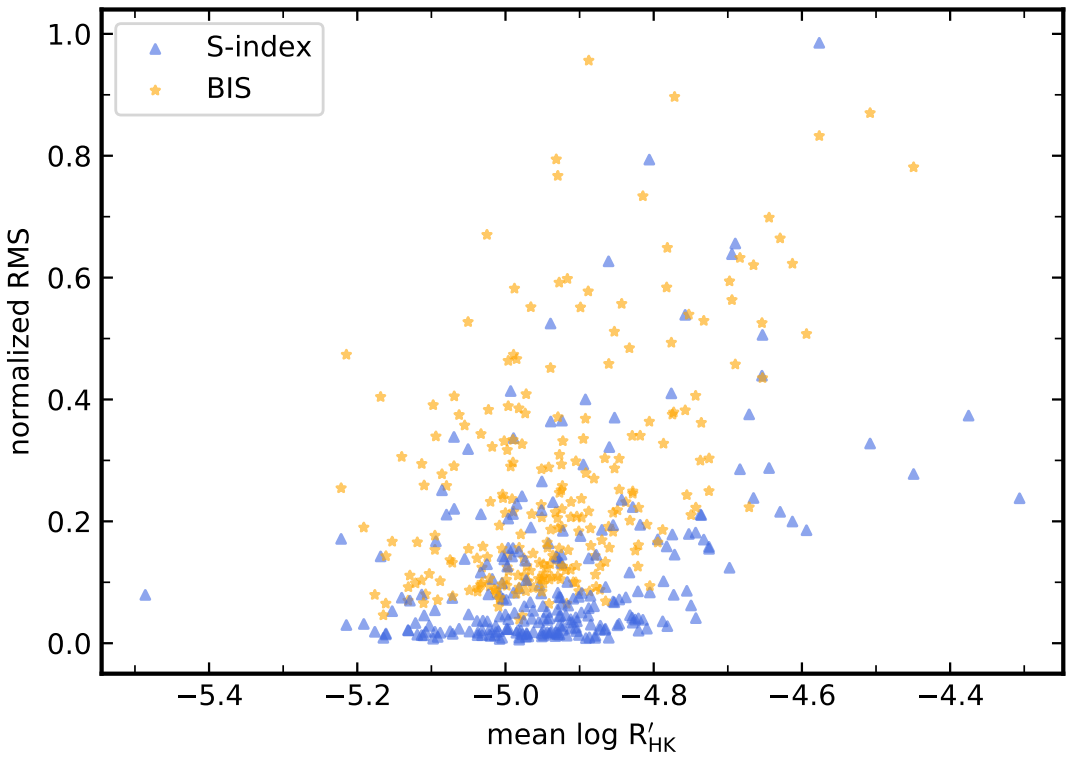


Figure 2.3: The figure shows the normalised root mean square (RMS) of the time-series for both S-index and BIS versus the mean $\log R'_{\rm HK}$.

2014a; Thompson et al., 2017). Hence we exclude $\dot{G}(t)$ term for S-index. In terms of choices of kernels for $G(t)$, we opted for MEP kernels available in `S+LEAF` introduced earlier, which is a faster approximation of the QP kernel. a and b are free parameters to be determined by the model fit.

To explore the posterior of free parameters, we used nested sampling through `PolyChord`[4] (Handley et al., 2015a; Handley et al., 2015b), with all priors on the free parameters set as uninformative. This nested sampling approach is favoured here as it demonstrates better performance when compared to MCMC approaches, especially in datasets exhibiting multi-modality.

We show an example of implementing the framework using the * 61 Vir dataset from our broader sample. In Table 2.1, we show the inferred values of the free parameters in the model based on the posterior distributions, alongside the priors designated for the parameters during the sampling. The inferred values are the

[4] https://github.com/PolyChord/PolyChordLite

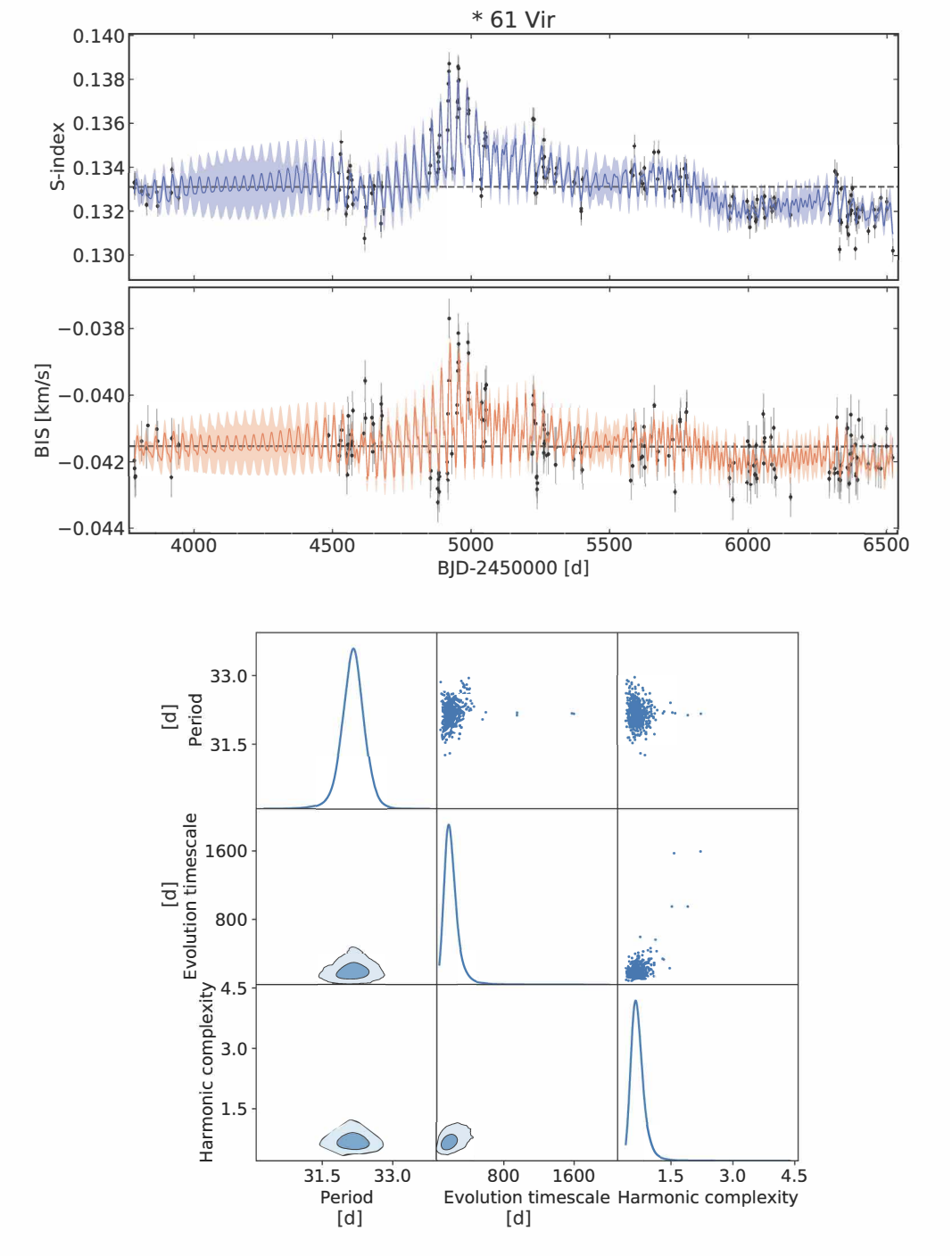


Figure 2.4: Demonstration of multi-dimensional GP fit on the * 61 Vir (HD 115617) dataset. In the upper panel, the black markers show the time-series of S-index and BIS, and their associated uncertainties. The coloured (blue and coral) lines show the GP predictions over the observations, and the corresponding shaded areas of the lines show one σ uncertainty of predicted distributions. The horizontal dashed grey lines show the means of the time-series. The grey extensions on the errorbars show the jitter terms. The lower panel show posterior distributions of selected model parameters, including period, evolution timescale and harmonic complexity.

Table 2.1: Priors and inferred value of selected parameters in the model for the example on the * 61 Vir (HD 115617) dataset.

Parameter	Prior *	Inferred value
Period P (d)	$\mathcal{U}[0.1, 60.0]$	$32.1^{+0.2}_{-0.2}$
Evolution timescale l (d)	$\mathcal{U}[1, 5000]$	186^{+78}_{-50}
Harmonic complexity f	$\mathcal{U}[0.01, 5.00]$	$0.68^{+0.16}_{-0.12}$
a_{BIS}(m s^{-1})	$\mathcal{U}[-\mathrm{rms(BIS)}, \mathrm{rms(BIS)}]$	$0.55^{+0.15}_{-0.08}$
b_{BIS}(m s^{-1} d^{-1})	$\mathcal{U}[-10 * \mathrm{rms(BIS)}, 10 * \mathrm{rms(BIS)}]$	$-3.5^{+0.7}_{-0.9}$
$a_{S-\mathrm{index}}$(d^{-1})	$\mathcal{U}[-\mathrm{rms(S-index)}, \mathrm{rms(S-index)}]$	$0.0013^{+0.0003}_{-0.0002}$

* $\mathcal{U}$[a,b] refers to a uniform prior between a and b.

median of the posterior distributions, and the statistical uncertainties are given by the 16th to 84th percentile range.

Figure 2.4 shows a multi-GP fit to the time-series of * 61 Vir (HD 115617) dataset. In the top panel, the black markers show the time-series of S-index and BIS, and their associated uncertainties. The coloured (blue and coral) lines show the GP predictions over the observations, and the corresponding shaded areas of the lines show one σ uncertainty of predicted distributions. The horizontal dashed grey lines show the means of the time-series. The grey extensions on the errorbars show the jitter terms. The bottom panel shows posterior distributions of selected model parameters, including period, evolution timescale and harmonic complexity.

2.4 Stellar rotation

2.4.1 Rotation period measurement

The variations induced by active regions in activity indicators, modulated by the rotation of the star, are believed to closely represent the star's true rotation period (e.g., Angus et al., 2018; Nicholson and Aigrain, 2022). As such, the rotation period is captured by the hyper-parameter period in our multi-dimensional GP model when the activity-induced signal is well-modelled.

We applied the multi-dimensional GP framework with an MEP kernel to each star in our sample. To determine whether the period is securely detected, on a separate

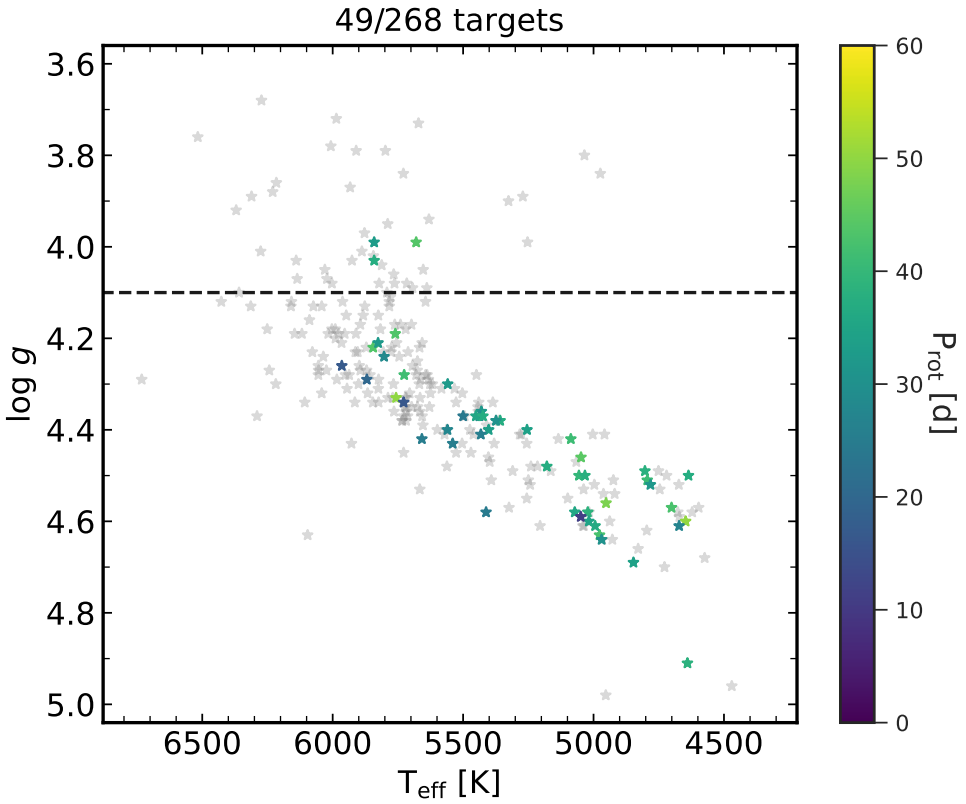


Figure 2.5: HR diagram of our sample with the coloured points showing the measured rotation period for the 49 targets which passed our vetting procedure. The grey markers denote the other targets for which the rotation periods were not recovered or were not considered sufficiently robust to include in subsequent analysis. The dotted line at log g = 4.1 approximately separates main-sequence stars from evolved stars, i.e., sub-giants.

analysis, we applied the same framework but used a Matern 3/2 kernel instead with GP, representing an aperiodic model in comparison to the above periodic model (with the MEP kernel). The evidence Z of each model is estimated through nested sampling to calculate $\Delta\ln \mathrm{Z} = \ln \mathrm{Z}_{\mathrm{periodic}} - \ln \mathrm{Z}_{\mathrm{aperiodic}}$. A positive value indicated that the periodic model is preferred. We then used $\Delta\ln$ Z as our starting criterion to select targets where the periodic model is preferred. Additionally, we rejected cases through visual inspection where the $\Delta\ln$ Z was unreliable because of two possible scenarios. Firstly, in cases where neither model fits well, the $\Delta\ln$ Z criterion could be biased as the Bayesian model comparison is only meaningful when both models adequately fit the data. Secondly, the MEP kernel (or the quasi-periodic

kernel in general) we employed has regions in the parameter space that could behave aperiodically, for example, when the evolution timescale is shorter than the rotation period. In such cases, even if the periodic model appears preferable based on evidence difference, secure detection of the period is not guaranteed. Upon thorough vetting, we identified 49 targets for which we believe the model adequately captured the activity-induced variations, resulting in an accurate recovery of their true rotation periods.

We also want to emphasise that in comparison to conventional methods, such as periodograms or the autocorrelation function (ACF), the GP approach displays greater resilience against harmonic misdetection when determining rotation periods (e.g., Angus et al., 2018). This robustness arises from the utilisation of the MEP or QP kernel, where potential harmonic components are inherently considered, i.e., the 'harmonic complexity' hyperparameter serves as an approximate gauge of the harmonic contribution. This ensures the algorithm finds the true rotation period which is associated with the lowest base frequency.

The selected 49 targets are shown in an HR diagram in Figure 2.5, with the colour of the markers representing their recovered rotation periods. All other targets of which the periods have not been successfully recovered are denoted with grey markers. The dotted line at log g = 4.1 approximately separates main-sequence stars from evolved stars, i.e., sub-giants. Notably, we find that the rotation period broadly decreases as effective temperature rises (or for earlier spectral types), although there is a considerable amount of scatter at any given temperature. This is broadly consistent with both theoretical expectations (e.g., Skumanich, 1972a) and preceding rotation period surveys (e.g., McQuillan et al., 2014). Further insights on age-activity-rotation relations are discussed in section 2.4.3.

Regarding the detection rate, there are approximately 23% of main-sequence G and K stars of which the rotation period is well recovered. However, this rate decreases significantly for earlier type stars such as F stars, and for sub-giants. The rapid rotation of F stars, typically with periods under 10 days, necessitates denser time sampling than what is currently available to accurately determine

their rotation periods. In the case of sub-giants, their evolving nature results in diminished or altered activity patterns, deviating from the assumption of our activity model and thus reducing the detection rate. It is important to note that the detection rate for the rotation period can be considered as an indirect measure of the model's efficacy at mitigating activity signals across different stellar types. In this context, the model demonstrates the best performance for G and K stars. For F stars, the model's efficacy could be enhanced with optimised sampling, tailored to capture their rapid rotation. However, it is essential to recognise that effective activity modelling does not necessarily equate to a target's appropriateness for planet detection. For example, challenges in modelling activity could arise due to a lack of pronounced active regions on the stellar surface. In such cases, the amplitude of the activity signal remains minimal and as such the star could still possibly be an attractive candidate for exoplanet detection.

Last but not least, it is important to highlight that compared to traditional methods like the Lomb-Scargle periodogram, the GP method holds an advantage in determining the signal's period when the rotational modulation signal evolves over time (Angus et al., 2018). Additionally, with the multi-GP method used in this work, we search for and identify common periodic signatures among multiple time-series, thereby increasing the confidence in the measured rotation period. Moreover, the multi-GP framework offers insights beyond just the rotation periods of the stars. Specifically, it provides a quantitative understanding of the covariance structure of the activity signals. Such information can be integrated into statistical frameworks to disentangle activity from planetary signals in the RVs. This can be achieved through either simple linear regression models, or a more complex framework that explicitly takes the covariance structure into account, such as the L1 periodogram of Hara et al. (2017). We defer further exploration of this potential to future work.

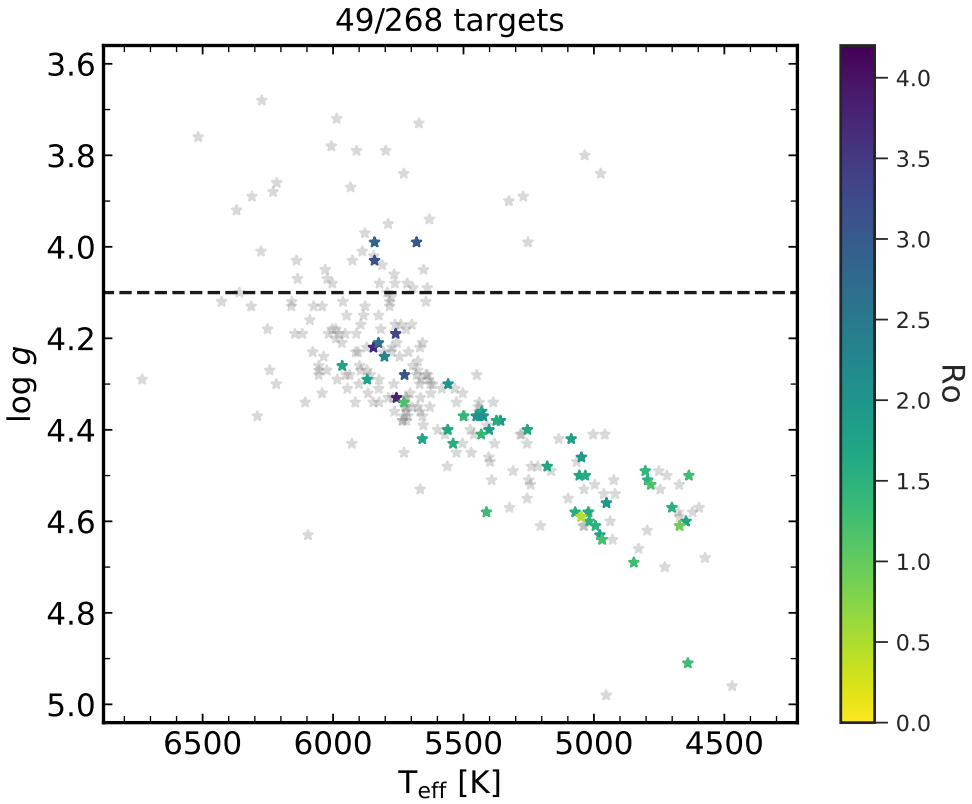


Figure 2.6: As Figure 2.5, but the colour scale now shows the Rossby number.

2.4.2 Rossby number

The Rossby number Ro, which is given by the ratio of the rotation period $P_{\rm rot}$ to the convective turnover timescale $\tau_{\rm c}$:

$$\mathrm{Ro} = \mathrm{P}_{\rm rot}/\tau_{\rm c}, \tag{2.3}$$

is a key parameter of the dynamo process that gives rise to large-scale magnetic fields among the stars we are considering in this study. Numerous empirical studies have shown that Ro is a key parameter for the angular momentum evolution, as well as for magnetic field strength and topology of late-type stars (see e.g. Noyes et al., 1984b).

To first order, the Rossby number provides a robust indicator of a star's magnetic activity level: stars with lower $R_{\rm o}$ values tend to be more magnetically active. This arises because for FGK stars, faster rotation enhances the efficiency of the convective dynamo: the Coriolis force organises convective motions into large-scale, helical

flows that more effectively generate magnetic fields (e.g. Christensen et al., 2009; Brun and Browning, 2017).

We evaluated the Rossby number for the stars in our sample with rotation period measurements, using the empirical relations from Cranmer and Saar (2011) to calculate τ_c:

$$\tau_c\,(T_{\rm eff}\,) = 314.24 \exp\left[-\left(\frac{T_{\rm eff}}{1952.5\ \rm K}\right) - \left(\frac{T_{\rm eff}}{6250\ \rm K}\right)^{18}\right] + 0.002. \tag{2.4}$$

. Note that this relation is applicable only to main-sequence stars.

The Rossby number of the sample is shown in the right panel of Figure 2.6. The coloured points show the selected 49 targets from the visual vetting. Unlike $\log R'_{\rm HK}$, the correlation between the Rossby number and the effective temperature is relatively weak, and can be explained mostly by the fact that the hotter stars in our sample are starting to evolve off the main sequence.

2.4.3 Implications for spin-down of middle-aged stars

Sun-like stars are expected to undergo spin-down following a simple power law (e.g., Skumanich, 1972b) to the first order during their main-sequence stage. This spin-down emerges from the continuous loss of angular momentum, a consequence of torques generated from interactions between the star's surface magnetic field and its stellar wind. However, subsequent observations from open clusters spanning a wide age range challenged the Skumanich-like spin-down. Specifically, these results have brought to light two phases in a star's evolution where the spin-down process appears to 'stall' (Meibom et al., 2009; Agüeros et al., 2018; Douglas et al., 2019).

To account for these stalling effects, several scenarios have been proposed. The two main hypotheses are the core-envelope coupling (e.g., Spada et al., 2011; Lanzafame and Spada, 2015; Spada and Lanzafame, 2020; Johnstone et al., 2021) happening at a few Myrs to a Gyr, and the weakened magnetic braking (e.g., van Saders et al., 2016) happening at a few Gyrs. The former suggests angular momentum can be transferred from the radiative core to the convective envelope

due to internal differential rotation, attenuating the spin-down of the envelope. It should be noted that this internal structure applies only to Sun-like stars. The latter indicates that magnetic braking is substantially weakened once the star achieves a critical Rossby number, possibly due to a change in the magnetic field's morphology. This alteration – shifting from a dipole field to a quadrupole field or an even more intricate configuration – makes it less effective at shedding angular momentum (Réville et al., 2015).

Gaps in the rotation period versus effective temperature diagram have been found with Kepler stars (e.g., McQuillan et al., 2013; McQuillan et al., 2014; Gordon et al., 2021). These gaps can potentially be explained by the aforementioned stalling effects. David et al. (2022) provide stronger evidence by combining the *Kepler* rotation periods with more accurate spectroscopic effective temperatures from, e.g., the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) (Cui et al., 2012; Zhao et al., 2012). Their findings indicate two prominent pile-ups in the period–effective temperature diagram. The short-period pile-up is attributed to the core-envelope coupling mechanism, while the long-period pile-up is linked to the phenomenon of weakened magnetic braking. This latter association is primarily because it aligns seamlessly with a consistent Rossby number curve.

We constructed a similar diagram, by combining rotation periods measured in *Kepler* data by McQuillan et al. (2014) with effective temperatures from LAMOST. This resulted in a blue Kernel Density Estimation (KDE) plot, displayed in the background of Figure 2.7, generated with the codes [5] provided by David et al. (2022). Empirical cluster sequences of different ages are shown in grey lines, derived from Curtis et al. (2020). The purple lines correspond to constant Rossby numbers, following

$$P_{\rm rot}\left(\mathrm{Ro}, \mathrm{T_{eff}}\right) = \mathrm{Ro} \times \tau_{\rm c}\left(\mathrm{T_{eff}}\right), \tag{2.5}$$

with convective turnover timescales from Cranmer and Saar (2011) as Equation 2.4.

[5] https://github.com/trevordavid/rossby-ridge

In the KDE, the long-period pile-up is visible below the Ro = 1.45 curve, while the short-period pile-up appears above the Ro = 0.4 curve. The Sun is shown as a reference in red, with $T_{\mathrm{eff},\odot} = 5772$ K (Prša et al., 2016) and $P_{\mathrm{rot},\odot} = 27 \pm 2$ d estimated from Snodgrass and Ulrich (1990). We have overlaid the asteroseismic sample from Hall et al. (2021), depicted in grey, and our HARPS sample of 49 well-determined rotation periods (detailed in section 2.4.1) in orange. Both samples clearly occupy a distinct period-temperature space to the much larger *Kepler*, predominantly located above the long-period pile-up. We interpret this primarily as a consequence of the different selection effects for the three samples. The *Kepler* sample, while much larger and benefiting from tight, regular time sampling, relies on detecting the rotational modulation of active regions in broad-band optical light curves. This becomes increasingly difficult for less active stars as the fraction of the stellar disk covered by active regions drops, and for slower rotators as the rotation period becomes comparable to both the lifetime of the active regions and the duration of individual *Kepler* 'quarters'. The asteroseismic sample uses p-modes to measure internal and surface rotation rates, and is strongly biased towards hotter and lower gravity (evolved) stars, for which the amplitude of the p-modes is larger and their frequencies lower than for later type main-sequence stars. The HARPS sample spans a longer baseline and uses spectroscopic activity indicators rather than broad-band photometry to detect the rotational modulation of active regions. It is reasonable to expect that targeted indicators should be more sensitive to active regions than photometry, particularly as the active regions decrease in size and number and the rotation slows. A possible additional contributing factor is the fact that, while dark spots dominate the activity patterns for active stars, less active stars tend to be more faculae-dominated (Meunier et al., 2010a). The photometric signatures of faculae are rather subtle, whereas they have strong chromospheric signatures, which are probed by spectroscopic indicators such as $\log R'_{\mathrm{HK}}$ and also have a strong signature in line-shape indicators such as the BIS via their local dampening effect on convective flows. A more in-depth discussion on this topic can be found in section 2.5.

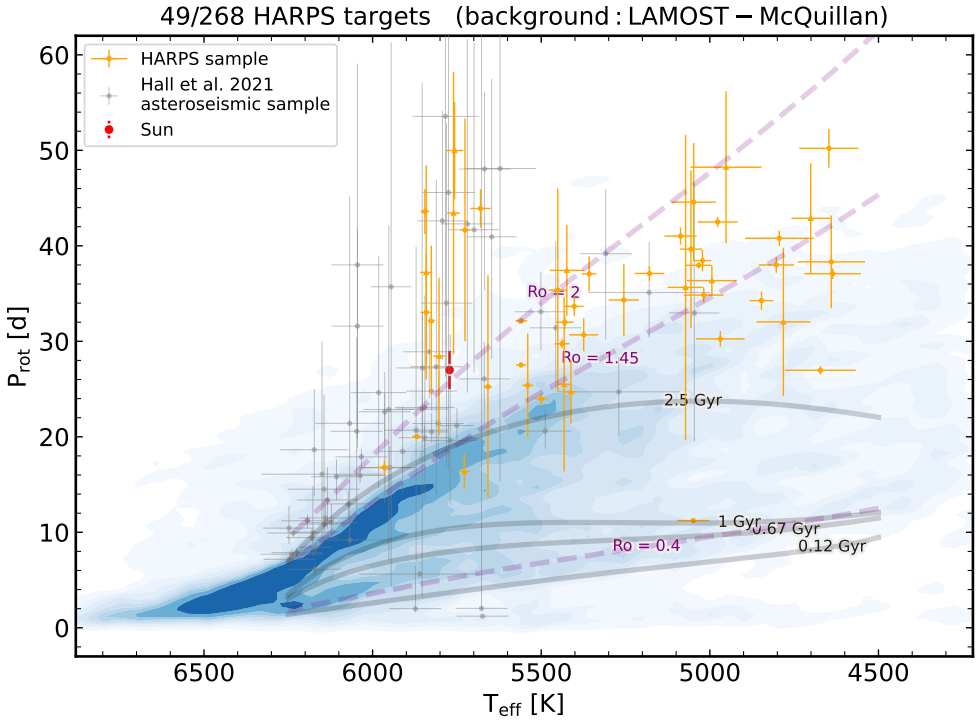


Figure 2.7: Rotation period versus effective temperature for our sample, compared to previous results from the literature. The blue shading in the background represents a kernel density estimate of the *Kepler*/LAMOST period-temperature results. Empirical cluster sequences of different ages are shown in grey lines, derived from Curtis et al. (2020). Lines of constant Rossby number are shown in purple. The asteroseismic sample of Hall et al. (2021) is plotted in grey and our new results based on HARPS data are plotted in orange.

Considering the combined HARPS and asteroseismic samples, we see that two distinct populations emerge. The first is composed of hotter stars (with $T_{\rm eff} \geq 5700\,{\rm K}$), mostly from the asteroseismic sample, though with a handful of HARPS measurements. This population follows a very steep period-temperature relation, consistent with the rapid spin-down expected as stars leave the main sequence and start to evolve towards the red giant branch.

The second population, predominantly from our sample, consists of cooler (mid-G to mid-K) main-sequence stars. It follows a shallower period-temperature relation, and its upper envelope corresponds to the Ro = 2 curve. This second population could possibly be the extension of the long-period pile-up observed in the *Kepler* sample, but skewed towards cooler stars. In the *Kepler* sample, the long-period

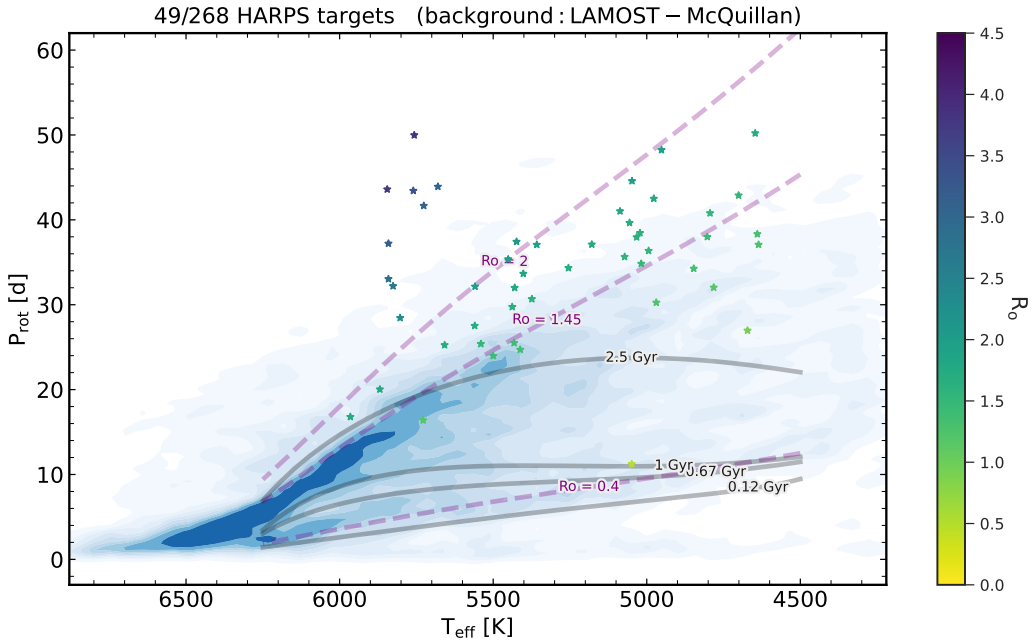


Figure 2.8: As Figure 2.7, but the points now show our sample colour-coded according to Rossby number.

pile-up is less pronounced at cooler temperatures, but this could be a selection effect: at lower temperatures, the pile-up moves to longer periods, which correspond to lower photometric amplitudes, and the stars also become intrinsically fainter. In Figure 2.8, we show the Rossby numbers computed for our sample compared to the KDE of the *Kepler* sample. They fall in the range of 1.5–2.5, consistent with the critical Rossby number inferred from the photometric long-period pile-up. A slight discrepancy between this tail and the anticipated long-period pile-up extension might arise from biases in effective temperature measurements from different sources, as argued in David et al. (2022) in the context of the asteroseismic sample.

2.5 Faculae-to-spots ratio

Spots and faculae (or their chromospheric counterpart, plages) are the two main types of active regions that are expected to induce variations in the RVs and activity indicators used in planet searches. These variations are induced by two simultaneous processes. The first, known as the photometric effect, is thought to be mainly induced by dark spots, which locally reduce (or even suppress entirely)

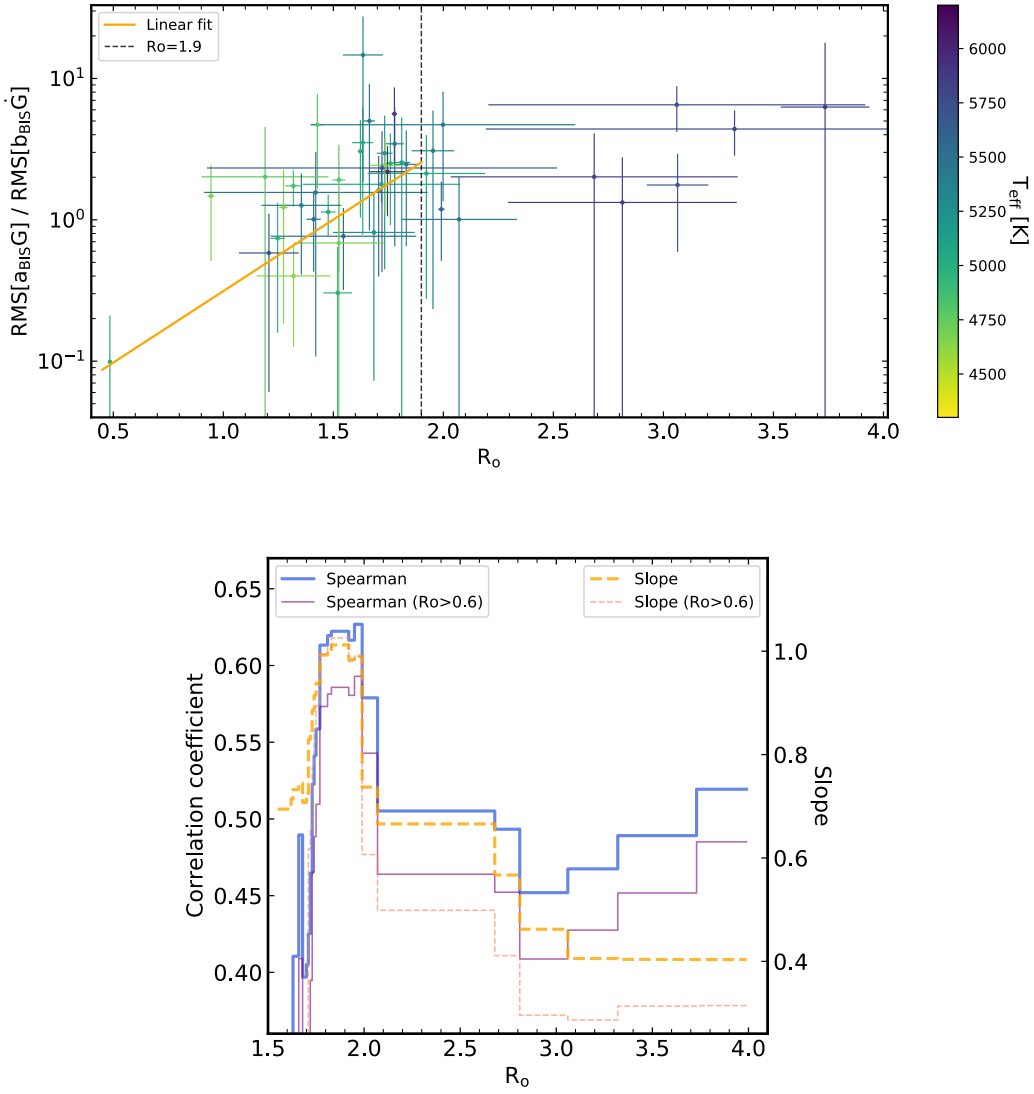


Figure 2.9: The upper panel shows the RMS ratio A/B (a proxy for the faculae-to-spot ratio) versus the Rossby number Ro (more magnetically active stars have smaller Ro). The error bars show the derived statistical uncertainties, and the colour scheme indicates the effective temperature. The dotted grey line indicates Ro = 1.9, and the orange line shows the linear fit to the data within the range of Ro < 1.9. The lower panel shows the estimated Spearman correlation coefficient between A/B and Ro, as well as the slope of the linear fit to log(A/B) as a function of Ro. The two proxies are plotted against the cut-off value of Ro, ranging from 1.6 to 4.1. The blue and orange lines show the analysis of the full data, while the thinner purple and coral lines show the analysis excluding the data point at Ro $\approx$ 0.5 for comparison.

the local flux emerging from the stellar surface. The other is due to the fact that convective up-flows are suppressed in regions of enhanced magnetic flux density, which includes both spots and faculae. This results in a local reduction in the net blue-shift caused by convective flows relative to the "clear" photosphere, known as the Inhibition of Convective Blueshift (ICB) effect. Both processes cause distortions to the disk-integrated line profiles, resulting in variations in the RV, FWHM and BIS, while chromospheric activity indicators such as the S-index or $\log R'_{\rm HK}$ are sensitive to the fractional coverage of the active regions themselves.

The differentiating factor between spots and faculae lies in the processes through which they induce variations. While spots cause variations via both these processes, faculae primarily influence the ICB effect (e.g., Dumusque et al., 2014a). Within the context of our multi-dimensional GP framework, the terms $aG(t)$ and $b\dot{G}(t)$ distinctively represent these two effects (e.g., Aigrain et al., 2012; Rajpaul et al., 2015). Hence, the ratio of the RMS values of these two terms, RMS$[aG(t)]$ / RMS$[b\dot{G}(t)]$, which we denote as A/B in the following text, can be interpreted as an approximate representation of the relative faculae-to-spots ratio. This potentially offers an interesting way to measure the faculae-to-spots ratio in the sample of stars where we consider the results of the GP modelling to be sufficiently robust.

In turn, doing this would allow us, in principle, to test the widely held paradigm that more magnetically active stars are more spot-dominated, while quieter stars have fewer spots and more faculae (e.g., Meunier et al., 2010b; Amazo-Gómez et al., 2020). Of course, direct measurements of the star's magnetic field necessitate polarimetric observations, and few stars in our sample have such observations. However, the Rossby number Ro derived in Section 2.4.2 is known to correlate strongly with the strength of the magnetic field $<$ B $>$ (e.g., Reiners et al., 2014; Vidotto et al., 2014; Reiners et al., 2022), with an exception for very inactive stars (i.e., Ro $>$ 2).

For the targets of which we have detected their rotation periods securely, we examined the relationship between the RMS ratio A/B and the derived Rossby

number Ro in the top panel of Figure 2.9. The error bars show the propagated statistical uncertainties, and the colour scheme indicates the effective temperature of the stars. We excluded samples exhibiting uncertainties in the A/B exceeding 200%.

To evaluate the monotonic relationship between the A/B and the Ro, we calculated the Spearman correlation coefficient, considering an upper cut-off value for Ro in the range of 1.6 to 4.0. The result is shown in the bottom panel of Figure 2.9. We found that the Spearman coefficient peaks between $1.8 < \mathrm{Ro} < 2.0$ with a value of 0.62, after which the value drops rapidly. In addition, we conducted linear fits to log(A/B) as a function of Ro, and showed the slope of the fit versus the Ro cut-off value in the same panel. We found that the slope also peaks between $1.8 < \mathrm{Ro} < 2.0$, behaving similarly to the Spearman coefficient. Both proxies indicate a relatively strong positive correlation between A/B and Ro until $\mathrm{Ro} \approx 1.9$, after which the correlation significantly weakens.

Additionally, we conducted analyses to assess the impact of the data point at $\mathrm{Ro} \approx 0.5$ by excluding it from the correlation analysis. The results are represented as thinner lines in the same panel, labelled as $\mathrm{Ro} > 0.6$. Both the Spearman coefficient and the slope showed behaviors remarkably similar to those observed in the full analysis. This similarity suggests that the overall results are not predominantly influenced by this single data point. Despite this, we assert that the detection at $\mathrm{Ro} \approx 0.5$ is robust. The lack of data in the $0.5 < \mathrm{Ro} < 1.0$ range is likely attributed to the HARPS sample's bias towards inactive stars.

The relationship between A/B and Ro is aligned with theoretical predictions: a lower Ro implies a stronger magnetic field and consequently, a star that is more spot-dominated. This leads to a decreased facular-to-spot ratio, mirrored by a lower A/B, which is exactly the trend we have seen in the $\mathrm{Ro} < 1.9$ regime. However, when $\mathrm{Ro} > 1.9$, the correlation disappears, and there is no longer any obvious relationship between the A/B and the Rossby number. We speculate that the disappearance of the correlation might arise from a morphological transition in the stellar magnetic field beyond $\mathrm{Ro} \approx 1.9$, away from a predominantly dipolar

configuration. If the B-field structure becomes more complex, the correlation between B-field strength and Rossby number would not be expected to persist.

Intriguingly, the Rossby number at which the correlation with the A/B disappears, Ro $\approx$ 1.9, is close to the Ro that marks the upper envelope of the *Kepler* main-sequence rotation period-temperature distribution in Figures 2.7 and 2.8. It also corresponds approximately to the critical Rossby number in the range $1.4 \leq \mathrm{Ro} \leq 2.0$ postulated in stellar spin-down theories (e.g., van Saders et al., 2016; van Saders et al., 2019; David et al., 2022), where the efficiency of stellar angular momentum dissipation is expected to drop significantly due to the same alterations in magnetic field morphology.

A more detailed investigation of the faculae-to-spot ratio and its dependence on magnetic field strength would certainly be valuable, but would likely require more sophisticated activity indicators that can help disentangle the contributions of the different types of active regions (see e.g. Crétignier et al. 2023) as well as a larger sample of stars with direct, polarimetric magnetic field measurements, and is therefore beyond the scope of this work.

2.6 Conclusions

This chapter presents the first results from the activity properties of F, G and K stars using archival HARPS data. We applied the multi-dimensional GP framework developed by Rajpaul et al. (2015) and Barragán et al. (2022) for activity mitigation in RV planet searches to activity indicators extracted from HARPS spectra of 268 well-observed targets with precisely determined stellar parameters. Applying the GP framework to such a large dataset was made possible by using the efficient implementation of our framework provided by Delisle et al. (2022). While we do plan to perform a joint analysis of the activity indicators and the RVs in the future, this is made challenging by the presence in the RV time-series of residual systematic effects and an unknown number of planetary signals. A new version of the HARPS Data Reduction Software is under development (Dumusque et al. priv. comm.), which is expected to lead to significantly reduced systematics and has already

allowed new planet candidates to be identified (with further HARPS observations being gathered to confirm them, ESO program 1110.C-4043, PI Hara). In the meantime, applying the multi-dimensional GP framework to the activity indicators alone already reveals a wealth of useful information. Our key findings are as follows.

We successfully recovered rotation periods for 49 slow rotators in our sample. We found that the rotation period decreases as effective temperature rises, i.e. for earlier spectral types, which is broadly consistent with both theoretical expectations (e.g., Skumanich, 1972a) and preceding rotation period surveys (e.g., McQuillan et al., 2014). One limitation of our approach is that we do not consider the effects of differential rotation, primarily due to the sparse sampling of the majority of the targets in our sample. A possible extension of this work would be to analyse the data season by season and investigate seasonal variations in the period as well as in the other parameters of the model.

We placed our results in the context of existing, photometric estimates of field star rotation periods from *Kepler*, and discussed their implications for the age-activity-rotation relations of F, G & K stars. Our samples typically have longer periods than those derived from photometry, with the exception of the asteroseismic sample. We ascribe this to two factors. First, spectroscopic surveys like HARPS have a much longer baseline compared to photometric surveys. This allows for multi-seasonal monitoring, increasing sensitivity to slow rotators. Second, spectroscopic indicators are more sensitive to active regions, particularly the faculae that tend to dominate in slowly-rotating, magnetically inactive stars, than broad-band optical photometry.

Taken together with the pre-existing asteroseismic sample, our new sample of slowly-rotating ($P >\approx 30\,\mathrm{d}$) FGK stars consists of two, distinct sub-populations. One is a potential extension of the long-period pileup towards cooler stars, with Rossby numbers in the range $1.4 < \mathrm{Ro} < 2.0$. The other consists of 'hot slow rotators' with $T_{\mathrm{eff}} > 5700\,\mathrm{K}$, which we interpret as sub-giants which are rapidly spinning down as they start to expand.

Overall, our results broadly agree with the findings of David et al. (2022), but this is the first time that the population of late G and K-type slow rotators is

observed so clearly, above the 'critical' Rossby number of Ro $\approx$ 1.5 that marks the upper envelope of the period-temperature distributions for *Kepler* stars with photometrically derived surface rotation periods. We infer that angular momentum loss via a magnetised wind continues beyond this critical value, which instead marks the point where the active region covering fraction becomes too low to allow for photometric rotation period estimates.

We also explored indirectly how the ratio of faculae to spots in active regions varies across our sample, using the RMS ratio of $aG(t)$ to $b\dot{G}(t)$ as a proxy. We find that this ratio is positively correlated with the Rossby number up to Ro $\approx$ 1.9, in accordance with theoretical expectations, since a lower Ro implies a stronger magnetic field and consequently, a more spot-dominated star. The correlation disappears for larger values of Ro, indicating a possible transition in the stars' magnetic field morphology away from large-scale, predominantly dipolar fields that give rise to clearly detectable rotational modulation in the activity indicators.

Thus, both our rotation period measurements and the relative contribution of higher-order versus lower-order behaviour in the activity indicator time-series are consistent with the idea that a significant transition in magnetic field morphology occurs around a critical Rossby number in the range $1.4 \leq \text{Ro} \leq 2.0$. Notably, this transition aligns well with the critical Ro values suggested in stellar spin-down theories.

This chapter demonstrates that rich information can be extracted from spectroscopic activity indicators, which sheds light on the structure and evolution of stars. Our intention is to extend this methodology to the RV once there is existing progress on the systemic correction and strategy for a thorough search of planetary signals. We note that the covariance structure of the activity signals learnt from this work can be effectively used to separate stellar and planetary components in the RVs.

3

A Gaussian process model for stellar activity in 2-D line profile time-series

Contents

3.1 Introduction 59
3.2 A physics-based model for line profile variations 61
3.2.1 Preliminaries 61
3.2.2 Extension to line-profile variations 61
3.2.3 Disentangling activity signals from planets 67
3.2.4 Model inputs 69
3.3 Tests and applications 71
3.3.1 Simulated data from SOAP 2.0 71
3.3.2 Injection and recovery test on HARPS-N solar data . . 77
3.3.3 Application to real observations 81
3.4 Discussion 83
3.5 Conclusions 84

3.1 Introduction

Stellar variability is currently the limiting factor on the detection and characterisation of many exoplanets via the RV method (e.g. Fischer et al., 2016; Crass et al., 2021b). This variability arises from the interplay between magnetic fields, convection and rotation. The most obvious component of this variability is associated with evolving magnetically active regions (spots and plage) on the rotating stellar surface,

though (super-) granulation in the quiet photosphere also plays a significant role (e.g. Meunier et al., 2015b; Meunier and Lagrange, 2019b; Al Moulla et al., 2023; O'Sullivan and Aigrain, 2024). Over the past decade, considerable effort has been devoted to developing methods to model and mitigate these stellar RV signals in order to disentangle them from planetary signals (see Zhao et al., 2022, for a review).

As introduced in Chapter 1 and shown in Figure 1.10, a typical workflow for RV extraction begins with a series of high-dispersion spectra. Each spectrum is cross-correlated with a template to produce a CCF, which is then fitted to derive the RV and other 1-D proxies such as FWHM and BIS. In this chapter, we chose to start from an intermediate stage in the workflow, the line-profile or CCF stage. As the CCF combines information from thousands of spectral lines, its SNR is much higher than that of the spectra, so such a method should in principle be applicable to moderate SNR data. Most existing methods to model activity at the CCF level (Jones et al. 2017, Collier Cameron et al. 2021, John et al. 2022, Zhao et al. 2022, Klein et al. 2024) work by separating the CCF variations which can be explained by pure Doppler shifts from those which correspond to changes in the profile shape. However, despite their diminished amplitude, notable quasi-periodic variations tied to stellar activity still persist in the shift-driven RVs.

Most existing methods in general do not explicitly model the time-dependence of the line-profile variations. Doing this seems like a natural progression, as the geometry of the problem (evolving active regions on a rotating sphere) imparts fairly strong constraints on the joint time- and velocity-dependence of the activity signal. These constraints form the basis for the FF' method of Aigrain et al. (2012) and its extension to modelling RV and activity indicator time-series using GPs (Rajpaul et al., 2015), and here we extend that model to a 2–D time-series of line-profiles or CCFs. The hope is that the resulting model will provide a more complete description of the activity signal, while exploiting more of the information contained in the spectra, and being less sensitive to the time-sampling, than pure time-domain methods.

In this chapter, we introduce the CCF-GP framework, a novel approach for modelling stellar activity signals within time-series of line-profiles or CCFs in Section 3.2. We then present tests and applications of the framework on both synthetic and real data in Section 3.3. We present our conclusions in Section 3.5.

3.2 A physics-based model for line profile variations

We seek to model the perturbation to the mean profile of the absorption lines in a star's spectrum induced by evolving active regions on the rotating stellar surface. Unlike previous works, which modelled either the time- or the velocity(wavelength)-dependence of activity signals, but not both, we construct a 2-D model and apply it directly to the time-series of relative line profiles.

3.2.1 Preliminaries

Our model can be viewed as an extension of the FF' framework introduced by Aigrain et al. (2012), which employs simple geometric arguments to describe the relationship between stellar activity signals observed in photometry and those seen in RV measurements. This framework was introduced and discussed in detail in Chapter 1.

Building on this, Rajpaul et al. (2015) adapted the FF' formalism to jointly model the impact of stellar active regions on RV time series and on various spectroscopic activity indicators extracted from the same set of spectra. This approach is often referred to as a "multi-GP" framework and has likewise been presented in Chapter 1. Several extensions and publicly available implementations of this framework have since been developed (e.g., Jones et al., 2017; Barragán et al., 2022; Delisle et al., 2022; Nardiello et al., 2022; Hara and Delisle, 2023).

3.2.2 Extension to line-profile variations

We now consider, instead of a discrete set of time-series (RV and one or more activity indicators), a 2-D time-series of line profiles. More precisely, our starting

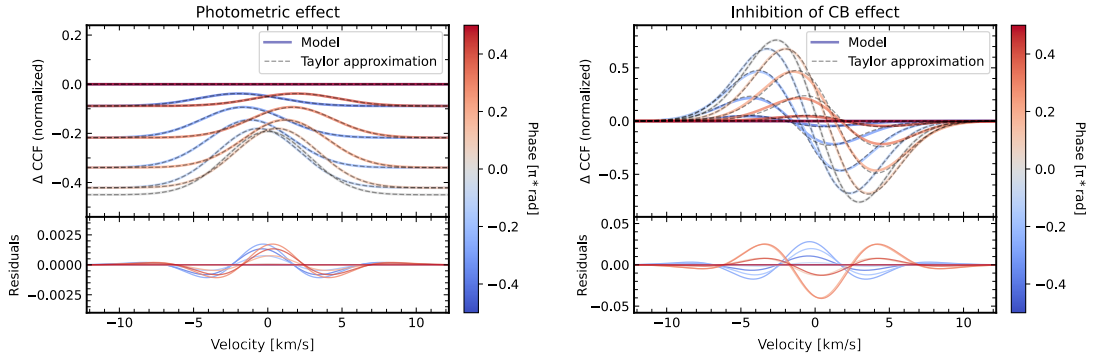


Figure 3.1: Perturbation of the line profile due to the photometric effect of a dark spot (left, $\Delta C_{\rm phot}$) and the ICB effect of a facula (right, $\Delta C_{\rm icb}$) for a single active region traversing the stellar disk along its equator. The coloured curves were evaluated according to Equations 3.1 and 3.5 respectively, at different rotational phases. The dashed grey lines show the corresponding Taylor approximations in Equations 3.2 and 3.6, truncated to 3$^{\rm rd}$-order. The lower panels show the residuals between the line profiles evaluated from the original model and the Taylor approximations. In the left panel, the vertical offsets between ΔCCFs from different rotation phases arise from the fact that we are estimating the line-profile using CCFs that are not continuum-normalised, and the continuum level therefore reflects the total flux variations induced by the active region.

point is a time-series of relative line-profiles, $\Delta C(v,t)$, where v denotes velocity and t denotes time, obtained by subtracting a reference line profile (representing the line profile in the absence of any active regions) $C_{\rm ref}(v,t)$ from the original time-series. In other words, we seek to model the perturbation to the line profile induced by (one or more) active regions. We do not specify at this stage how the line-profiles are extracted from the spectra, or how the reference line profile is constructed; this will be discussed in detail in Section 3.2.4.

As in the FF' framework, we treat the effect on the stellar line profile of an active region moving across the stellar disk as the sum of two distinct terms: the photometric effect and the ICB effect.

The photometric effect

To first order, an active region (e.g., a dark spot) alters the mean stellar line profile by suppressing the flux emerging from a small region of the stellar surface that rotates with the star. The resulting distortion of the line-profile, $\Delta C_{\rm phot}(v,t)$,

can be modelled as:

$$\Delta C_{\rm phot}(v,t) \propto F(t)\; C_{\rm loc}\left(v - V_{\rm loc}(t)\right), \tag{3.1}$$

where $C_{\rm loc}(v)$ is the intrinsic line profile of the quiet photosphere (before rotational broadening). In Equation 3.1, it is evaluated at the location of the spot, i.e. Doppler-shifted by $V_{\rm loc}(t)$, the line-of-sight component of the stellar surface velocity at the location of the line profile. As in the FF' framework, $F(t)$ is a time-dependent term that accounts for the spot's photometric contrast, fore-shortening, and limb-darkening. It can be thought of as the relative photometric perturbation caused by the spot, in the bandpass probed by the spectrograph.

Following Equation 3.1, we introduce a Taylor-series expansion of $C_{\rm loc}$ about $V_{\rm loc} = 0$:

$$\Delta C_{\rm phot}(v,t) \propto F(t) \sum_{k=0}^{\infty} \frac{V_{\rm loc}^k(t)}{k!}\, C_{\rm loc}^{(k)}(v), \tag{3.2}$$

where $C_{\rm loc}^{(k)}(v)$ is the $k^{\rm th}$ derivative of $C_{\rm loc}$ with respect to v. The coloured curves in the left panel of Figure 3.1 show $\Delta C_{\rm phot}$ computed according to Equation 3.1 for a dark spot traversing the stellar disk along its equator. In this case, $F(t)$ is approximated as a pure sine function. The dashed grey lines show the Taylor approximation in Equation 3.2, truncated to 3$^{\rm rd}$ order.

As we have already seen, we can approximate $V_{\rm loc}(t)$ as $\gamma\, F'(t)$, where γ is a constant coefficient. Therefore, the line-profile perturbation due to the photometric effect can be represented as

$$\Delta C_{\rm phot}(v,t) = \sum_{k=0}^{\infty} u_k(t)\, p_k(v), \tag{3.3}$$

where the vectors $p_k(v) \equiv {\rm C_{loc}}^{\rm (k)}(v)$ are given by the local, quiet line-profile and its derivatives, while the vectors u are proportional to a combination of F and its derivatives, similar to the original FF' framework:

$$\begin{aligned} u_0(t) &\propto F(t), \\ u_1(t) &\propto \gamma F(t) F'(t), \\ u_2(t) &\propto \frac{1}{2}\gamma^2 F(t) F'^2(t), \\ u_3(t) &\propto \frac{1}{6}\gamma^3 F(t) F'^3(t), ... \end{aligned} \tag{3.4}$$

Inhibition of convective blueshift effect

As previously mentioned, the magnetic field associated with active regions suppresses convection locally, leading to a local redshift in the line profile. Note that this is a very simplified representation: in reality, the magnetic field also alters the line profiles. For now, we model the ICB effect as a straightforward shift in the local line profile at the location of the active region.

Given this premise, the perturbation of the line-profile due to the ICB effect can be expressed as

$$\Delta C_{\rm cb}(v,t) \propto F(t)\left[C_{\rm loc}\left(v - V_{\rm loc}(t) - \delta V_{\rm c}\right) - C_{\rm loc}\left(v - V_{\rm loc}(t)\right)\right], \tag{3.5}$$

where $\delta V_{\rm c}$ denotes the local redshift induced by the magnetic field. As in the case of the photometric effect, we then employ a Taylor-series expansion:

$$\Delta C_{\rm cb}(v,t) \propto F(t)\sum_{k=0}^{\infty}\frac{\left(V_{\rm loc}(t) - \delta V_{\rm c}\right)^k - V_{\rm loc}^k(t)}{k!}C_{\rm loc}^{(k)}(v). \tag{3.6}$$

The right panel of Figure 3.1 shows the ICB perturbation to the line profile, $\Delta C_{\rm cb}$, caused by a facula crossing the stellar disk along its equator, as given by Equation 3.5. Again, here $F(t)$ is approximated as a pure sine function. The dashed grey lines show Taylor approximation of Equation 3.6, once again truncated to 3$^{\rm rd}$ order.

We can thus represent the line-profile perturbation due to the ICB effect in the same form as in Equation 3.3:

$$\Delta C_{\rm cb}(v,t) = \sum_{k=0}^{\infty} u_k(t)\, p_k(v), \tag{3.7}$$

but the vectors $u_k(t)$ are now given by

$$\begin{aligned}
&u_0(t) = 0,\\
&u_1(t) \propto -\delta V_{\rm c} F(t),\\
&u_2(t) \propto -\delta V_{\rm c}\gamma F(t)F'(t) + \frac{1}{2}\delta V_{\rm c}^2 F(t),\\
&u_3(t) \propto -\frac{1}{2}\delta V_{\rm c}\gamma^2 F(t)F'^2(t) + \frac{1}{2}\delta V_{\rm c}^2\gamma F(t)F'(t) - \frac{1}{6}\delta V_{\rm c}^3 F(t),\\
&...
\end{aligned} \tag{3.8}$$

Full model

For a single active region, we can combine the photometric and ICB components of the perturbation by adding them together. The full model takes the same form as Equations 3.3 and 3.7:

$$\Delta C(v,t) = \sum_{k=0}^{\infty} u_k(t)\, p_k(v), \tag{3.9}$$

and, as before, the vectors $p_k(v) \equiv \mathrm{C_{loc}}^{(\mathrm{k})}(v)$ are given by the local, quiet line-profile and its derivatives, but the vectors $u_k(t)$ are given by:

$$\begin{aligned}
u_0(t) &= \alpha\, F(t), \\
u_1(t) &= \alpha\, \gamma F(t)F'(t) - \beta\, \delta V_\mathrm{c} F(t), \\
u_2(t) &= \frac{\alpha}{2}\, \gamma^2 F(t)F'^2(t) \\
&\quad + \beta\, [-\delta V_\mathrm{c} \gamma F(t)F'(t) + \frac{1}{2}\delta V_\mathrm{c}^2 F(t)], \\
u_3(t) &= \frac{\alpha}{6}\, \gamma^3 F(t)F'^3(t) \\
&\quad + \beta\, [-\frac{1}{2}\delta V_\mathrm{c} \gamma^2 F(t)F'^2(t) + \frac{1}{2}\delta V_\mathrm{c}^2 \gamma F(t)F'(t) - \frac{1}{6}\delta V_\mathrm{c}^3 F(t)], \\
&\ldots
\end{aligned} \tag{3.10}$$

where α and β are constant coefficients.

For the case where the line profiles are continuum-normalized, we expect $\Delta C_\mathrm{norm}(v,t)$ described by:

$$\Delta C_\mathrm{norm}(v,t) = \frac{\Delta C(v,t) + C_\mathrm{ref}(v,t)}{1 + \alpha F(t)} - C_\mathrm{ref}(v,t), \tag{3.11}$$

assuming the $C_\mathrm{ref}(v,t)$ is continuum-normalized already.

Latent GP model for the $u_k(t)$

In the original multi-GP framework, which is tailored for 1D time-series data, each time-series is modelled as a sum of terms in F^2 and FF'. These are modelled by defining a latent variable $G \equiv F^2$, over which a GP prior is placed. Because $dG/dt = 2FF'$, the observed time series are linear combinations of the latent GP and its first derivative, and are thus GPs themselves.

In our 2-D model for the line profile variations, the expressions for the vectors $u_k(t)$ in Equation 3.10 involve combinations of F and F', F'^2, F'^3, ... These can no longer be expressed directly as linear combinations of a GP and its derivative(s). As a result, even if we assume that F, or some function of F (such as F^2, F' or F'^2) is a GP, the $u_k(t)$ are not GPs, and neither is the time-series of line-profile variations.

We explored two possible solutions to this problem. The first consists in placing a GP prior on F and jointly sampling F and F' from this prior, before computing F'^2, F'^3, ... These are then used to compute the $u_k(t)$ before the resulting full model is compared to the observed line-profile time-series. This approach allows us to reproduce the precise form of Equation 3.10, but involves drawing a GP sample for every likelihood evaluation. The computational cost of such an approach proved to be prohibitive.

We therefore adopted a less accurate, but cheaper solution, which consists in treating each $u_k(t)$ as a linear combination of some GP $G(t)$ and its time-derivative $G'(t)$, just as in the multi-GP framework described in Section 1.4.2. While this is clearly an approximation, our experience in modelling RV and activity indicator time-series with the multi-GP framework has shown that the combination of a GP and its first derivative is typically flexible enough to capture higher-order variations.

Therefore, in our final model each $u_k(t)$ is approximated as $a_k G(t) + b_k G'(t)$ and $p_k(v) \equiv \mathrm{C_{loc}}^{(\mathrm{k})}(v)$. We describe the procedure for deriving $p_k(t)$ from the observations later in Section 3.2.4. Consequently, the line profile residuals at a given velocity are also a linear combination of G and G':

$$\Delta C(v,t) = \sum_{k=0}^{r} p_k(v)\, a_k\, G(t) + \sum_{k=0}^{r} p_k(v)\, b_k\, G'(t), \tag{3.12}$$

where r is the number of terms used. We found that setting $r = 3$ provides a sufficiently accurate approximation for Sun-like stars while also efficiently reducing computational costs. Note that the value of r should be determined on a case-by-case basis: for stars with higher $v \sin i$, a larger r should be used.

Thus, the matrix $\mathbf{\Delta C}$ is modelled as a latent GP model with n outputs, where n is the number of velocity bins, and the relationship between the output time-series

is controlled by the $p_k(t)$, which corresponds to the emergent line profile of the quiet photosphere. Importantly, the total number of free parameters in the model is only, $2r + m$, where m is the number of parameters of the GP covariance function used.

Relationship to Principal Component Analysis

The key insight in our model is that the line-profile perturbations can be modelled as a sum of separable functions of time and velocity. This is in some sense analogous to performing Principal Component Analysis (PCA) on the matrix of line-profile residuals. Specifically, the Singular Value Decomposition (SVD) of this matrix is given by

$$\mathbf{\Delta C} = \mathbf{U\,S\,P}, \tag{3.13}$$

where $\mathbf{U}$ and $\mathbf{P}$ are rectangular matrices of left- and right-singular vectors (eigenvectors) which depend only on time and on velocity, respectively, while $\mathbf{S}$ is a diagonal matrix of singular values (square-root of eigenvalues). However, in SVD each set of singular vectors forms an orthonormal basis, whereas this isn't necessarily the case for the $u_k(t)$ and $p_k(v)$ vectors in our model.

3.2.3 Disentangling activity signals from planets

A key purpose of the model introduced above is to disentangle activity signals from the Doppler shifts caused by planetary companions, and thereby to improve the detectability and characterisation of the planets. To do this, we must explicitly account for the planet(s) when modelling the line-profile variations, as our model would otherwise have enough flexibility to absorb or modify some of the planet signals. The parameters of the planet(s) and the activity model must therefore be sampled jointly. In order to do this without escalating the computational cost, it is vital to keep the GP as the bottom layer of the workflow, to ensure the likelihood can still be computed analytically.

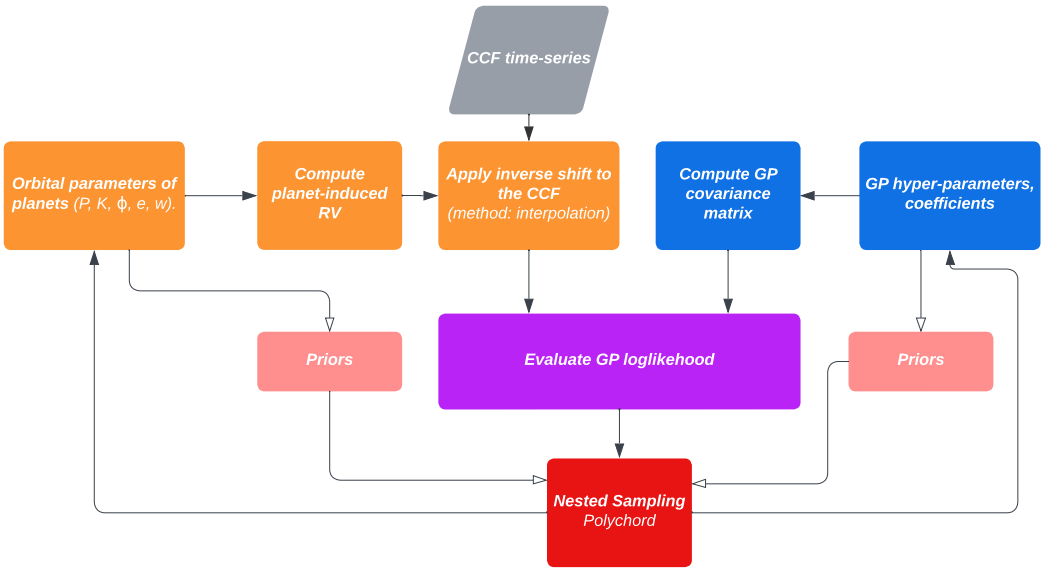


Figure 3.2: Flowchart illustrating the modelling workflow used in this work. The inputs are the observed time series of CCFs as well as the priors on all free parameters in the model. The output is the estimated posterior distribution of the free parameters.

Workflow

Figure 3.2 illustrates our overall workflow. At each step, we sample the parameters for the planet(s) and the activity model simultaneously. The $5p+1$ planet parameters (where p is the number of planets, and the remaining parameter defines the systemic RV) are used to evaluate the planet-induced RV signal, and the line profiles are then shifted inversely by the corresponding amount.

To apply a shift to the line profiles, we initially fit a non-parametric function to each profile using least-squares cubic interpolation, as implemented in `SciPy` (Virtanen et al., 2020)[1]. Subsequently, we apply an inverse shift to the velocity grid of each line profile to account for the RV signal induced by the planet. This shifted velocity grid is then input into the fitted non-parametric function to compute the shifted line profile for each time stamp.

In parallel, given the $2r + m$ activity parameters, a GP solver is used to evaluate the covariance matrix $\mathbf{K}$ of the model and compute a marginal likelihood according to Equation 1.9.

[1] https://scipy.org/

Practical implementation of GP

We use the fast GP solver `S+LEAF 2` (Delisle et al., 2022) with a Matérn 3/2 exponential periodic (MEP) kernel. This serves as a fast approximation of the quasi-periodic kernel frequently used to describe rotational modulation active regions in stellar light curves and RV time-series, allowing the likelihood evaluation to scale linearly with the number of observations and velocity bins. We also include a constant term on the diagonal of the covariance matrix to absorb any imperfections in our model, which results in $m = 4$ parameters for the GP covariance function.

As we are sampling over a relatively large number of parameters, and expect the likelihood hyper-surface to be multi-modal in some cases, we sample the joint posterior over the parameters using the `PolyChord`[2] nested sampler (Handley et al., 2015a; Handley et al., 2015b). Nested sampling is generally considered more adept at handling multi-modal problems, making it a more appropriate choice for this application than more standard Markov Chain Monte Carlo (MCMC) samplers, while the polychord implementation is known to perform especially well for large numbers of parameters.

3.2.4 Model inputs

The starting point for our method is a matrix $\mathbf{\Delta C}$ containing a time-series of relative line-profiles:

$$\Delta C_{ij} = C_{ij} - c_{\mathrm{ref,i}} = C\left(v_i, t_j\right) - c_{\mathrm{ref}}(v_i), \tag{3.14}$$

obtained by subtracting a reference line profile $\mathbf{c}_{\mathrm{ref}}$ (representing the line profile in the absence of any active regions) from each row of the matrix of line profiles $\mathbf{C}$.

Extracting the line profiles

Except where stated otherwise, we equate each row of $\mathbf{C}$ to the CCF produced by the standard "DRS" (Data Reduction Software, used for HARPS, HARPS-N and ESPRESSO data (e.g., Lovis 2007) for that spectrum. The spectrum is first

[2] https://github.com/PolyChord/PolyChordLite

corrected for instrumental effects and telluric contamination using the `YARARA` pipeline (Cretignier et al., 2021a; Cretignier et al., 2023). These CCFs are obtained by cross-correlating the spectra with a binary mask, which is zero everywhere except at the centre of individual spectral lines (Baranne et al. 1996, resulting in "inverted" CCFs, which resemble an absorption line profile, e.g. the top-middle panel in Figure 1.10). Note that if the CCFs are un-normalised, the continuum level (away from the centre of the line profile) varies with time.

In principle, it may be preferable to extract line profiles from the spectra using Least Squares Deconvolution (LSD, e.g., Donati et al. 1997; Kochukhov et al. 2010). However, we chose to use CCFs in this work because they are computed automatically by the DRS and thus readily available. In practice, we have found that, provided the line lists used in the creation of the CCF template and in the LSD process are similar, the resulting line profiles are also very similar.

Reference profile

Ideally, the reference profile $\mathbf{c}_{\rm ref}$ should be equal to the line profile in the absence of any active regions, so that the residuals $\mathbf{\Delta C}$ represent the perturbations induced by activity. In practice, we do not normally have spectra which are guaranteed to be activity-free. Therefore, the reference line profile is obtained by averaging the line profiles $\mathbf{C}$ over the time axis, optionally selecting a subset with a lower activity level (as measured, for example, by the calcium S or $\log R'_{\rm HK}$ indices).

Estimating the $p_k(v)$

We approximate the local, quiet line profile ${\rm C_{loc}}^{(k)}(v)$ as the deconvolution of the reference profile $\mathbf{c}_{\rm ref}$ with a rotational broadening kernel estimated from the $v \sin i$ of the star. The effect of the rotational broadening becomes negligible for slow rotators, when the $v \sin i$ is significantly smaller than the width of the instrumental profile.

Table 3.1: Key input parameters for simulating time-series of CCFs with `SOAP2`.

Case	$P_{\rm rot}$	$N_{\rm spots}$	Long	Lat	$N_{\rm faculae}$	Long	Lat
–	[d]	–	[deg]	[deg]	–	[deg]	[deg]
<i>	25.0	1	180	30	1	60	60
<ii>	25.0	2	180,-150	30,-15	2	60,-40	60,70

3.3 Tests and applications

In this section, we analysed both synthetic and real datasets. Our objectives are fourfold: a) to validate the framework by assessing its ability to model signatures from stellar spots and faculae; b) to determine if the framework can more effectively disentangle planetary and stellar signals than existing 1-D time series methods; c) to evaluate its efficiency in detecting planets with low-amplitude signals; and d) to demonstrate its capability to operate effectively at a Signal-to-Noise Ratio (SNR) typical of most real-world observations, approximately 100.

3.3.1 Simulated data from SOAP 2.0

Before implementing the suggested framework on real data, we tested it using simulated data from version 2 of the `SOAP` (Spot Oscillation and Planet) code (Dumusque et al., 2014b; Boisse et al., 2012). `SOAP2` [3] enables users to generate time-series of CCFs distorted by active regions, namely spots and faculae, on the stellar surface. This is achieved by specifying the latitude, longitude, and size of each region, alongside the star's basic properties such as its rotation period. The tests involved two configurations – cases <i> and <ii>, detailed in Table 3.1, including the simulated rotation period of the star, the numbers and locations of the spots and faculae. For both cases, the sampling rate is one point per day, and the 150-day duration corresponds to six times the rotation period of the star.

For the first case, we simulated a stellar surface featuring one spot and one facula, primarily to validate the framework's concept, specifically, its ability to track the signatures in the CCFs influenced by both the spot and facula. The simulated reference-subtracted CCFs (ΔCCF) are shown in Figure 3.3 in the left panel, with

[3] https://www.astro.up.pt/resources/soap2/

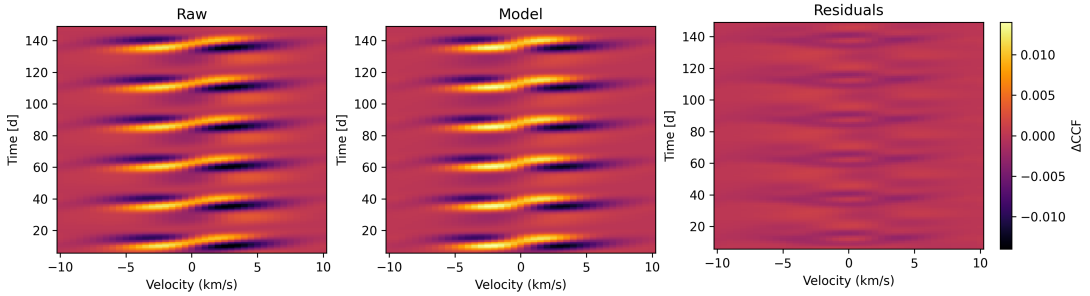


Figure 3.3: Left panel: simulated reference-subtracted CCF (ΔCCF) with `SOAP2` (case <I>). Middle panel: the best-fit CCF-GP model. Right panel: the residuals after subtracting the best-fit model.

the best-fit CCF-GP model in the middle panel and residuals in the right panel. The typical amplitude of the variations in the raw ΔCCF is around 1%, decreasing to approximately 0.3% in the residuals. Figure 3.4 shows a 3-D representation of ΔCCF versus velocity and time, enabling a clearer visualisation of the model fit. This figure highlights that the variations are correlated in both the velocity and time domains, and demonstrates that their correlations are effectively captured by our framework.

For the second case, our objective was to demonstrate the model's ability to limit the degeneracy between the signals induced by stellar activity and planets. For this test, we simulated a more intricate and realistic stellar surface with two spots and two faculae, with the properties listed in Table 3.1. To challenge the framework, we then injected a planetary signal into the simulated CCFs, with exactly the same orbital period as the star's rotation period (in this instance, 25 days). This planet signal had a semi-amplitude of 1.0 m/s and its phase was arbitrarily set to 0.3 π rad. The phase is defined with reference to T=0 and phase=0, corresponding to the position of inferior conjunction for the injected planet. Figure 3.5 displays the mean-subtracted 1-D proxies, including RV, BIS, and FWHM, derived from the planet-injected simulated CCFs. The first panel also shows the planet-induced RV in grey to enable its comparison to the overall RV signal.

We first computed the Lomb-Scargle periodogram (using `Astropy` (Astropy Collaboration et al., 2013; Astropy Collaboration et al., 2018; Astropy Collaboration et al., 2022) of each 1-D time series, as shown in Figure 3.6. The grey vertical

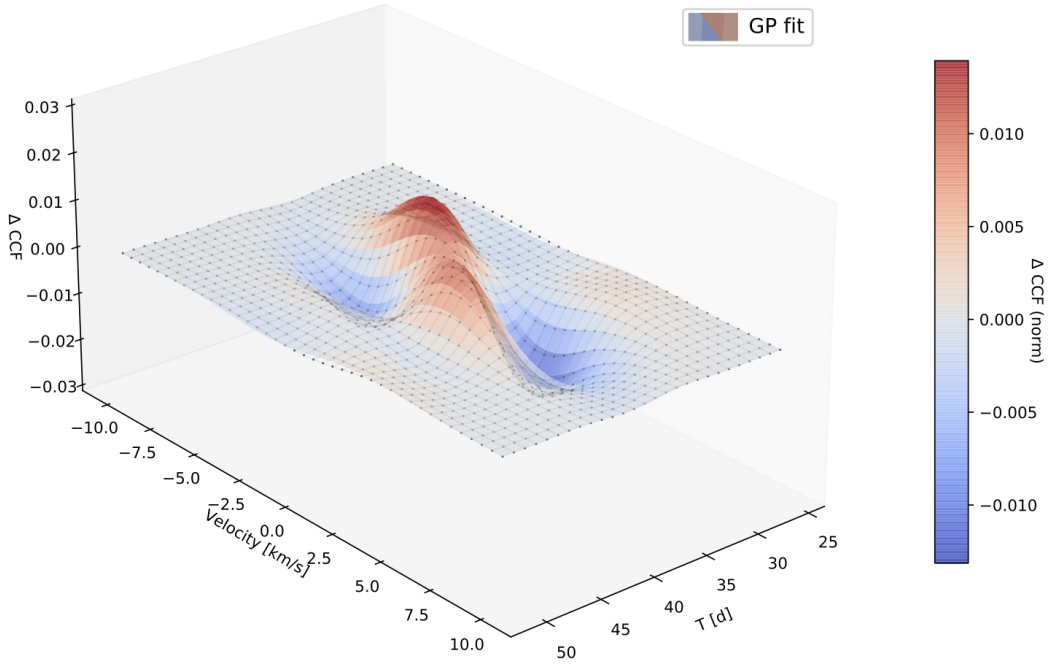


Figure 3.4: 3-D visualisation of the CCF-GP model fit to the referenced-subtracted CCF, ΔCCF simulated with `SOAP2` (case <i>), zooming into one rotation period. The black markers represent the simulated data, and the coloured surface illustrates the model fit, with the colour gradient indicating ΔCCF values (positive in red and negative in blue).

markers indicate the positions of 25-day signals and their harmonics. Using the traditional approach of comparing the periodogram of the RV and 1-D activity indicators, any peak in the RV periodogram that also appears in the periodograms of the activity indicators would be dismissed as likely arising from activity. Therefore, discerning signals induced by planets from those induced by stellar activity is impossible using periodograms alone when both share the same period.

We subsequently applied the CCF-GP framework to the CCF time series. For comparison, we also applied the multi-GP framework, to the RV, BIS, and FWHM extracted from the same CCF time series. Both models incorporated the same planet model for RVs, featuring a single Keplerian. Crucially, the parameter priors in both the multi-GP and CCF-GP were identical and remained uninformative. Figure 3.7 presents the posterior distribution of the recovered orbital parameters

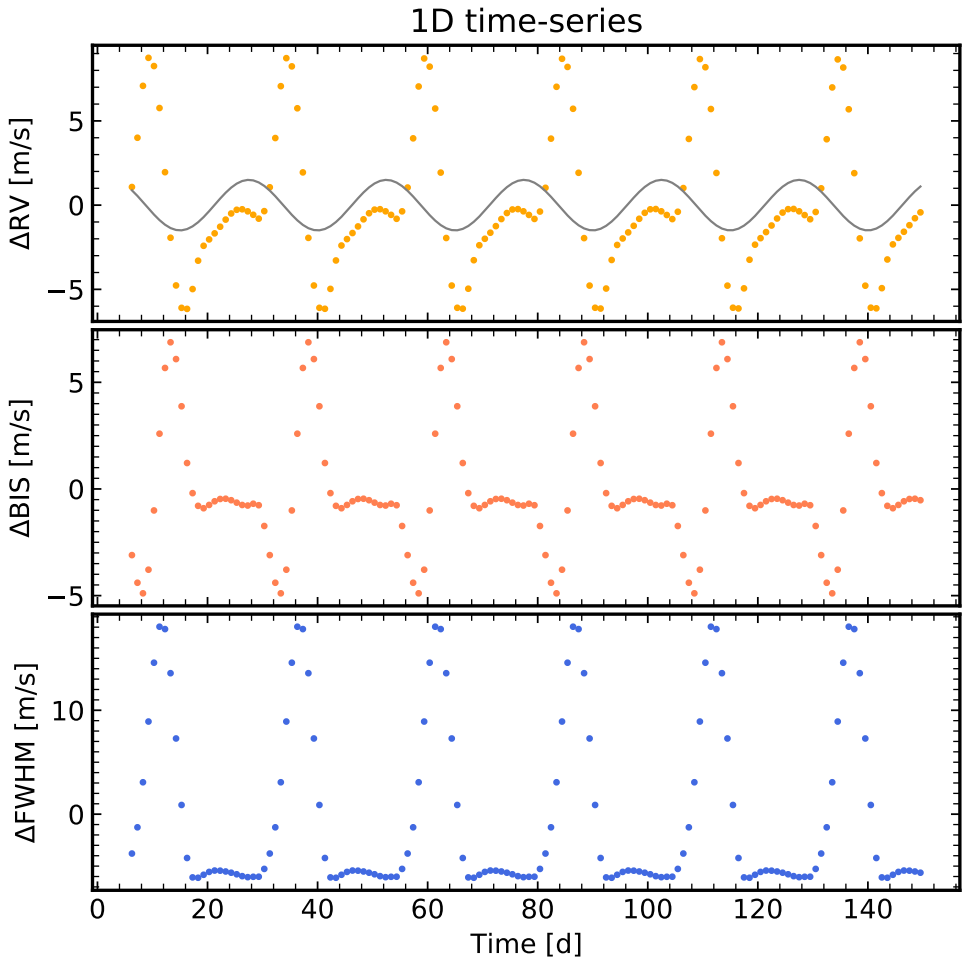


Figure 3.5: Mean-subtracted 1-D time-series, including RV, BIS, and FWHM, which are derived from the planet-injected simulated CCFs generated with `SOAP2` case <ii>. The first panel also shows the planet-induced RV in grey.

from both CCF-GP (in blue) and multi-GP (in orange) models. Vertical red lines indicate the true values from the injections.

It is evident that both models adeptly recover the 25-day period. However, as the stellar rotation period is also at 25 days, to be certain that the detected signal is indeed that of the planet, we also required that the phase of the planet signal be retrieved. Observing the posterior distribution over the phase reveals a notable advantage for the CCF-GP, which captures the truth within a 1-σ uncertainty range, while the phase determined via the multi-GP deviates by more

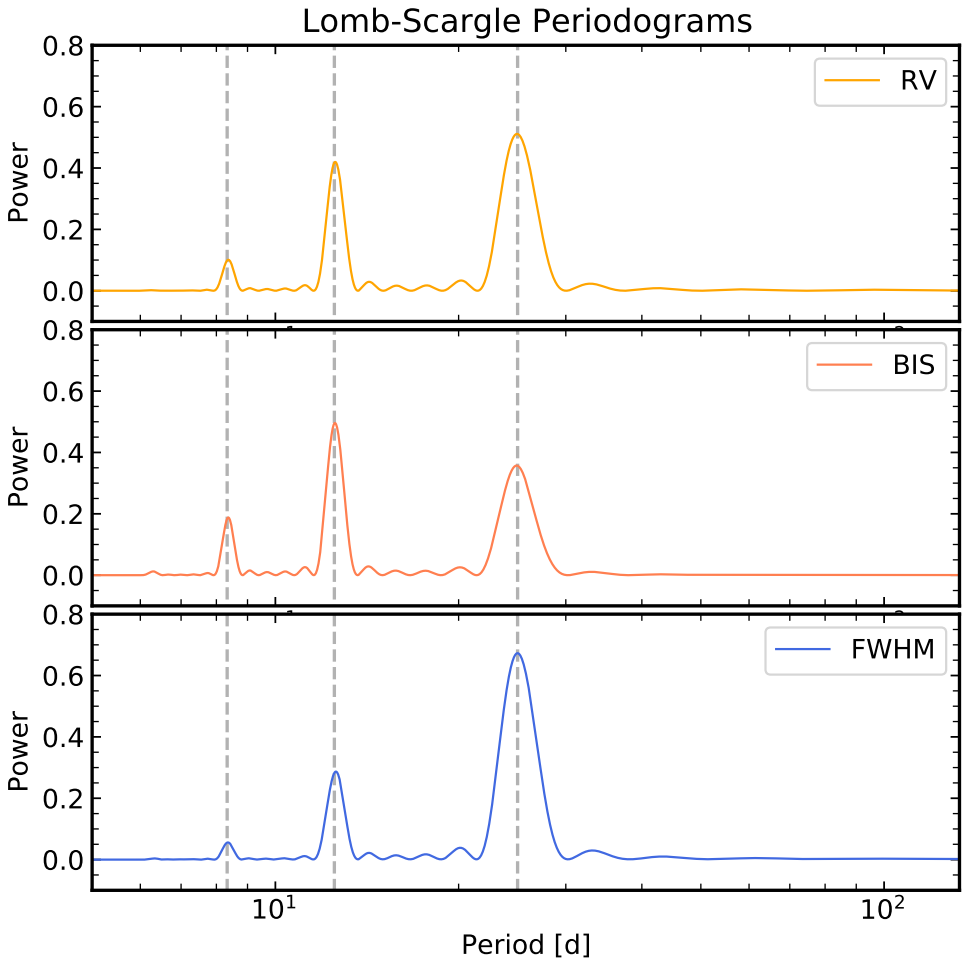


Figure 3.6: Lomb-Scargle periodogram of each 1-D time series (including RV, BIS, and FWHM) shown in Figure 3.5. The grey vertical markers indicate the positions of the 25-day rotation signal and its harmonics.

than 0.1 π rad from the true value. Regarding the semi-amplitude, both models exhibit a tendency to provide inflated estimates of $K = 1.86^{+0.05}_{-0.05}$ m/s (CCF-GP) and $K = 2.57^{+0.26}_{-0.24}$ m/s (multi-GP), though the CCF-GP derived value is closer to the truth of $K = 1.5$ m/s. This inflation implies that the models are conservative in removing the activity signals, leaving a fragment of these signals intertwined with the planet signals. In the case of the CCF-GP model, this slightly conservative behaviour might come from the simplification of modelling the CCFs as vanilla Gaussian processes (as discussed earlier in Section 3.2), potentially neglecting minor

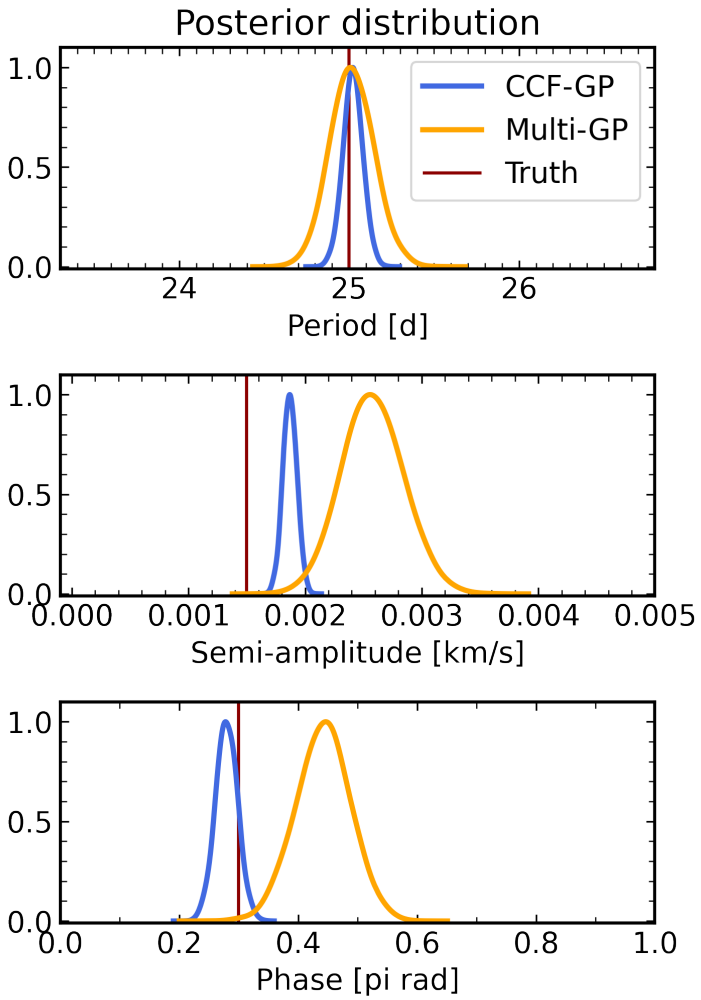


Figure 3.7: Posterior distributions of the recovered orbital parameters for the simulated dataset (`SOAP2` case <ii>) from both CCF-GP (in blue) and multi-GP (in orange) models. Vertical red lines indicate the true values from the injections.

yet pertinent non-Gaussian aspects of the signals.

With the same simulated dataset, we conduct evidence comparisons on combinations of model and simulated dataset with 0 or 1 planet with CCF-GP and multi-GP frameworks. The four cases are listed in Table 3.2, with their corresponding evidence $\log Z$ estimated from the nested sampling. We also provide $\Delta \log Z$, which is calculated as the evidence of the 'correct' model minus the 'incorrect' model. We find that for the data containing 1 injected planet, the model containing 1 planet (1-planet model) is always strongly preferred compared to the 0-planet model for

Table 3.2: Evidence for combinations of model and simulated dataset with 0 or 1 planet, generated from `SOAP2` case <ii>.

Case	CCF-GP $\log Z$	CCF-GP $\Delta \log Z$	Multi-GP $\log Z$	Multi-GP $\Delta \log Z$
0-planet data and 0-planet model	44232.20 ± 0.33	-45.44	2348.28 ± 0.30	-2.31
0-planet data and 1-planet model	44277.64 ± 0.32	–	2350.59 ± 0.29	–
1-planet data and 0-planet model	43941.19 ± 0.24	–	2292.24 ± 0.30	–
1-planet data and 1-planet model	44242.19 ± 0.34	301.00	2349.71 ± 0.29	57.47

both CCF-GP and multi-GP. However, for the data without planet injection, the 1-planet model is preferred to the 0-planet model for both CCF-GP and multi-GP, though such a preference for the latter one is less presented. Further development is needed to prevent false-positive detections.

Overall, we find that in this specific test, the CCF-GP approach generally outperforms the multi-GP approach in disentangling activity and planetary signals, especially when their periods closely match the star's rotation period or its harmonics.

3.3.2 Injection and recovery test on HARPS-N solar data

To assess the behaviour of the CCF-GP framework under more realistic conditions, and to determine its ability to detect low-amplitude planet signals (while having a benchmark truth), we conducted planet injection and recovery tests using HARPS-N solar data.

The spectra were collected using the HARPS-N solar telescope (Cosentino et al., 2012b; Dumusque et al., 2015; Phillips et al., 2016) with observations taken for several hours each day at 5-minute cadence since 2015. Initial processing of the raw data was performed by Dumusque et al. (2021b) employing the ESPRESSO data-reduction software (DRS; Pepe et al., 2021b). Following Collier Cameron et al. (2019), cloud-affected spectra were excluded, and the remainder were shifted to the solar rest frame through interpolation using the Solar barycentric RV solution derived with the `JPL Horizons` software (Giorgini et al., 1996). All sets of three consecutive 5-minute exposures are combined to make one 15-minute exposure for mitigating the effect of p-mode oscillations. We then simulated the temporal sampling one might expect for hypothetical stars observed from La Palma by adding 0.5 days to the observation times (to shift them to night times), then selected at

most one observation each night when a hypothetical star located at RA = 90°, DEC = +30° is observable at an airmass below 1.8. Our test utilised data from mid-2015 to mid-2017, corresponding to an active part of the solar magnetic activity cycle. There are 130 observations in total, with an average sampling rate of one point per 3.8 days, and a typical SNR of ∼370 per 0.82 km/s velocity bin in échelle order 60 of the spectra. Figure 3.8 shows the mean-subtracted RV measured from the selected data, which exhibits a root mean square (RMS) value of 1.64 m/s.

We first applied the CCF-GP framework to analyse the HARPS-N solar CCFs without injecting any planetary signals. This involved fitting the CCF-GP model to the data and subsequently subtracting it to get the CCF residuals. We then computed the RV residuals by fitting a Gaussian function to each of these residual CCFs, which are represented as coral markers in Figure 3.8. The RV residuals have an RMS value of 0.44 m/s, with an RMS reduction factor of around 4 when compared to the raw solar RVs.

We also calculated the Lomb-Scargle periodograms of the raw and the residual RVs, shown in Figure 3.9. The periodograms are normalised as half of the square root of the power spectrum density (PSD) values, and are thus called the semi-amplitude spectrum. Grey vertical markers indicate the positions of the 27-day solar-rotation signal and its harmonics. Most of the frequency components, especially the ones related to the solar rotation period, are significantly reduced in the residual RVs.

We tested two scenarios by injecting a planet into the HARPS-N solar CCFs with semi-amplitudes of 1.0 m/s and 0.3 m/s. In both cases, the orbital period was set to 33 days, and the phases were arbitrarily set to 0.7 π rad. The phase is defined with reference to BJD=2450000. These correspond to planets of masses around 5 and 1.5 Earth masses respectively. Figure 3.8 shows the RVs of the planet for both semi-amplitude settings in blue and orange. We then applied the CCF-GP framework on the planet-injected CCF. All priors on the free parameters are uninformative, which means this is a blind search (though we only tested a one-planet model).

Figure 3.10 shows the posterior distribution of the extracted planetary parameters, with vertical red lines indicating the true values from the injections. For the

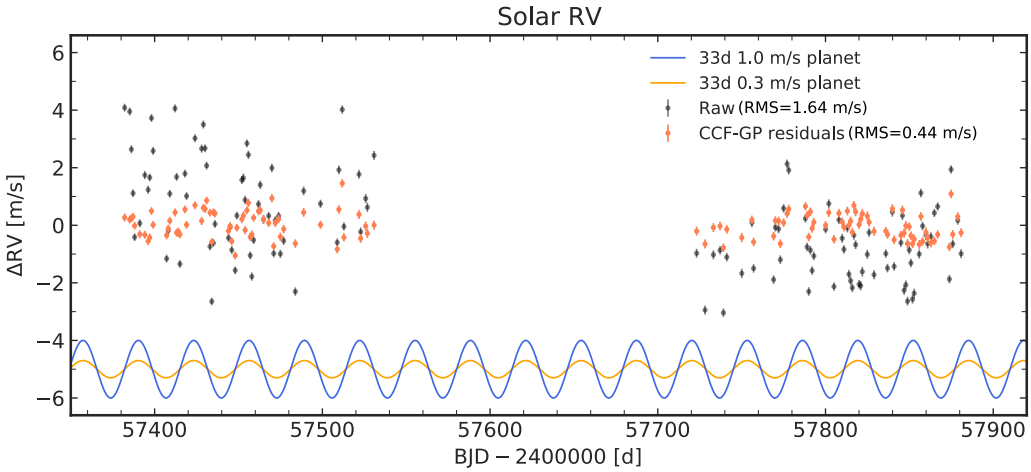


Figure 3.8: The HARPS-N solar RVs used for injection/recovery tests (black markers) and the residual RVs (coral markers), which were measured from the CCFs after subtracting the best-fit CCF-GP model. Also shown are the injected 33-d planet signals for both semi-amplitude settings in blue (1.0 m/s) and orange (0.3 m/s), offset by -5 m/s in the figure.

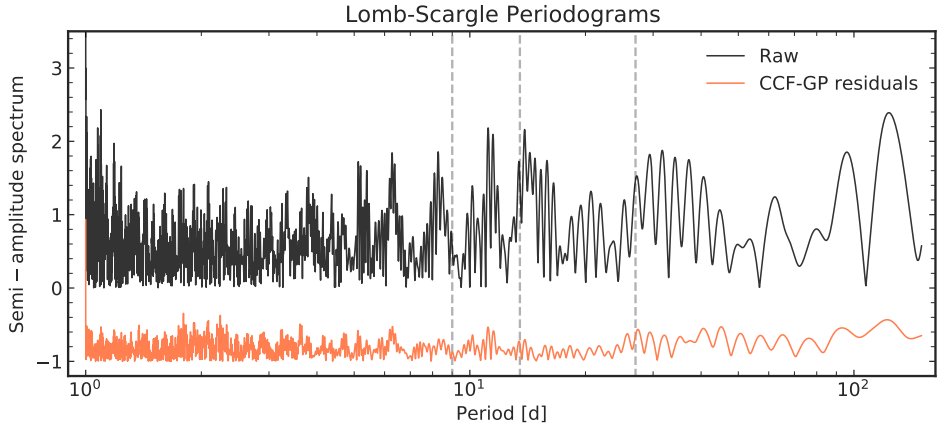


Figure 3.9: Lomb-Scargle periodogram of the raw solar RV and the RV residuals shown in Figure 3.8. The grey vertical markers indicate the positions of the 27-day solar-rotation signal and its harmonics.

1.0 m/s case, all the planet parameters are accurately recovered. For the 0.3 m/s case, while the period and phase are well-retrieved, the semi-amplitude starts to be overestimated, with a derived value of $K = 0.42^{+0.04}_{-0.04}$ m/s. This discrepancy can be attributed to multiple factors. As already discussed in Section 3.3.1, one plausible explanation is the overlooked non-Gaussianity of the CCF time series. Another factor is the presence of physical effects in the solar data that are not

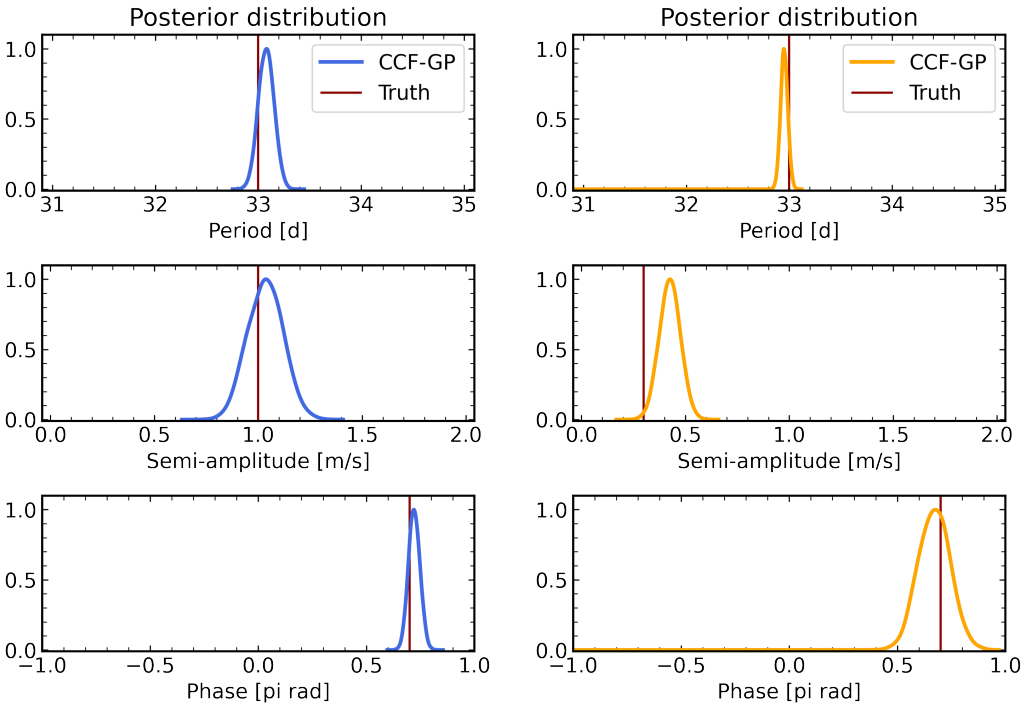


Figure 3.10: Posterior distribution of the extracted planetary parameters for the HARPS-N injection/recovery tests for planets with 33-d period and semi-amplitudes of 1.0 m/s (left, blue) or 0.3 m/s (right, orange). Vertical red lines indicate the true values from the injections.

accounted for in our model, such as the bisector shape alterations at active regions, and the effects of super-granulation, etc.

We also tested a case specifically targeting long-period planets. For this case, we utilized the same HARPS solar RV dataset as in the previous cases, but without accounting for the hypothetical star's observability constraints. The data contains 290 observations, with an average sampling rate of one data point per 1.7 days. This number of observations would take around four or five years to gather on the hypothetical star considered in the previous test. Once a full decade of HARPS-N solar data becomes publicly available, we will be able to simulate observations from a survey such as the Terra Hunting Experiment (Hall et al., 2018) more realistically.

For this test, we injected a planet with an orbital period of 101.2 days and a semi-amplitude of 0.5 m/s, corresponding to a planet of 3.6 Earth masses. The corresponding raw and planet RVs are shown in the left panel of Figure 3.11. The

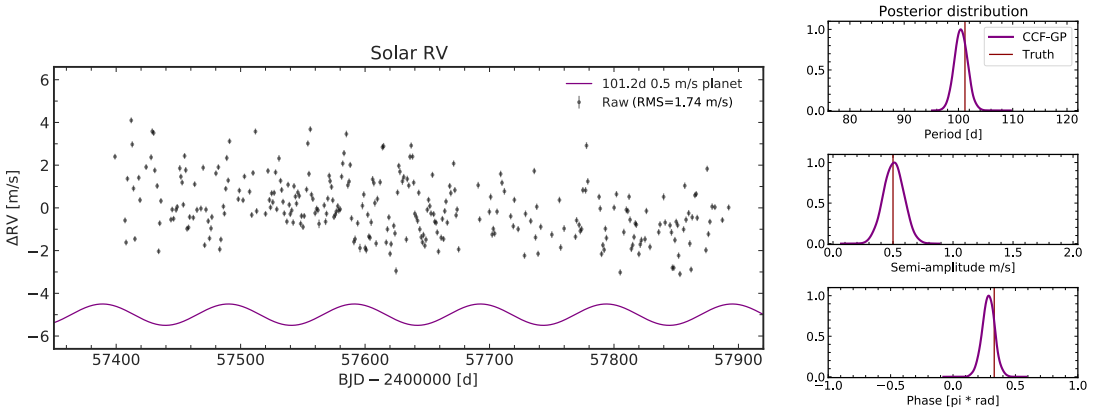


Figure 3.11: As Fig. 3.8 and 3.10, but the figure shows the injection/recovery test of the long-period planet at 101.2-d orbital period, and of a semi-amplitude of 0.5 m/s.

results of applying the CCF-GP model, as shown in the right panel of the same figure, reveal that the recovered planet parameters align well with the true values.

3.3.3 Application to real observations

We applied the CCF-GP framework to HARPS observations of HD 13808 to demonstrate its efficacy on real-world observations with an average SNR of ∼100. HD 13808 is a system consisting of a K2V dwarf star orbited by two confirmed non-transiting Neptune-mass planets (Ahrer et al., 2021). The data, which are queried from the ESO archive, include 246 spectroscopic measurements taken with HARPS, mostly between 2005 and 2016. The data underwent pre-processing via the automated DRS pipeline. Figure 3.12 presents the series of reference-subtracted CCFs and the associated SNR per pixel (in echelle order 60). The SNR for these observations varies between approximately 50 and 180, with a median value of around 100. We applied the CCF-GP framework, incorporating a two-planet RV model. As in previous tests, all priors on free parameters remained uninformative, implying that this was a blind detection test.

In Figure 3.13, the posterior distributions of the extracted planetary parameters are displayed. For real observations, a definitive 'true' value remains elusive. Therefore, we compared the extracted values with those reported in two recent publications, Ahrer et al. (2021) and Delisle et al. (2022). Delisle et al. (2022)

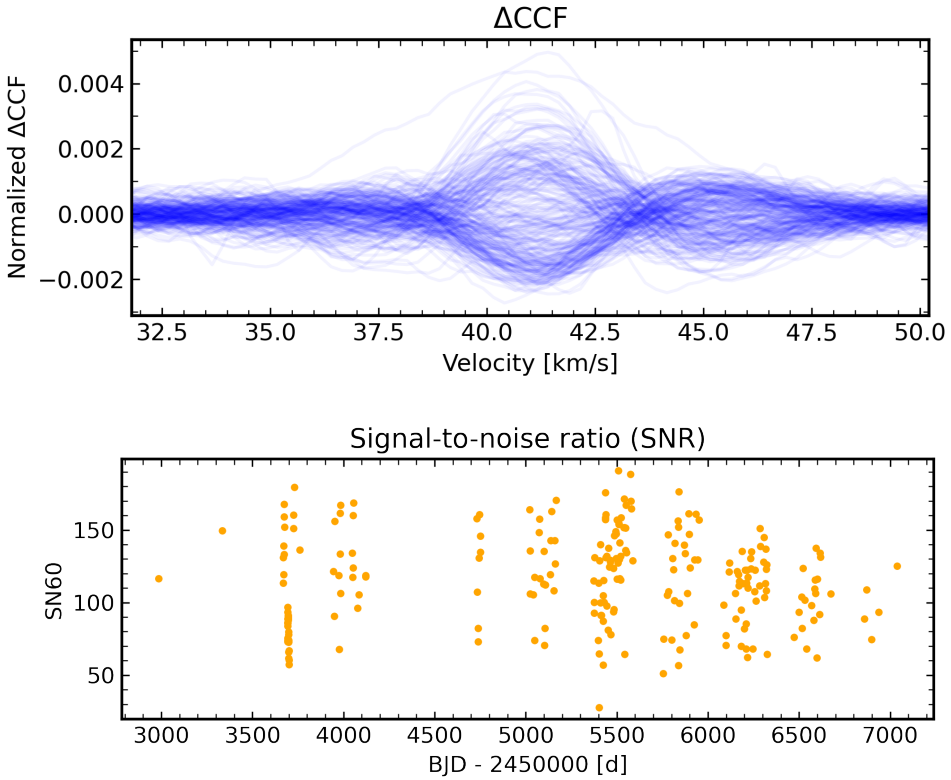


Figure 3.12: Reference-subtracted CCFs and the associated SNR per pixel (in échelle order 60) plotted over time. The SNR for these observations varies between approximately 50 and 180, with a median value of 100.

implemented a modified version of the multi-GP with a fast MEP kernel. Their best-fit values for the planet parameters are indicated by the orange markers. Conversely, Ahrer et al. (2021) evaluated various models for activity, endorsing that the harmonic activity model (model 1 in Figure 3.13), which simultaneously fits the 1st, 2nd, and 3rd harmonic of rotation period to the BIS and RVs, and the multi-GP model (model 2) as two of the most robust models; their results are indicated by purple and coral markers. The derived semi-amplitudes for planets b and c are $K_{\rm b} = 3.60^{+0.12}_{-0.11}$ m/s and $K_{\rm c} = 1.94^{+0.12}_{-0.12}$ m/s, which are within one sigma deviation from the values reported in Ahrer et al. (2021) and Delisle et al. (2022). This example demonstrates that the CCF-GP framework can effectively work with data of an averaged SNR of around 100, accommodating the majority of current observations and producing results consistent with existing approaches.

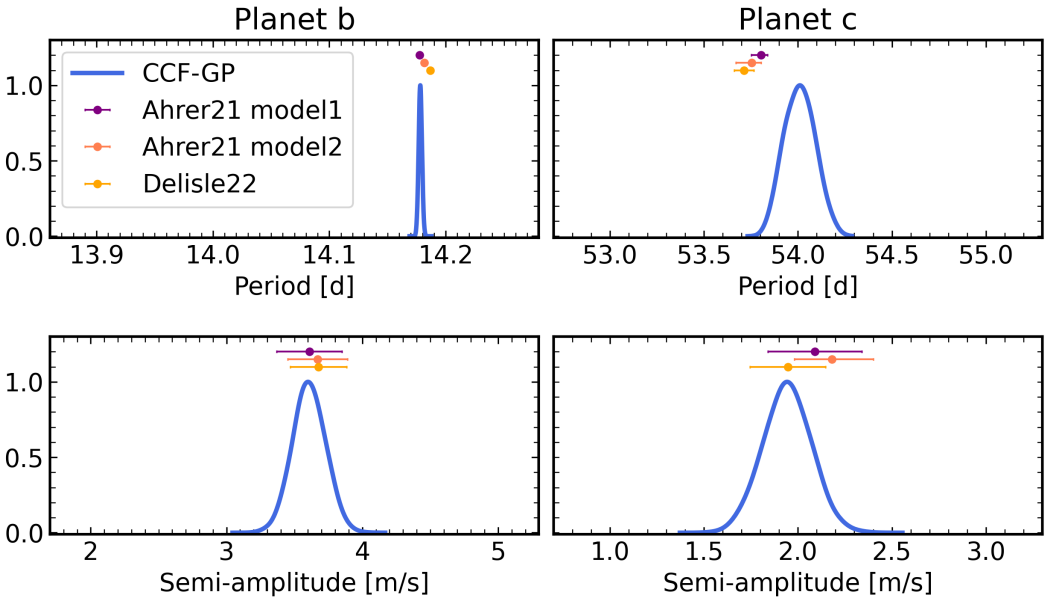


Figure 3.13: Posterior distributions of the derived planetary parameters from the CCF-GP analysis on the HD 13808 data. The derived values and associated uncertainties of the corresponding parameters from Ahrer et al. (2021) and Delisle et al. (2022) are indicated by the markers with errorbars.

3.4 Discussion

The CCF-GP model has demonstrated strong performance in disentangling stellar activity from planetary signals, achieving more reliable recovery of planetary parameters compared to previous techniques such as multi-GP approaches. This improvement arises because stellar activity and planetary signals are less degenerate in the two-dimensional line-profile time series than in one-dimensional time series of RV or activity indicators. In contrast to other methods that operate directly on the full spectrum or line-profile time series, the CCF-GP framework also incorporates temporal information through a GP, which increases its robustness when dealing with realistic data of relatively low S/N.

There are still several limitations, including a tendency to overestimate amplitudes and the occasional occurrence of false positives in tests. These issues stem from the current model's simplifications and inherent limitations in its present implementation.

The physical model we have outlined in Section 3.2.2 describes the perturbation to the profile of a single absorption line caused by a single active region crossing the stellar disk, and is in itself an approximation. Furthermore, there may be multiple active regions present at the same time at different locations on the stellar disk. The ratio of the a and b coefficients in Equation 3.10 depends on the relative area and contrast spot and plage, which can differ between active regions, while the local redshift caused by plage (δV_c in Equation 3.5) is likely to depend on both the strength of the magnetic field in the plage and the viewing angle. Thus, treating the α and β coefficients in Equation 3.10 as independent of time is not strictly correct.

Another aspect that we have overlooked in our physically motivated model is that the observed line profile is not equal to the true line profile, but is convolved with the instrumental line profile. For instance, when we incorporate the distortion of the line profile ΔC into the reference profile $c_{\rm ref}$, both profiles have already been convolved with the instrumental profile. This can be slightly biased from reality when the two unconvolved components are first summed, and the convolution is performed after. To some extent, we expect these imperfections in our physical model to be absorbed by the flexible latent GP formulation we have adopted in Section 3.2.2. Conversely, this means that the physical interpretation of the model becomes less tractable.

In Section 5.3, I discuss potential future developments aimed at addressing these limitations.

3.5 Conclusions

In this chapter, we introduced the CCF-GP framework, a novel approach developed to model stellar activity signals within time-series of CCFs using GPs. This model addresses the distortions in CCFs due to activity regions on the stellar surface, i.e., spots and faculae, stemming from both the photometric effect and the inhibition of convective blueshift effect.

Upon testing on both synthetic and real datasets – with SNRs down to around 100 – the framework showcased its robustness in distinguishing between activity and planetary signals, even when they share identical periods. Planet injection/recovery

tests on the realistically sampled HARPS-N solar data highlight the framework's potential to push the detection limit of planetary signals down to $K = 0.3$ m/s at $P = 33$ d (corresponding to a planet of 1.5-Earth mass) during the Sun's high activity phase in its magnetic cycle. The framework surpasses the performance of multi-GP models on RV and activity indicators in the tests conducted. We note that similar to the multi-GP approach, CCF-GP also requires adequate temporal sampling relative to the rotation period of the star for optimal performance.

4

A possible misaligned orbit for the young planet AU Mic c

Contents

4.1 Introduction 88
4.2 Observations and data reduction 92
4.2.1 CHEOPS photometry 92
4.2.2 NGTS photometry 94
4.2.3 ESPRESSO spectroscopy 95
4.3 Flare mitigation 97
4.3.1 Identifying and masking flares in the CHEOPS lightcurve 97
4.3.2 Extracting a flare proxy from the ESPRESSO spectra . 97
4.3.3 Extracting flare-resistant CCFs and RVs 99
4.4 Joint modelling of the photometry and RVs 100
4.4.1 The model 100
4.4.2 Results 101
4.5 Discussion and conclusions 102
4.5.1 Limitations of our measurement 107
4.5.2 The importance of joint modelling 107
4.5.3 Dynamic stability and transit probability 108
4.5.4 Possible explanations for a misaligned orbit 110

4.1 Introduction

Transiting planets around young stars are useful laboratories for studying the formation and evolution of planetary systems. This is particularly true for systems younger than 100 Myr. At such young ages, the key processes shaping the final system are at their strongest, including interactions between the planets, the disk they form in, and the host star (Burrows et al., 1997; Mordasini et al., 2012; Owen, 2019).

In particular, information about the early dynamical history of the system is encoded in the orientation of the orbits of young planets and the spin of the host star (Winn and Fabrycky, 2015). Planets might undergo scattering events, collisions (i.e., giant impacts), or secular perturbations after their formation within the protoplanetary disk, which can lift their orbital inclinations (Chatterjee et al., 2008; Wu and Murray, 2003). Subsequently, the orbital plane might be altered by tidal dissipation (see, e.g., Bolmont and Mathis, 2016). Obliquity measurements for young planets with well-measured ages are vital to constrain the key timescales for these mechanisms. Specifically, measurements for systems that are too young to have undergone significant tidal alteration of obliquity allow us to gain insight into their initial configurations (Mantovan et al., 2024). However, these measurements are still scarce. Only nine such measurements have been published to date (see Albrecht et al., 2022b, for a review). Some practical obstacles to performing these measurements are the rapid rotation and intense magnetic activity of the host stars, which severely hamper the detection and characterisation of young planets.

The AU Microscopii system illustrates these points. It is one of the most interesting young multi-planet systems known to exist, and is favourable for observations because it is only 9.7 pc away (Gaia Collaboration et al., 2018c) and is one of the brightest M dwarfs in the sky ($V = 8.7$; Torres et al. 2006). As a member of the Beta Pictoris moving group (Barrado y Navascués et al., 1999), it has a well-constrained age of $24.3^{+0.3}_{-0.3}$ Myr (e.g., Malo et al., 2014; Mamajek and Bell, 2014; Messina et al., 2016; Miret-Roig et al., 2020; Galindo-Guil et al., 2022)

and has long been known to host an edge-on debris disk (Kalas et al., 2004), in which fast-moving structures of unclear origin have been imaged with SPHERE on the Very Large Telescope and *Hubble* Space Telescope (HST) (Boccaletti et al., 2015; Boccaletti et al., 2018).

However, the star's photometry and RV have been intensively monitored and found to be highly variable and complex. The light curve displays quasi-periodic brightness fluctuations that indicate a 4.85-day rotation period (Rodono et al., 1986; Hebb et al., 2007), along with frequent flares (Robinson et al., 2001). The spin of the star appears to be closely aligned with the debris disk, based on the comparison between the inclinations of the stellar rotation axis, $i_{\rm star} = 87.3^{+1.9}_{-2.8}$ degrees, and the disk axis, $i_{\rm disk} = 89.4^{+0.1}_{-0.1}$ degrees (Hurt and MacGregor, 2023).

Two close-in transiting Neptune-sized planets were discovered around AU Mic using photometry from the Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2015). The planets were reported by Plavchan et al., 2020 and Martioli et al., 2021 with orbital periods of $P_b = 8.46\,$d and $P_c = 18.86\,$d, respectively. Ongoing monitoring of the transits using the TESS, Spitzer (Werner et al., 2004) and CHaracterising ExOPlanet Satellite (CHEOPS; Benz et al., 2021) space telescopes, and ground-based observatories, revealed significant transit timing variations (TTVs) in the system (Szabó et al., 2021; Szabó et al., 2022; Wittrock et al., 2022; Wittrock et al., 2023). To explain the observed TTVs, Wittrock et al. (2023) invoked an additional, non-transiting planet "d" located between planets b and c, with a mass comparable to the Earth and an orbital period of $P_d = 12.73\,$d.

Donati et al. (2023) reported a candidate non-transiting planet, designated planet e, orbiting further out in the system, based on near-infrared RV observations obtained with the Spectro-Polarimetre Infra-Rouge (SPIRou) (Artigau et al., 2014) at the 3.6-m Canada–France–Hawaii Telescope (CFHT). This planet may also contribute to the observed TTVs. The detection was later confirmed by Donati et al. (2025) based on six years of SPIRou monitoring, which yielded a refined orbital period of $33.11 \pm 0.06\,$days and a mass estimate of $21.1^{+5.4}_{-4.3}\,M_\oplus$. The study

also placed a 90% confidence upper limit of $4.9\,M_\oplus$ on the mass of the candidate planet d, which had previously been suggested based on TTVs.

Several studies have reported masses for the two transiting planets based on optical and near-infrared RV monitoring (Cale et al., 2021; Zicher et al., 2022b; Donati et al., 2023). The star's rapid rotation and strong magnetic activity induce RV variations that can reach up to 1 km/s in amplitude, dwarfing the planet-induced signals and making the resulting mass measurements highly dependent on the modelling of stellar activity. In this work we adopt the mass and radius estimates from Donati et al. (2023) of $R_b = 3.55 \pm 0.13\,R_\oplus$, $M_b = 10.2^{+3.9}_{-2.7}\,M_\oplus$, $R_c = 2.56 \pm 0.12\,R_\oplus$ and $M_c = 14.2^{+4.8}_{-3.5}\,M_\oplus$.

The fact that planet c is denser than planet b ($\rho_b = 1.26^{+0.68}_{-0.43}$ g cm^{-3}, $\rho_c = 4.7^{+2.5}_{-1.6}$ g cm^{-3}; Donati et al. 2023) suggests that planet c has a more massive solid core and a smaller H/He gaseous envelope (Zicher et al., 2022b). This can be in tension with the general theoretical expectation that a more massive solid core should be able to accrete a larger mass of gas from the protoplanetary disk (Owen and Wu, 2013). Possible explanations include: (a) planet c could be the result of a giant impact between two outer planets in the system, or (b) planets b and c could have formed in starkly different conditions or at different times within the same disk (Zicher et al., 2022b). In the latter case, for example, planet c may have formed after the gas disk had largely dissipated and therefore never reached the critical core mass in time. Measuring the orientation of AU Mic c's orbit might be helpful to differentiate between these scenarios.

When a transiting planet crosses over the disk of the host star, a portion of the blue or red-shifted starlight (due to stellar rotation) is blocked from the line of sight. This causes distortions of the disk-integrated stellar spectrum and an "anomalous" apparent radial velocity, which reflects the trajectory of the planet. This is the so-called Rossiter-McLaughlin (RM) effect, which allows for the measurement of the projected angle between the planet's orbit and the stellar rotation axis, also known as the projected spin-orbit angle (Queloz et al., 2000; Winn et al., 2005). Multiple studies (Palle et al., 2020; Martioli et al., 2021; Hirano et al., 2020; Addison et al.,

2021), have used RM observations to constrain the projected spin-orbit angle of AU Mic b to be consistent with zero ($\lambda_b = -2.96^{+10.44}_{-10.30}$; Palle et al. 2020).

In this chapter, we report an observing campaign aiming to make the first same type of measurement for planet c. A misaligned orbit for planet c would suggest that the planet has undergone significant dynamical interactions after the dispersal of the gas disk, thus favouring the giant-impact scenario (a), or another type of dynamical disruption. In contrast, an aligned orbit would indicate a more sedate dynamical history, compatible with scenario (b).

The measurement of the projected spin-orbit angle λ from RM observations is known to be affected by significant degeneracies with other parameters that are also related to the local radial velocity of the star along the transit chord. These include the time of the midpoint of the transit, the planetary orbital inclination i (or impact parameter, $b = a\cos i/R_\star$) and the star's projected rotation velocity $v \sin i_\star$ (e.g., Gaudi and Winn, 2007). Thanks to the wealth of existing space-based photometry and high-resolution spectroscopy, most of these parameters are relatively well-known for AU Mic and its transiting planets (e.g., $P_{\rm rot} = 4.856 \pm 0.003$ d and $v \sin i_\star = 8.5 \pm 0.2$ km/s; Donati et al. 2023). However, both planets display TTVs, with TTV amplitudes of 20–30 minutes reported for planet b, and 5–10 minutes for planet c (Szabó et al., 2021; Szabó et al., 2022; Wittrock et al., 2022; Wittrock et al., 2023). These TTVs make it challenging to predict transit midpoints precisely enough to plan and interpret our new spectroscopic observations; for planet c, the most recent published observations were conducted two years before our ESPRESSO observations. We therefore arranged for simultaneous photometric observations with CHEOPS and the Next-Generation Transit Survey (NGTS; Wheatley et al., 2018) to constrain the times of the transit midpoints.

In this chapter, we present the first measurement of the projected spin-orbit angle for the young sub-Neptune AU Mic c, using simultaneous observations with VLT/ESPRESSO, CHEOPS, and NGTS. The remainder of this chapter is structured as follows. We present the observations and basic data reduction steps in Section 4.2, and describe the steps we took to mitigate the impact of flares in Section 4.3.

We then present our joint model of the CHEOPS photometry and ESPRESSO RVs, and the resulting constraints on the obliquity of planet c, in Section 4.4. We discuss the limitations and potential implications of our results, and outline our key conclusions in Section 4.5.

4.2 Observations and data reduction

In this section, we present the observations used in this work and the basic data reduction steps.

4.2.1 CHEOPS photometry

We observed the two transits of AU Mic c that occurred on the nights of the 7$^{\text{th}}$ and the 26$^{\text{th}}$ of July 2023 (UTC) using the CHEOPS space telescope (Benz et al., 2021) simultaneously with the ESPRESSO observations, in order to refine the mid-transit times and characterise the photometric variability of the host star. CHEOPS's payload consists of a 32 cm diameter Ritchey-Chrétien telescope and a single-CCD camera covering the wavelength range 330–1100 nm with a field of view of 0.32 deg. Each transit was covered by a CHEOPS visit[1] of AU Mic, taken as part of the CHEOPS Guaranteed Time Observations (GTO) programme number CH_PR140071, see Table 4.1.

We extracted lightcurves using Point-Spread-Function (PSF) fitting on circular "imagettes" of 30-pixel radius centred on the target using the purpose-built software `PIPE`[2] (Szabó et al., 2021; Szabó et al., 2022; Brandeker et al., 2024), which are downloaded for each 5-second exposure. This approach yields similar photometric precision, but better time-sampling than the standard CHEOPS `Data Reduction Pipeline` (`DRP`)[3]. The latter performs instead aperture photometry on a 200 × 200-pixel subarrays of the full detector, that, due to their larger size, must be co-added in batches of 7 exposures before downloading. The top panel of Figure 4.1 shows

[1] A visit is a sequence of successive CHEOPS orbits about the Earth that are devoted to observing a single target. CHEOPS revolves around the Earth every ∼99 minutes in a Sun-synchronous, low-Earth orbit (700 km altitude).

[2] See https://github.com/alphapsa/PIPE

[3] Tested using version 14.1.3 of the DRP (Hoyer et al., 2020).

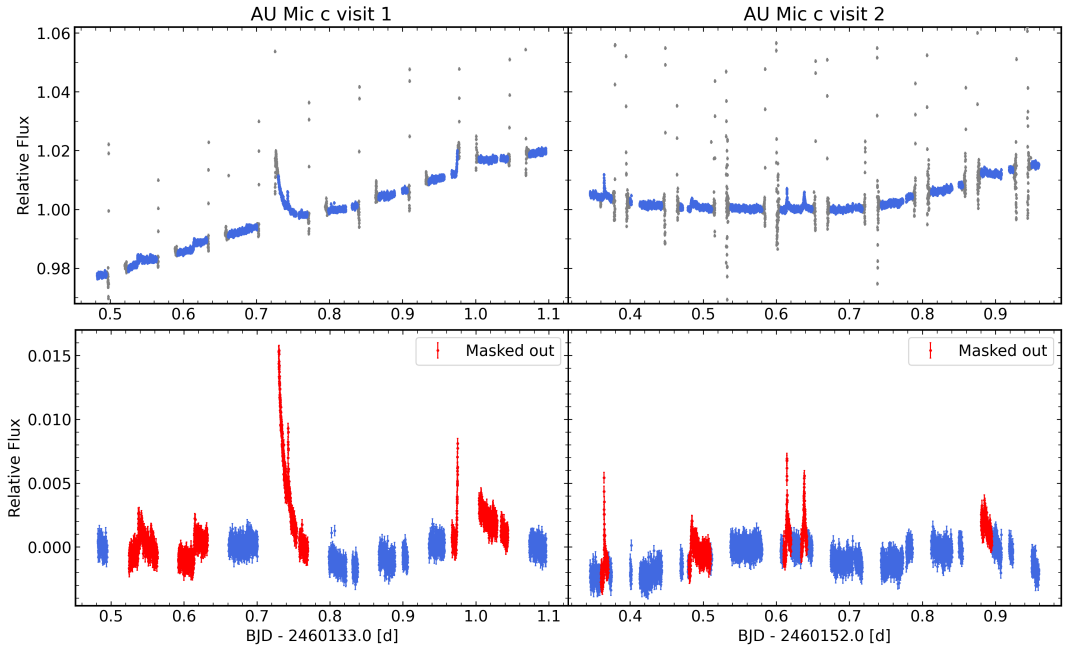


Figure 4.1: CHEOPS photometric observations of the AU Mic c transits. The left and right column show the first and second visit, respectively. The top row shows the lightcurves as extracted by the `PIPE` software after normalising each visit to unity, with the outliers caused by enhanced straylight at the start and end of each orbit marked in grey. The second row shows the data after subtracting a low-order polynomial trend (for visualisation purposes only), with the observations affected by flares highlighted in red.

Table 4.1: Log of the CHEOPS photometric observations of AU Mic c used in this work.

Visit No.	Start date [UTC]	End date [UTC]	File key	CHEOPS product	Integration time [s]	Number of frames
1	2023-07-07 23:25	2023-07-08 14:09	`CH_PR140071_TG000901`	Subarray	7×5.0	919
1	2023-07-07 23:25	2023-07-08 14:09	`CH_PR140071_TG000901`	Imagettes	5.0	6433
2	2023-07-26 20:08	2023-07-27 10:51	`CH_PR140071_TG000602`	Subarray	7×5.0	1052
2	2023-07-26 20:08	2023-07-27 10:51	`CH_PR140071_TG000602`	Imagettes	5.0	7364

CHEOPS observations log. The table lists the time interval of individual observations (time notation follows the ISO-8601 convention), the file key, which can be used to locate the observations in the CHEOPS archive, the type of the photometric product, the integration time, and the total number of frames in each visit.

the `PIPE` CHEOPS lightcurve after normalising each visit to unity. The average uncertainty of each photometric data point in the `PIPE` lightcurve, corresponding to a 5-second integration, is $\sim$ 400 ppm.

Outliers are visible just before and after each interruption within a given visit[4], caused by the increased background associated with scattered light from Earth. To

[4]These interruptions happen once per CHEOPS orbit when the target is occulted by the Earth, or it is barely visible due to stray light from the illuminated Earth limb or particle hits during passages above the South Atlantic anomaly.

remove these outliers, we masked all data taken when the background flux measured by the PIPE software exceeded 20 electrons/pixel. Outliers identified in this way are marked in grey in the top panel of Figure 4.1. This resulted in the removal of 1062 out of 6433 frames (16.5%) for the case of the first visit, and 1592 out of 7364 (21.6%) for the second, leaving a total of 11143 data points for the remainder of the analysis.

4.2.2 NGTS photometry

The NGTS is a ground-based photometric facility consisting of twelve independently steerable 0.2 m diameter robotic telescopes, located at the ESO Paranal Observatory in Chile. AU Mic was observed by NGTS simultaneously with the ESPRESSO observations of the two transits of AU Mic c. The observations were performed using the NGTS multi-telescope mode in which multiple NGTS telescopes are used to simultaneously observe the same star. This observing mode has been shown to significantly improve the photometric precision of NGTS data (e.g. Smith et al., 2020; Bryant et al., 2020).

Nine telescopes were used on the night of the 7$^{\text{th}}$ of July 2023, and ten on the night of of the 26$^{\text{th}}$ of July 2023. For both nights, the observations were performed using the custom NGTS filter (520–890 nm) and an exposure time of 5 seconds. The observations were performed with the telescopes heavily defocussed in order to avoid saturation and reduce effects of per-pixel sensitivity systematics and X-Y systematics (Southworth et al., 2009a; Southworth et al., 2009b; Nascimbeni et al., 2011). A custom aperture photometry pipeline was used to reduce the NGTS data. This pipeline uses the SEP Python library (Bertin and Arnouts, 1996; Barbary, 2016) to perform the source and photometry and also utilises *Gaia* DR2 information (Gaia Collaboration et al., 2018a) to identify comparison stars similar to AU Mic in terms of brightness, colour, and CCD position.

Both nights of observations were severely impacted by clouds, which caused large spurious fluctuations in the photometry (see top row of Figure B.2). The data from the second transit also show ramps at the beginning and end of the night, which are most likely systematic trends due to a lack of insufficient high

quality comparison stars from these observations for such a bright and red star. Even after masking out the most severely affected data (middle row of Figure B.2), the transit signals cannot be easily distinguished from the strong out-of-transit trend and residual short-term systematics. Consequently, the NGTS data were not used in our final analysis, though we did compare them to our fitted model as a consistency check (see Appendix B.2 for details).

4.2.3 ESPRESSO spectroscopy

We observed two transits of AU Mic c during the nights of July 7$^{\text{th}}$ 2023 and July 26$^{\text{th}}$ 2023, using the high precision spectrograph ESPRESSO (Pepe et al., 2014; Pepe et al., 2021b) mounted on the Very Large Telescope (VLT) (Program ID: 111.24VF, PI: H. Yu). For both runs, we used one unit telescope and employed the high-resolution mode (HR; R $\sim$ 13800) and the 2x1/SLOW binning/readout mode. Each exposure lasted 300 seconds. On the first night, 85 spectra were obtained, and on the second night, 66 spectra were obtained, with mean signal-to-noise ratios (SNRs) at 550 nm of 110.2 and 115.3, respectively. During both runs, the seeing ranged between 0.7 and 1.4 arc seconds. During the second run, there was a problem with the control unit of the UT2 telescopes that interrupted the observations at $\text{BJD}_{\text{TDB}} = 2460152.67160$. A decision was made to switch telescopes to UT1, and the observations resumed at $\text{BJD}_{\text{TDB}} = 2460152.73378$. UT1 was used for the remainder of the night.

The spectra were first reduced using the official ESPRESSO data reduction software (DRS) version 3.0.0 (Pepe et al., 2021b). The DRS data products include reduced 2-D and 1-D (order-merged) spectra as well as Cross-Correlation Functions (CCFs) and RVs extracted by fitting a Gaussian function to each CCFs (along with the other parameters of the Gaussian fit). A preliminary inspection of these products revealed that the spectra, CCFs and RVs were all severely affected by the flares and other stellar activity. We therefore decided to perform our own post-processing and RV extraction, as previous studies (e.g., Bourrier et al., 2021a) have shown that this can yield significant gains over the DRS data products.

We used the YARARA package (Cretignier et al., 2021b), to perform post-processing of the 1D order-merged (S1D) spectra in order to remove or reduce instrumental and telluric signals. As YARARA is a data-driven pipeline, it benefits from using as many observations of a given star with a given instrument as possible. We therefore also re-processed the transit of AU Mic b observed in 2019 with ESPRESSO (Program ID: 103.2010, PI: J. Strachan) and published by Palle et al., 2020. This has the added benefit of broadening the range of Barycentric Earth Radial Velocity (BERV) spanned by the observations, which helps with the telluric correction step of YARARA.

Before entering the YARARA pipeline, the spectra were continuum subtracted using the RASSINE package (Cretignier et al., 2020b). Outliers due to cosmic rays or dead pixels were then masked out via sigma clipping and replaced by the time-median value. The spectra were then corrected for contamination by microtelluric lines, as described in Cretignier et al. (2021b). The resulting, post-processed S1D spectra were used in the remainder of the analysis.

After combining all the spectra into a single, master-spectrum, we used this to construct a list of non-blended absorption lines (Cretignier et al., 2020b). The so-called KITCAT line-list was used to compute a CCF for each spectrum, which contains in total around 5000 lines, from which RVs were extracted. However, the resulting RVs were severely affected by flares, so we later re-extracted the RVs using only the subset of lines least sensitive to the flares (see Section 4.3.3).

It is worth noting that the ESPRESSO and NGTS observations were taken simultaneously, and the two telescopes are only a few km apart. The same weather pattern that affected the NGTS observations likely also impacted the ESPRESSO spectra. Specifically, cloud cover is expected to affect primarily the overall light level, without significantly altering the spectral shape, and should therefore have a negligible impact on the extracted CCFs and RVs.

4.3 Flare mitigation

Several flares occurred during and around each of the two observed transits of AU Mic c. Together with the longer-term trends caused by star-spots, they dominate both the CHEOPS photometry and ESPRESSO RVs. We therefore adopted the following procedure to mitigate their impact on the final analysis.

4.3.1 Identifying and masking flares in the CHEOPS lightcurve

The impact of the flares on the CHEOPS lightcurves is easier to see after subtracting a second-order polynomial fit, intended to represent the long-term trend due to active regions rotating on the stellar surface (see bottom row of Figure 4.1).

While it would in principle be possible to model the flares, this would result in a very large number of free parameters: each flare would require a minimum of three parameters (i.e., the occurrence time, amplitude, and decay rate), and there are at least 5 flares per visit. We therefore opted to mask out the data that were clearly affected by flares (with a typical minimum equivalent duration of 10 minutes) before modelling the spot and transit signals. The data points affected by flares were identified by visual inspection, cross-checking against the flare proxy derived from the ESPRESSO spectra (see Section 4.3.2) whenever the two time-series overlap, and are marked in red in Figure 4.1. Although the onset of a flare is relatively straightforward to identify, the tail end of a flare is much less well defined, so we opted to exclude data after each flare up to the end of the affected CHEOPS orbit.

4.3.2 Extracting a flare proxy from the ESPRESSO spectra

We first characterized the flare signal in the `YARARA-processed` S1D spectra. The flares manifest mainly as emission lines that decay both in width and intensity after the flare impulse on a timescale of around an hour. We started by selecting two spectra occurring right before and after the start of the largest flare (occurring just before the ingress of the first transit of AU Mic c, on the night of the 7^{th} of July 2023) and computing the difference between them. We then visually inspected this "flare spectrum" in order to identify atomic line transitions and selected a total

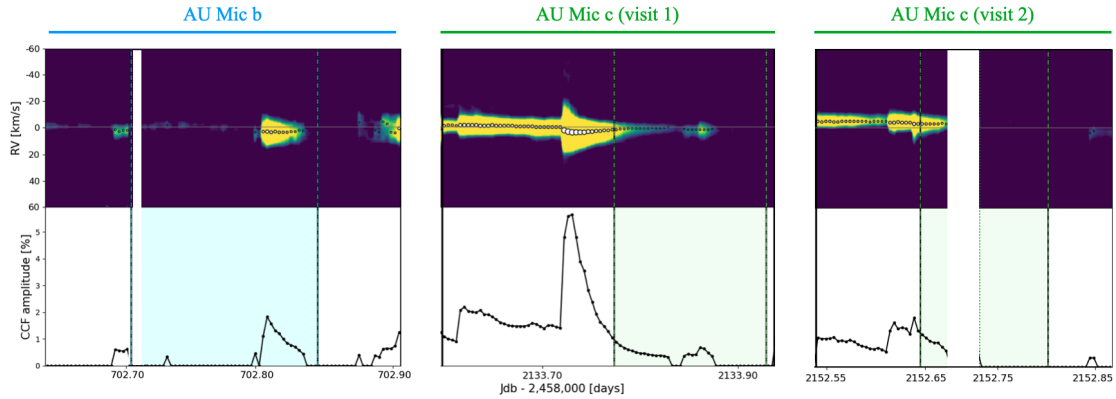


Figure 4.2: Extraction of the flare proxy. The top panel shows the time-series of "flare CCFs" (see text for details), from left to right: the 2021 transit of AU Mic b, then the two transits of AU Mic c observed in 2023. The white markers indicate the mean of a Gaussian fit to each CCF, while their size reflects the amplitude (or contrast of this Gaussian), which is also displayed in the bottom panel, and consists of our flare proxy. The shaded area in each panel in the bottom row indicates the transits of planets b and c. The gap during the second transit of planet c corresponds to the interruption due to the change of the UT telescope.

of 33 well-defined lines between 4046 and 5434 Å. These lines were cross-matched with values in the National Institute for Standards and Technology (NIST) atomic spectra database[5] to determine their rest wavelengths. The resulting line-list was then used to construct a binary mask, which was cross-correlated with each of the observed spectra, after subtracting a reference spectrum. The latter was constructed separately for each season (2021 and 2023), by averaging a small number of spectra taken well away from any flare signals. The resulting "flare CCF" time-series for the three visits is shown in the top row of Figure 4.2. Each CCF in this time-series was fit with a Gaussian function, and the amplitude (or contrast) of this fit – shown in the bottom row of Figure 4.2 – constitutes our flare proxy.

In this figure, the massive flare that occurred in the middle of the night for visit 1 of AU Mic c is clearly visible. In comparison, this flare was 5 times larger and twice as long in duration as the ones that occurred for visit 2 or the visit of AU Mic b. Despite our attempts to mitigate it, the flare occurring in this first visit was too strong and made the production of out-of-transit products impossible. It can also be noted that the RVs associated with flares were generally between −5

[5] https://www.nist.gov/pml/atomic-spectra-database

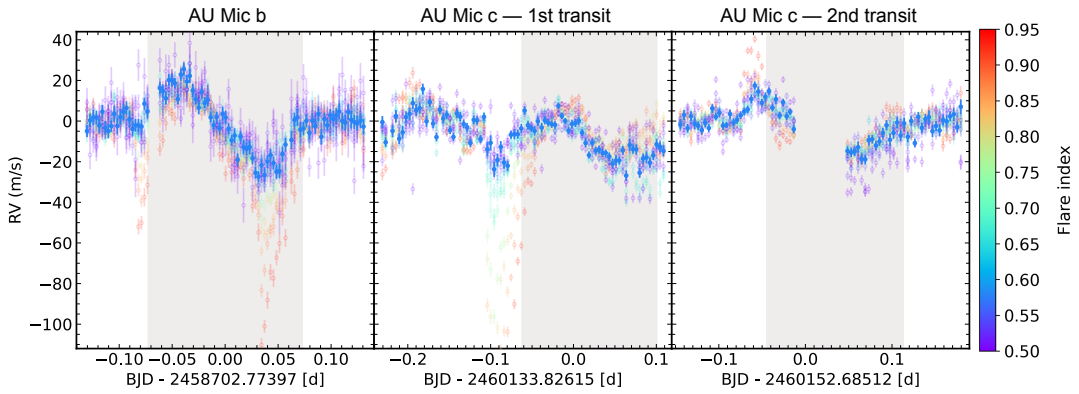


Figure 4.3: RV time-series extracted with progressively fewer flare sensitive lines (from red to blue) for the archival transit of planet b (left) and the two transits of AU Mic c (middle and right). The colour represents the mean flare index of all lines in the line list. Note that the flare just before the first transit of AU Mic c caused such a large RV perturbation in the RVs extracted with all the lines (red points) that it goes beyond the y-axis range shown here. A quadratic RV trend caused by rotational modulation of active regions has been fit to the out-of-transit data and subtracted prior to plotting. The grey shaded area in each panel indicates the duration of the transit.

and 5 km/s, likely due to the local radial velocity of the surface where they were produced and the extra radial velocity of the ejected matter.

4.3.3 Extracting flare-resistant CCFs and RVs

The CCFs and RVs produced using the `KITCAT` mask, shown in red on Figure 4.3, are heavily affected by the flares. We therefore sought to identify the subset of lines that are least sensitive to the flares while retaining sufficient RV information content. To quantify the flare-sensitivity of each line, we first modelled the time evolution of flux recorded in each pixel (wavelength) of the residual spectrum as the sum of a low-order polynomial (intended to account for spots on the rotating stellar surface) and a term proportional to the flare proxy derived in Section 4.3.2. We then defined a flare index that quantifies the sensitivity of each line to flares by computing the average of the correlation coefficient over a line profile in a given velocity window of $\pm$ 10 km s^{-1} that contains most of the Doppler information of a spectral line.

We used this coefficient to rank the lines in order of flare sensitivity, and computed CCFs and RVs with a progressively lower limiting sensitivity to flares until only the 250 least sensitive lines remained. While using fewer stellar lines

decreases the impact of the flares, it increases the photon noise. By examining the transit of AU Mic b, we found that the mask with 2000 lines (out of the total ∼5000 lines) gave the best trade-off between flare suppression and RV precision. The RVs extracted with this mask, shown in dark blue on Figure 4.3, were used in the remainder of the analysis.

The final RVs for the archival transit of planet b are very similar to those reported by Palle et al., 2020, which strengthens our confidence in our reduction and RV extraction methodology. The RV uncertainties are somewhat larger, though, because we are using fewer lines. The flare correction described above is effective for most of the archival and new ESPRESSO observations. However, careful examination of Figure 4.3 shows that the impact of the strongest flares on the RVs are not completely corrected when the flare proxy (displayed in the bottom row of Figure 4.2) exceeds a value of 3. We therefore excluded the RVs taken at those times from the remainder of our analysis.

4.4 Joint modelling of the photometry and RVs

We jointly modelled the CHEOPS lightcurve and the ESPRESSO flare-resistant RVs obtained during both transits of AU Mic c in order to derive constraints on the transit times and the projected orbital obliquity simultaneously, while accounting for the uncertainties arising from the star's intrinsic variability and instrumental noise.

4.4.1 The model

For both types of observation (photometry and RVs), our model consists of three components. The first component represents the planetary transits. We used the `batman` package (Kreidberg, 2015) to model the photometric transits, and the `ARoME` package (Boué et al., 2013) to model the RM effect in the RVs (this model is optimised for CCF-derived RVs). The transit models used for both transits and both types of observations share the same star and planet parameters, including: the sky-projected stellar obliquity (λ), inclination (i), scaled semi-major axis ($a/R_{\rm s}$),

planet-to-star radius ratio (R_p/R_s), and limb darkening coefficients (u_1 and u_2)[6]. To account for possible TTVs, each transit was modelled using a separate mid-transit time (T_c). We placed uninformative uniform priors on the sky-projected stellar obliquity as well as the mid-transit times, and Gaussian priors on all other orbital parameters based on the best-fit model from the transit analysis in Wittrock et al. (2023). The RM model also depends on the projected stellar rotational velocity, $v \sin i_\star$; we placed a Gaussian prior on this parameter based on the spectroscopic measurement of Donati et al. (2023).

The second component of our model consists of a second-order polynomial to account for the variations caused by active regions on the rotating stellar surface. Each dataset for each transit was modelled using a separate set of 3 polynomial coefficients. Finally, the third component is a Gaussian Process (GP) with a Matérn 5/2 kernel, which absorbs short-term variations, whether caused by the star's activity or by instrumental systematics. The GP hyperparameters were shared across the two transits, but differed between the lightcurves and RVs. We adopted uninformative (broad and uniform) priors for all parameters associated with the polynomial and GP components, except for the GP timescale ρ that represents how quickly the correlation between points decreases with distance, which was restricted to the range 10–60 minutes. All the priors used are reported in Table B.1. We used nested sampling with `PolyChord`[7](Handley et al., 2015a; Handley et al., 2015b) to explore the posterior distribution of the free parameters in the model.

4.4.2 Results

The best-fit models are shown over-plotted on the data for both the lightcurves and RVs in Figure 4.4. From top to bottom, the different rows of panels show: the input data with lines and contours corresponding to the full model (GP + polynomial + transits), the data and models after subtracting the polynomial component, and

[6]The flux-weighted mean wavelength of the customized CCF mask used for the ESPRESSO data is 624.4 nm. This closely aligns with the centre of mass of the CHEOPS passband, which ranges from 400 to 1100 nm. Therefore, we applied the same limb darkening coefficients to both the CHEOPS and ESPRESSO data.

[7]https://github.com/PolyChord/PolyChordLite

after subtracting both the polynomial and the GP components. We also show a phase-folded version of the two transits after subtracting the polynomial and GP components of the model in Figure 4.5. The posterior distributions for the key parameters of the fit are plotted Figure 4.6, while the priors and inferred 68.3% and 95% credible intervals in Table 4.2. The priors and inferred values of all free parameters are listed in Table B.1, and the full posterior distributions are shown in Figure B.1 in the appendix.

The best-fit model shown in the bottom-panel of Figure 4.4 and in Figure 4.5 displays the one-sided RM signature characteristic of a misaligned orbit (as opposed to a prograde or retrograde one), and the best-fit (maximum a-posteriori) value of the projected spin-orbit angle is $\lambda \approx 87$ degrees. However, the posterior distribution for λ is bimodal, with a second subsidiary maximum at $\lambda \approx 34$ degrees, and a width that appears to encompass 0 degrees. Formally, we measure $\lambda = 67.8^{+31.7}_{-49.0}$ degrees, where the central value represents the median and the uncertainties encompass the central 68.3% of the posterior distribution. After taking into account their weights from the nested sampling, 89% of the posterior samples correspond to somewhat misaligned orbits ($\lambda_c \geq 10$ degrees).

The bimodality in the projected spin-orbit angle is partly explained by its strong correlation with the mid-transit times, particularly for the second transit. The latter is itself multi-modal, owing to the gaps in the CHEOPS lightcurve, and this has a knock-on effect on the spin-orbit angle. The derived mid-transit times correspond to TTVs of 27 and 49 minutes relative to the published ephemerides – somewhat larger than the TTVs previously reported for planet c, though the uncertainties are also large ($\sim$ 15 minutes at 1-σ).

4.5 Discussion and conclusions

We have used simultaneous high-precision photometry from CHEOPS and precise RVs from ESPRESSO to measure the projected obliquity of AU Mic relative to the orbital plane of the young, dense sub-Neptune AU Mic c. While inconclusive, our results provide a tentative indication (89% confidence) that the orbit of AU Mic c

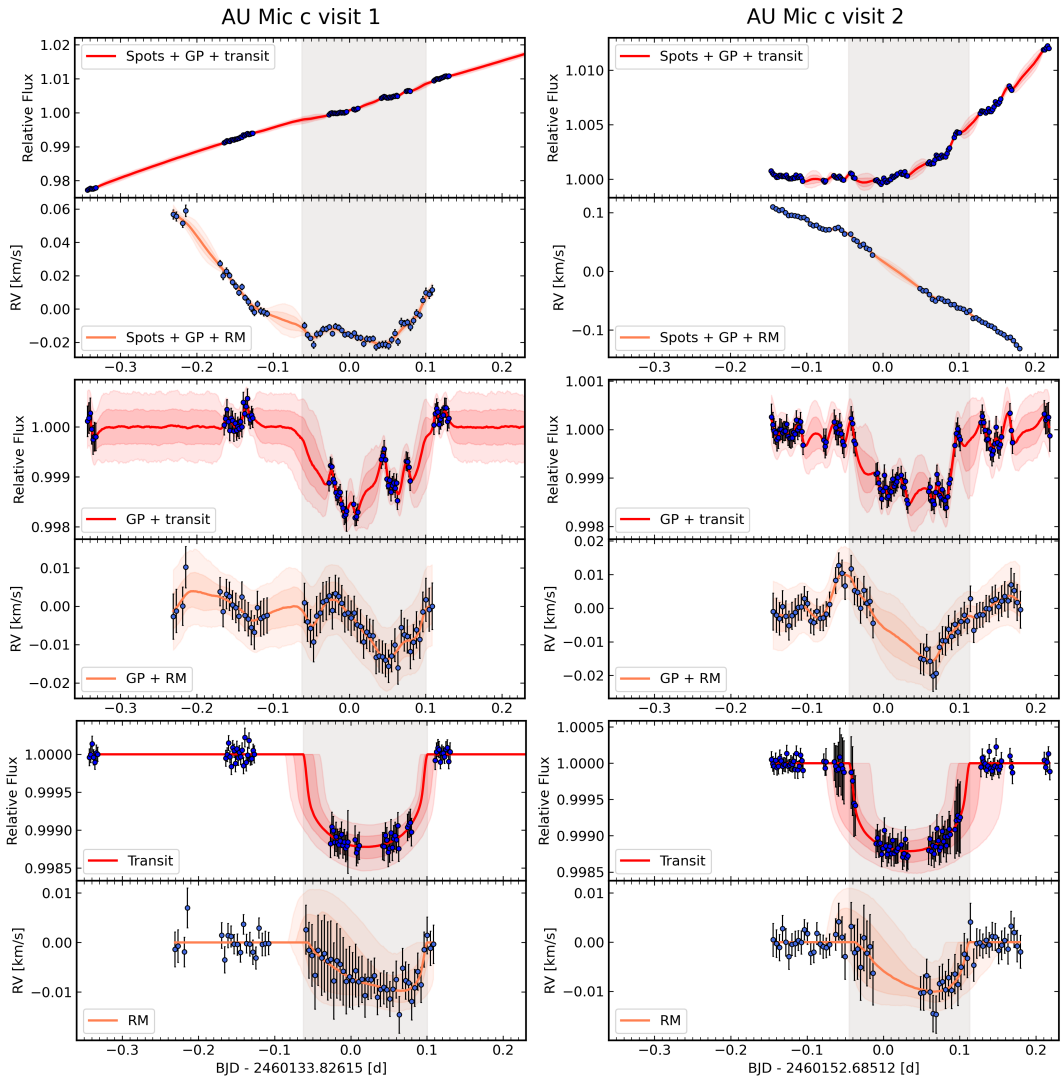


Figure 4.4: Joint analysis of the first and second transits of AU Mic c. The upper plots show the raw data with lines and contours corresponding to the full model (GP + polynomial + transits). The middle plots show the data and models subtracting the polynomial component. The lower plots show the model and data further subtracting the GP component. For each plot, the model is estimated through 10,000 randomly selected samples with corresponding weights generated from the nested sampling. The solid lines correspond to the weighted median of the samples, while the contours correspond to the 68.3% (which is equivalent to 1-σ for a Gaussian distribution) and 95% credible intervals (CIs) of the samples. Each of the samples obtained via nested sampling is associated with a weight, which was taken into account when evaluating the percentiles used to compute the credible intervals. In the middle and lower plots, the data points are the original data minus the weighted median of the samples for the components of the model subtracted, evaluated at the location of the observations, and their error bars are calculated as the quadrature sum of the original error bars and the weighted standard deviation of these samples. The solid lines and the surrounding shaded areas represent the median, 68.3% and 95% credible intervals of the transit model. The grey shaded area in each panel indicates the duration of the transit.

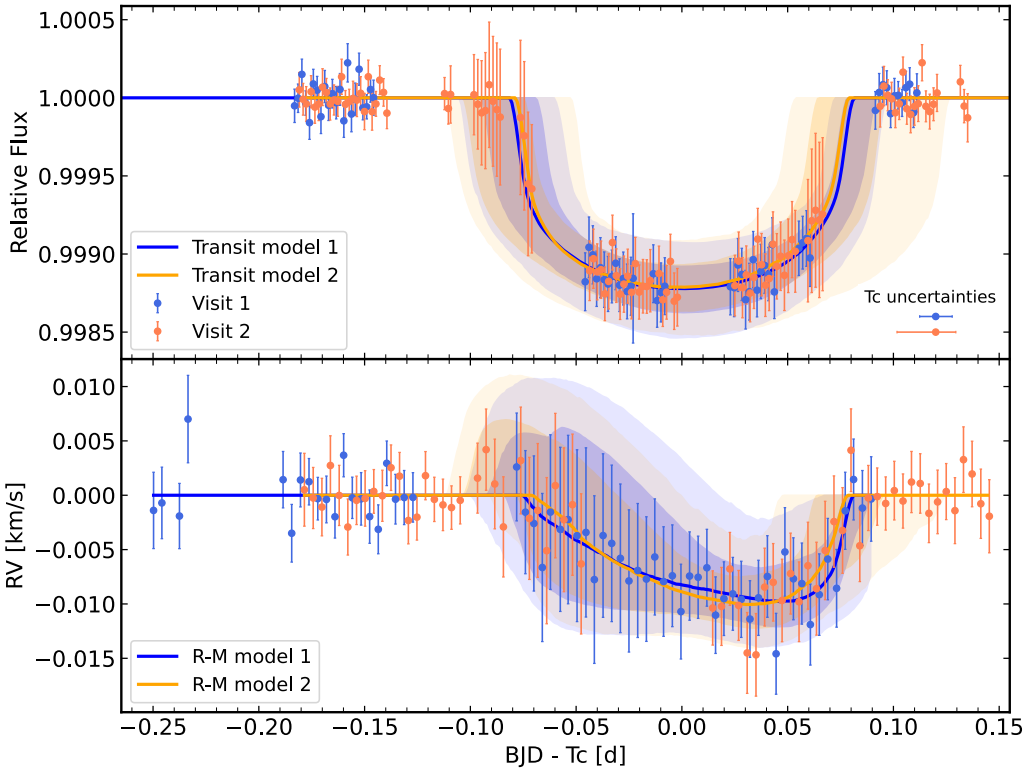


Figure 4.5: Phase-folded lightcurve (top) and RV (bottom) observations of the transits of AU Mic b (transit 1 in blue, transit 2 in orange). The polynomial and GP terms used to model the spot signals and residual short-term variations respectively have been subtracted from the data, and the vertical error bars incorporate the uncertainty in those model components. The solid line and shaded areas represent the median, 68.3% (1σ) and 95% ($\approx 2\sigma$) credible intervals of the transit component of the model. The horizontal errorbars shown in the top panel represent the 1σ uncertainties on mid-transit times Tc of each transit.

Table 4.2: Inferred values for the key parameters from the joint analysis of the first and second transits of AU Mic c. The full set of parameters included in the fit is reported in Table B.1.

Parameter	Value[(a)]	95% credible interval
λ_c (deg)	$67.8^{+31.7}_{-49.0}$	−23.2 – 120.7
i_c (deg)	$89.09^{+0.18}_{-0.17}$	88.77 – 89.47
$T_{c,1} - 2460133.82615$[(b)] (days)	$0.0191^{+0.0078}_{-0.0076}$	0.0046 – 0.0343
$T_{c,2} - 2460152.68512$[(b)] (days)	$0.0341^{+0.0095}_{-0.0183}$	0.0073 – 0.0703

[(a)] Median and 68.3% credible interval

[(b)] Transit times are quoted relative to the published ephemerides of (Wittrock et al., 2023), namely $Tc = 2458342.2240$ and $P_c = 18.858970$ d.

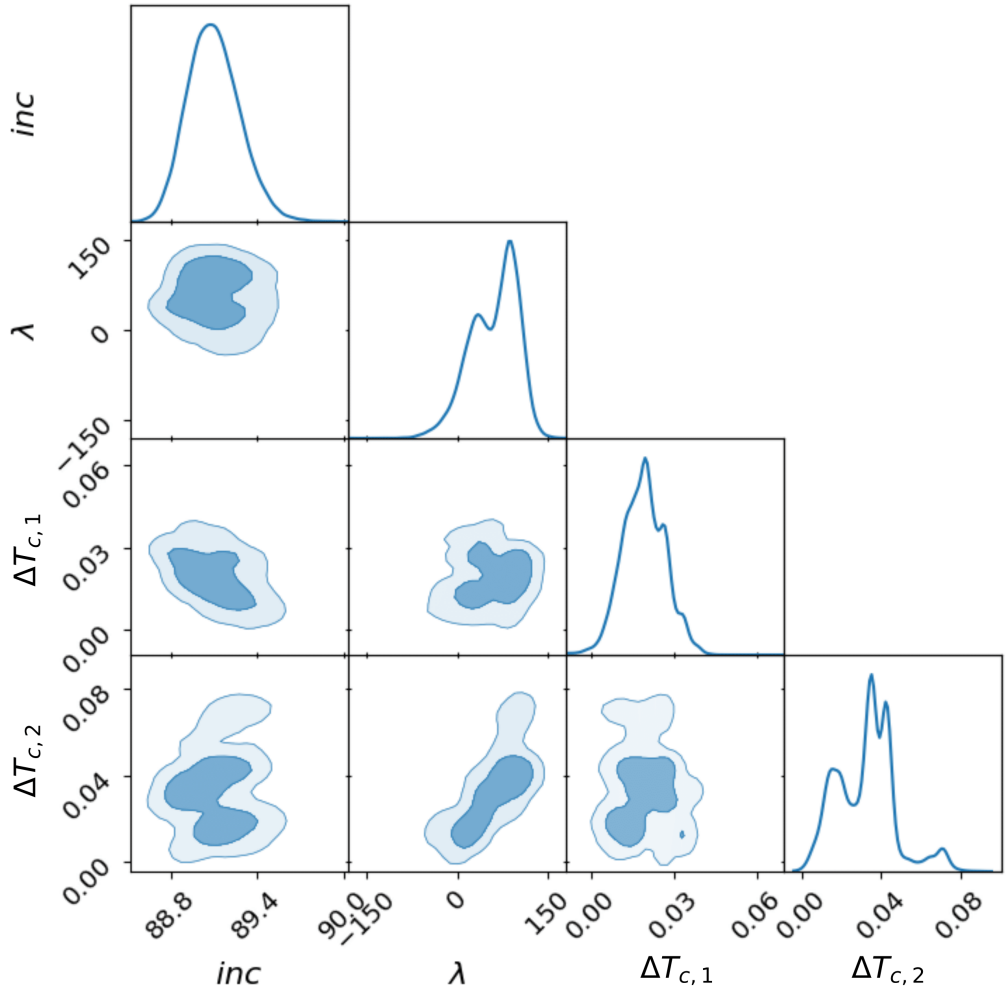


Figure 4.6: Posterior distribution of key parameters in the joint analysis of the first and second transits of AU Mic c.

is misaligned with the spin axis of its host star, and hence also with planet b's orbit and the debris-disk seen at wider separation. Figure 4.7 shows a schematic representation of the orbital configuration of planet c relative to the host star and the debris disk, based on the posterior samples from our joint fit of the CHEOPS lightcurve and ESPRESSO RVs.

If this result is confirmed, it will be the first observation of a young ($\leq$ 100 Myr) planet on a misaligned orbit. We note that, as the star is seen essentially equator-on, the true obliquity for both planets' orbits is within a few degrees of the sky-projected value. Only 7 other planets with ages $\leq$ 100 Myr have

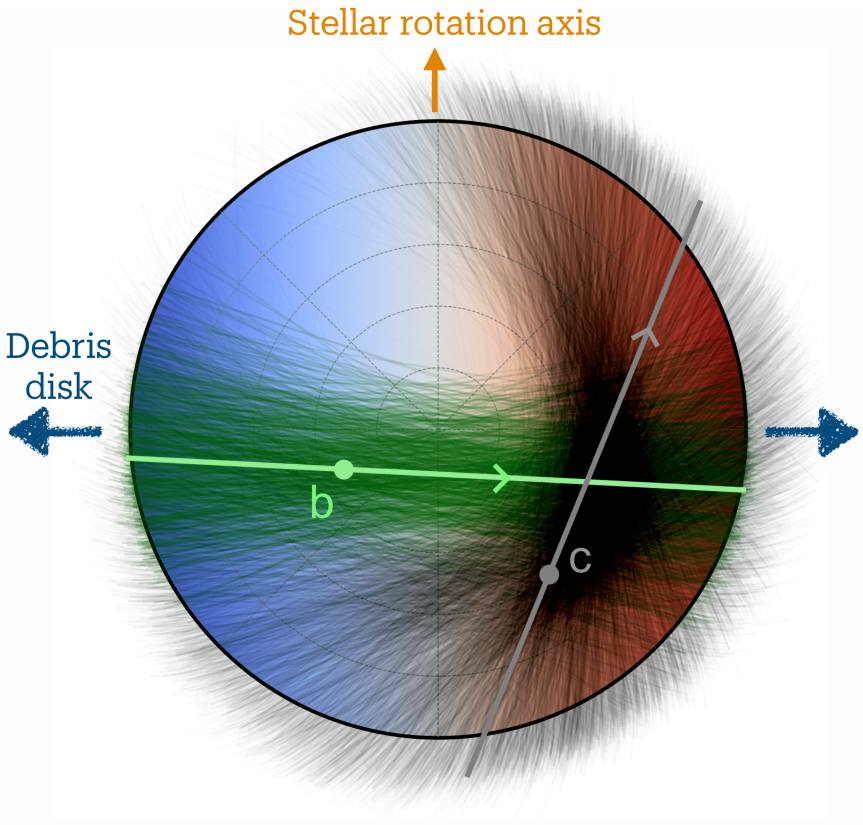


Figure 4.7: Schematic representation of the spin-orbit alignment of planets b and c within the AU Mic system. The stellar disk is colour-coded to represent local RVs, which vary as the star rotates, transitioning from blue-shifted on the left to red-shifted on the right. The bold green and grey lines represent the inferred orbits for planet b from Palle et al. (2020) and planet c from this work, with arrows indicating the directions of motion for the planets. Thin green and grey lines represent 50,000 randomly selected samples for the orbits of planets b and c, respectively. For the orbit of planet c, the transparency of the thin grey lines scales as the weight of each sample in the nested sampling. The orange arrow marks the axis of stellar rotation. The dark blue arrow marks the orientation of the debris disk.

measured obliquities to date. Moreover, very few systems with two transiting planets exhibit large mutual inclinations, irrespective of age. One comparable example is the HD 3167 system, where the close-in super-Earth appears to be on an aligned orbit (Bourrier et al., 2021b), while the outer mini-Neptune is on a nearly perpendicular orbit (Dalal et al., 2019).

4.5.1 Limitations of our measurement

The large uncertainties on our reported value of the projected spin-orbit angle λ for planet c preclude any claim of a significant misalignment at this stage (which would require at least a 3σ departure from $\lambda = 0$).

Degeneracies with other parameters, such as orbital inclination i and especially mid-transit times, are known to affect the measurement of the projected spin-orbit angle, λ (e.g., Narita et al., 2010; Triaud et al., 2010). This highlights the critical role played by CHEOPS in constraining the transit times. On the other hand, even with simultaneous photometry and RVs, we were not able to constrain the transit times as well as we might have hoped given the precision and cadence of our observations, and this had a knock-on effect on the obliquity measurements. This was partly because strong flares coincided with the ingress and egress of both transits, making key parts of the data unusable (in addition to the gaps in each CHEOPS orbit caused by the occultations of the target by the Earth).

4.5.2 The importance of joint modelling

Before we developed the joint model described in Section 4.4, our initial attempts at modelling the data involved fitting the CHEOPS lightcurve first, then using the posteriors from that fit as priors on the RV fit. When fitting the CHEOPS lightcurves alone, we were able to model the flares explicitly, rather than masking them out, resulting in somewhat tighter constraints on the mid-transit times. The subsequent fit to the ESPRESSO RVs was restricted to the second transit observation, as the lack of post-transit baseline precluded the use of the first. This fit to the second transit yielded a nominally much more significant detection of a misaligned orbit, with $\lambda = 108.7 \pm 12.5$ degrees. However, the fit of the CHEOPS lightcurves including flares implied anomalously large TTVs of order 50 minutes – much larger than the 5–10 minute TTVs reported for planet c to date (Szabó et al., 2021; Szabó et al., 2022; Wittrock et al., 2022; Wittrock et al., 2023). In addition, there was then a significant tension (around 2-σ) between the mid-transit time obtained from the second CHEOPS lightcurve fit and from the RV fit (despite the

use of the CHEOPS-based prior): the sequential approach does not guarantee a self-consistent fit between the lightcurve and the RVs. This motivated our decision to model both types of observations simultaneously, rather than sequentially, as reported in Section 4.4. Doing this required masking the flares, which had the effect of considerably increasing the uncertainty on the mid-transit times, and hence on the projected spin-orbit angle, but allowed us to include the ESPRESSO observations of the first transit into the fit. The result is less precise, but in our view more robust, than the sequential approach. Nonetheless, we note here that the sequential fit supports a significantly misaligned orbit.

Modelling the two observed transits jointly was also important, as the first ESPRESSO observation covered the full transit, but no post-transit baseline, while the gap in the second observation fell during the transit.

4.5.3 Dynamic stability and transit probability

The predicted sky-projected spin-orbit angle of planet c (Table 4.2) may correspond to a high mutual inclination between the two planets, which raises concerns about the dynamic stability and transit probability of the system.

Petrovich et al. (2020) suggests that large mutual inclinations could lead to eccentric instability, resulting in increased eccentricity, orbit crossings, and overall instability. Conversely, strong planet-planet interactions between planets b and c might stabilise the system, as indicated by Denham et al. (2019), particularly in the presence of general relativity (GR) precession. Due to the significant mutual inclination, planet c would induce a nodal precession on planet b, causing variations in planet b's impact parameter and leading to transit duration variations (TDVs) as well as the TTVs that have been observed. More accurate predictions of the TTVs and TDVs, requiring detailed N-body calculations, could provide a further test of the mutual inclination of the system.

To check if the system can be stable, we performed a dynamical analysis in a similar way as for other planetary systems (eg. Correia et al., 2005; Correia et al., 2010). The system is integrated on a regular 2D grid of initial conditions around

the best-fit solution from Table 5 in Wittrock et al. (2023). Each initial condition is integrated for 5 000 yr, using the symplectic integrator SABAC4 (Laskar and Robutel, 2001), with a step size of 5×10^{-4} yr and GR corrections. We then performed a frequency analysis (Laskar, 1990; Laskar, 1993) of the mean longitude of planet c over two consecutive time intervals of 2 500 yr, and determined the main frequency, n and n', respectively. The stability is measured by $\Delta = \log|1 - n'/n|$, the stability index, which estimates the chaotic diffusion of the orbits. A higher absolute value of the stability index indicates that the system is more stable. We note that the stability analysis here does not consider the impact of the debris disk.

In the lower panel of Figure 4.8, we present the stability of the system with varying the projected spin-orbit angle[8] and the eccentricity of planet c. The colour represents the stability index, with orange and red indicating strongly chaotic unstable trajectories. A stability index below -4 indicates that the system is stable on time scales around 20 Myr, corresponding to the current age of the AU Mic system. We observe that orbits with eccentricities higher than 0.3 or spin-orbit angles higher than 50 degrees are completely unstable.

In comparison, we show the posterior distribution of the measured sky-projected spin-orbit angle of AU Mic c λ_c in the upper panel of Figure 4.8. We adopted an eccentricity for planet c of $0.041^{+0.047}_{-0.026}$ measured from RV modelling in Zicher et al. (2022b). The posterior distribution of λ_c reveals a median value around 67 degrees and a maximum a-posteriori value around 87 degrees, which lies in an unstable region. However, stability is possible within the uncertainty of λ_c for lower values, especially around the second subsidiary maximum in the posterior distribution around 34 degrees.

To check the transit probability of the two planets, we used the semi-analytical transit probability model presented in Ragozzine and Holman (2010), Brakensiek and Ragozzine (2016) and Read et al. (2017). The input parameters of the model include the radius of the star and the semi-major axis of the two planets, for which we use values reported in Donati et al. (2023) and Wittrock et al. (2023). By assuming

[8]We assume $i_* \approx i_b \approx i_c \approx 90°$ (Wittrock et al., 2023), and thus $\lambda_c \approx \Omega$ (Fabrycky and Winn, 2009b), where λ_c is the spin-orbit angle and Ω is the longitude of the node of planet c.

planet b is on a well-aligned orbit, we estimated the double transit probability versus the projected spin-orbit angle of planet c, shown in the middle panel of Figure 4.8. Compared to the case where both planets are well-aligned, the transit probability is about 1/17 of that scenario when planet c is misaligned with $\lambda_c = 67.8^{+31.7}_{-49.0}$ degrees.

In summary, a significantly misaligned orbit for planet c is in some degree of tension with dynamical stability arguments, and with the fact that we see both planets in transit. While these arguments alone do not preclude such an orbit, they tend to favour the lower end of the the range of mutual inclinations allowed by our analysis and presented in Section 4.4. Ideally, repeat observations should be carried out to settle the matter.

4.5.4 Possible explanations for a misaligned orbit

Despite the tentative nature of our result, it is interesting to speculate as to the mechanisms that could lead to a misaligned orbit for AU Mic c, without disrupting the mutual alignment between the star, disk, and planet b.

One such mechanism is planet-planet scattering. Interactions between planets can lead to significant orbital changes. When these interactions drive relative velocities up to the escape velocity of the most massive body in the system, collisions or giant impacts become likely (Goldreich et al., 2004; Schlichting, 2014). Thus, these collisions typically occur at escape velocity and can result in inclination changes proportional to the ratio of the escape velocity to the local Keplerian velocity. Thus, an order of magnitude estimate of the inclination change is $\sim v_{\rm col}/v_{\rm Kep}$. Since $v_{\rm esc}/v_{\rm Kep} \sim 1/2$ then large inclination changes, similar to the one we have observed are plausible (e.g., Rasio and Ford, 1996; Weidenschilling and Marzari, 1996; Lin and Ida, 1997; Davies et al., 2014). In the event of a collision, such a scenario would likely result in planet c losing its atmosphere aligning with the abnormal density trend observed between planets b and c derived from mass measurements. However, a collision causing a large inclination is also likely to result in high eccentricity unless it occurred along a specific trajectory that resulted in high inclination and low eccentricity. Nonetheless, the eccentricity of AU Mic c is weakly constrained

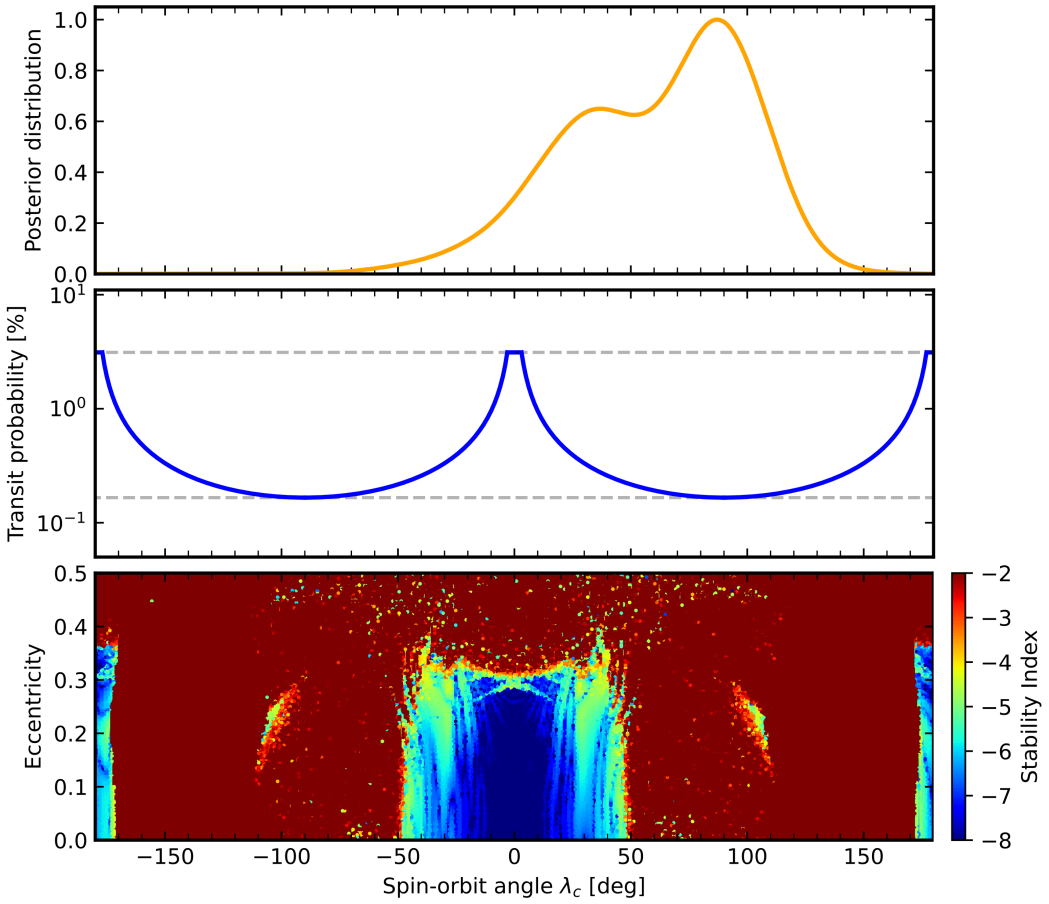


Figure 4.8: Upper panel: Posterior distribution of the sky-projected spin-orbit angle of AU Mic c λ_c. Middle panel: Transit probability versus λ_c, assuming AU Mic b is well aligned, estimated following Read et al. (2017). Lower panel: Stability analysis of the AU Mic planetary system. For fixed initial conditions (Table 5 in Wittrock et al. (2023)), the parameter space of the system is explored by varying the projected spin-orbit angle and the eccentricity of planet c. The step size is 0.9° in the spin-orbit angle and 0.0025 in the eccentricity. For each initial condition, the system is integrated over 5 000 yr and a stability indicator is calculated, which involves a frequency analysis of the mean longitude of planet c. The chaotic diffusion is measured by the variation in the main frequency (see text). Red and orange points correspond to highly unstable orbits.

and may still be relatively significant. Additionally, debris from such a collision could potentially form a disk, allowing the planet to circularise.

Another possibility is a nodal secular resonance, which can drive large misalignments (Ward et al., 1976; Lai, 2014; Owen and Lai, 2017). Given the possible presence of an exterior planetary companion (e.g., Wittrock et al., 2023; Donati et al., 2023), this mechanism is a viable explanation for the planet c's misalignment. In this scenario, the nodal precession frequency of the exterior companion is driven by a dispersing protoplanetary disk, which slowly evolves as the disc disperses. If, during this evolution, the frequency aligns with the nodal precession frequency of planet c, the exterior planet could drive planet c into a highly inclined orbit.(Petrovich et al., 2020). In particular, this mechanism drives planets towards polar orbits, broadly consistent with the obliquity we measure. Such secular resonances do not necessarily result in large eccentricities, although they can be induced; however, they do require the system to be configured to allow the precession frequencies to become commensurate during disc dispersal.

A third possibility is Kozai-type interactions with the disk. Kozai-type dynamics (e.g., Lidov, 1962; Kozai, 1962) cause inclination and eccentricity to oscillate in phases that are 90 degrees out of sync. Since RV monitoring of the system is consistent with a circular orbit for planet c (Zicher et al., 2022b; Donati et al., 2023), we might currently be observing AU Mic c in a high inclination, low eccentricity phase. Moreover, if the event that caused the tilt in planet c's orbit happened before the dispersal of the disk, the eccentricity might have been dampened as the planet repeatedly passed through the disk (e.g., Teyssandier et al., 2013; Teyssandier and Terquem, 2014).

5 Conclusions and Future Work

Contents

5.1 Thesis summary 113
5.2 Ongoing work 117
5.2.1 2-D spot model 117
5.2.2 Preliminary study of the obliquity of HD 3167 b 120
5.3 Future work 128

5.1 Thesis summary

The thesis investigated the challenges that stellar activity posed for the detection and characterisation of exoplanets, particularly in the context of the RV method. By combining high-precision observational data with advanced statistical techniques, I developed methods to improve the understanding, modelling, and mitigation of stellar activity. These advances enabled more accurate measurements of planet masses and stellar spin–orbit alignments. The key contributions of this work are summarised below.

In Chapter 2, I investigated the rotational and magnetic activity characteristics of FGK-type stars with a multi-dimensional Gaussian process (GP) framework, using archival data from the HARPS spectrograph. The analysis focused on 268

well-observed stars with accurately determined stellar parameters, with particular attention to two spectroscopic activity indicators: the S-index and the Bisector Inverse Slope (BIS). Rotation periods were successfully identified for 49 slow rotators within the sample. A clear trend was observed, in which the rotation period decreased with increasing effective temperature, i.e. stars of earlier spectral types tended to rotate more rapidly. This behaviour was broadly consistent with theoretical predictions (e.g. Skumanich, 1972a) and previous empirical studies (e.g. McQuillan et al., 2014).

I placed these results in the broader context of rotation period measurements from *Kepler*, which are primarily derived from photometric variability, to examine their implications for age–activity–rotation relations and the spin-down of middle-aged late F-, G-, and K-type stars. The sample provided a valuable complement to these photometric studies, as it predominantly included stars with longer rotation periods, which are often missed or underestimated in light curve analyses.

When combined with existing asteroseismic constraints, the newly identified sample of slowly rotating FGK stars ($P \gtrsim 30\,\mathrm{d}$) revealed two distinct sub-populations. The first appeared to be an extension of the known long-period pile-up toward cooler temperatures, occupying a Rossby number range of $1.4 < \mathrm{Ro} < 2.0$. The second group consisted of relatively hot slow rotators ($T_{\mathrm{eff}} > 5700\,\mathrm{K}$), which I interpreted as subgiant stars undergoing rapid spin-down as they began to evolve off the main sequence.

These findings were broadly consistent with those of David et al. (2022), but presented the clearest observational evidence to date of a population of slowly rotating late G- and K-type stars lying above the so-called critical Rossby number ($\mathrm{Ro} \approx 1.5$) that defined the upper boundary in the period–temperature plane for *Kepler* stars with detectable surface rotation. These results suggested that angular momentum loss via magnetised stellar winds continued beyond this threshold.

I also examined how the spot-to-facular ratio varied across the sample. This ratio increased with Rossby number up to $\mathrm{Ro} \approx 1.9$, consistent with stronger magnetic activity and more spot-dominated surfaces at lower Ro. Beyond this

point, the correlation disappeared, likely indicating a transition from large-scale dipolar fields to more complex magnetic geometries, weakening the connection between field strength and Rossby number.

Both our rotation period measurements and the changing balance of low- versus high-order variability in the activity indicator time series supported the presence of a magnetic morphology transition near a critical Rossby number in the range $1.4 \leq \mathrm{Ro} \leq 2.0$. This aligned well with theoretical predictions for the onset of spin-down behaviour in stellar evolution models.

The findings in Chapter 2 offered useful guidance for the design of future RV planet searches, such as the Terra Hunting Experiment and the ground-based follow-up programme for PLATO, and helped refine models of stellar magnetic evolution and internal structure.

In Chapter 3, I introduced a novel, physics-informed GP framework to model stellar activity directly in time series of line profiles or CCFs. This model captured the distortions in CCFs caused by active regions on the stellar surface, such as spots and faculae, accounting for both the photometric effect and the inhibition of convective blueshift effect. Unlike existing approaches, this method exploited the temporal correlations across velocity (or wavelength) channels in a way that was grounded in physical principles.

I evaluated the performance of the framework using both synthetic and real observational datasets, including cases with signal-to-noise ratios as low as $\sim$100. The model demonstrated strong performance in disentangling stellar activity from planetary signals, even when both shared identical periods, highlighting its robustness in challenging regimes.

To further assess its capabilities, I carried out injection-recovery tests on two years of HARPS-N solar observations, sampled realistically across the solar magnetic cycle. The method successfully recovered a signal corresponding to a 1.5-Earth-mass planet with a semi-amplitude of $0.3\,\mathrm{m\,s^{-1}}$ and a 33-day orbital period. These results pushed the detection threshold for small planets to $\mathrm{K} = 0.3\,\mathrm{m\,s^{-1}}$ at $\mathrm{P} = 33\,\mathrm{d}$.

In these benchmark tests, the framework outperformed multi-GP models applied to RV and activity indicators.

In Chapter 4, I presented the first measurement of the stellar spin–orbit alignment for AU Mic c, a warm sub-Neptune in the 24 Myr-old AU Microscopii system. This analysis combined high-precision photometry from *CHEOPS* with radial velocity measurements from *ESPRESSO* to determine the projected obliquity of the host star with respect to the orbital plane of AU Mic c, a young and compact Neptune-sized planet.

After accounting for stellar flares, magnetic activity, and transit-timing variations, we estimated the sky-projected obliquity of AU Mic c to be $\lambda_c = 67.8^{+31.7}_{-49.0}$ degrees (1σ). Although the result was not definitive, it suggested—at an 89% confidence level—that the planet's orbit might be misaligned with the stellar spin axis. This would have implied a non-coplanar configuration relative to both AU Mic b and the system's outer debris disk. If confirmed, AU Mic c would represent the first known case of a misaligned orbit in a planetary system younger than 100 Myr.

Several scenarios could potentially have accounted for both the tentative misalignment of AU Mic c and its unusually high density. It was particularly intriguing to consider mechanisms capable of tilting the orbit of planet c while preserving the alignment between the star, planet b, and the outer debris disk. One possible explanation was planet-planet scattering, where gravitational interactions between planets could lead to significant changes in orbital inclination. Alternatively, a nodal secular resonance might have been responsible, gradually exciting the inclination of planet c without strongly affecting planet b. A third possibility involved Kozai-like interactions with the circumstellar disk, which could have induced coupled oscillations in inclination and eccentricity that evolved out of phase.

We noted that a significant misalignment in the orbit of planet c raised some tension with the system's overall dynamical stability and the observed geometry—that both planets were seen in transit. While these considerations did not rule out a misaligned configuration, they suggested it might have been dynamically constrained.

Additional observations would have been essential to refine the spin-orbit angle of planet c and assess the robustness of this result.

5.2 Ongoing work

In this section, I outline two ongoing projects initiated during my DPhil, which remain preliminary and unpublished.

The first involves the development of a two-dimensional spot model aimed at correcting spot-induced distortions in CCFs during planet transit. This model was created during the analysis of the spin-orbit alignment of AU Mic c presented in Chapter 4, and has shown improved performance compared to the model used in the published version, which addressed activity effects solely in the RV domain. At the time, this more advanced approach was not adopted due to the technical challenges of jointly fitting the spot model with the RM signal and the photometric transit data, particularly because precise transit timing is essential and the correction cannot be treated as a separate step. Future work will aim to fully integrate this model into a joint photometric and spectroscopic fitting framework.

The second project is a preliminary study of the spin-orbit alignment of planets in the HD 3167 system, conducted in collaboration with Professor Joshua Winn at Princeton University. Like AU Mic, this system is thought to host planets on mutually misaligned orbits. Our analysis focuses on planet b, with the goal of verifying and refining previous obliquity measurements and better understanding the system's dynamical architecture.

5.2.1 2-D spot model

We model the spot signature in the 2-D time series of flare-insensitive CCFs for the second transit of AU Mic c, and for the transit of AU Mic b as a validation. Compared to RVs, CCFs retain more information from the spectrum, thereby having potentially lower degeneracy between planetary signals and spot signals. Specifically, in the CCFs the spot signal behaves as a slowly moving signature with varying amplitude through the full observations. In contrast, the planet

signals only appear in the in-transit frames and move at a usually much larger speed compared to the spot signal.

To model the spot signal, we first calculate a reference CCF by averaging the first 10 out-of-transit frames before the ingress. We then subtracted the reference CCF from all the CCFs to get ΔCCFs. These are shown in the first column of Figure 5.1. The ΔCCFs are plotted with an arbitrary shift regarding time, and the blue and grey colours correspond to the out-transit and in-transit frames. We can see clearly a smoothly varying signature in both velocity and time domains among the full observations.

Given the stochastic nature of the spots, and the limited available information for exacting the exact distribution of spots on the stellar disk, the deterministic form of the spot-induced signal is not possible to derive. However, the covariance structure of the data is rather clear. The covariance of the signal follows approximately a Gaussian function in the velocity domain with a width corresponding to the *vsini* of the star. This incorporates the information that spots are distributed on the projected stellar disk. On top of this, the signal evolves in both velocity and time domains with approximately constant evolution time scales when the time span of the observation is much smaller than the rotation period of the star, e.g., a few hours, as in the case of our transit observations.

Hence, assuming the spot signal follows a multi-variate Guassian distribution, we model the signal using a 2-D GP. The full covariance matrix is modelled as the product of a 1-D local Gaussian kernel in the velocity domain and a 2-D isotropic squared exponential kernel in both time and velocity domains. The local Gaussian kernel is defined as,

$$k\left(\mathbf{x}_i, \mathbf{x}_j\right) = \exp\left(-\frac{\left(x_i - x_0\right)^2 + \left(x_j - x_0\right)^2}{2w}\right), \tag{5.1}$$

where x_0 corresponds to the location of the Gaussian, and w corresponds to the width of the Gaussian, which relates to the *vsini* of the star. The 2-D isotropic squared exponential kernel is defined as

$$k\left(r^2\right) = \exp\left(-\sqrt{r^2}\right) \tag{5.2}$$

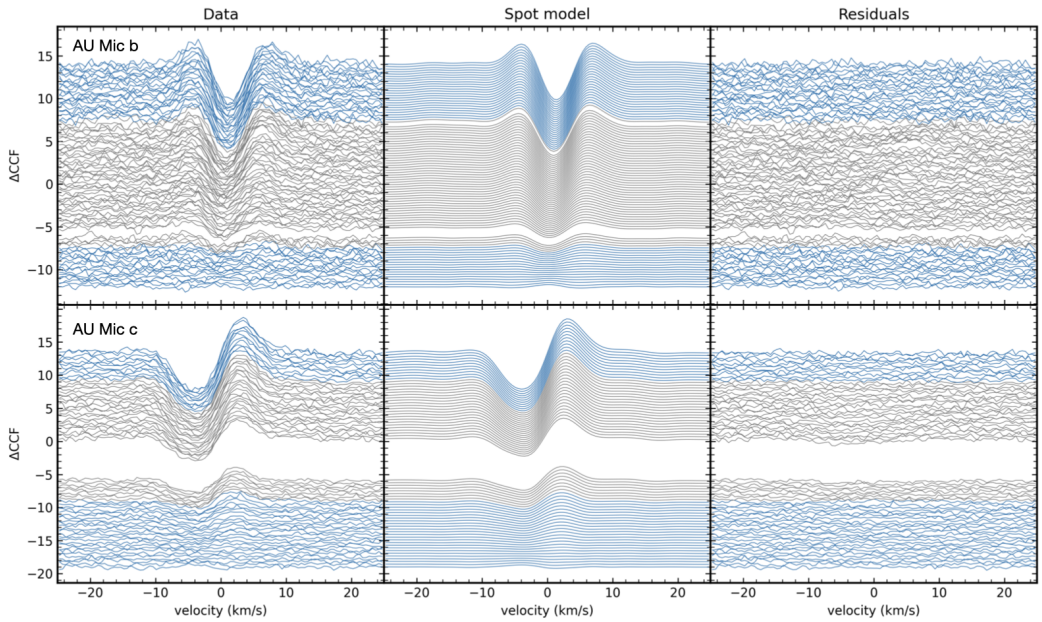


Figure 5.1: Spot correction for the flare insensitive CCFs for the transit of AU Mic b (upper), and the second transit of AU Mic c (lower). The left panel shows the reference-subtracted CCFs. The middle panel shows the best-fit spot model to the reference-subtracted CCFs. The right panel shows the residuals after subtracting the best-fit models.

where r is the squared distance between the input coordinates x_i and x_j,

$$r^2 = (\mathbf{x}_i - \mathbf{x}_j)^{\mathrm{T}} C^{-1} (\mathbf{x}_i - \mathbf{x}_j) . \tag{5.3}$$

In our case, C is a two-dimensional isotropic metric with evolution timescales in time and velocity domains, $l1$ and $l2$, respectively, on the diagonal terms,

$$C = \begin{pmatrix} l_1 & 0 \\ 0 & l_2 \end{pmatrix} . \tag{5.4}$$

To further take into account the small drift of spots in the velocity domain as they co-rotate with the stellar surface, we apply an inverse shift, which is a linear function of time, to the velocity grid to compensate for such a drift. The slope of the linear function relates to the $v$sini represented by w in the local Gaussian kernel in the model.

To avoid potential overfitting of the 2-D GP model that results in affecting the embedded planetary signal, we train our 2-D GP model only on the out-of-transit

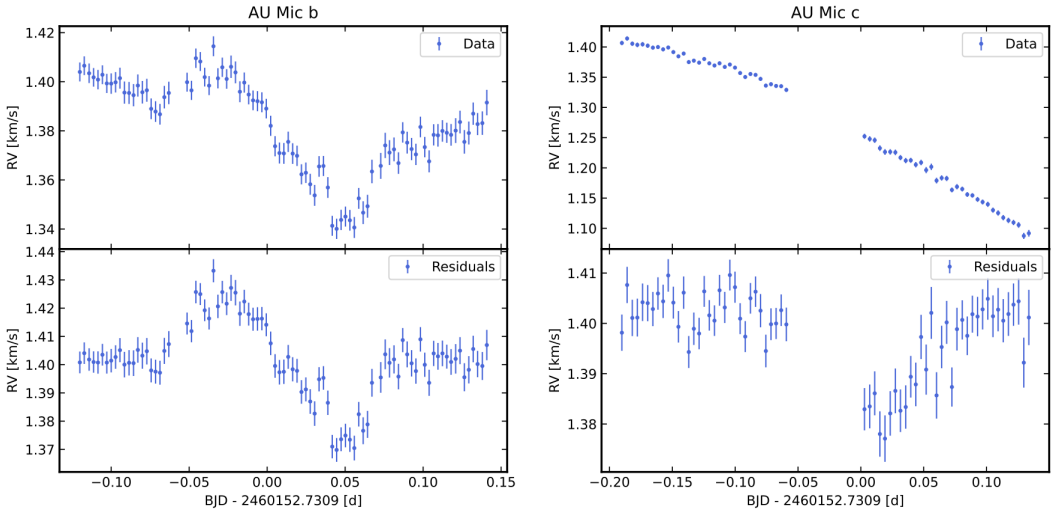


Figure 5.2: Comparison of RVs derived before (upper) and after (lower) subtracting the best-fit spot model in the CCFs, for AU Mic b (left) and c (right).

frames (blue lines) based on the CHEOPS transit analysis and then predict the model on the full time grid that includes the in-transit frames (grey lines). We implement the 2-D GP model using `george` (Ambikasaran et al., 2015) and obtain the best-fit models using `scipy.minimize`. The best-fit models are shown in the middle column in Figure 5.1, and the residuals of subtracting the best-fit models are shown in the right columns. By adding the reference CCF back, we can measure the RVs of both CCFs before and after subtracting the spot model (corresponding to data and residuals in the figure). The RVs are measured by fitting a Gaussian function to each CCFs. A comparison of the RVs before and after subtracting the spot model in the CCFs are shown in Figure 5.2 for both AU Mic b and c. The spot-free RVs, labelled as residuals in the figure, are then passed to the next step for RM modelling.

5.2.2 Preliminary study of the obliquity of HD 3167 b

Background and motivation

HD 3167 is a bright ($V = 9$) K0-type star located at a distance of 47 pc. It is known to host three planets. Planets b and c are transiting planets first reported by Vanderburg et al. (2016). Planet b is an ultra-short-period (USP) planet with a period of $P = 0.96$ days, a radius of $R_p = 1.70^{+0.18}_{-0.15}\,R_\oplus$, and a mass of $M_p =$

$5.02 \pm 0.38\,M_\oplus$. Planet c is a mini-Neptune with $P = 29.84$ days, $R_p = 3.01^{+0.42}_{-0.28}\,R_\oplus$, and $M_p = 9.80^{+1.30}_{-1.24}\,M_\oplus$. The third planet, HD 3167 d, was detected via RV measurements by Christiansen et al. (2017), with a period of $P = 8.51$ days and a minimum mass of $M_p \sin i = 6.90 \pm 0.71\,M_\oplus$. Gandolfi et al. (2017) also reported two additional signals at periods of 6 and 10.7 days, although their planetary nature remains unconfirmed.

The orbital architecture of the HD 3167 system remains both puzzling and intriguing. Dalal et al. (2019) conducted the first obliquity measurement of planet c and found it to be nearly polar. This result was later confirmed by Bourrier et al. (2021a), who used their Rossiter-McLaughlin effect Revolutions (RMR) technique to obtain a more precise measurement, reporting a sky-projected spin–orbit angle of $\lambda = -108.9^{+5.4}_{-5.5}$ degrees and a 3D spin–orbit angle of $107.7^{+5.1}_{-4.9}$ degrees. As noted by Bourrier et al. (2021a), the RMR technique enhances the extraction of information from transit time series by jointly modelling the spectral profiles of all planet-occulted regions through Bayesian inference. It also allows for the integration of transit data from multiple planets and instruments within a consistent modelling framework. Using the RMR technique, Bourrier et al. (2021a) found that planet b has an orbit closely aligned with the stellar spin axis, reporting a projected spin–orbit angle of $\lambda = -6.6^{+6.6}_{-7.9}$ degrees and a 3D obliquity of $29.5^{+7.2}_{-9.4}$ degrees.

These measurements reveal a striking configuration in which the two planets are nearly perpendicular to each other, with a mutual inclination of $102.3^{+7.4}_{-8.0}$ degrees. Bourrier et al. (2021a) suggested that the observed architecture may result from the planet b maintaining its original alignment due to strong coupling with the stellar equator, whereas the planet c could have been tilted into a near-polar orbit as a consequence of long-term gravitational interactions with a more distant companion.

The motivation for refining the obliquity measurement of planet b stems from two key considerations. First, the initial detection by Bourrier et al. (2021a) was not sufficiently precise to conclusively rule out a coplanar configuration between planets b and c. Second, while such an extreme mutual inclination is theoretically possible, its geometric transit probability is low, suggesting it may be rare to observe.

Teng et al. (2025) further carried out a detailed dynamical study and found that observing both HD 3167 b and c in transit from a single observer point is less likely due to the system's geometry. Moreover, even in cases where both planets are initially observable in transit, the mutual inclination leads to rapid orbital evolution, causing one or both planets to cease transiting within a decade-to-century timescale.

Observations and initial RV analysis

Winn et al. observed three transits of HD 3167 b with the *Keck Planet Finder* (KPF; Gibson et al. 2016) on the nights of September 21, 22, and October 19, 2023 (Program: C316, PI: J. Winn). Each exposure was set to 300 s, followed by a 60 s readout time. The final RV uncertainty per exposure is approximately 0.35 $\mathrm{m\,s^{-1}}$. The publicly available `KPF` Data Reduction Pipeline (DRP) was used to calibrate the raw data and produce order-by-order extracted spectra. CCFs and RVs were then computed using the G9 ESPRESSO weighted binary mask (Pepe et al., 2002).

Stefansson et al. conducted the initial RV analysis for the three transits. In their analysis, the projected obliquity λ_b was estimated by fitting an RM model jointly with a polynomial to account for smooth baseline variations likely caused by stellar activity. I replicated this analysis using the same RV dataset, and the conclusions remain broadly unchanged.

We were unable to confidently detect the RM signal or place meaningful constraints on λ_b; the result remains inconclusive. This is primarily due to the faintness of the expected RM signal. Using the planetary parameters derived from the transit analysis (Vanderburg et al., 2016), the maximum RM amplitude (defined as the absolute deviation from the baseline) is expected to be less than 0.7 $\mathrm{m\,s^{-1}}$, and it drops to around 0.3 $\mathrm{m\,s^{-1}}$ in strongly misaligned configurations. The median RV uncertainty is approximately 0.35 $\mathrm{m\,s^{-1}}$, and additional systematics are also present. Given these limitations, the lack of robust detection is not surprising. This motivates a closer examination of the CCFs, which may provide stronger evidence for the RM effect.

Doppler tomography model

To model the Doppler tomography signals induced by the planets, we follow the methodologies of Collier Cameron et al. (2010a) and Cegla et al. (2016). The signals represent the local line profile of the star blocked by the transiting planet. Assuming the local line profile is Gaussian, the signal will also be Gaussian with a centre velocity $v_{\rm p}(t)$, width v_{res} (which in most cases is directly tied to the resolution of the instrument), and a varying amplitude. The signal can be expressed as

$$A(t) * G(v_{\rm p}(t), v_{res}). \tag{5.5}$$

The centre velocity can be derived from

$$v_{\rm p}(t) = u_{\rm p}(t) v \sin I, \tag{5.6}$$

where $u_{\rm p}(t)$ corresponds to the normalised perpendicular distance from planet to the stellar rotation axis, $v \sin I$ is the projected rotational velocities of the star. The amplitude $A(t)$ is related to $1 - LC(t)$, where $LC(t)$ is a standard transit light curve model that accounts for the limb-darkening effect. The exact relationship depends on the specific normalisation procedure applied to the CCFs.

The model is implemented in Python using `JAX`, enabling full GPU acceleration and significantly enhancing computational efficiency. I employ `jaxoplanet` (Hattori et al., 2024) to compute the planet's position relative to the star and to generate limb-darkened transit light curve models. These are then passed to the models described above to compute the precise planet-induced signals in the CCFs. The framework is fully compatible with Bayesian inference, with posterior sampling carried out via MCMC using `NumPyro` (Phan et al., 2019; Bingham et al., 2019). Notably, this model can be combined with the 2-D GP spot model described in Section 5.2.1 to jointly fit planetary signals and stellar activity induced by spots.

Preliminary CCF analysis

The CCFs output by the pipeline are provided order by order. I co-add these to produce a combined CCF, as well as separate CCFs for the Green CCD channel

(wavelength range 440–590 nm) and the Red CCD channel (590–860 nm), followed by normalisation using the CCF wings. The Green and Red channels are checked separately because they can exhibit different systematics.

Reference CCFs are then computed from the out-of-transit frames, with the out-of-transit windows determined using the transit parameters reported in Vanderburg et al. (2016). These reference CCFs are subtracted from each exposure to highlight the variations, denoted as ΔCCF. The resulting maps are shown for the three visits in Figures 5.3, 5.4, and 5.5. The transit duration is marked by white dotted lines in each panel.

In this analysis, we do not detect a significant planetary signal within the transit window. However, we do observe uneven background structure and systematics at the $\sim$0.1% level. The origin of these systematics is not yet fully understood. For example, in visit 2, a clear pattern of two bands with alternating signs of variation is visible. This could potentially be caused by stellar surface inhomogeneities such as spots, but it might also result from instrumental effects, such as wavelength drift. To distinguish between these possibilities, a detailed investigation of the full spectrum is required.

We then combined the CCFs from all three visits to investigate whether stacking could enhance any potential low-amplitude planetary signals. To do this, the CCFs were first interpolated onto a common time grid to ensure temporal alignment across visits. The combined CCFs were then co-added and normalised. As before, a reference CCF was constructed by averaging the out-of-transit frames and subtracted to highlight residual variations. The result is shown in the upper panel of Figure 5.6.

We do not detect a clear planetary signature in the combined ΔCCF. In the lower panel of Figure 5.6, we overplot the expected planetary signal generated using the `JAX`-based forward model mentioned above, adopting the planetary parameters and projected obliquity from Bourrier et al. (2021b). This model does not appear to match any feature in the observations.

Although the model framework supports full Bayesian inference, we chose not to proceed with it in this case. Given the presence of systematics at the 0.1% level

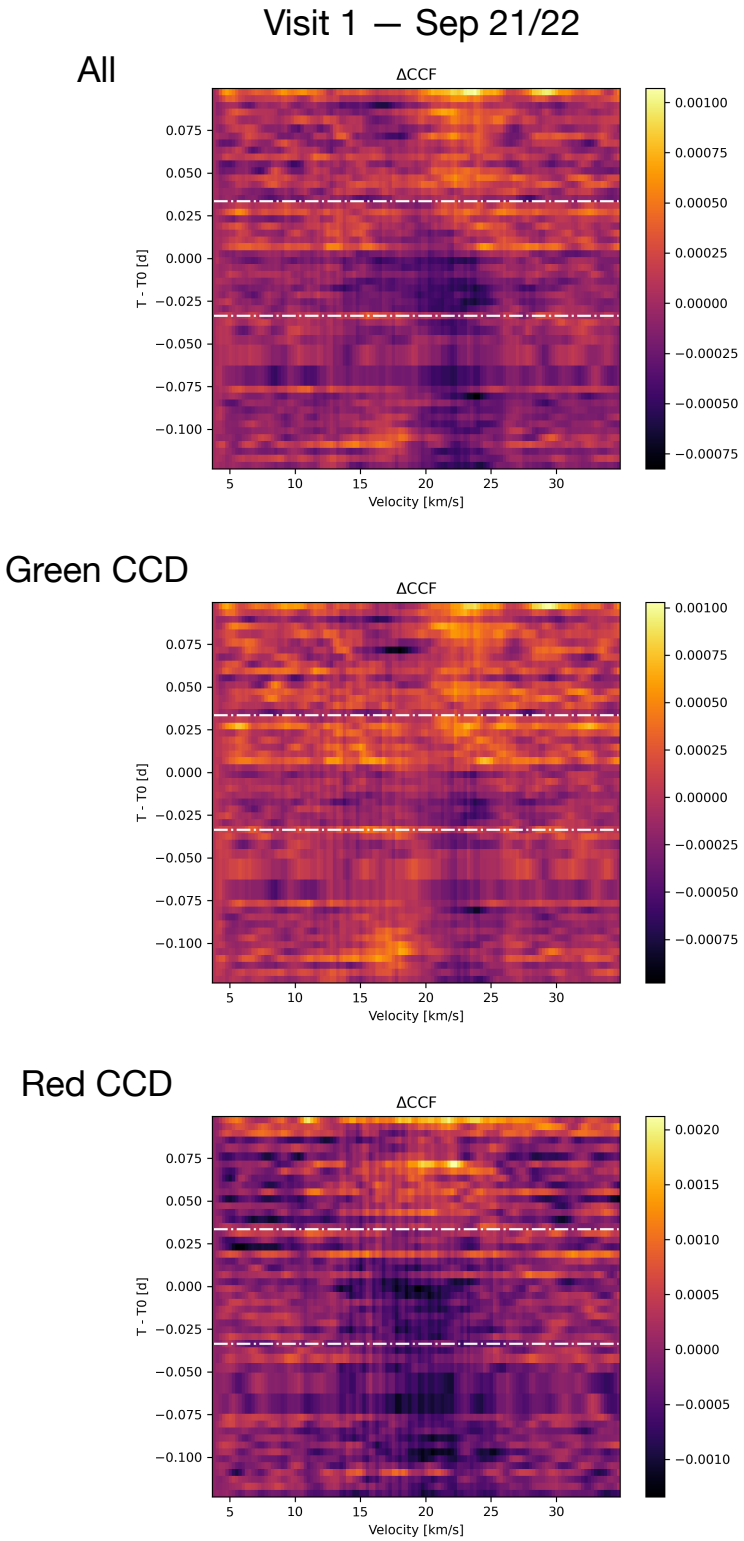


Figure 5.3: Reference-subtracted CCFs for the first transit of HD 3167 b. The upper panel shows the CCFs computed using all spectral orders. The middle and lower panels correspond to CCFs derived from the Green and Red CCDs, respectively. The transit duration is marked by white dotted lines in each panel.

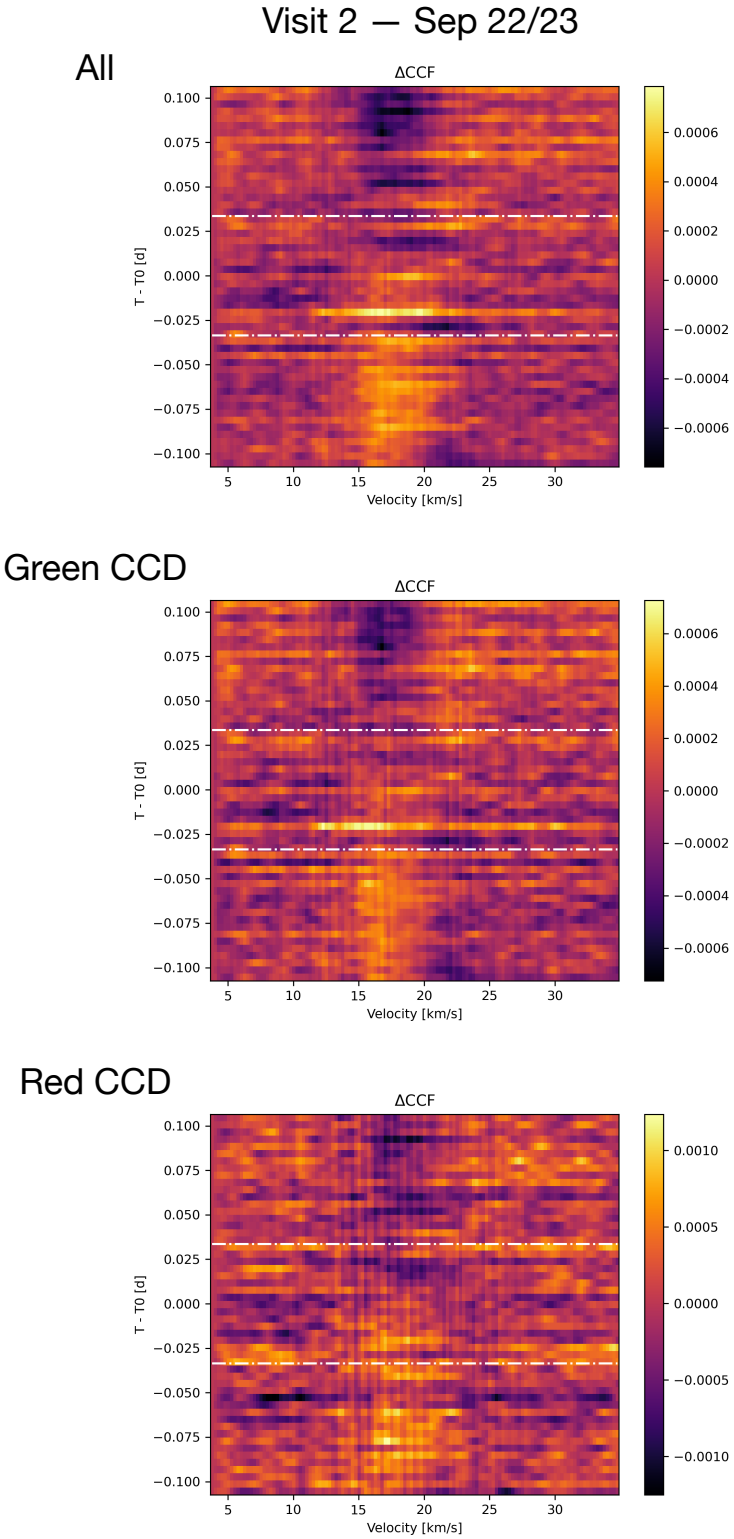


Figure 5.4: Reference-subtracted CCFs for the second transit of HD 3167 b, similar to those shown in Fig. 5.3.

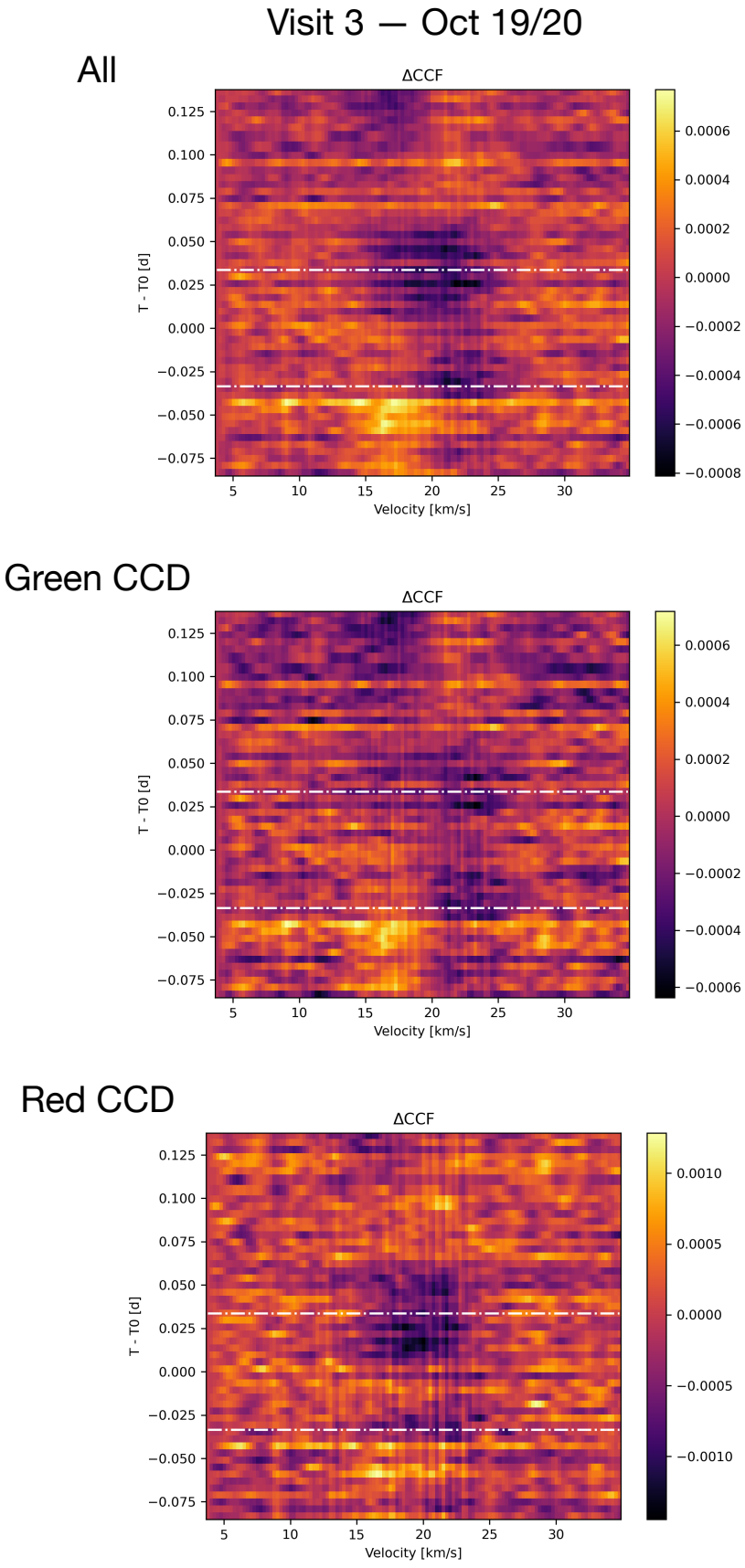


Figure 5.5: Reference-subtracted CCFs for the third transit of HD 3167 b, similar to those shown in Fig. 5.3.

in the CCFs, attempting to fit a model without first understanding and mitigating these systematics would likely lead to overfitting. We therefore conclude that this analysis does not yield a conclusive constraint on the obliquity of HD 3167 b.

Further investigation is needed to better characterise the systematics at the spectrum level. One possible avenue is to apply tools such as `YARARA` (Cretignier et al., 2021a) to correct for instrumental effects directly in the spectra. However, this would require adapting the KPF data format to be compatible with a toolset currently tailored for HARPS, HARPS-N, and ESPRESSO. Another possibility is to obtain repeated observations of future transits with KPF or next-generation spectrographs offering higher sensitivity. Although the current analysis is inconclusive, HD 3167 remains a highly compelling system for future exploration.

5.3 Future work

In this section, I outline aspects of future work that build on the research conducted during my DPhil and discussed in the preceding chapters.

For Chapter 2, one limitation of the analysis of rotational and magnetic activity in FGK-type stars is the neglect of differential rotation, primarily due to the sparse temporal sampling for most stars in the sample. A potential extension of this work would be to perform a season-by-season analysis, allowing for the investigation of temporal variations in both the rotation period and other model parameters.

The long-term evolution of the activity indicators examined here also offers an opportunity to search for magnetic activity cycles among our targets. This will be the focus of an upcoming paper in our series (Crétignier et al. in prep.). In the longer term, we also plan to incorporate the RV time-series, as well as other activity indicators, into our modelling, in order to improve our understanding of the strengths and limitations of the multi-dimensional GP framework for activity mitigation in RV planet searches, as a function of the stellar properties.

For Chapter 3, there are plans to extend the application of the CCF-GP model to broader contexts. One such application involves analysing archival AU Mic data to refine the mass estimates of the known planets in the system and to search for

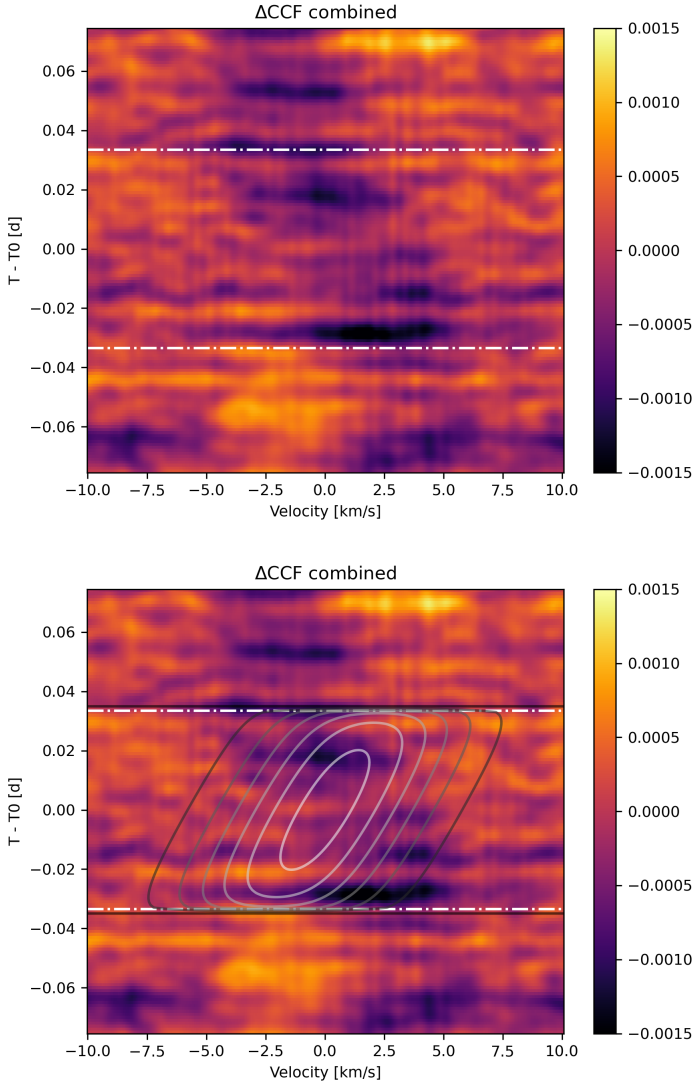


Figure 5.6: Combined CCFs from the three transits. The lower panel additionally shows the expected planet track during transit, overplotted as contours, based on the obliquity measurement from Bourrier et al. (2021b) and a forward model.

potential additional companions (Klein et al. in prep.). Another application is the *PLATO* RV Challenge, a blind test in which participants attempt to recover planetary signals from synthetic datasets. The results are then compared across different techniques to benchmark their performance in realistic conditions.

In the future development of the CCF-GP model, I plan to explicitly address the issue of non-Gaussianity. One approach as a substitute for using the current approximate GP model is to place a GP prior on F and use, e.g., Markov chain Monte Carlo (MCMC) methods to sample from it and then propagate the samples to approximate the exact stochastic process of the CCF model. However, this comes at the cost of increased computational demands.

Additionally, there is significant potential in enriching the model with more detailed physics. One would be the replacement of the current QP or MEP kernels with kernels that are obtained from physical parameters, such as the 'impulse response' GP kernel developed in Hara and Delisle (2023). Furthermore, the model could benefit from incorporating the differential responses of spectral lines to activity, as well as the effects of super-granulation. The inherent flexibility of the model's mathematical framework ensures that these integrations can be implemented seamlessly.

For the obliquity analysis of AU Mic c presented in Chapter 4, follow-up observations are highly desirable given the importance of the system and the significant impact of stellar activity on the measurements. Repeated transit observations would help to confirm the tentative misalignment and potentially reveal evidence of orbital precession.

Professor Vincent Bourrier (University of Geneva) re-observed the AU Mic system during the summer of 2024 and is open to collaboration on a comprehensive reanalysis of the system. This effort will integrate archival ESPRESSO data for both planets b and c, the new ATREIDES transit observation of planet c, simultaneous NGTS and CHEOPS photometry, and long-term light curves to better characterise stellar spot properties. The analysis will employ the ANTARESS pipeline to reduce all ESPRESSO epochs, generate CCFs using active/quiet line selection tools

developed with RASSINE, and perform a joint multi-planet, multi-epoch fit using an updated spotted RM model. A promising by-product of this approach is the potential extraction of high-resolution spot spectra, offering deeper insight into stellar surface inhomogeneities and further enhancing the scientific value of the study.

The dataset presented in this work contains rich spectroscopic information on stellar flares, which we plan to investigate further. This analysis will help shed light on the underlying physics of flares, which remains poorly understood.

Appendices

A

Rotation and activity in HARPS data

A.1 Confirmed rotation period measurements

From Chapter 2, Table A.1 shows the measured rotation periods P_{rot} for the selected 49 targets. In addition to P_{rot}, we have provided information about the time span of the data T_{span}, as well as all other hyperparameters from our model, including the evolution timescale l and the harmonic complexity Γ. We have also included the derived Rossby number Ro and the RMS$[aG(t)]$ / RMS$[b\dot{G}(t)]$ (A/B; representing the facular-to-spot ratio). All values, except for T_{span}, come with associated uncertainties.

Table A.1: Confirmed rotation period measurements of 49 targets.

Simbad name	T_{span} [yr]	P_{rot} [d]	l [d]	Γ	Ro	A/B
BD-08 2823	2.42	$27.0^{+0.4}_{-0.4}$	966^{+559}_{-462}	$1.86^{+0.91}_{-0.64}$	$0.94^{+0.02}_{-0.02}$	$1.5^{+1.0}_{-1.0}$
HD 104067	2.38	$30.3^{+0.8}_{-0.8}$	386^{+379}_{-230}	$1.21^{+0.66}_{-0.51}$	$1.25^{+0.03}_{-0.03}$	$0.7^{+0.6}_{-0.6}$
HD 109200	7.24	$39.6^{+8.3}_{-8.3}$	57^{+32}_{-38}	$0.76^{+0.34}_{-0.12}$	$1.72^{+0.36}_{-0.36}$	$1.8^{+0.5}_{-0.5}$
HD 115617	7.49	$32.1^{+0.2}_{-0.2}$	186^{+78}_{-50}	$0.68^{+0.16}_{-0.12}$	$1.99^{+0.02}_{-0.02}$	$1.2^{+0.7}_{-0.7}$
HD 125072	7.44	$40.8^{+0.8}_{-0.8}$	482^{+455}_{-227}	$1.14^{+0.47}_{-0.32}$	$1.52^{+0.03}_{-0.03}$	$1.9^{+1.5}_{-1.5}$
HD 125184	9.38	$43.9^{+2.0}_{-2.0}$	200^{+151}_{-78}	$0.59^{+0.19}_{-0.14}$	$3.06^{+0.14}_{-0.14}$	$1.8^{+1.2}_{-1.2}$
HD 125595	3.44	$37.1^{+0.5}_{-0.5}$	804^{+485}_{-410}	$2.82^{+1.78}_{-1.33}$	$1.27^{+0.02}_{-0.02}$	$1.2^{+1.0}_{-1.0}$
HD 13060	0.39	$34.3^{+3.8}_{-3.8}$	538^{+420}_{-272}	$3.91^{+3.06}_{-1.75}$	$1.68^{+0.19}_{-0.19}$	$0.8^{+0.7}_{-0.7}$
HD 136713	9.49	$36.4^{+1.6}_{-1.6}$	395^{+527}_{-231}	$0.87^{+0.48}_{-0.35}$	$1.52^{+0.07}_{-0.07}$	$0.3^{+0.3}_{-0.3}$
HD 13808	8.08	$38.0^{+0.3}_{-0.3}$	379^{+197}_{-106}	$1.20^{+0.28}_{-0.20}$	$1.62^{+0.01}_{-0.01}$	$3.1^{+2.0}_{-2.0}$
HD 144628	9.44	$38.5^{+1.1}_{-1.1}$	184^{+135}_{-64}	$1.59^{+0.64}_{-0.38}$	$1.63^{+0.05}_{-0.05}$	$3.5^{+2.7}_{-2.7}$
HD 15337	10.94	$37.1^{+0.8}_{-0.8}$	1330^{+456}_{-513}	$5.43^{+2.72}_{-2.48}$	$1.73^{+0.04}_{-0.04}$	$3.0^{+2.5}_{-2.5}$
HD 154088	5.25	$30.7^{+1.7}_{-1.7}$	150^{+66}_{-38}	$1.53^{+0.54}_{-0.35}$	$1.64^{+0.09}_{-0.09}$	$14.7^{+12.8}_{-12.8}$
HD 154577	7.44	$34.3^{+1.0}_{-1.0}$	104^{+42}_{-30}	$1.00^{+0.17}_{-0.14}$	$1.32^{+0.04}_{-0.04}$	$1.7^{+0.5}_{-0.5}$
HD 157172	5.40	$35.4^{+10.6}_{-10.6}$	498^{+478}_{-247}	$4.23^{+2.96}_{-2.16}$	$2.00^{+0.60}_{-0.60}$	$4.7^{+3.4}_{-3.4}$
HD 157830	2.56	$25.4^{+5.4}_{-5.4}$	566^{+722}_{-332}	$3.46^{+2.45}_{-1.48}$	$1.55^{+0.33}_{-0.33}$	$0.8^{+0.4}_{-0.4}$
HD 161098	9.43	$27.5^{+0.2}_{-0.2}$	323^{+163}_{-94}	$1.00^{+0.37}_{-0.22}$	$1.71^{+0.01}_{-0.01}$	$1.6^{+1.2}_{-1.2}$
HD 16417	3.33	$33.0^{+6.1}_{-6.1}$	102^{+90}_{-38}	$0.83^{+0.51}_{-0.23}$	$2.81^{+0.52}_{-0.52}$	$1.3^{+1.4}_{-1.4}$
HD 168863	2.36	$32.0^{+7.8}_{-7.8}$	1000^{+605}_{-529}	$2.58^{+2.71}_{-1.28}$	$1.19^{+0.29}_{-0.29}$	$2.0^{+2.5}_{-2.5}$
HD 171587	2.56	$24.7^{+3.3}_{-3.3}$	650^{+583}_{-390}	$2.76^{+1.95}_{-1.18}$	$1.35^{+0.18}_{-0.18}$	$1.3^{+0.9}_{-0.9}$
HD 172513	1.18	$24.0^{+0.6}_{-0.6}$	276^{+173}_{-127}	$1.14^{+0.40}_{-0.34}$	$1.41^{+0.03}_{-0.03}$	$1.0^{+0.6}_{-0.6}$
HD 176986	10.06	$34.8^{+0.8}_{-0.8}$	95^{+34}_{-24}	$0.65^{+0.12}_{-0.11}$	$1.48^{+0.03}_{-0.03}$	$1.1^{+0.4}_{-0.4}$
HD 183658	3.40	$28.4^{+8.2}_{-8.2}$	803^{+535}_{-460}	$3.57^{+2.85}_{-2.03}$	$2.300^{+0.66}_{-0.66}$	–
HD 189567	9.04	$41.7^{+11.7}_{-11.7}$	75^{+52}_{-62}	$0.64^{+0.89}_{-0.11}$	$3.06^{+0.86}_{-0.86}$	$6.5^{+2.3}_{-2.3}$
HD 192310	7.23	$41.0^{+0.9}_{-0.9}$	222^{+175}_{-64}	$1.23^{+0.37}_{-0.21}$	$1.81^{+0.04}_{-0.04}$	$2.5^{+2.8}_{-2.8}$
HD 202206	1.29	$50.0^{+5.1}_{-5.1}$	918^{+406}_{-352}	$5.54^{+2.27}_{-2.04}$	$3.810^{+0.39}_{-0.39}$	–
HD 208704	2.28	$32.2^{+7.8}_{-7.8}$	878^{+615}_{-507}	$3.10^{+2.29}_{-1.47}$	$2.69^{+0.65}_{-0.65}$	$2.0^{+2.1}_{-2.1}$
HD 215152	8.10	$38.0^{+0.8}_{-0.8}$	182^{+70}_{-42}	$2.59^{+0.86}_{-0.59}$	$1.43^{+0.03}_{-0.03}$	$4.7^{+3.1}_{-3.1}$
HD 216770	9.26	$37.4^{+4.7}_{-4.7}$	314^{+221}_{-149}	$1.57^{+0.81}_{-0.51}$	$2.07^{+0.26}_{-0.26}$	$1.0^{+1.0}_{-1.0}$
HD 21693	5.47	$32.0^{+0.7}_{-0.7}$	321^{+294}_{-119}	$0.84^{+0.31}_{-0.17}$	$1.78^{+0.04}_{-0.04}$	$3.4^{+2.8}_{-2.8}$
HD 21749	0.39	$38.3^{+4.9}_{-4.9}$	382^{+353}_{-222}	$1.45^{+0.60}_{-0.47}$	$1.32^{+0.17}_{-0.17}$	$0.4^{+0.3}_{-0.3}$
HD 22049	0.23	$11.2^{+0.1}_{-0.1}$	128^{+74}_{-56}	$1.70^{+0.58}_{-0.58}$	$0.48^{+0.01}_{-0.01}$	$0.1^{+0.1}_{-0.1}$
HD 223171	12.48	$37.2^{+11.2}_{-11.2}$	85^{+163}_{-64}	$0.98^{+2.82}_{-0.38}$	$3.170^{+0.95}_{-0.95}$	–
HD 26965	12.41	$35.6^{+16.0}_{-16.0}$	51^{+52}_{-40}	$1.64^{+3.74}_{-0.77}$	$1.560^{+0.70}_{-0.70}$	–
HD 27894	13.56	$48.2^{+8.0}_{-8.0}$	827^{+762}_{-635}	$3.34^{+2.80}_{-1.60}$	$1.968^{+0.32}_{-0.32}$	–
HD 36003	11.21	$50.2^{+2.0}_{-2.0}$	214^{+129}_{-78}	$0.85^{+0.22}_{-0.15}$	$1.74^{+0.07}_{-0.07}$	$2.4^{+1.3}_{-1.3}$
HD 40307	7.42	$42.5^{+0.5}_{-0.5}$	494^{+273}_{-139}	$1.21^{+0.29}_{-0.20}$	$1.76^{+0.02}_{-0.02}$	$2.5^{+1.6}_{-1.6}$

HD 45184	15.01	$20.0^{+0.1}_{-0.1}$	278^{+115}_{-78}	$1.22^{+0.27}_{-0.21}$	$1.78^{+0.01}_{-0.01}$	$5.6^{+3.0}_{-3.0}$
HD 4915	13.33	$25.2^{+11.7}_{-11.7}$	180^{+326}_{-163}	$2.49^{+2.46}_{-1.28}$	$1.72^{+0.80}_{-0.80}$	$2.3^{+1.9}_{-1.9}$
HD 51608	7.43	$37.1^{+1.8}_{-1.8}$	845^{+762}_{-527}	$1.60^{+1.10}_{-0.52}$	$1.95^{+0.10}_{-0.10}$	$3.1^{+2.8}_{-2.8}$
HD 63765	12.33	$25.5^{+9.1}_{-9.1}$	95^{+82}_{-42}	$3.20^{+3.15}_{-1.59}$	$1.42^{+0.51}_{-0.51}$	$1.6^{+1.5}_{-1.5}$
HD 65277A	2.43	$42.9^{+5.8}_{-5.8}$	1368^{+426}_{-501}	$5.28^{+2.56}_{-2.04}$	$1.52^{+0.20}_{-0.20}$	$0.7^{+0.9}_{-0.9}$
HD 68978	8.24	$16.8^{+0.8}_{-0.8}$	63^{+30}_{-20}	$0.73^{+0.20}_{-0.13}$	$1.75^{+0.08}_{-0.08}$	$2.2^{+1.1}_{-1.1}$
HD 69830	7.50	$33.7^{+1.0}_{-1.0}$	234^{+418}_{-106}	$1.31^{+0.63}_{-0.31}$	$1.83^{+0.06}_{-0.06}$	$2.5^{+1.8}_{-1.8}$
HD 71835	7.47	$29.7^{+0.5}_{-0.5}$	1123^{+583}_{-595}	$2.06^{+1.80}_{-0.70}$	$1.66^{+0.03}_{-0.03}$	$5.0^{+4.2}_{-4.2}$
HD 78429	13.37	$43.4^{+14.8}_{-14.8}$	19^{+58}_{-8}	$0.84^{+5.90}_{-0.33}$	$3.32^{+1.13}_{-1.13}$	$4.4^{+1.5}_{-1.5}$
HD 89454	1.43	$16.4^{+1.9}_{-1.9}$	327^{+278}_{-187}	$1.07^{+0.47}_{-0.36}$	$1.21^{+0.14}_{-0.14}$	$0.6^{+0.5}_{-0.5}$
HD 93083	10.26	$44.6^{+6.2}_{-6.2}$	321^{+448}_{-213}	$1.75^{+1.06}_{-0.66}$	$1.92^{+0.27}_{-0.27}$	$2.1^{+1.9}_{-1.9}$
HD 96700	5.33	$43.6^{+2.3}_{-2.3}$	194^{+177}_{-72}	$0.79^{+0.46}_{-0.21}$	$3.73^{+0.20}_{-0.20}$	$6.3^{+11.6}_{-11.6}$

A.2 Basic properties of the sample

From Chapter 2, Table A.2 shows the basic properties of all 268 targets in the sample, including the effective temperature $T_{\rm eff}$, surface gravity log g, and mean log $R'_{\rm HK}$ from the catalogue provided by Gomes da Silva et al. (2021), with associated uncertainties. Additionally, we have included details about the temporal coverage of observations, denoted as $\rm T_{range}$, as well as the temporal span of the data processed using DRS 3.5. Any temporal span not covered by the DRS 3.5 processed data is filled by the DRS 3.8 processed data.

Table A.2: Basic properties of all 268 targets in the sample.

Simbad name	$T_{\rm eff}$ [K]	$T_{\rm eff}$ err [K]	log g	log g err	Mean $\log R'_{\rm HK}$	Mean $\log R'_{\rm HK}$ err	$\rm T_{range}$	$\rm T_{range}$ (DRS 3.5)
BD-08 2823	4672.0	105.0	4.61	0.26	-4.6944	0.0039	2004-01 to 2018-04	2004-01 to 2010-01
CD-23 395	4673.0	175.0	4.52	0.47	-4.8528	0.0040	2004-07 to 2016-10	2004-07 to 2013-01
CD-24 10619	4574.0	155.0	4.68	0.44	-4.5766	0.0036	2005-06 to 2021-02	2005-06 to 2011-07
CD-26 2288	4924.0	94.0	4.51	0.28	-4.6539	0.0037	2004-01 to 2019-10	2004-01 to 2011-12
HD 10180	5911.0	19.0	4.19	0.03	-4.9957	0.0043	2003-11 to 2017-08	2003-11 to 2015-01
HD 101930	5164.0	61.0	4.49	0.11	-5.0048	0.0033	2004-02 to 2016-06	2004-02 to 2015-05
HD 102117	5657.0	24.0	4.21	0.04	-5.1215	0.0050	2004-01 to 2016-05	2004-01 to 2015-05
HD 102365	5629.0	29.0	4.35	0.03	-4.9489	0.0030	2003-11 to 2017-05	2003-11 to 2015-05
HD 103197	5250.0	60.0	4.51	0.11	-5.0796	0.0059	2004-02 to 2018-04	2004-02 to 2010-04
HD 103720	5017.0	88.0	4.58	0.16	-4.4493	0.0038	2005-02 to 2021-02	2005-02 to 2015-05
HD 103774	6732.0	56.0	4.29	0.06	-4.7718	0.0073	2004-12 to 2018-04	2004-12 to 2013-02
HD 104067	4969.0	72.0	4.64	0.13	-4.7370	0.0017	2004-02 to 2010-04	2004-02 to 2010-04
HD 104800	5697.0	25.0	4.35	0.02	-4.8877	0.0058	2004-02 to 2015-01	2004-02 to 2015-01
HD 105690	5666.0	38.0	4.53	0.07	-4.3063	0.0012	2009-04 to 2017-06	2009-04 to 2014-07
HD 106116	5680.0	15.0	4.28	0.03	-5.0317	0.0041	2004-02 to 2017-06	2004-02 to 2015-05
HD 10647	6218.0	20.0	4.30	0.04	-4.7431	0.0024	2003-11 to 2020-01	2003-11 to 2012-01
HD 10700	5310.0	17.0	4.49	0.03	-4.9773	0.0025	2003-09 to 2012-10	2003-10 to 2012-10
HD 107094	5562.0	17.0	4.48	0.03	-4.8283	0.0055	2004-01 to 2021-02	2004-01 to 2015-03
HD 109200	5056.0	33.0	4.50	0.08	-4.9542	0.0027	2004-02 to 2018-05	2004-02 to 2015-05
HD 109271	5783.0	18.0	4.13	0.02	-4.9963	0.0074	2005-02 to 2018-04	2005-02 to 2014-07
HD 111232	5460.0	21.0	4.41	0.03	-4.9832	0.0041	2004-02 to 2021-03	2004-02 to 2015-05
HD 111515	5398.0	18.0	4.47	0.02	-4.9528	0.0043	2004-02 to 2017-07	2004-02 to 2009-03

HD 111777	5666.0	19.0	4.36	0.03	-4.9115	0.0048	2004-02 to 2018-06	2004-02 to 2014-05
HD 11397	5564.0	26.0	4.40	0.04	-4.8975	0.0047	2003-10 to 2019-09	2003-10 to 2014-08
HD 114076	5066.0	25.1	4.47	0.07	-4.9911	0.0058	2004-02 to 2015-04	2004-02 to 2015-04
HD 114613	5729.0	17.0	3.84	0.02	-5.1529	0.0041	2004-01 to 2019-04	2004-01 to 2015-05
HD 114729	5844.0	12.0	4.02	0.02	-5.0046	0.0044	2004-02 to 2021-03	2004-02 to 2014-08
HD 114853	5705.0	14.0	4.32	0.02	-4.9435	0.0037	2004-01 to 2017-08	2004-01 to 2015-05
HD 115617	5559.0	17.0	4.30	0.03	-5.0132	0.0029	2004-01 to 2021-03	2004-01 to 2015-05
HD 117207	5667.0	21.0	4.22	0.04	-5.0881	0.0045	2004-02 to 2021-03	2004-02 to 2015-05
HD 117618	5990.0	13.0	4.18	0.02	-4.9790	0.0040	2004-02 to 2019-08	2004-02 to 2010-05
HD 119173	5779.0	44.0	4.11	0.04	-4.8290	0.0050	2006-02 to 2020-02	2006-02 to 2014-03
HD 11964	5326.0	19.0	3.90	0.04	-5.1763	0.0037	2003-10 to 2018-11	2003-10 to 2015-01
HD 119949	6359.0	36.0	4.10	0.04	-4.9050	0.0046	2004-02 to 2014-01	2004-02 to 2014-01
HD 124292	5443.0	22.0	4.35	0.04	-5.0128	0.0037	2004-02 to 2017-08	2004-02 to 2015-05
HD 125072	4794.0	102.0	4.51	0.24	-4.9238	0.0018	2004-02 to 2017-08	2004-02 to 2015-05
HD 125184	5680.0	30.0	3.99	0.05	-5.0952	0.0039	2004-02 to 2017-08	2004-02 to 2015-05
HD 125595	4636.0	83.9	4.50	0.26	-4.7575	0.0030	2004-05 to 2019-08	2004-05 to 2010-02
HD 125612A	5913.0	17.0	4.23	0.03	-4.8988	0.0083	2004-02 to 2021-03	2004-02 to 2015-03
HD 125881	6036.0	17.0	4.24	0.03	-4.8726	0.0036	2004-02 to 2013-05	2004-02 to 2013-05
HD 126525	5638.0	13.0	4.28	0.02	-4.9978	0.0045	2004-06 to 2019-05	2004-06 to 2015-05
HD 126793	5904.0	33.0	4.23	0.03	-4.8659	0.0048	2004-05 to 2015-02	2004-05 to 2015-02
HD 126803	5470.0	18.0	4.45	0.04	-4.9348	0.0050	2004-02 to 2019-08	2004-02 to 2014-05
HD 129642	4919.0	65.0	4.54	0.16	-5.0077	0.0037	2004-02 to 2017-08	2004-02 to 2015-05
HD 13060	5255.0	45.0	4.40	0.09	-4.8547	0.0051	2003-10 to 2016-09	2003-10 to 2015-01
HD 131653	5324.0	26.0	4.57	0.04	-4.9976	0.0077	2004-05 to 2015-07	2004-05 to 2015-05
HD 131664	5901.0	26.0	4.30	0.03	-4.8432	0.0076	2004-05 to 2021-02	2004-05 to 2014-07

HD 132569	4967.0	44.7	4.64	0.11	-4.6443	0.0043	2004-05 to 2011-08	2004-05 to 2011-08
HD 133633	5571.0	19.0	4.41	0.04	-4.9437	0.0053	2005-05 to 2018-07	2005-05 to 2014-05
HD 134060	5966.0	14.0	4.21	0.03	-4.9983	0.0038	2004-02 to 2017-05	2004-02 to 2015-05
HD 134088	5675.0	22.0	4.35	0.03	-4.8092	0.0041	2004-05 to 2015-06	2004-05 to 2015-04
HD 134606	5633.0	28.0	4.29	0.05	-5.1077	0.0042	2004-07 to 2017-05	2004-07 to 2015-05
HD 134987	5740.0	23.0	4.17	0.04	-5.1106	0.0043	2004-02 to 2016-04	2004-02 to 2015-05
HD 135625	6003.0	14.0	4.08	0.04	-5.0250	0.0087	2004-02 to 2021-03	2004-02 to 2015-05
HD 136352	5664.0	14.0	4.29	0.02	-4.9412	0.0031	2004-05 to 2017-08	2004-05 to 2015-05
HD 136713	4994.0	74.0	4.61	0.14	-4.7748	0.0020	2005-05 to 2016-09	2005-05 to 2015-05
HD 137388	5240.0	53.0	4.48	0.11	-4.8949	0.0040	2005-07 to 2016-08	2005-07 to 2014-06
HD 137676	5253.0	18.0	3.99	0.03	-5.1132	0.0044	2004-02 to 2018-06	2004-02 to 2011-08
HD 13808	5033.0	38.0	4.50	0.08	-4.8921	0.0034	2003-12 to 2016-02	2003-12 to 2015-01
HD 1388	5954.0	10.0	4.19	0.02	-4.9714	0.0038	2003-10 to 2017-09	2003-10 to 2014-12
HD 141624	5871.0	30.0	4.22	0.03	-4.9127	0.0050	2005-05 to 2018-08	2005-05 to 2014-07
HD 142709	4728.0	65.0	4.70	0.18	-4.9360	0.0023	2004-07 to 2016-04	2004-07 to 2015-02
HD 143361	5503.0	36.0	4.32	0.06	-5.1130	0.0072	2007-05 to 2020-03	2007-05 to 2014-09
HD 144628	5022.0	26.0	4.58	0.08	-4.9259	0.0025	2004-07 to 2019-04	2004-07 to 2015-05
HD 145377	6054.0	16.0	4.27	0.03	-4.6128	0.0051	2005-06 to 2011-04	2005-06 to 2011-04
HD 145417	4953.0	48.4	4.98	0.08	-4.8971	0.0032	2004-05 to 2014-08	2004-05 to 2014-08
HD 14745	6290.0	39.0	4.37	0.04	-4.8876	0.0073	2003-11 to 2021-02	2003-11 to 2014-03
HD 148156	6251.0	25.0	4.18	0.05	-4.9315	0.0087	2005-04 to 2019-07	2005-04 to 2013-07
HD 148211	5948.0	22.0	4.15	0.02	-4.9037	0.0050	2004-05 to 2016-04	2004-05 to 2015-03
HD 148303	4829.0	84.0	4.66	0.22	-4.6532	0.0022	2004-07 to 2013-05	2004-07 to 2013-05
HD 149396	5657.0	23.0	4.37	0.03	-4.6836	0.0056	2004-09 to 2016-05	2004-09 to 2013-07
HD 150177	6216.0	28.0	3.86	0.03	-4.8915	0.0041	2004-05 to 2013-07	2004-05 to 2013-07

HD 150433	5665.0	12.0	4.33	0.02	-4.9512	0.0038	2005-08 to 2017-08	2005-08 to 2014-09
HD 15337	5179.0	44.0	4.48	0.09	-4.9226	0.0039	2003-12 to 2019-09	2003-12 to 2015-01
HD 153950	6074.0	15.0	4.13	0.03	-4.9721	0.0082	2005-05 to 2018-03	2005-05 to 2014-07
HD 154088	5374.0	43.0	4.38	0.07	-5.0718	0.0029	2006-04 to 2017-08	2006-04 to 2015-05
HD 154577	4847.0	35.0	4.69	0.07	-4.8699	0.0022	2004-05 to 2017-08	2004-05 to 2015-05
HD 156098	6517.0	44.0	3.76	0.05	-4.7816	0.0045	2005-05 to 2021-03	2005-05 to 2015-05
HD 15612	5256.0	34.0	4.55	0.06	-4.5080	0.0044	2003-11 to 2018-02	2003-11 to 2014-03
HD 156411	5910.0	16.0	3.79	0.01	-5.1400	0.0096	2005-05 to 2019-07	2005-05 to 2013-07
HD 157172	5451.0	27.0	4.37	0.05	-5.0040	0.0037	2005-07 to 2017-03	2005-07 to 2015-05
HD 157347	5676.0	16.0	4.27	0.03	-5.0213	0.0038	2006-03 to 2019-08	2006-03 to 2015-05
HD 157830	5540.0	16.0	4.43	0.02	-4.7868	0.0030	2004-05 to 2012-09	2004-05 to 2012-09
HD 1581	5951.0	13.0	4.28	0.03	-4.9174	0.0031	2003-10 to 2020-12	2003-10 to 2014-12
HD 16008	5770.0	14.0	4.33	0.03	-4.8465	0.0049	2003-10 to 2021-03	2003-10 to 2015-01
HD 161098	5560.0	15.0	4.40	0.02	-4.9240	0.0037	2006-04 to 2017-09	2006-04 to 2015-05
HD 161566	6230.0	24.0	3.88	0.04	-4.9879	0.0077	2005-06 to 2019-07	2005-06 to 2013-03
HD 16160	4796.0	109.0	4.62	0.27	-4.8649	0.0020	2003-10 to 2008-09	2003-10 to 2008-09
HD 16280	4677.0	83.5	4.58	0.31	-4.6712	0.0039	2003-10 to 2021-03	2003-10 to 2014-12
HD 16417	5841.0	17.0	3.99	0.02	-5.1028	0.0043	2003-11 to 2017-09	2003-11 to 2015-01
HD 165131	5870.0	15.0	4.27	0.03	-4.9658	0.0086	2006-05 to 2021-03	2006-05 to 2015-05
HD 166724	5099.0	36.0	4.55	0.09	-4.7253	0.0033	2004-05 to 2011-09	2004-05 to 2011-09
HD 167677	5474.0	20.0	4.40	0.03	-4.9899	0.0074	2005-05 to 2019-08	2005-05 to 2015-05
HD 168863	4782.0	80.0	4.52	0.23	-4.7762	0.0041	2006-07 to 2021-03	2006-07 to 2014-07
HD 170493	4751.0	108.0	4.49	0.25	-4.8776	0.0024	2005-07 to 2015-09	2005-07 to 2015-05
HD 171028	5671.0	16.0	3.73	0.03	-5.1098	0.0071	2004-10 to 2017-09	2004-10 to 2011-03
HD 171587	5412.0	15.0	4.58	0.02	-4.7553	0.0035	2004-11 to 2014-08	2004-11 to 2014-08

HD 172513	5500.0	18.0	4.37	0.03	-4.7730	0.0025	2004-05 to 2007-05	2004-05 to 2007-05
HD 172568	5728.0	22.0	4.45	0.03	-4.9046	0.0051	2005-05 to 2018-08	2005-05 to 2013-10
HD 175607	5392.0	17.0	4.51	0.03	-4.9298	0.0052	2004-07 to 2015-04	2004-07 to 2015-04
HD 176354	5271.0	40.0	3.89	0.07	-5.0505	0.0052	2006-05 to 2019-07	2006-05 to 2011-09
HD 176986	5018.0	59.0	4.60	0.11	-4.8283	0.0028	2004-07 to 2017-08	2004-07 to 2015-04
HD 177565	5627.0	19.0	4.30	0.03	-4.9085	0.0027	2003-10 to 2017-09	2003-10 to 2015-05
HD 17865	5877.0	24.0	4.13	0.03	-4.9272	0.0051	2003-11 to 2014-09	2003-11 to 2014-09
HD 17970	5038.0	31.0	4.53	0.06	-5.0043	0.0037	2003-10 to 2017-09	2003-10 to 2015-01
HD 181433	4962.0	134.0	4.54	0.27	-5.1290	0.0036	2003-10 to 2017-08	2003-10 to 2015-05
HD 181720	5792.0	17.0	4.10	0.02	-4.9792	0.0051	2005-05 to 2017-11	2005-05 to 2014-08
HD 183658	5803.0	17.0	4.24	0.02	-4.9407	0.0043	2005-08 to 2019-05	2005-08 to 2015-05
HD 185283	4746.0	79.3	4.53	0.20	-4.9514	0.0046	2005-07 to 2019-09	2005-07 to 2015-05
HD 189567	5726.0	15.0	4.28	0.01	-4.9120	0.0031	2003-10 to 2017-08	2003-10 to 2015-05
HD 190248	5644.0	30.0	4.12	0.05	-5.1315	0.0034	2003-10 to 2016-10	2003-10 to 2015-05
HD 190647	5639.0	24.0	4.09	0.04	-5.1613	0.0064	2004-09 to 2021-03	2004-09 to 2014-10
HD 190984	6007.0	25.0	3.78	0.03	-5.0626	0.0073	2004-05 to 2013-08	2004-05 to 2013-08
HD 191797	5037.0	48.5	4.61	0.11	-4.3751	0.0037	2004-07 to 2019-08	2004-07 to 2014-07
HD 192310	5087.0	48.0	4.42	0.10	-4.9509	0.0019	2003-11 to 2016-10	2003-11 to 2014-10
HD 19467	5720.0	10.0	4.18	0.01	-5.0079	0.0040	2003-10 to 2017-10	2003-10 to 2015-01
HD 195145	5625.0	20.0	4.32	0.03	-4.9820	0.0081	2006-07 to 2018-06	2006-07 to 2013-07
HD 19641	5806.0	14.0	4.23	0.02	-4.9734	0.0082	2003-11 to 2021-01	2003-11 to 2014-12
HD 197027	5694.0	28.0	4.09	0.04	-5.0001	0.0067	2009-05 to 2019-05	2009-05 to 2014-11
HD 197197	5812.0	16.0	4.04	0.02	-5.0203	0.0061	2003-11 to 2018-10	2003-11 to 2015-04
HD 199288	5765.0	19.0	4.36	0.02	-4.8868	0.0033	2003-11 to 2019-11	2003-11 to 2015-05
HD 199289	5928.0	37.0	4.43	0.03	-4.8466	0.0050	2003-11 to 2018-08	2003-11 to 2014-10

HD 199604	5817.0	22.0	4.18	0.03	-4.9118	0.0057	2003-11 to 2018-10	2003-11 to 2010-11
HD 199847	5763.0	20.0	4.08	0.02	-5.0033	0.0067	2003-11 to 2018-10	2003-11 to 2014-09
HD 20003	5494.0	27.0	4.37	0.05	-4.9891	0.0050	2003-12 to 2017-08	2003-12 to 2015-01
HD 200538	6042.0	18.0	4.13	0.03	-5.0017	0.0085	2004-09 to 2017-04	2004-09 to 2012-09
HD 202206	5757.0	25.0	4.33	0.03	-4.7434	0.0028	2004-05 to 2019-06	2004-05 to 2012-06
HD 202871	6055.0	16.0	4.28	0.04	-4.7825	0.0075	2005-06 to 2012-09	2005-06 to 2012-09
HD 20407	5866.0	14.0	4.32	0.01	-4.8786	0.0034	2003-10 to 2017-11	2003-10 to 2015-02
HD 204313	5776.0	22.0	4.23	0.02	-5.0384	0.0046	2006-05 to 2017-09	2006-05 to 2014-10
HD 204941	4997.0	36.0	4.52	0.10	-4.9275	0.0034	2004-12 to 2016-09	2004-12 to 2014-10
HD 205294	6370.0	32.0	3.92	0.04	-4.8884	0.0072	2003-10 to 2018-09	2003-10 to 2011-08
HD 205536	5442.0	23.0	4.36	0.04	-5.0377	0.0036	2003-10 to 2020-11	2003-10 to 2009-11
HD 206998	5822.0	26.0	4.08	0.03	-4.9271	0.0058	2003-11 to 2018-10	2003-11 to 2014-10
HD 2071	5719.0	14.0	4.35	0.02	-4.9186	0.0036	2003-10 to 2017-10	2003-10 to 2014-11
HD 207129	5937.0	13.0	4.28	0.02	-4.8853	0.0030	2003-10 to 2020-11	2003-10 to 2015-05
HD 20781	5256.0	29.0	4.43	0.05	-5.0498	0.0043	2003-12 to 2017-09	2003-12 to 2015-02
HD 20782	5774.0	14.0	4.22	0.01	-4.9044	0.0036	2003-10 to 2017-10	2003-10 to 2015-01
HD 207832	5718.0	27.0	4.33	0.04	-4.7324	0.0052	2003-10 to 2019-09	2003-10 to 2013-10
HD 207869	5527.0	21.0	4.45	0.05	-4.9486	0.0060	2003-11 to 2017-08	2003-11 to 2013-08
HD 20807	5824.0	15.0	4.29	0.03	-4.8720	0.0029	2003-10 to 2021-02	2003-10 to 2015-02
HD 208487	6146.0	19.0	4.19	0.03	-4.9230	0.0038	2004-09 to 2019-09	2004-09 to 2007-06
HD 20868	4720.0	91.0	4.50	0.22	-4.9930	0.0050	2003-11 to 2019-08	2003-11 to 2008-08
HD 208704	5826.0	11.0	4.21	0.01	-4.9438	0.0037	2004-07 to 2017-10	2004-07 to 2014-11
HD 210918	5755.0	12.0	4.21	0.02	-5.0101	0.0037	2003-10 to 2019-12	2003-10 to 2014-09
HD 211038	4974.0	17.0	3.84	0.05	-5.1615	0.0030	2003-11 to 2017-08	2003-11 to 2014-10
HD 21132	6243.0	34.0	4.27	0.05	-4.8531	0.0058	2004-02 to 2020-01	2004-02 to 2015-01

HD 21209	4671.0	65.0	4.59	0.15	-4.7945	0.0026	2003-10 to 2016-09	2003-10 to 2015-01
HD 215152	4803.0	52.0	4.49	0.15	-4.8599	0.0022	2003-12 to 2016-09	2003-12 to 2014-08
HD 215456	5789.0	15.0	3.95	0.03	-5.1311	0.0049	2003-10 to 2017-05	2003-10 to 2014-10
HD 215497	5003.0	103.0	4.41	0.26	-5.0695	0.0050	2004-10 to 2018-08	2004-10 to 2009-11
HD 216770	5424.0	51.0	4.37	0.07	-4.9645	0.0031	2003-10 to 2019-09	2003-10 to 2013-11
HD 21693	5430.0	26.0	4.36	0.04	-4.9244	0.0035	2003-10 to 2017-09	2003-10 to 2015-01
HD 21749	4640.0	100.0	4.91	0.07	-4.6899	0.0019	2003-11 to 2016-12	2003-11 to 2009-12
HD 218504	5962.0	29.0	4.12	0.03	-4.9327	0.0045	2003-11 to 2014-09	2003-11 to 2014-09
HD 219828	5888.0	14.0	4.01	0.02	-5.1190	0.0064	2005-05 to 2015-08	2005-05 to 2013-11
HD 220339	4938.0	32.0	4.60	0.08	-4.8213	0.0024	2003-10 to 2020-12	2003-10 to 2014-12
HD 22049	5049.0	48.0	4.59	0.09	-4.4937	0.0007	2003-11 to 2020-01	2003-11 to 2007-08
HD 220507	5698.0	17.0	4.17	0.03	-5.0747	0.0049	2003-10 to 2017-10	2003-10 to 2014-12
HD 221580	5322.0	24.0	2.71	0.04	-5.4860	0.0158	2003-11 to 2007-10	2003-11 to 2007-10
HD 222669	5894.0	17.0	4.27	0.02	-4.8376	0.0038	2003-12 to 2016-10	2003-12 to 2014-12
HD 223171	5841.0	18.0	4.03	0.02	-5.0245	0.0042	2003-10 to 2018-11	2003-10 to 2014-12
HD 224685	5504.0	30.0	4.43	0.06	-4.9248	0.0064	2004-12 to 2018-08	2004-12 to 2014-12
HD 224817	5894.0	22.0	4.17	0.02	-4.9232	0.0052	2003-11 to 2015-07	2003-11 to 2014-11
HD 22879	5857.0	27.0	4.28	0.02	-4.8791	0.0035	2003-10 to 2017-10	2003-10 to 2015-01
HD 23249	5035.0	39.0	3.80	0.08	-5.1648	0.0024	2003-10 to 2017-03	2003-10 to 2015-01
HD 24062	6107.0	60.0	4.34	0.06	-5.0123	0.0086	2003-11 to 2021-03	2003-11 to 2013-11
HD 24633	5276.0	43.0	4.41	0.07	-4.9850	0.0071	2003-11 to 2021-03	2003-11 to 2013-12
HD 25171	6160.0	24.0	4.13	0.03	-4.9322	0.0068	2003-11 to 2020-12	2003-11 to 2015-02
HD 26965	5072.0	53.0	4.58	0.19	-4.9495	0.0022	2003-10 to 2021-02	2003-10 to 2015-01
HD 27063	5767.0	14.0	4.30	0.03	-4.7496	0.0029	2003-11 to 2008-02	2003-11 to 2008-02
HD 27894	4952.0	105.0	4.56	0.20	-4.9392	0.0065	2003-10 to 2019-03	2003-10 to 2013-11

HD 28254A	5653.0	33.0	4.05	0.05	-5.1685	0.0094	2003-10 to 2021-03	2003-10 to 2014-09
HD 28471	5745.0	14.0	4.24	0.01	-4.9963	0.0043	2003-11 to 2018-01	2003-11 to 2013-01
HD 297396	4622.0	114.0	4.58	0.36	-4.8059	0.0039	2004-01 to 2021-02	2004-01 to 2015-05
HD 31103	6078.0	16.0	4.23	0.02	-4.6294	0.0055	2003-11 to 2020-11	2003-11 to 2013-11
HD 31128	6096.0	67.0	4.63	0.06	-4.8146	0.0057	2003-11 to 2015-04	2003-11 to 2015-04
HD 31527	5898.0	13.0	4.26	0.02	-4.9395	0.0038	2003-10 to 2017-08	2003-10 to 2015-02
HD 31822	6042.0	16.0	4.32	0.03	-4.8204	0.0035	2003-10 to 2016-03	2003-10 to 2015-02
HD 3220	5846.0	15.0	4.34	0.02	-4.8183	0.0079	2003-11 to 2010-11	2003-11 to 2010-11
HD 32564	5533.0	29.0	4.31	0.06	-5.0305	0.0052	2009-11 to 2016-03	2009-11 to 2015-01
HD 330075	4958.0	52.0	4.41	0.13	-4.9777	0.0041	2004-02 to 2015-04	2004-02 to 2012-02
HD 35854	4928.0	56.0	4.64	0.11	-4.8052	0.0021	2003-10 to 2018-05	2003-10 to 2015-02
HD 36003	4647.0	88.0	4.60	0.21	-4.8609	0.0016	2003-12 to 2016-03	2003-12 to 2015-03
HD 36379	6030.0	14.0	4.05	0.02	-4.9482	0.0040	2003-10 to 2017-04	2003-10 to 2015-03
HD 3823	6022.0	14.0	4.07	0.02	-4.9692	0.0039	2003-10 to 2017-09	2003-10 to 2015-01
HD 38858	5733.0	12.0	4.38	0.01	-4.9051	0.0029	2003-10 to 2019-01	2003-10 to 2015-02
HD 39091	6003.0	17.0	4.18	0.03	-4.9815	0.0037	2003-12 to 2020-12	2003-12 to 2015-01
HD 39194	5205.0	23.0	4.61	0.05	-4.9628	0.0041	2003-11 to 2017-08	2003-11 to 2015-02
HD 3964	5729.0	19.0	4.37	0.04	-4.8328	0.0073	2003-11 to 2020-12	2003-11 to 2014-09
HD 40307	4977.0	59.0	4.63	0.16	-4.9389	0.0021	2003-10 to 2019-04	2003-10 to 2015-02
HD 40397	5527.0	20.0	4.34	0.04	-5.0316	0.0037	2003-10 to 2016-09	2003-10 to 2015-04
HD 40865	5719.0	16.0	4.38	0.03	-4.9144	0.0053	2003-10 to 2017-12	2003-10 to 2014-12
HD 41248	5713.0	21.0	4.37	0.03	-4.8806	0.0053	2003-10 to 2017-12	2003-10 to 2014-01
HD 4308	5644.0	16.0	4.28	0.03	-4.9553	0.0035	2003-10 to 2021-01	2003-10 to 2015-01
HD 43197	5449.0	42.0	4.28	0.08	-5.0857	0.0070	2003-12 to 2021-03	2003-12 to 2013-03
HD 44219	5766.0	18.0	4.06	0.03	-5.0551	0.0083	2003-11 to 2021-03	2003-11 to 2013-03

HD 45184	5869.0	14.0	4.29	0.02	-4.9011	0.0031	2003-10 to 2021-03	2003-10 to 2015-03
HD 45364	5434.0	20.0	4.37	0.03	-4.9753	0.0041	2003-12 to 2017-09	2003-12 to 2015-02
HD 457	6089.0	23.0	4.16	0.03	-5.0230	0.0085	2003-10 to 2020-11	2003-10 to 2014-12
HD 47186	5675.0	21.0	4.25	0.04	-5.0712	0.0043	2003-12 to 2017-09	2003-12 to 2015-04
HD 48115	5825.0	12.0	4.31	0.02	-4.7526	0.0070	2004-01 to 2020-11	2004-01 to 2011-01
HD 48265	5798.0	29.0	3.79	0.14	-5.1912	0.0063	2007-09 to 2021-03	2007-09 to 2015-02
HD 4915	5658.0	13.0	4.42	0.03	-4.7879	0.0025	2003-10 to 2020-12	2003-10 to 2008-09
HD 51608	5358.0	22.0	4.38	0.05	-5.0082	0.0043	2003-12 to 2017-09	2003-12 to 2015-04
HD 52265	6136.0	31.0	4.07	0.03	-4.9915	0.0043	2004-02 to 2021-02	2004-02 to 2010-04
HD 5388	6311.0	33.0	3.89	0.03	-4.9165	0.0082	2003-11 to 2019-07	2003-11 to 2010-01
HD 56274	5734.0	22.0	4.38	0.03	-4.8320	0.0034	2003-11 to 2015-04	2003-11 to 2015-04
HD 59468	5618.0	20.0	4.30	0.03	-5.0095	0.0034	2003-10 to 2018-05	2003-10 to 2015-05
HD 59711A	5722.0	13.0	4.33	0.02	-4.9295	0.0041	2003-10 to 2017-11	2003-10 to 2015-01
HD 60532	6273.0	37.0	3.68	0.04	-5.0976	0.0050	2006-02 to 2011-06	2006-02 to 2011-06
HD 61051	5363.0	27.0	4.38	0.05	-5.0692	0.0079	2004-01 to 2021-01	2004-01 to 2015-03
HD 61383	5716.0	14.0	4.08	0.02	-5.0179	0.0088	2003-12 to 2021-03	2003-12 to 2015-03
HD 61986	5725.0	20.0	4.35	0.04	-4.8898	0.0045	2003-12 to 2017-05	2003-12 to 2010-12
HD 63765	5432.0	19.0	4.41	0.03	-4.7358	0.0026	2003-12 to 2019-04	2003-12 to 2010-04
HD 65277A	4701.0	57.0	4.57	0.16	-5.0334	0.0025	2003-12 to 2017-01	2003-12 to 2015-05
HD 65907A	5945.0	16.0	4.31	0.02	-4.9009	0.0032	2003-10 to 2021-01	2003-10 to 2015-04
HD 68146	6427.0	44.0	4.12	0.04	-4.8606	0.0033	2006-02 to 2014-05	2006-02 to 2014-05
HD 68284	5933.0	26.0	3.87	0.03	-5.0952	0.0062	2003-11 to 2015-01	2003-11 to 2015-01
HD 68607	5215.0	45.0	4.48	0.08	-4.7254	0.0027	2003-12 to 2007-04	2003-12 to 2007-04
HD 68978	5965.0	22.0	4.26	0.02	-4.8548	0.0029	2003-10 to 2016-03	2003-10 to 2015-05
HD 69611	5762.0	25.0	4.17	0.03	-4.9562	0.0045	2003-11 to 2017-04	2003-11 to 2015-01

HD 69830	5402.0	28.0	4.40	0.04	-4.9989	0.0027	2003-10 to 2019-12	2003-10 to 2015-05
HD 70642	5668.0	22.0	4.30	0.04	-5.0217	0.0046	2003-12 to 2021-03	2003-12 to 2015-03
HD 71334	5694.0	13.0	4.26	0.03	-4.9853	0.0045	2003-12 to 2017-02	2003-12 to 2015-01
HD 71835	5438.0	22.0	4.37	0.04	-4.9417	0.0037	2003-12 to 2017-05	2003-12 to 2015-05
HD 7199	5386.0	45.0	4.34	0.08	-4.9891	0.0037	2003-11 to 2017-09	2003-11 to 2015-01
HD 72659	5926.0	12.0	4.03	0.01	-4.9963	0.0086	2004-02 to 2021-03	2004-02 to 2014-06
HD 72673	5243.0	22.0	4.52	0.04	-4.9259	0.0022	2003-12 to 2019-04	2003-12 to 2015-05
HD 73267	5373.0	30.0	4.38	0.05	-5.0939	0.0067	2004-01 to 2021-03	2004-01 to 2015-05
HD 73524	6017.0	13.0	4.19	0.03	-4.9962	0.0040	2003-12 to 2019-04	2003-12 to 2015-05
HD 73583	4597.0	69.0	4.57	0.28	-4.4603	0.0032	2004-02 to 2020-03	2004-02 to 2005-12
HD 7449	6024.0	13.0	4.27	0.03	-4.8206	0.0032	2003-11 to 2016-01	2003-11 to 2015-01
HD 74698	5783.0	19.0	4.12	0.02	-5.0331	0.0090	2004-02 to 2021-03	2004-02 to 2015-05
HD 74957	5915.0	20.0	4.34	0.03	-4.9276	0.0081	2004-01 to 2021-03	2004-01 to 2015-02
HD 76151	5788.0	23.0	4.33	0.02	-4.7266	0.0022	2003-11 to 2021-02	2003-11 to 2012-11
HD 77110	5717.0	20.0	4.36	0.02	-4.9256	0.0050	2004-01 to 2017-05	2004-01 to 2014-02
HD 77338	5440.0	52.0	4.34	0.11	-5.0587	0.0047	2004-12 to 2019-06	2004-12 to 2013-09
HD 78429	5760.0	19.0	4.19	0.02	-4.9518	0.0037	2003-12 to 2017-06	2003-12 to 2015-05
HD 79601	5825.0	25.0	4.15	0.03	-4.9291	0.0047	2004-01 to 2015-02	2004-01 to 2015-02
HD 82342	4470.0	21.0	4.96	0.05	-4.9746	0.0030	2003-12 to 2016-05	2003-12 to 2015-05
HD 82516	5041.0	57.0	4.61	0.12	-4.9816	0.0031	2003-12 to 2017-07	2003-12 to 2015-05
HD 82943	5989.0	20.0	4.20	0.02	-4.9623	0.0034	2004-01 to 2017-06	2004-01 to 2015-05
HD 8389A	5283.0	64.0	4.41	0.12	-5.0311	0.0028	2003-11 to 2020-12	2003-11 to 2013-09
HD 8535	6158.0	13.0	4.12	0.02	-4.9292	0.0081	2003-11 to 2019-08	2003-11 to 2014-10
HD 85390	5135.0	45.0	4.42	0.12	-4.9666	0.0037	2003-12 to 2017-05	2003-12 to 2015-05
HD 85725	5986.0	26.0	3.72	0.04	-5.2147	0.0093	2004-02 to 2021-02	2004-02 to 2015-05

HD 87838	6118.0	33.0	4.19	0.03	-4.8489	0.0045	2004-01 to 2017-07	2004-01 to 2015-05
HD 88218	5878.0	14.0	3.97	0.02	-5.0915	0.0045	2003-12 to 2016-05	2003-12 to 2015-05
HD 8828	5403.0	25.0	4.46	0.03	-5.0145	0.0042	2003-10 to 2017-09	2003-10 to 2015-01
HD 88725	5654.0	17.0	4.39	0.03	-4.8848	0.0039	2004-02 to 2015-06	2004-02 to 2014-12
HD 89454	5728.0	17.0	4.34	0.03	-4.6977	0.0024	2004-01 to 2008-01	2004-01 to 2008-01
HD 89839	6314.0	24.0	4.13	0.06	-4.9294	0.0080	2004-02 to 2021-02	2004-02 to 2015-05
HD 90156	5599.0	12.0	4.40	0.02	-4.9509	0.0034	2004-01 to 2017-07	2004-01 to 2015-05
HD 91889	6140.0	22.0	4.03	0.03	-4.8652	0.0031	2006-02 to 2011-02	2006-02 to 2011-02
HD 92719	5824.0	16.0	4.34	0.03	-4.8249	0.0027	2004-01 to 2016-05	2004-01 to 2015-04
HD 93083	5048.0	66.0	4.46	0.16	-4.9899	0.0031	2004-01 to 2017-07	2004-01 to 2015-04
HD 93385	5977.0	18.0	4.19	0.02	-4.9714	0.0043	2003-12 to 2017-05	2003-12 to 2015-04
HD 94151	5583.0	19.0	4.31	0.02	-4.9724	0.0038	2003-12 to 2019-04	2003-12 to 2013-03
HD 94771	5631.0	21.0	3.94	0.03	-5.2218	0.0090	2004-02 to 2021-03	2004-02 to 2015-05
HD 95456	6276.0	22.0	4.01	0.04	-4.9432	0.0038	2003-12 to 2017-06	2003-12 to 2015-05
HD 95542	5984.0	15.0	4.29	0.03	-4.6653	0.0068	2004-01 to 2020-12	2004-01 to 2015-03
HD 9578	6055.0	14.0	4.26	0.03	-4.5938	0.0055	2003-11 to 2020-11	2003-11 to 2012-11
HD 96423	5711.0	18.0	4.23	0.02	-5.0486	0.0042	2003-12 to 2019-05	2003-12 to 2015-01
HD 96700	5845.0	13.0	4.22	0.02	-4.9378	0.0035	2004-01 to 2017-07	2004-01 to 2015-04
HD 97037	5883.0	14.0	4.15	0.02	-4.9815	0.0039	2004-01 to 2017-02	2004-01 to 2015-01
HD 97343	5410.0	20.0	4.39	0.03	-5.0374	0.0033	2004-01 to 2018-05	2004-01 to 2015-05
HD 98281	5381.0	23.0	4.43	0.04	-4.9154	0.0029	2004-01 to 2016-03	2004-01 to 2015-03

B

Obliquity measurement of AU Mic c

B.1 Table for derived values of all free parameters and their posterior distributions

From Chapter 4, the parameters of the joint fit described in Section 4.4 are listed in Table B.1, which gives the priors adopted over each parameter and the values inferred from the fit. LC-sig, LC-rho, LC-jit, RV-sig, RV-rho, RV-jit are GP hyperparameters in the model. sig is the overall variance, rho is the timescale in the Matérn 5/2 kernel representing how quickly the correlation between points decreases with distance, and jit is the jitter term. The 1- and 2-D posterior distributions are shown in Figure B.1.

Table B.1: Priors and inferred values for all free parameters from the joint analysis (CHEOPS photometry and ESRPESSO RVs) of the first and second transits of AU Mic c.

Parameter	Prior	Value[(a)]	95% credible interval
i	$\mathcal{N}[89.22, 0.21]$	$89.09^{+0.18}_{-0.17}$	88.77 – 89.47
λ	$\mathcal{U}[-180.0, 180.0]$	$67.8^{+31.7}_{-49.0}$	−23.2 – 120.7
$v$sini	$\mathcal{N}[8.5, 0.2]$	$8.49^{+0.19}_{-0.19}$	8.10 – 8.88
$R_\mathrm{p}/R_\mathrm{s}$	$\mathcal{N}[0.0311, 0.0028]$	$0.0320^{+0.0019}_{-0.0019}$	0.0281 – 0.0356
a/R_s	$\mathcal{N}[32.05, 1.0]$	$32.35^{+0.96}_{-0.95}$	30.49 – 34.24
u_1	$\mathcal{N}[0.47, 0.1]$	$0.51^{+0.09}_{-0.09}$	0.33 – 0.70
u_2	$\mathcal{N}[0.30, 0.1]$	$0.32^{+0.09}_{-0.09}$	0.13 – 0.51
$T_{c,1} - 2460133.82615$[(b)] (days)	$\mathcal{U}[-0.02, 0.10]$	$0.0191^{+0.0078}_{-0.0076}$	0.0046 – 0.0343
$T_{c,2} - 2460152.68512$[(b)] (days)	$\mathcal{U}[-0.02, 0.10]$	$0.0341^{+0.0095}_{-0.0183}$	0.0073 – 0.0703
RV1-b0	$\mathcal{U}[-3.0, 3.0]$	$1.5982^{+0.2411}_{-0.2583}$	1.0643 – 2.0453
RV1-b1	$\mathcal{U}[-1.0, 1.0]$	$0.0540^{+0.0312}_{-0.0366}$	−0.0225 – 0.1170
RV1-b2	$\mathcal{U}[-3.0, 3.0]$	$-0.0130^{+0.0040}_{-0.0044}$	−0.0218 – −0.0047
RV2-b0	$\mathcal{U}[-3.0, 3.0]$	$-0.7591^{+0.2436}_{-0.2538}$	−1.2881 – −0.2426
RV2-b1	$\mathcal{U}[-1.0, 1.0]$	$-0.7207^{+0.0204}_{-0.0205}$	−0.7683 – −0.6750
RV2-b2	$\mathcal{U}[-3.0, 3.0]$	$0.0228^{+0.0042}_{-0.0043}$	0.0136 – 0.0315
LC-sig	$\mathcal{U}[0.0, 0.01]$	$0.00033^{+0.00004}_{-0.00004}$	0.00026 – 0.00044
LC-rho	$\mathcal{U}[0.006, 0.042]$	$0.0064^{+0.0008}_{-0.0003}$	0.0060 – 0.0087
LC-jit	$\mathcal{U}[0.0, 0.001]$	$0.0001^{+0.0000}_{-0.0000}$	0.0000 – 0.0001
RV-sig	$\mathcal{U}[0.0, 0.1]$	$0.0050^{+0.0018}_{-0.0013}$	0.0028 – 0.0092
RV-rho	$\mathcal{N}[0.018, 0.01]$	$0.0234^{+0.0071}_{-0.0063}$	0.0119 – 0.0383
RV-jit	$\mathcal{U}[0.0, 0.01]$	$0.0011^{+0.0004}_{-0.0004}$	0.0002 – 0.0018
LC1-b0	$\mathcal{U}[-2.0, 2.0]$	$0.0456^{+0.0167}_{-0.0168}$	0.0125 – 0.0795
LC1-b1	$\mathcal{U}[-2.0, 2.0]$	$-0.0078^{+0.0034}_{-0.0033}$	−0.0145 – −0.0011
LC1-b2	$\mathcal{U}[-2.0, 2.0]$	$0.0649^{+0.0013}_{-0.0013}$	0.0623 – 0.0675
LC1-b3	$\mathcal{U}[-2.0, 2.0]$	$1.0022^{+0.0001}_{-0.0002}$	1.0019 – 1.0025
LC2-b0	$\mathcal{U}[-2.0, 2.0]$	$0.0408^{+0.1210}_{-0.1055}$	−0.1928 – 0.2729
LC2-b1	$\mathcal{U}[-2.0, 2.0]$	$0.1271^{+0.0141}_{-0.0149}$	0.0972 – 0.1560
LC2-b2	$\mathcal{U}[-2.0, 2.0]$	$0.0213^{+0.0024}_{-0.0029}$	0.0158 – 0.0271
LC2-b3	$\mathcal{U}[-2.0, 2.0]$	$1.0010^{+0.0002}_{-0.0002}$	1.0007 – 1.0014

[(a)] Median and 68.3% credible interval

[(b)] Transit times are quoted relative to the published ephemerides of (Wittrock et al., 2023), namely $Tc = 2458342.2240$ and $P_c = 18.858970$ d.

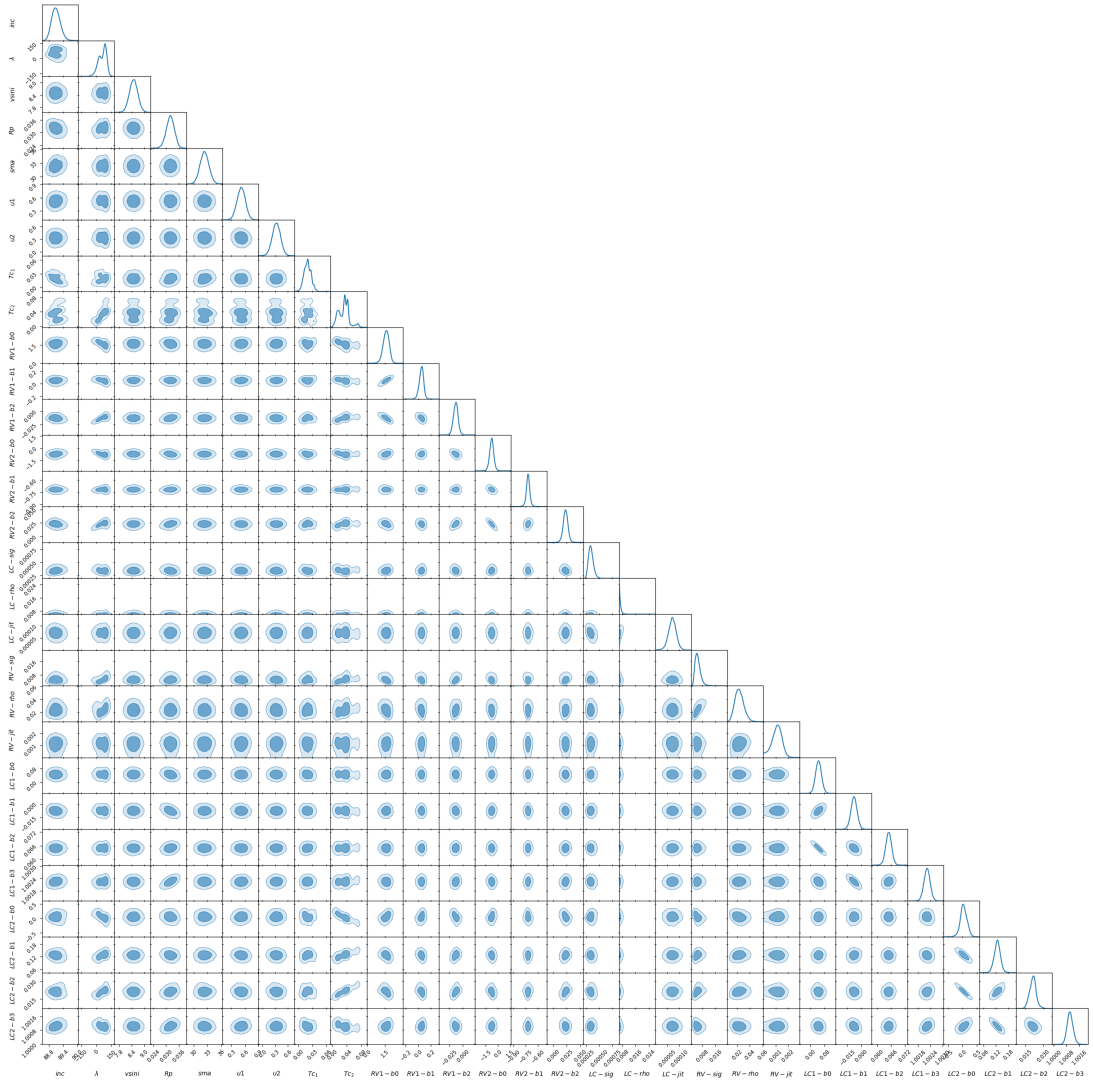


Figure B.1: Posterior distribution of all free parameters in the joint analysis (CHEOPS photometry and ESRPESSO RVs) of the first and second transits of AU Mic c.

B.2 NGTS light curve analysis

From Chapter 4, we performed a stand-alone analysis of the NGTS data collected on the nights of 2023 July 7 and 26. Initially we attempted an agnostic analysis where the transit mid-point times and the planet-to-star radius ratio were allowed to vary freely. However, the NGTS data are significantly impacted by the presence of clouds during both observations. The light curves also display significant short-term variability, which may be due to flares or spots, and ramps at the beginning and the end of the observations, which are most likely due to insufficient comparison stars of a brightness and colour similar to AU Mic. Due to this the results of this agnostic analysis are highly sensitive to the out-of-transit flux baseline model used and the priors chosen. Therefore, in this instance we feel that these NGTS observations are unable to provide independent constraints on the planet parameters or transit timing. However, the photometric precision of the NGTS measurements is high enough that, dependent on the exact transit timing, the data may provide evidence against the transit timing reported by the CHEOPS-ESPRESSO observations. We therefore perform a transit analysis to check for any inconsistency between the NGTS observations and the results of the CHEOPS-ESPRESSO joint analysis.

For this analysis, we compute the transit models using the BATMAN Python package (Kreidberg, 2015). The only planet parameter we allow to vary is the transit mid-point time, $T_{\rm C}$. For this parameter, we use a truncated Gaussian prior with a mean and standard deviation taken from the median and 68.3% credible interval from the CHEOPS-ESPRESSO joint analysis and a strict upper limit set by the 95% credible interval from the joint analysis. The values of these are provided in Table B.1. We fix the remaining planet parameters – orbital period, P, planet-to-star radius ratio, $R_{\rm P}/R_*$, semi-major axis scaled to the stellar radius, a/R_*, orbital inclination, i, eccentricity, e, argument of periastron, ω, and the quadratic limb-darkening coefficients – to the median values from the joint analysis (Table B.1). Simultaneously to varying the transit mid-point time we fit the out-of-transit flux baseline using a quadratic polynomial trend in time.

Table B.2: Results obtained from the analysis of the NGTS observations. The parameter values are the median posterior values and the uncertainties give the 16 and 84 percentile ranges of the posteriors

Name	Symbol	Value	
		2023 July 07	2023 July 26
Transit Mid-Point Time	T_C	$2460133.8485^{+0.0061}_{-0.0090}$	$2460152.7360^{+0.0013}_{-0.0014}$
Constant Detrending Coefficient	c_0	$1.1738^{+0.0043}_{-0.0044}$	1.1968 ± 0.0031
Quadratic Detrending Coefficient	c_1	-0.524 ± 0.012	$-0.5876^{+0.0086}_{-0.0087}$
Quadratic Detrending Coefficient	c_2	$0.3819^{+0.0079}_{-0.0080}$	$0.4278^{+0.0060}_{-0.0058}$

We explored the joint posterior over the fit parameters using Markov Chain Monte Carlo (MCMC) using the EMCEE (Foreman-Mackey et al., 2013) package. Twelve walkers were run for a burn-in phase of 3000 steps followed by a further 15000 steps "production" run. The resulting parameters from this analysis are provided in Table B.2 and the best-fit models are shown in Figure B.2. The resulting transit mid-point times from this analysis are $2460133.8485^{+0.0061}_{-0.0090}$ BJD-TDB and $2460152.7360^{+0.0013}_{-0.0014}$ BJD-TDB for visits one and two respectively. For the night 2023 July 7 the NGTS data are fully consistent with the results of the CHEOPS-ESPRESSO joint analysis. For the night 2023 July 26 the NGTS observations suggest a slightly later transit than the 68.3% credible interval of the CHEOPS-ESPRESSO joint analysis but fully support the presence of the transit occuring during the 95% credible interval.

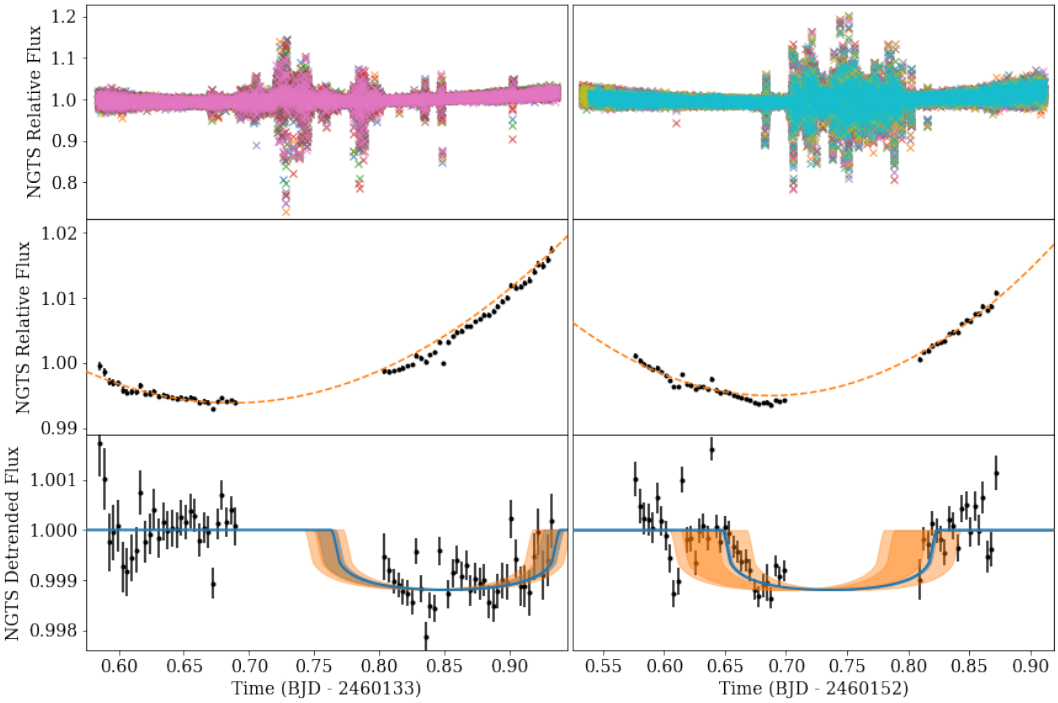


Figure B.2: NGTS photometric observations of the AU Mic c transits. The left and right columns show the first and second visits, respectively. The top row shows the full observations for all telescopes used, with data from each individual telescope plotted using a different colour. The middle row shows the data with the cloud-affected portion removed, and with the data from all telescopes binned together at a timescale of 5 minutes (the binning is performed for visualisation purposes only, the analysis was performed on the unbinned data). The orange dashed line in the middle row shows the best-fit polynomial component to the NGTS data. The bottom row shows the data after subtracting the polynomial component. The blue line shows the best fit model from the NGTS analysis with the blue shaded region showing the 68% credible interval. The dark and light orange shaded regions give the 67.3% and 95% credible intervals respectively from the CHEOPS-ESPRESSO joint analysis.

References

Addison, B. C. et al. (2021). AJ 162.4, 137, p. 137.

Agüeros, M. A. et al. (2018). ApJ 862.1, 33, p. 33.

Ahrer, E. et al. (2021). MNRAS 503.1, pp. 1248–1263.

Aigrain, S., Pont, F., and Zucker, S. (2012). MNRAS 419.4, pp. 3147–3158.

Aigrain, S. and Foreman-Mackey, D. (2023). ARA&A 61, pp. 329–371.

Akeson, R. L. et al. (2013). PASP 125.930, p. 989.

Al Moulla, K. et al. (2023). A&A 669, A39, A39.

Albrecht, S. H., Dawson, R. I., and Winn, J. N. (2022a). PASP 134.1038, 082001, p. 082001.

Albrecht, S. H., Dawson, R. I., and Winn, J. N. (2022b). PASP 134.1038, 082001, p. 082001.

Amazo-Gómez, E. M. et al. (2020). A&A 642, A225, A225.

Ambikasaran, S. et al. (2015). *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38, p. 252.

Angus, R. et al. (2018). MNRAS 474.2, pp. 2094–2108.

Arentoft, T. et al. (2008). ApJ 687.2, pp. 1180–1190.

Artigau, É. et al. (2014). In: *Ground-based and Airborne Instrumentation for Astronomy V*. Ed. by S. K. Ramsay, I. S. McLean, and H. Takami. Vol. 9147. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 914715, p. 914715.

Astropy Collaboration et al. (2013). A&A 558, A33, A33.

Astropy Collaboration et al. (2018). AJ 156.3, 123, p. 123.

Astropy Collaboration et al. (2022). *apj* 935.2, 167, p. 167.

Ballard, S. et al. (2011). ApJ 743.2, 200, p. 200.

Baranne, A. et al. (1996). A&AS 119, pp. 373–390.

Barbary, K. (2016). *Journal of Open Source Software* 1.6, p. 58.

Barrado y Navascués, D. et al. (1999). ApJ 520.2, pp. L123–L126.

Barragán, O. et al. (2018). A&A 612, A95, A95.

Barragán, O. et al. (2022). MNRAS 509.1, pp. 866–883.

Beaulieu, J. .-. et al. (2006). Nature 439.7075, pp. 437–440.

Benedict, G. F. et al. (2002). ApJ 581.2, pp. L115–L118.

Benz, W. et al. (2021). *Experimental Astronomy* 51.1, pp. 109–151.

Bertin, E. and Arnouts, S. (1996). A&AS 117, pp. 393–404.

Bingham, E. et al. (2019). *J. Mach. Learn. Res.* 20, 28:1–28:6.

Boccaletti, A. et al. (2018). A&A 614, A52, A52.

Boccaletti, A. et al. (2015). Nature 526.7572, pp. 230–232.

Boisse, I., Bonfils, X., and Santos, N. C. (2012). A&A 545, A109, A109.

Boisse, I. et al. (2011). A&A 528, A4, A4.

Bolmont, E. and Mathis, S. (2016). *Celestial Mechanics and Dynamical Astronomy* 126.1-3, pp. 275–296.

Boué, G. et al. (2013). A&A 550, A53, A53.

Bourrier, V. et al. (2021a). A&A 654, A152, A152.

Bourrier, V. et al. (2021b). A&A 654, A152, A152.

Brakensiek, J. and Ragozzine, D. (2016). ApJ 821.1, 47, p. 47.

Brandeker, A., Patel, J. A., and Morris, B. M. (2024). *PIPE: Extracting PSF photometry from CHEOPS data*. Astrophysics Source Code Library, record ascl:2404.002.

Brun, A. S. and Browning, M. K. (2017). *Living Reviews in Solar Physics* 14.1, 4, p. 4.

Bryant, E. M. et al. (2020). MNRAS 494.4, pp. 5872–5881.

Burrows, A. et al. (1997). ApJ 491.2, pp. 856–875.

Cale, B. L. et al. (2021). AJ 162.6, 295, p. 295.

Cegla, H. (2019). *Geosciences* 9.3, p. 114.

Cegla, H. M. et al. (2013). ApJ 763.2, 95, p. 95.

Cegla, H. M. et al. (2016). A&A 588, A127, A127.

Cegla, H. M. et al. (2018). ApJ 866.1, 55, p. 55.

Cegla, H. M. et al. (2019). ApJ 879.1, 55, p. 55.

Chaplin, W. J. et al. (2019). AJ 157.4, 163, p. 163.

Chatterjee, S. et al. (2008). ApJ 686.1, pp. 580–602.

Christensen, U. R., Holzwarth, V., and Reiners, A. (2009). Nature 457.7226, pp. 167–169.

Christiansen, J. L. et al. (2017). AJ 154.3, 122, p. 122.

Christiansen, J. L. et al. (2025). *arXiv e-prints*, arXiv:2506.03299, arXiv:2506.03299.

Collier Cameron, A. et al. (2010a). MNRAS 403.1, pp. 151–158.

Collier Cameron, A. et al. (2010b). MNRAS 407.1, pp. 507–514.

Collier Cameron, A. et al. (2019). MNRAS 487.1, pp. 1082–1100.

Collier Cameron, A. et al. (2021). MNRAS 505.2, pp. 1699–1717.

Correia, A. C. M. et al. (2005). A&A 440, pp. 751–758.

Correia, A. C. M. et al. (2010). A&A 511, A21.

Cosentino, R. et al. (2012a). In: *Ground-based and Airborne Instrumentation for Astronomy IV*. Ed. by I. S. McLean, S. K. Ramsay, and H. Takami. Vol. 8446. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 84461V, p. 84461V.

Cosentino, R. et al. (2012b). In: *Ground-based and Airborne Instrumentation for Astronomy IV*. Ed. by I. S. McLean, S. K. Ramsay, and H. Takami. Vol. 8446. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 84461V, p. 84461V.

Cranmer, S. R. and Saar, S. H. (2011). ApJ 741.1, 54, p. 54.

Crass, J. et al. (2021a). *arXiv e-prints*, arXiv:2107.14291, arXiv:2107.14291.

Crass, J. et al. (2021b). *arXiv e-prints*, arXiv:2107.14291, arXiv:2107.14291.

Cretignier, M., Dumusque, X., and Pepe, F. (2022). A&A 659, A68, A68.

Cretignier, M. et al. (2020a). A&A 633, A76, A76.

Cretignier, M. et al. (2020b). A&A 640, A42, A42.

Cretignier, M. et al. (2021a). A&A 653, A43, A43.

Cretignier, M. et al. (2021b). A&A 653, A43, A43.

Cretignier, M. et al. (2023). A&A 678, A2, A2.

Crétignier, M., Pietrow, A. G. M., and Aigrain, S. (2023). MNRAS submitted.

Cui, X.-Q. et al. (2012). *Research in Astronomy and Astrophysics* 12.9, pp. 1197–1242.

Curtis, J. L. et al. (2020). ApJ 904.2, 140, p. 140.

Dalal, S. et al. (2019). A&A 631, A28, A28.

David, T. J. et al. (2022). ApJ 933.1, 114, p. 114.

Davies, M. B. et al. (2014). In: *Protostars and Planets VI*. Ed. by H. Beuther et al., pp. 787–808.

de Beurs, Z. L. et al. (2022). AJ 164.2, 49, p. 49.

Delisle, J. .-. et al. (2022). A&A 659, A182, A182.

Denham, P. et al. (2019). MNRAS 482.3, pp. 4146–4154.

Di Maio, C. et al. (2024). A&A 683, A239, A239.

Donati, J. .-. et al. (1997). MNRAS 291.4, pp. 658–682.

Donati, J. .-. et al. (2014). MNRAS 444.4, pp. 3220–3229.

Donati, J. .-. et al. (2023). MNRAS 525.1, pp. 455–475.

Donati, J. .-. et al. (2025). *arXiv e-prints*, arXiv:2507.01746, arXiv:2507.01746.

Douglas, S. T. et al. (2019). ApJ 879.2, 100, p. 100.

Dumusque, X. (2018). A&A 620, A47, A47.

Dumusque, X., Boisse, I., and Santos, N. C. (2014a). ApJ 796.2, 132, p. 132.

Dumusque, X., Boisse, I., and Santos, N. C. (2014b). ApJ 796.2, 132, p. 132.

Dumusque, X. et al. (2011a). A&A 525, A140, A140.

Dumusque, X. et al. (2011b). A&A 535, A55, A55.

Dumusque, X. et al. (2021a). A&A 648, A103, A103.

Dumusque, X. et al. (2021b). A&A 648, A103, A103.

Dumusque, X. et al. (2015). ApJ 814.2, L21, p. L21.

Esposito, M. et al. (2014). A&A 564, L13, p. L13.

Fabrycky, D. C. and Winn, J. N. (2009a). ApJ 696.2, pp. 1230–1240.

Fabrycky, D. C. and Winn, J. N. (2009b). ApJ 696.2, pp. 1230–1240.

Faria, J. P. et al. (2018). *Journal of Open Source Software* 3.26, p. 487.

Faria, J. P. et al. (2022). A&A 658, A115, A115.

Fischer, D. A. et al. (2016). PASP 128.964, p. 066001.

Foing, B. H. (1988). *Irish Astronomical Journal* 18, p. 257.

Foreman-Mackey, D. et al. (2013). PASP 125.925, p. 306.

Gaia Collaboration et al. (2016). A&A 595, A1, A1.

Gaia Collaboration et al. (2018a). *ArXiv e-prints*.

Gaia Collaboration et al. (2018b). A&A 616, A1, A1.

Gaia Collaboration et al. (2018c). A&A 616, A1, A1.

Galindo-Guil, F. J. et al. (2022). A&A 664, A70, A70.

Gandolfi, D. et al. (2017). AJ 154.3, 123, p. 123.

Gaudi, B. S. et al. (2008). *Science* 319.5865, p. 927.

Gaudi, B. S. and Winn, J. N. (2007). ApJ 655.1, pp. 550–563.

Gibson, S. R. et al. (2016). In: *Ground-based and Airborne Instrumentation for Astronomy VI*. Ed. by C. J. Evans, L. Simard, and H. Takami. Vol. 9908. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 990870, p. 990870.

Gillon, M. et al. (2017). Nature 542.7642, pp. 456–460.

Ginsburg, A. et al. (2019). AJ 157, 98, p. 98.

Giorgini, J. D. et al. (1996). In: *AAS/Division for Planetary Sciences Meeting Abstracts #28*. Vol. 28. AAS/Division for Planetary Sciences Meeting Abstracts, 25.04, p. 25.04.

Goldreich, P., Lithwick, Y., and Sari, R. (2004). *Annu. Rev. Astron. Astrophys.* 42.1, pp. 549–601.

Gomes da Silva, J. et al. (2021). A&A 646, A77, A77.

Gomes da Silva, J. et al. (2018). *The Journal of Open Source Software* 3.31, p. 667.

Gordon, T. A. et al. (2021). ApJ 913.1, 70, p. 70.

Gray, D. F. (2008). *The Observation and Analysis of Stellar Photospheres*.

Gupta, A. and Schlichting, H. E. (2020). MNRAS 493.1, pp. 792–806.

Hall, O. J. et al. (2021). *Nature Astronomy* 5, pp. 707–714.

Hall, R. D. et al. (2018). MNRAS 479.3, pp. 2968–2987.

Handley, W. J., Hobson, M. P., and Lasenby, A. N. (2015a). MNRAS 450, pp. L61–L65.

Handley, W. J., Hobson, M. P., and Lasenby, A. N. (2015b). MNRAS 453.4, pp. 4384–4398.

Hara, N. C. and Delisle, J.-B. (2023). *arXiv e-prints*, arXiv:2304.08489, arXiv:2304.08489.

Hara, N. C. and Ford, E. B. (2023). *Annual Review of Statistics and Its Application* 10.1, pp. 623–649.

Hara, N. C. et al. (2017). MNRAS 464.1, pp. 1220–1246.

Hattori, S. et al. (2024). *exoplanet-dev/jaxoplanet: Astronomical time series analysis with JAX*. Version v0.0.2.

Haywood, R. D. et al. (2014). MNRAS 443.3, pp. 2517–2531.

Haywood, R. D. et al. (2016). MNRAS 457.4, pp. 3637–3651.

Haywood, R. D. et al. (2022). ApJ 935.1, 6, p. 6.

Hebb, L. et al. (2007). MNRAS 379.1, pp. 63–72.

Hébrard, G. et al. (2020). A&A 640, A32, A32.

Hirano, T. et al. (2020). ApJ 899.1, L13, p. L13.

Hoyer, S. et al. (2020). A&A 635, A24, A24.

Hurt, S. A. and MacGregor, M. A. (2023). ApJ 954.1, 10, p. 10.

Isaacson, H. and Fischer, D. (2010). ApJ 725.1, pp. 875–885.

John, A. A., Collier Cameron, A., and Wilson, T. G. (2022). MNRAS 515.3, pp. 3975–3995.

Johnstone, C. P., Bartel, M., and Güdel, M. (2021). A&A 649, A96, A96.

Jones, D. E. et al. (2017). *arXiv e-prints*, arXiv:1711.01318, arXiv:1711.01318.

Jurgenson, C. et al. (2016). In: *Ground-based and Airborne Instrumentation for Astronomy VI*. Ed. by C. J. Evans, L. Simard, and H. Takami. Vol. 9908. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 99086T, 99086T.

Kalas, P., Liu, M. C., and Matthews, B. C. (2004). *Science* 303.5666, pp. 1990–1992.

Kjeldsen, H. and Bedding, T. R. (1995). A&A 293, pp. 87–106.

Klein, B. et al. (2021). MNRAS 500.2, pp. 1844–1850.

Klein, B. et al. (2022). MNRAS 512.4, pp. 5067–5084.

Klein, B. et al. (2024). MNRAS 531.4, pp. 4238–4262.

Koch, D. et al. (1996). *JGR* 101.E4, pp. 9297–9302.

Kochukhov, O., Makaganiuk, V., and Piskunov, N. (2010). A&A 524, A5, A5.

Kozai, Y. (1962). AJ 67, pp. 591–598.

Kreidberg, L. (2015). PASP 127.957, p. 1161.

Kurucz, R. L. (1993). *Physica Scripta Volume T* 47, pp. 110–117.

Lagrange, A. .-. et al. (2009). A&A 493.2, pp. L21–L25.

Lai, D. (2014). MNRAS 440.4, pp. 3532–3544.

Lanzafame, A. C. and Spada, F. (2015). A&A 584, A30, A30.

Laskar, J. (1990). Icarus 88, pp. 266–291.

Laskar, J. (1993). *Physica D Nonlinear Phenomena* 67, pp. 257–281.

Laskar, J. and Robutel, P. (2001). *Celestial Mechanics and Dynamical Astronomy* 80, pp. 39–62.

Liang, Y., Winn, J. N., and Melchior, P. (2024). AJ 167.1, 23, p. 23.

Lidov, M. L. (1962). Planet. Space Sci. 9.10, pp. 719–759.

Lienhard, F. et al. (2023). MNRAS 522.4, pp. 5862–5878.

Lin, D. N. C. and Ida, S. (1997). ApJ 477.2, pp. 781–791.

Lomb, N. R. (1976). Ap&SS 39.2, pp. 447–462.

Lovis, C. (2007). "Recherche de systèmes planétaires aux limites de la spectroscopie DopplerRecherche de systèmes planétaires aux limites de la spectroscopie DopplerSearch for planetary systems at the limits of Doppler spectroscopy;" PhD thesis. University of Geneva, Switzerland.

Malo, L. et al. (2014). ApJ 792.1, 37, p. 37.

Mamajek, E. E. and Bell, C. P. M. (2014). MNRAS 445.3, pp. 2169–2180.

Mantovan, G. et al. (2024). A&A 684, L17, p. L17.

Marois, C. et al. (2008). *Science* 322.5906, p. 1348.

Martioli, E. et al. (2021). A&A 649, A177, A177.

Mayor, M. (1985). In: *Stellar Radial Velocities*. Ed. by A. G. D. Philip and D. W. Latham, pp. 21–34.

Mayor, M. et al. (2003). *The Messenger* 114, pp. 20–24.

Mayor, M. and Queloz, D. (1995). *Nature* 378.6555, pp. 355–359.

McQuillan, A., Aigrain, S., and Mazeh, T. (2013). MNRAS 432.2, pp. 1203–1216.

McQuillan, A., Mazeh, T., and Aigrain, S. (2014). ApJS 211.2, 24, p. 24.

Meibom, S., Mathieu, R. D., and Stassun, K. G. (2009). ApJ 695.1, pp. 679–694.

Messina, S. et al. (2016). A&A 596, A29, A29.

Meunier, N., Desort, M., and Lagrange, A. .-. (2010a). A&A 512, A39, A39.

Meunier, N., Desort, M., and Lagrange, A. .-. (2010b). A&A 512, A39, A39.

Meunier, N. and Lagrange, A. .-. (2019a). A&A 625, L6, p. L6.

Meunier, N. and Lagrange, A. .-. (2019b). A&A 625, L6, p. L6.

Meunier, N., Lagrange, A. .-., and Cuzacq, S. (2019). A&A 632, A81, A81.

Meunier, N. et al. (2015a). A&A 583, A118, A118.

Meunier, N. et al. (2015b). A&A 583, A118, A118.

Miret-Roig, N. et al. (2020). A&A 642, A179, A179.

Mordasini, C. et al. (2012). A&A 547, A111, A111.

Murray, C. D. and Dermott, S. F. (1999). *Solar system dynamics*.

Nardiello, D. et al. (2022). A&A 664, A163, A163.

Narita, N. et al. (2010). PASJ 62.6, pp. L61–L65.

Nascimbeni, V. et al. (2011). A&A 527, A85, A85.

Nicholson, B. A. and Aigrain, S. (2022). MNRAS 515.4, pp. 5251–5266.

Noyes, R. W. et al. (1984a). ApJ 279, pp. 763–777.

Noyes, R. W. et al. (1984b). ApJ 279, pp. 763–777.

O'Sullivan, N. K. and Aigrain, S. (2024). MNRAS 531.4, pp. 4181–4202.

Owen, J. E. (2019). *Annual Review of Earth and Planetary Sciences* 47, pp. 67–90.

Owen, J. E. (2020). MNRAS 498.4, pp. 5030–5040.

Owen, J. E. and Lai, D. (2017). MNRAS 469.3, pp. 2834–2844.

Owen, J. E. and Wu, Y. (2013). ApJ 775.2, 105, p. 105.

Owen, J. E. and Wu, Y. (2017). ApJ 847.1, 29, p. 29.

Palle, E. et al. (2020). A&A 643, A25, A25.

Pepe, F. et al. (2002). A&A 388, pp. 632–638.

Pepe, F. et al. (2014). *Astronomische Nachrichten* 335.1, p. 8.

Pepe, F. et al. (2021a). A&A 645, A96, A96.

Pepe, F. et al. (2021b). A&A 645, A96, A96.

Pepe, F. A. et al. (2010). In: *Ground-based and Airborne Instrumentation for Astronomy III*. Ed. by I. S. McLean, S. K. Ramsay, and H. Takami. Vol. 7735. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 77350F, 77350F.

Petit, P. et al. (2015). A&A 584, A84, A84.

Petrovich, C. et al. (2020). ApJ 902.1, L5, p. L5.

Phan, D., Pradhan, N., and Jankowiak, M. (2019). *arXiv preprint arXiv:1912.11554*.

Phillips, D. F. et al. (2016). In: *Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation II*. Ed. by R. Navarro and J. H. Burge. Vol. 9912. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 99126Z, 99126Z.

Plavchan, P. et al. (2020). Nature 582.7813, pp. 497–500.

Prša, A. et al. (2016). AJ 152.2, 41, p. 41.

Queloz, D. et al. (2000). A&A 359, pp. L13–L17.

Queloz, D. et al. (2001). A&A 379, pp. 279–287.

Ragozzine, D. and Holman, M. J. (2010). *arXiv e-prints*, arXiv:1006.3727, arXiv:1006.3727.

Rajpaul, V., Aigrain, S., and Roberts, S. (2016). MNRAS 456.1, pp. L6–L10.

Rajpaul, V. et al. (2015). MNRAS 452.3, pp. 2269–2291.

Rajpaul, V. M. (2017). *PhD thesis, University of Oxford, United Kingdom*.

Rajpaul, V. M., Aigrain, S., and Buchhave, L. A. (2020). MNRAS 492.3, pp. 3960–3983.

Rasio, F. A. and Ford, E. B. (1996). *Science* 274, pp. 954–956.

Rasmussen, C. E. and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. MIT Press.

Read, M. J., Wyatt, M. C., and Triaud, A. H. M. J. (2017). MNRAS 469.1, pp. 171–192.

Reiners, A., Schüssler, M., and Passegger, V. M. (2014). ApJ 794.2, 144, p. 144.

Reiners, A. et al. (2022). A&A 662, A41, A41.

Réville, V. et al. (2015). ApJ 814.2, 99, p. 99.

Ricker, G. R. et al. (2014). In: *Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave*. Ed. by J. Oschmann Jacobus M. et al. Vol. 9143. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 914320, p. 914320.

Ricker, G. R. et al. (2015). *Journal of Astronomical Telescopes, Instruments, and Systems* 1, 014003, p. 014003.

Roberts, S. et al. (2017). *Phil. Trans. R. Soc. A: Math. Phys. Eng. Sci.* 371, 20110550, p. 20110550.

Robinson, R. D. et al. (2001). ApJ 554.1, pp. 368–382.

Rodono, M. et al. (1986). A&A 165, pp. 135–156.

Scargle, J. D. (1982). ApJ 263, pp. 835–853.

Schlichting, H. E. (2014). ApJ 795.1, L15, p. L15.

Schwab, C. et al. (2018). In: *Ground-based and Airborne Instrumentation for Astronomy VII*. Ed. by C. J. Evans, L. Simard, and H. Takami. Vol. 10702. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1070271, p. 1070271.

Shahaf, S. and Zackay, B. (2023). MNRAS 525.4, pp. 6223–6236.

Skumanich, A. (1972a). ApJ 171, p. 565.

Skumanich, A. (1972b). ApJ 171, p. 565.

Smith, A. M. S. et al. (2020). *Astronomische Nachrichten* 341.3, pp. 273–282.

Sneden, C. A. (1973). PhD thesis. USA: The University of Texas at Austin.

Snodgrass, H. B. and Ulrich, R. K. (1990). ApJ 351, p. 309.

Sousa, S. G. et al. (2007). A&A 469.2, pp. 783–791.

Sousa, S. G. et al. (2015). A&A 577, A67, A67.

Southworth, J. et al. (2009a). MNRAS 396.2, pp. 1023–1031.

Southworth, J. et al. (2009b). MNRAS 399.1, pp. 287–294.

Spada, F. and Lanzafame, A. C. (2020). A&A 636, A76, A76.

Spada, F. et al. (2011). MNRAS 416.1, pp. 447–456.

Spiegel, D. S. and Burrows, A. (2012). ApJ 745.2, 174, p. 174.

Strassmeier, K. G., Ilyin, I., and Steffen, M. (2018). A&A 612, A44, A44.

Suárez Mascareño, A. et al. (2020). A&A 639, A77, A77.

Szabó, G. M. et al. (2021). A&A 654, A159, A159.

Szabó, G. M. et al. (2022). A&A 659, L7, p. L7.

Teng, H.-Y. et al. (2025). AJ 170.1, 51, p. 51.

Teyssandier, J. and Terquem, C. (2014). MNRAS 443.1, pp. 568–583.

Teyssandier, J., Terquem, C., and Papaloizou, J. C. B. (2013). MNRAS 428.1, pp. 658–669.

Thompson, A. P. G. et al. (2017). MNRAS 468.1, pp. L16–L20.

Thompson, S. J. et al. (2016). In: *Ground-based and Airborne Instrumentation for Astronomy VI*. Ed. by C. J. Evans, L. Simard, and H. Takami. Vol. 9908. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 99086F, 99086F.

Torres, C. A. O. et al. (2006). A&A 460.3, pp. 695–708.

Triaud, A. H. M. J. et al. (2010). A&A 524, A25, A25.

Triaud, A. H. M. J. (2018). "The Rossiter-McLaughlin Effect in Exoplanet Research". In: *Handbook of Exoplanets*. Ed. by H. J. Deeg and J. A. Belmonte, 2, p. 2.

van Saders, J. L., Pinsonneault, M. H., and Barbieri, M. (2019). ApJ 872.2, 128, p. 128.

van Saders, J. L. et al. (2016). Nature 529.7585, pp. 181–184.

Vanderburg, A. et al. (2016). ApJ 829.1, L9, p. L9.

Vidotto, A. A. et al. (2014). MNRAS 441.3, pp. 2361–2374.

Virtanen, P. et al. (2020). *Nature Methods* 17, pp. 261–272.

Ward, W. R., Colombo, G., and Franklin, F. A. (1976). Icarus 28.4, pp. 441–452.

Weidenschilling, S. J. and Marzari, F. (1996). Nature 384.6610, pp. 619–621.

Wenger, M. et al. (2000). A&AS 143, pp. 9–22.

Werner, M. W. et al. (2004). ApJS 154.1, pp. 1–9.

Wheatley, P. J. et al. (2018). MNRAS 475.4, pp. 4476–4493.

Winn, J. N. and Fabrycky, D. C. (2015). ARA&A 53, pp. 409–447.

Winn, J. N. et al. (2005). ApJ 631.2, pp. 1215–1226.

Winn, J. N. et al. (2009). ApJ 703.2, pp. L99–L103.

Wittrock, J. M. et al. (2022). AJ 164.1, 27, p. 27.

Wittrock, J. M. et al. (2023). AJ 166.6, 232, p. 232.

Wright, J. T. (2018). "Radial Velocities as an Exoplanet Discovery Method". In: *Handbook of Exoplanets*. Ed. by H. J. Deeg and J. A. Belmonte, 4, p. 4.

Wu, Y. and Murray, N. (2003). ApJ 589.1, pp. 605–614.

Yu, H. et al. (2025). MNRAS 536.3, pp. 2046–2063.

Yu, H. et al. (2024a). MNRAS 535.1, pp. 634–646.

Yu, H. et al. (2024b). MNRAS 528.4, pp. 5511–5527.

Yu, L. et al. (2017). MNRAS 467.2, pp. 1342–1359.

Zeng, L., Sasselov, D. D., and Jacobsen, S. B. (2016). ApJ 819.2, 127, p. 127.

Zhao, G. et al. (2012). *Research in Astronomy and Astrophysics* 12.7, pp. 723–734.

Zhao, L. L. et al. (2022). AJ 163.4, 171, p. 171.

Zhao, Y. et al. (2024). *arXiv e-prints*, arXiv:2405.13247, arXiv:2405.13247.

Zicher, N. et al. (2022a). MNRAS 512.2, pp. 3060–3078.

Zicher, N. et al. (2022b). MNRAS 512.2, pp. 3060–3078.